



Published as: *Int J Parasitol.* 2011 May ; 41(6-10): 609–614.

## Extraordinary *Trypanosoma cruzi* diversity within single mammalian reservoir hosts implies a mechanism of diversifying selection

Martin S. Llewellyn<sup>a,\*</sup>, John B. Rivett-Carnac<sup>b</sup>, Sinead Fitzpatrick<sup>a</sup>, Michael D. Lewis<sup>a</sup>, Matthew Yeo<sup>a</sup>, Michael W. Gaunt<sup>a</sup>, and Michael A. Miles<sup>a</sup>

<sup>a</sup>The London School of Hygiene and Tropical Medicine, Keppel Street, London WC1e 7HT, UK

<sup>b</sup>Lion House, 23 Onslow Road, Richmond, Surrey TW10 6QH, UK

### Abstract

Graphical abstract—■■■



**Research highlights**—► We report extraordinary genetic diversity within single reservoir hosts of *Trypanosoma cruzi*. ► Two hundred and thirty-three biological clones were taken from eight mammals. ► Nine polymorphic microsatellite markers were amplified. ► Forty-nine distinct multilocus genotypes were defined. ► Widespread multiclonality contrasts with the precarious nature of *T. cruzi* vectorial transmission. ► We propose that non-neutral processes could account for the diversity observed.

### Abstract

*Trypanosoma cruzi* is an evolutionarily ancient parasitic protozoan endemic to the Americas. Multiple genetic and phenotypic markers indicate that this parasite is highly diverse, with several divergent and discrete major genotypes reported. Infection multiclonality has been observed among numerous metazoan and unicellular endoparasitic species. However, few studies report the complexity of mixed infections within an individual host in any detail or consider their ecological and biological implications. Here we report extraordinary genetic diversity within single reservoir hosts of *T. cruzi* I using nine polymorphic microsatellite markers across 211 clones from eight mammals from three different sylvatic foci in South America. Forty-nine distinct multilocus

© 2011 Elsevier Ltd.

\*Corresponding author. Tel.: +44 207 927 2319; fax: +44 207 636 8739. martin.llewellyn@lshtm.ac.uk.

This document was posted here by permission of the publisher. At the time of deposit, it included all changes made during peer review, copyediting, and publishing. The U.S. National Library of Medicine is responsible for all links within the document and for incorporating any publisher-supplied amendments or retractions issued subsequently. The published journal article, guaranteed to be such by Elsevier, is available for free, on ScienceDirect.

genotypes were defined, with as many as 10 isolated from the same host. We discuss our data in the light of previous population genetic studies of this and related parasitic protozoa and contrast high levels of diversity within each host with the precarious nature of *T. cruzi* contaminative vectorial transmission. Finally, we propose that non-neutral processes could easily account for the diversity we observe and suggest a functional link with survival in the host.

## Keywords

*Trypanosoma cruzi*; Chagas disease; Microsatellite; Infrapopulation; Multiclonality; Didelphis

## 1 Introduction

*Trypanosoma cruzi*, the etiological agent of Chagas disease, is ancient and endemic in the Americas. The progenitors of this digenic kinetoplastid are thought to date their origins to the breakup of Gondwana and were, therefore, probably parasites of South America's earliest mammalian fauna (Stevens et al., 1999). As such, *T. cruzi* is common in sylvatic cycles of transmission throughout the continent where numerous species of native mammal and haematophagous triatomine vector maintain the parasite.

*Trypanosoma cruzi* is highly genetically diverse. Six major genotypes or 'Discrete Typing Units' (DTUs) are recognised to date, newly dubbed TcI–TcVI (Zingales et al., 2009). TcI, TcIII and TcIV are by far the most common in wild transmission cycles (Yeo et al., 2005; Miles et al., 2009). Among domestic transmission cycles, TcI predominates in northern South America, where it may be associated with less severe human disease, while TcII, TcV and TcVI are more common in the Southern Cone, where Chagasic megasyndromes are more frequent (Miles et al., 2009). Among TcI and TcIII isolates at least, genetic diversity at the sub-species level in *T. cruzi* can be extrapolated to considerable diversity at the sub-lineage level based on microsatellite and sequence markers (Herrera et al., 2007; Spotorno et al., 2008; Llewellyn et al., 2009b; Miles et al., 2009). TcV and TcVI lineages, on the other hand, show minimal diversity across multiple markers and a large geographic area, consistent with a recent hybrid origin and rapid expansion into domestic transmission cycles (Lewis et al., unpublished data). TcII, sparsely isolated, is poorly defined in both genetic and epidemiological terms (Lewis et al., unpublished data), however, a growing collection of isolates from sylvatic cycles across Brazil looks set to change this (Lima et al., personal communication). Among TcI isolated from lowland Bolivia, Venezuela and Brazil, Poisson-distributed allele frequencies point to large, stable effective population sizes (Llewellyn et al., 2009b)

The occurrence of multiclonal infections in a single host is an almost universal expectation given measurable levels of genetic diversity in a species of endoparasite. In *T. cruzi*, hosts and vectors have occasionally been identified with mixed infections of different DTUs (e.g. (Bosseno et al., 1996; Herrera et al., 2005; Yeo et al., 2007; Burgos et al., 2008)). Preliminary data demonstrate that multiple variants of the same DTU could also be present (Macedo et al., 2001; Yeo et al., 2007). Sources of intra-host diversity include superinfection from discrete sources as well as the simultaneous transmission of multiclonal parasite populations. In both cases, clonal multiplicity of infection is likely to be related to both the intensity and efficiency of transmission, as is the case with other vector-borne protozoa, most notably *Plasmodium falciparum* (e.g. Schoepflin et al., 2009). However, genetic exchange between individuals within the host or vector is another potential source of multiclonality. One would expect de novo mutation to account only minimally for intra-host genetic diversity, given an appreciable rate of horizontal transmission. On the other hand, genetic exchange will act rapidly to generate multiple non-identical genotypes (Halkett et

al., 2005), even if the overall level of genetic diversity remains low. A recent review of the evidence for genetic exchange in *Leishmania* described the diversity of clones within an individual host or vector as the possible location of the ‘demographic unit’ in a parasite population (Rougeron et al., 2010). This implies that individual clones in close physical proximity are those that are most likely to exchange genetic material.

A group of parasites of the same species inhabiting a given host has also been termed an ‘infrapopulation’ (Bush et al., 1997). Infrapopulations, including all different lifecycle stages of the parasite, can be grouped to form the ‘component’ population (Bush et al., 1997) – the level at which most micro-parasite population genetic studies operate. However, instead of examining different clones from each host separately, difficulties associated with live parasite isolation and large scale biological cloning means that the ‘individual’ in such studies is often simply DNA extracted from host tissue or primary parasite culture. This extract, therefore, represents a composite of a subset of the clones present in each vector or host (e.g. Nebavi et al., 2006; Koffi et al., 2009; Llewellyn et al., 2009a,b; Morrison et al., 2009; Rougeron et al., 2009). The theoretical consequences of sampling a single clone per host have recently been discussed and novel study designs proposed to improve the estimation of key population statistics and processes (Prugnolle and De Meeus, 2010). Lamentably, such designs are not always possible in the field as sample sizes may be inflated beyond the scope of (and available funding for) the study. The theoretical consequences of analysing the clonal composites, a more commonly encountered problem, warrants further exploration.

Here we present multilocus microsatellite data for 211 biological clones taken from eight TcI reservoir host strains. We demonstrate substantial multilocus genotype (MLG) diversity within these strains, even where multiple ( $\geq 3$ ) alleles at individual loci are not present in the uncloned profile and despite multiple repassages, long-term cryopreservation and different isolation techniques. We compare the resultant MLGs with those from the original, uncloned microsatellite profile with the aim of qualitatively assessing bias associated with using such profiles in wider population studies. We show that the uncloned profile commonly represents the most common MLG in the infrapopulation with lesser, proportionate additions from less frequent genotypes. Finally we evaluate the implications of within-host diversity in TcI in terms of our understanding of parasite transmission, population genetic structure and the selective forces acting on the parasite.

## 2 Materials and methods

### 2.1 Isolation of *T. cruzi* clones

A total of 211 biological clones were derived from eight isolates (Table 1) using a solid phase medium cloning technique as described in Yeo et al. (2007). Log-phase uncloned culture density was measured using a haemocytometer, and  $10^3$ – $10^4$  cells inoculated from cultures that showed little or no clumping. Isolates were selected on the basis of multiclinality inferred from a multiple ( $\geq 3$ ) alleles at individual loci in their uncloned state (Llewellyn et al., 2009b). Four isolates were selected from the population showing most multiple alleles (*VEN*<sub>Silv</sub>) and two from populations showing fewer (*BOL*<sub>North</sub> and *BRAZ*<sub>North-East</sub>) (Llewellyn et al., 2009b). Among these eight, two isolates were selected that showed no multiple alleles (XE5167 and M16) to evaluate the potential for ‘hidden multiclinality’ in apparently diploid uncloned profiles.

### 2.2 PCR amplification of microsatellite loci

Nine previously described microsatellite loci were amplified across all clones (Llewellyn et al., 2009b (Supplementary Table S1)). Markers were selected on the basis of the frequency

of multi-allelic samples at that locus present in the original study. The following reaction cycle was implemented: a denaturation step of 4 min at 95 °C, 30 amplification cycles (95 °C for 20 s, 57 °C for 20 s, 72 °C for 20 s) and a final 20 min elongation step at 72 °C. With a final volume of 10 µl, 1× ThermoPol Reaction Buffer (New England Biolabs (NEB), UK), 4 mM MgCl<sub>2</sub>, 34 µM dNTPs; 0.75 pmols of each primer, 1 unit of *Taq* polymerase (NEB, UK) and 1 ng of genomic DNA were added. Five fluorescent dyes were used to label forward primers – 6-FAM and TET (Proligo, Germany), NED, PET and VIC (Applied Biosystems, UK). Allele sizes were determined using an automated capillary sequencer (AB3730, Applied Biosystems, UK), manually checked for errors and typed “blind” to control for user bias.

### 2.3 Analysis of multilocus microsatellite profiles

Pair-wise genetic distance between all clone corrected MLGs within each infrapopulation was estimated using  $D_{AS}$  as described previously (Llewellyn et al., 2009b) and calculated in MICROSAT (Minch et al., 1995). Multi-allelic loci among aneuploid clones were also accommodated as in Llewellyn et al., 2009b. Individual-level genetic distances were calculated as the mean across 1000 re-sampled datasets and the overall  $D_{AS}$  calculated as the arithmetic mean over all pair-wise distances in the infrapopulation. Among clone corrected diploid MLGs a number of parameters were estimated.  $F_{IS}$ , a measure of the distribution of heterozygosity within and between individuals, was estimated per locus in each group of unique MLGs from each population in FSTAT 2.9.3.2 (Goudet, 1995), as well as over populations over loci using Weir and Cockerman’s estimator  $f$  in the same package. Among Venezuelan MLGs, subdivision between clone corrected infrapopulations was assessed in Arelquin 3.1 (Excoffier et al., 2005) using an Analysis of Molecular Variance (AMOVA) to determine whether a significant proportion of variance could be attributed to grouping MLGs by host origin. Finally, the extent of multilocus linkage disequilibrium was assessed using the Index of Association ( $I_A$ ) in each population.

## 3 Results

### 3.1 Infrapopulation diversity

Forty-nine unique MLGs were identified among 211 clones analysed (Table 1). A broadly proportional relationship existed between the number of multi-allelic loci in the non-cloned profile (Original MAL) and the number of distinct MLGs in the sample of the resultant infrapopulation (G), and all infrapopulations contained at least two distinct clones. Importantly, however, multi-allelic profiles in the uncloned isolate were not a prerequisite to observing diversity in the infrapopulation. Among XE5167 clones, for example, 10 distinct MLGs were observed, whilst no multi-allelic loci were observed in the uncloned sample. Additionally, neither time since isolation nor method of isolation seemed to have an observable effect on the number of distinct MLGs present in the sample. Considerable mean genetic divergence ( $D_{AS} > 0.5$ ) between unique MLGs was observed in some isolates (Table 1) and broadly corresponded to the number of distinct MLGs present.

### 3.2 Infrapopulation clonal composition

Fig. 1 demonstrates the relative abundance of distinct MLGs within each infrapopulation sample. Among each group of MLGs a majority genotype was easily identifiable. Abundant secondary genotypes occurred only in M13 and M18. However, evenly balanced MLG frequencies were not observed within any infrapopulation samples. With the aim of evaluating how multiclonality affects the composite genotype derived from the primary culture, an estimate of the relative contribution of each MLG in each infrapopulation to the original, uncloned microsatellite profile was derived. This value was defined as the proportion of alleles shared ( $P_S$ ) between each MLG and the composite genotype, as a

function of those not shared ( $P_S \times (1 - P_{NS})$ ). A significant linear relationship existed between this estimate and the relative frequency of the corresponding MLG in the infrapopulation sample ( $R_{XY} = 0.367$ ,  $P < 0.012$ ), indicating that the composite genotype was generally most similar to the most abundant MLG in the infrapopulation. Neighbour-Joining (NJ) trees drawn from pair-wise  $D_{AS}$  values (Fig. 2) also revealed that the most frequent MLGs (Supplementary Table S2) within an infrapopulation were generally highly divergent and that infrapopulations commonly represented a mixture of both closely and distantly related MLGs (Fig. 2).

### 3.3 Aneuploidy, fusion and linkage

Multiple aneuploid clones were isolated across those infrapopulations studied (Table 1). Among these were a number of putative hybrids with identifiable parental types in the same infrapopulation (Supplementary Fig. S1). These were examined under the proposed ‘fusion-then-loss’ model for genetic exchange in *T. cruzi* (Gaunt et al., 2003). However, we could not rule out artefactual hybrid-like profiles associated with plate cloning. In practice this meant that microsatellite peak intensity deviation from hybrid copy number expectations led to the exclusion of all ‘hybrids’ as potential cloning artefacts i.e. mixes of two divergent clones (Supplementary Fig. S1). This phenomenon also meant that MLG diversity in infrapopulations containing ‘aneuploid’ clones may have been marginally over-estimated. In Fig. 2 these isolates cluster between the two most common diploid MLGs, as would be expected if they represent a mixture or hybrid between two distinct clones. Importantly, strong linkage disequilibrium among diploid MLGs (Supplementary Table S2) within each population, as well as, crucially, multiple identical genotypes, supported widespread clonality, not frequent recombination ( $VEN_{Silv} I_A = 2.54$ ,  $P < 0.001$ ;  $BOL_{North} I_A = 6.07$ ,  $P < 0.001$ ;  $BRAZ_{North-East} I_A = 3.14$ ,  $P < 0.001$ ).

### 3.4 Heterozygosity and infrapopulation subdivision

Groups of clone corrected diploid MLGs from each population all demonstrated excess homozygosity, as indicated by positive  $F_{IS}$  values over loci (Mean  $F_{IS} \pm SE$ :  $VEN_{Silv} = 0.111 \pm 0.07$ ;  $BOL_{North} = 0.325 \pm 0.14$ ;  $BRAZ_{North-East} = 0.335 \pm 0.15$ ) as well as by bootstrapping over populations (Weir and Cockerman’s unbiased estimator  $f = 0.191$  (95% Confidence Interval (95% C.I.) 0.038–0.333)). An AMOVA across infrapopulations from  $VEN_{Silv}$  did demonstrate significant structuring by individual host ( $F_{ST} = 0.175$ ,  $P < 0.0001$ , 1023 random permutations), and correspondingly a Wahlund effect may be depressing levels of heterozygosity in this population.

## 4 Discussion

We believe this study demonstrates for the first time the extraordinary wealth of parasite genetic diversity that can exist within a single wild reservoir host of *T. cruzi* I. Surprisingly, on the basis of the current dataset, the method of parasite isolation as well as the time since isolation has had little bearing on the number of different MLGs present in each sample. Population statistics, including linkage and heterozygosity indices, when calculated across clones, approximate those observed in the same populations of uncloned isolates (Llewellyn et al., 2009b). The presence of a majority genotype in each sample, and the correlation between the abundance of a genotype in an infrapopulation and its contribution to the uncloned microsatellite profile, to some extent validate the conclusions drawn from population genetic studies that involve uncloned isolates or ‘clonal composites’.

It is reasonable to assume that the parasite diversity represented here comprises barely a fraction of the total potentially present within each individual host. Presumably, continual cloning and genotyping would uncover novel genotypes at a rate of approximately one in

every four (49/211 – Table 1) clones analysed. Serial sampling of individual hosts at different time intervals and from different tissues (e.g. Valadares et al., 2008) would likely reveal further genotypes. Solid phase cloning may be too imprecise to evaluate such diversity. Whilst ‘fusion-recombinant’ artefacts can be identified and excluded, other phenomena, especially differential *in vitro* growth rates between clones, introduce unquantifiable sample error. There are a number of promising new technologies to address these issues, particularly Whole Genome Amplification, which can now be effectively targeted at individual cells (e.g. Kwon and Cox, 2004). However, the problem of isolating individual parasite cells directly from blood and tissue remains.

Culture bias is a ubiquitous issue in all microbial sampling. In *T. cruzi* such bias is aggravated in mammalian samples by low circulating parasitemia. From the correlation between MLG abundance and contribution to the uncloned microsatellite profile, we can at least infer the same *in vitro* growth bias between MLGs in strains grown in liquid culture as those grown on solid phase medium. Thus each uncloned genotype represents most closely the majority genotype in the infrapopulation, not a balanced mix between MLGs, as in models proposed for *Trypanosoma congolense* (Holzmüller et al., 2010). In population genetic terms, this implies that a ‘population’ comprising multiple uncloned isolates approximates a population of clones, each clone taken from a different host. This also supports pioneering work by Finley et al. (1987) who modeled growth of *T. cruzi* clonal mixtures in culture and found that a single clone predominated, although Finley did not predict that divergent strains could also persist at lower frequencies (Finley et al., 1987).

Using simulated samples containing 10 infrapopulations from which 20 clones each were drawn, Prugnolle and De Meeus (2010) suggest that, by comparison, the inclusion of only a single clone per host or vector in a population study (modeled as 20 clones, one each from 20 infrapopulations) can artefactually lower observed heterozygosity (increase  $F_{IS}$ ) and reduce multilocus linkage ( $r_D$  which approximates  $I_A$ ) compared with treating each infrapopulation separately. Our sample sizes are too low to test this assertion empirically, largely because the high frequency of repeated genotypes vastly inflates the sample effort required to generate more than a single unique MLG per sample, an illustration of how difficult it is to test model-derived assertions in the field properly. However, when we grouped multiple clones from different hosts from the same study sites together, both linkage and heterozygosity statistics mirrored those which we observed previously (Llewellyn et al., 2009b).

Reconciling the maintenance of high infrapopulation diversity with what is understood about the transmission dynamics, genetics and biology of *T. cruzi* is a complex task. *De novo* mutation within the host is an unlikely explanation. The presence of highly divergent genotypes within the same host suggests frequent introduction of diversity into each infrapopulation via migration (i.e. transmission), not mutation. Interestingly, however, we cannot rule out that groups of closely related MLGs also observed within each infrapopulation (Fig. 2) may share an ancestor within the same host, or perhaps the same host pedigree, if congenital infection represents an important means of transmission between sylvatic reservoirs, as is the case in some human populations (e.g. Bern et al., 2009).

Genetic recombination, now known to occur in *T. cruzi* *in vitro* (Gaunt et al., 2003), as well as at both species (Machado and Ayala, 2001) and population (Ocaña-Mayorga et al., 2010) levels, does not seem a relevant source of multiclonality in the context of these data. Linkage disequilibrium among diploid MLGs from each population, as well as multiple instances of identical MLGs, suggest a predominantly clonal mode of reproduction. Furthermore, the putative fusion recombinants identified could not be confidently distinguished from cloning artefacts.

Criscione and Blouin (2006) suggest that clonal diversity can be maintained at the infrapopulation level given high rates of parasite transmission and correspondingly minimal MLG extinction. Poisson-distributed allele frequencies in  $VEN_{Silv}$ ,  $BOL_{North}$  and  $BRAZ_{North}$  are also consistent with low rates of clonal extinction (Llewellyn et al., 2009b). Subdivision by host among infrapopulations from  $VEN_{Silv}$ , while statistically significant in this study and consistent with low transmission, receives little support from previous work on *T. cruzi* at the intra-DTU level (Spotorno et al., 2008; Llewellyn et al., 2009b; Miles et al., 2009), where diversification by host is not observed. Instead this result is likely to reflect error associated with small infrapopulation sample sizes. Thus the balance of evidence still favours high sylvatic transmission rates in *T. cruzi*. Paradoxically, the stercorarian route via which *T. cruzi* infects its hosts is highly inefficient (estimated at one infectious contact in 650 triatomine feeds (Rabinovich et al., 1990)). However, the associated signature of serial population reductions and clonal extinctions is absent from our data. High clonal and genetic diversity are instead consistent with en masse transmission of parasite populations from vector to host. Oral transmission, increasingly reported in human populations (Pereira et al., 2009), could play a role, perhaps during opportunistic predation of triatomines. Recent experimental work in *T. brucei* shows that bottlenecks do occur during transmission within the tsetse vector (Oberle et al., 2010). Similar experimental study of the impact of stercorarian transmission on *T. cruzi* clonal diversity is required, including a comparison with the oral route.

To maintain the level of diversity and multiclonality observed in sylvatic TcI, oral transmission would need to occur at considerable frequency in both space and time. Instead, interaction between parasite and host could play a role. Microsatellite diversity in a clonal organism cannot be considered strictly neutral, even it occurs outside protein coding regions. This derives from the observation that microsatellite loci can appear non-neutral through genetic linkage, the so-called 'hitch-hiking' effect. Thus the entire *T. cruzi* genome is the selective unit in clonal populations. It is, therefore, conceivable that high infrapopulation diversity observed in this study at microsatellite loci may be indirectly maintained through selection. Such diversification reflects the heterogeneity of the host environment and the numerous cell types which *T. cruzi* infects, and there is limited evidence that different DTUs can sequester in different tissue types within the same host (e.g. Burgos et al., 2008; Valadares et al., 2008). Host immune pressure could also contribute, although *T. cruzi* may not experience such intense exposure as extracellular trypanosomes such as *Trypanosoma brucei*, where the mechanism for immune avoidance is more clearly defined.

Resolving the interplay between selection and transmission, as well the underlying genetic mechanism accounting for the coexistence of so many unique clones within individual *T. cruzi* reservoir hosts is beyond the scope of this study. Other work suggests that mismatch repair inefficiency may play an important role in generating antigenic variation in *T. cruzi* (Machado et al., 2006). If so, hypervariability at microsatellite loci could represent molecular fallout from such an adaptive strategy. Fortunately, many of the tools are now in place to enable carefully designed field and experimental work to test such assertions. This study provides a qualitative insight into the implications multiclonality has for population genetic studies as well as the first glimpse at the extraordinary levels of diversity that exist among *T. cruzi* clones within their ancient reservoir hosts.

## References

- Bern C. Verastegui M. Gilman R.H. Lafuente C. Galdos-Cardenas G. Calderon M. Pacori J. Del Carmen Abastoflor M. Aparicio H. Brady M.F. Ferrufino L. Angulo N. Marcus S. Sterling C. Maguire J.H. Congenital *Trypanosoma cruzi* transmission in Santa Cruz, Bolivia. *Clin. Infect. Dis.* 2009; 49:1667–1674. [PubMed: 19877966]

- Bossono M.F. Telleria J. Vargas F. Yaksic N. Noireau F. Morin A. Breniere S.F. *Trypanosoma cruzi*: study of the distribution of two widespread clonal genotypes in Bolivian *Triatoma infestans* vectors shows a high frequency of mixed infections. *Exp. Parasitol.* 1996; 83:275–282. [PubMed: 8823244]
- Burgos J.M. Begher S. Silva H.M. Bisio M. Duffy T. Levin M.J. Macedo A.M. Schijman A.G. Molecular identification of *Trypanosoma cruzi* I tropism for central nervous system in Chagas reactivation due to AIDS. *Am. J. Trop. Med. Hyg.* 2008; 78:294–297. [PubMed: 18256432]
- Bush A.O. Lafferty K.D. Lotz J.M. Shostak A.W. Parasitology meets ecology on its own terms: Margolis et al. revisited. *J. Parasitol.* 1997; 83:575–583. [PubMed: 9267395]
- Criscione C.D. Blouin M.S. Minimal selfing, few clones, and no among-host genetic structure in a hermaphroditic parasite with asexual larval propagation. *Evolution*. 2006; 60:553–562. [PubMed: 16637500]
- Excoffier L. Laval G. Schneider S. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol. Bioinform. Online*. 2005; 1:47–50. [PubMed: 19325852]
- Finley R.W. Dvorak J.A. *Trypanosoma cruzi*: analysis of the population dynamics of heterogeneous mixtures. *J. Protozool.* 1987; 34:409–415. [PubMed: 3323478]
- Gaunt M.W. Yeo M. Frame I.A. Stothard J.R. Carrasco H.J. Taylor M.C. Mena S.S. Veazey P. Miles G.A. Acosta N. de Arias A.R. Miles M.A. Mechanism of genetic exchange in American trypanosomes. *Nature*. 2003; 421:936–939. [PubMed: 12606999]
- Goudet J. FSTAT Version 1.2: a computer program to calculate *F*-statistics. *J. Heredity*. 1995; 86:485–486.
- Halkett F. Simon J.-C. Balloux F. Tackling the population genetics of clonal and partially clonal organisms. *Trends Ecol. Evol.* 2005; 20:194–201. [PubMed: 16701368]
- Herrera C. Bargues M.D. Fajardo A. Montilla M. Triana O. Vallejo G.A. Guhl F. Identifying four *Trypanosoma cruzi* I isolate haplotypes from different geographic regions in Colombia. *Infect. Genet. Evol.* 2007; 7:535–539. [PubMed: 17287152]
- Herrera L. D'Andrea P.S. Xavier S.C. Mangia R.H. Fernandes O. Jansen A.M. *Trypanosoma cruzi* infection in wild mammals of the National Park 'Serra da Capivara' and its surroundings (Piauí, Brazil), an area endemic for Chagas disease. *Trans. R. Soc. Trop. Med. Hyg.* 2005; 99:379–388. [PubMed: 15780345]
- Holzmueller P. Herder S. Cuny G. De Meeus T. From clonal to sexual: a step in *T. congolense* evolution? *Trends Parasitol.* 2010; 26:56–60. [PubMed: 20006549]
- Koffi M. De Meeus T. Bucheton B. Solano P. Camara M. Kaba D. Cuny G. Ayala F.J. Jamonneau V. Population genetics of *Trypanosoma brucei* gambiense, the agent of sleeping sickness in Western Africa. *Proc. Natl. Acad. Sci. USA*. 2009; 106:209–214. [PubMed: 19106297]
- Kwon Y.M. Cox M.M. Improved efficacy of whole genome amplification from bacterial cells. *Biotechniques*. 2004; 37(40):42–44.
- Llewellyn M.S. Lewis M.D. Acosta N. Yeo M. Carrasco H.J. Segovia M. Vargas J. Torrico F. Miles M.A. Gaunt M.W. *Trypanosoma cruzi* IIc: phylogenetic and phylogeographic insights from sequence and microsatellite analysis and potential impact on emergent Chagas disease. *PLoS Negl. Trop. Dis.* 2009; 3:e510. [PubMed: 19721699]
- Llewellyn M.S. Miles M.A. Carrasco H.J. Lewis M.D. Yeo M. Vargas J. Torrico F. Diosque P. Valente V. Valente S.A. Gaunt M.W. Genome-scale multilocus microsatellite typing of *Trypanosoma cruzi* discrete typing unit I reveals phylogeographic structure and specific genotypes linked to human infection. *PLoS Pathog.* 2009; 5:e1000410. [PubMed: 19412340]
- Macedo A.M. Pimenta J.R. Aguiar R.S. Melo A.I. Chiari E. Zingales B. Pena S.D. Oliveira R.P. Usefulness of microsatellite typing in population genetic studies of *Trypanosoma cruzi*. *Mem. Inst. Oswaldo Cruz*. 2001; 96:407–413. [PubMed: 11313654]
- Machado C.A. Ayala F.J. Nucleotide sequences provide evidence of genetic exchange among distantly related lineages of *Trypanosoma cruzi*. *Proc. Natl. Acad. Sci. USA*. 2001; 98:7396–7401. [PubMed: 11416213]
- Machado C.R. Augusto-Pinto L. McCulloch R. Teixeira S.M. DNA metabolism and genetic diversity in Trypanosomes. *Mutat. Res.* 2006; 612:40–57. [PubMed: 16040270]



- Miles M.A. Llewellyn M.S. Lewis M.D. Yeo M. Baleela R. Fitzpatrick S. Gaunt M.W. Mauricio I.L. The molecular epidemiology and phylogeography of *Trypanosoma cruzi* and parallel research on *Leishmania*: looking back and to the future. *Parasitology*. 2009; 136:1509–1528. [PubMed: 19691868]
- Minch, E.; Ruíz-Linares, A.; Goldstein, D.; Feldman, M.; Cavalli-Sforza, L. Stanford University Press; Stanford: 1995. MICROSAT – the Microsatellite Distance Program.
- Morrison L.J. Tweedie A. Black A. Pinchbeck G.L. Christley R.M. Schoenefeld A. Hertz-Fowler C. Macleod A. Turner C.M. Tait A. Discovery of mating in the major African livestock pathogen *Trypanosoma congolense*. *PLoS One*. 2009; 4:e5564. [PubMed: 19440370]
- Nebavi F. Ayala F.J. Renaud F. Bertout S. Eholie S. Moussa K. Mallie M. de Meeus T. Clonal population structure and genetic diversity of *Candida albicans* in AIDS patients from Abidjan (Cote d’Ivoire). *Proc. Natl. Acad. Sci. USA*. 2006; 103:3663–3668. [PubMed: 16501044]
- Oberle M. Balmer O. Brun R. Roditi I. Bottlenecks and the maintenance of minor genotypes during the life cycle of *Trypanosoma brucei*. *PLoS Pathog.*. 2010; 6:e1001023. [PubMed: 20686656]
- Ocaña-Mayorga S. Llewellyn M.S. Costales J.A. Miles M.A. Grijalva M.J. Sex, subdivision, and domestic dispersal of *Trypanosoma cruzi* lineage I in southern Ecuador. *PLoS Negl. Trop. Dis.*. 2010; 4(12):e915. [PubMed: 21179502]
- Pereira K.S. Schmidt F.L. Guaraldo A.M. Franco R.M. Dias V.L. Passos L.A. Chagas’ disease as a foodborne illness. *J. Food Prot.*. 2009; 72:441–446. [PubMed: 19350996]
- Prugnolle F. De Meeus T. Apparent high recombination rates in clonal parasitic organisms due to inappropriate sampling design. *Heredity*. 2010; 104:135–140. [PubMed: 19812614]
- Rabinovich J.E. Wisnivesky-Colli C. Solarz N.D. Gurtler R.E. Probability of transmission of Chagas disease by *Triatoma infestans* (Hemiptera: Reduviidae) in an endemic area of Santiago del Estero, Argentina. *Bull. World Health Organ*. 1990; 68:737–746. [PubMed: 2127382]
- Rougeron V. De Meeus T. Hide M. Walecx E. Bermudez H. Arevalo J. Llanos-Cuentas A. Dujardin J.C. De Doncker S. Le Ray D. Ayala F.J. Banuls A.L. Extreme inbreeding in *Leishmania braziliensis*. *Proc. Natl. Acad. Sci. USA*. 2009; 106:10224–10229. [PubMed: 19497885]
- Rougeron, V., De Meeus, T., Kako Ouraga, S., Hide, M., Banuls, A.L., 2010. “Everything you always wanted to know about sex (but were afraid to ask)” in *Leishmania* after two decades of laboratory and field analyses. *PLoS Pathog.* 6.
- Schoepflin S. Valsangiacomo F. Lin E. Kiniboro B. Mueller I. Felger I. Comparison of *Plasmodium falciparum* allelic frequency distribution in different endemic settings by high-resolution genotyping. *Malar J.*. 2009; 8:250. [PubMed: 19878560]
- Spotorno O.A. Córdova L. Solari I.A. Differentiation of *Trypanosoma cruzi* I subgroups through characterization of cytochrome b gene sequences. *Infect. Genet. Evol.*. 2008; 8:898–900. [PubMed: 18793754]
- Stevens J.R. Noyes H.A. Dover G.A. Gibson W.C. The ancient and divergent origins of the human pathogenic trypanosomes, *Trypanosoma brucei* and *T. cruzi*. *Parasitology*. 1999; 118:107–116. [PubMed: 10070668]
- Valadares H.M. Pimenta J.R. de Freitas J.M. Duffy T. Bartholomeu D.C. Oliveira Rde P. Chiari E. Moreira Mda C. Filho G.B. Schijman A.G. Franco G.R. Machado C.R. Pena S.D. Macedo A.M. Genetic profiling of *Trypanosoma cruzi* directly in infected tissues using nested PCR of polymorphic microsatellites. *Int. J. Parasitol.*. 2008; 38:839–850. [PubMed: 18154957]
- Yeo M. Acosta N. Llewellyn M. Sanchez H. Adamson S. Miles G.A. Lopez E. Gonzalez N. Patterson J.S. Gaunt M.W. de Arias A.R. Miles M.A. Origins of Chagas disease: *Didelphis* species are natural hosts of *Trypanosoma cruzi* I and armadillos hosts of *Trypanosoma cruzi* II, including hybrids. *Int. J. Parasitol.*. 2005; 35:225–233. [PubMed: 15710443]
- Yeo M. Lewis M.D. Carrasco H.J. Acosta N. Llewellyn M. da Silva Valente S.A. de Costa Valente V. de Arias A.R. Miles M.A. Resolution of multiclinal infections of *Trypanosoma cruzi* from naturally infected triatomine bugs and from experimentally infected mice by direct plating on a sensitive solid medium. *Int. J. Parasitol.*. 2007; 37:111–120. [PubMed: 17052720]
- Zingales B. Andrade S.G. Briones M.R. Campbell D.A. Chiari E. Fernandes O. Guhl F. Lages-Silva E. Macedo A.M. Machado C.R. Miles M.A. Romanha A.J. Sturm N.R. Tibayrenc M. Schijman A.G.

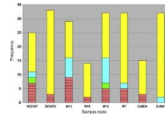
A new consensus for *Trypanosoma cruzi* intraspecific nomenclature: second revision meeting recommends TcI to TcVI. Mem. Inst. Oswaldo Cruz. 2009; 104:1051–1054. [PubMed: 20027478]

## Appendix A Supplementary data

Refer to Web version on PubMed Central for supplementary material.

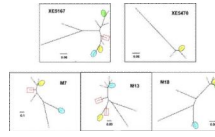
## Acknowledgments

Special thanks to H. Carrasco, M.D. Felciangeli and M. Segovia in Venezuela; J. Vargas in Bolivia; A and V Valente in Brazil. We gratefully acknowledge funding from the De Laszlo Foundation, UK, the Swire Charitable Trust, UK; the Wellcome Trust, UK and The European Union Seventh Framework Program Grant 223034.



**Fig. 1.**

Clonal composition of different intra-host *Trypanosoma cruzi* infrapopulation samples. Composite bars indicate frequency of different multilocus genotypes (MLGs) per isolate. Each division represents a different MLG frequency class. Yellow (light grey) bars indicate dominant/majority MLGs ( $n = 2-30$ ), blue and green (dark grey and black) bars indicate MLGs at lower frequency ( $n = 2-9$ ), and pink bars (medium grey) indicate MLGs represented by a single clone. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 2.**

Un-rooted Neighbour-Joining (NJ) trees based on pair-wise  $D_{AS}$  values between multilocus genotypes (MLGs) from five parasite infrapopulations where  $>4$  MLGs were identified.  $D_{AS}$  values were calculated as in Llewellyn et al., 2009b in MICROSAT and PHYLIP v3.6 (272, 280). Sample codes correspond to those in Table 1. Coloured ellipses represent genotypes at a frequency  $>1$  and correspond to Fig. 1. Dashed boxes indicate multilocus aneuploid clones.

**Table 1**

Samples cloned in this study, multiallelic loci and multiclonality.

Sample code	Host <sup>a</sup>	Date	Locality <sup>b</sup>	Isolation method <sup>c</sup>	Original MAL	<i>n</i>	<i>G</i>	Mean <i>D</i> <sub>AS</sub> ± S.E.	Clone MAL
XE5167	Dm	14.09.99	Pa/Br	XE	0	25	10	0.399 ± 0.035	4
XE5740	Dm	18.6.02	Pa/Br	XE	2	33	4	0.295 ± 0.10	0
M13	Dm	12.6.04	Ba/Ve	B	7	29	10	0.509 ± 0.042	5 <sup>d</sup>
M16	Dm	13.6.04	Ba/Ve	B	0	14	3	0.067 ± 0.017	0
M18	Dm	13.6.04	Ba/Ve	XE	5	32	8	0.523 ± 0.053	1
M7	Dm	14.5.04	Ba/Ve	B	4	32	7	0.591 ± 0.074	2 <sup>d</sup>
SJM34	Dm	7.9.04	Be/Bo	XE	4	14	3	0.583 ± 0.267	1
SJM41	Po	9.9.04	Be/Bo	B	2	32	2	0.056	0
Total:						211	49		

<sup>a</sup>Dm – *Didelphis marsupialis*; Po – *Philander opossum*.

<sup>b</sup>Pa/Br – Para, Brazil (VEN<sub>SilV</sub>); Ba/Ve – Barinas, Venezuela (BRAZ<sub>North-East</sub>); Be/Bo – Beni, Bolivia (BOL<sub>North</sub>). Population names correspond to those in Llewellyn et al. (2009b).

<sup>c</sup>XE – Xenodiagnosis, B – Blood agar culture. Original MAL – Number of multi-allelic ( $\geq 3$ ) loci present in uncloned sample. *n* – Number of clones analysed. *G* – Number of distinct multilocus genotypes (MLGs) present. Mean *D*<sub>AS</sub> – Mean pair-wise genetic distance between distinct MLGs from each host ± S.E. about the mean. Includes aneuploid MLGs calculated as a mean over 1000 diploid resamplings (see Materials and methods). Clone MAL – number of clones containing multi-allelic loci.

<sup>d</sup>Samples that include identical aneuploid genotypes represented more than once in the dataset.