Parasitology

cambridge.org/par

Evaluation of the multispecies coalescent method to explore intra-*Trypanosoma cruzi* I relationships and genetic diversity

César Gómez-Hernández[1,*], Sergio D. Pérez[2,*], Karine Rezende-Oliveira[3], Cecilia G. Barbosa[1], Eliane Lages-Silva[1], Luis Eduardo Ramírez[1] and Juan David Ramírez[2]

[1]Laboratorio de Parasitologia, Universidade Federal do Triângulo Mineiro, Uberaba, Minas Gerais, Brazil; [2]Grupo de Investigaciones Microbiológicas-UR (GIMUR), Programa de Biología, Facultad de Ciencias Naturales y Matemáticas, Universidad del Rosario, Bogotá, Colombia and [3]Laboratorio de Ciências Biomédicas, Universidade Federal de Uberlândia, Campus do Pontal, Ituiutaba, Minas Gerais, Brazil

Research Article

*These authors made equal contributions.

Cite this article: Gómez-Hernández C, Pérez SD, Rezende-Oliveira K, Barbosa CG, Lages-Silva E, Ramírez LE, Ramírez JD (2019). Evaluation of the multispecies coalescent method to explore intra-*Trypanosoma cruzi* I relationships and genetic diversity. *Parasitology* **146**, 1063–1074. https://doi.org/10.1017/S0031182019000428

Received: 14 November 2018
Revised: 20 February 2019
Accepted: 4 March 2019
First published online: 3 May 2019

Key words:
Chagas disease; discrete typing unit; genotype; housekeeping genes; lineage; multispecies coalescent

Author for correspondence:
Juan David Ramírez,
E-mail: juand.ramirez@urosario.edu.co

Abstract

Chagas Disease is a zoonosis caused by the parasite *Trypanosoma cruzi*. Several high-resolution markers have subdivided *T. cruzi* taxon into at least seven lineages or Discrete Typing Units (DTUs) (TcI-TcVI and TcBat). *Trypanosoma cruzi* I is the most diverse and geographically widespread DTU. Recently a TcI genotype related to domestic cycles was proposed and named as $TcI_{DOM.}$ Herein, we combined traditional markers and housekeeping genes and applied a Multispecies Coalescent method to explore intra-TcI relationships, lineage boundaries and genetic diversity in a random set of isolates and DNA sequences retrieved from Genbank from different countries in the Americas. We found further evidence supporting TcI_{DOM} as an independent and emerging genotype of TcI at least in Colombia and Venezuela. We also found evidence of high phylogenetic incongruence between parasite's gene trees (including introgression) and embedded species trees, and a lack of genetic structure among geography and hosts, illustrating the complex dynamics and epidemiology of TcI across the Americas. These findings provide novel insights into *T. cruzi* systematics and epidemiology and support the need to assess parasite diversity and lineage boundaries through hypothesis testing using different approaches to those traditionally employed, including the Bayesian Multispecies coalescent method.

Introduction

Establishing species and lineage boundaries by phylogenetics and population genetics is a central task in evolutionary biology and has become mandatory when studying emerging pathogens and neglected infectious diseases (Yang and Rannala, 2010). How parasite lineages and species are related to each other in space and time can provide valuable insights, such as pathogen adaptation to their hosts, the role of genetic exchange and/or clonality in emergence of novel genotypes, and spatial distribution of genetic variation; these features are often related to morbidity, transmissibility, or drug resistance (De Meeûs *et al.*, 2007).

Chagas disease (CD) is a complex zoonosis caused by the protozoan parasite *Trypanosoma cruzi* and infects nearly 8 million people in Latin America, with another 25 million people currently at risk of acquiring the infection (WHO, 2016). According to World Health Organization, CD remains the largest parasitic disease burden in the Western Hemisphere (WHO, 2016). In addition, natural *T. cruzi* transmission cycles are complex, with notable diversity of triatomine vectors and mammalian hosts of virtually all orders that interact between sylvatic and domestic cycles (Jansen *et al.*, 2017; Justi and Galvão, 2017).

T. cruzi is a successful parasite that displays extraordinary intraspecific genetic diversity, with at least six lineages, or discrete typing units (DTUs TcI–TcVI), currently recognized and distributed throughout the Americas (Zingales *et al.*, 2012), in addition an emergent and well-supported DTU closely related to TcI that is mostly found in bats (TcBat) (Marcili *et al.*, 2009; Lima *et al.*, *2015*; Ramírez *et al.*, 2014). Among these DTUs, TcV and TcVI are considered hybrids of parental groups TcII and TcIII, and are linked to domestic cycles and human infections in southern cone countries (Brisse *et al.*, 2003; Lewis *et al.*, 2011). Moreover, it was proposed TcI and TcII are the natural ancestors of the species and ancient recombination events that produced TcIII and TcIV (Westenberger *et al.*, 2005).

TcI is the most geographically widespread and diverse lineage, with overlapping distributions between domestic and sylvatic cycles (León, 2015; Ramírez and Hernández, 2018). Recently, based on multiloci approaches, a clearly divergent and homogeneous TcI genotype associated with human infections (TcI_{DOM}) was described (Llewellyn *et al.*, 2009*c*; Ramírez *et al.*, 2012; Ramírez *et al.*, 2013). Bayesian skyline plots proposed TcI_{DOM} divergence and expansion ~23–12 KYA, which is consistent with the first human settlements in America (Ramírez *et al.*, 2012). It was also suggested that TcI_{DOM} first made contact with humans in

CAMBRIDGE UNIVERSITY PRESS

Table 1. Cytb sequences used in this study

Reference strain	*Trypanosoma* Isolate/Species	Biological origin	Geographic origin	GenBank accession number
38	TCI	*Triatoma dimidiata* (Insect)	Guatemala (Jutiapa)	JX431260
46	TCI	*Triatoma dimidiata* (Insect)	Guatemala (Santa Rosa)	JX431261
66	TCI	*Triatoma dimidiata* (Insect)	Guatemala (Jalapa)	JX431262
67	TCI	*Triatoma dimidiata* (Insect)	Guatemala (Jutiapa)	JX431263
70	TCI	*Triatoma dimidiata* (Insect)	Guatemala (Jutiapa)	JX431264
71	TCI	*Triatoma dimidiata* (Insect)	Guatemala (Jalapa)	JX431265
83	TCI	*Triatoma dimidiata* (Insect)	Guatemala (Chiquimula)	JX431266
95	TCI	*Triatoma dimidiata* (Insect)	Guatemala (Chiquimula)	JX431267
100	TCI	*Triatoma dimidiata* (Insect)	Guatemala (Santa Rosa)	JX431268
113	TCI	*Triatoma dimidiata* (Insect)	Guatemala (Chiquimula)	JX431269
116	TCI	*Triatoma dimidiata* (Insect)	Guatemala (B/Verapaz)	JX431270
154	TCI	*Triatoma dimidiata* (Insect)	Guatemala (A/Verapaz)	JX431271
ANITAII	TCI	*Triatoma dimidiata* (Insect)	Mexico (Campeche)	JX431272
CAM6	TCI	*Triatoma dimidiata* (Insect)	Mexico (Campeche)	JX431273
CRISTY	TCI	*Triatoma dimidiata* (Insect)	Mexico (SL/Potosí)	JX431274
MICH1	TCI	*Triatoma dimidiata* (Insect)	Mexico (Michoacan)	JX431275
PLI	TCI	*Dipelogaster maxima* (Insect)	Mexico (B/California Sur)	JX431277
QROI	TCI	*Triatoma barberi* (Insect)	Mexico (Queretaro)	JX431278
TQI	TCI	*Triatoma pallidipennis* (Insect)	Mexico (Morelos)	JX431279
10462P2C3	TCI_{DOM}	*Homo sapiens*	Venezuela (Miranda)	JX431281
10462P2C7	TCI_{DOM}	*Homo sapiens*	Venezuela (Miranda)	JX431282
10968P1C1	TCI_{DOM}	*Homo sapiens*	Venezuela (Sucre)	JX431283
ANT3P1C6	TCI	*Homo sapiens*	Venezuela (DC)	JX431284
MG	TCI_{DOM}	*Homo sapiens*	Colombia (Arauca)	HQ713732
DA	TCI_{DOM}	*Homo sapiens*	Colombia (Boyacá)	HQ713730
Xchcl13	TCI_{DOM}	*Homo sapiens*	Colombia (Boyacá)	HQ713719
JEMcl2	TCI_{DOM}	*Homo sapiens*	Colombia (Putumayo)	HQ713720
FECcl15	TCI_{DOM}	*Homo sapiens*	Colombia (Boyacá)	HQ713722
CACQcl14	TCI_{DOM}	*Homo sapiens*	Colombia (Santander)	HQ713724
CGcl16	TCI_{DOM}	*Homo sapiens*	Colombia (Caquetá)	HQ713729
TryCC876	TCI	*Rhodnius pallescens* (Insect)	Panama	FJ555636
TryCC1107	TCI	*Rhodnius stali* (Insect)	Brazil (MS)	FJ555639
TryCC203	TCbat	*Myotis ruber* (Mammal)	Brazil (SP)	FJ002253
TryCC294	TCbat	*Myotis levis* (Mammal)	Brazil (SP)	FJ002254
TryCC597	TCbat	*Myotis nigricans* (Mammal)	Brazil (MS)	FJ002257
TryCC1089	*T.c. marinkellei*	*Artibeus planirostris* (Mammal)	Brazil (MS)	FJ900248
TryCC1093	*T.c. marinkellei*	*Artibeus planirostris* (Mammal)	Brazil (MS)	FJ002262
CGH1	TCI	*Meccus longipennis* (Insect)	Mexico	This work
CGH2	TCI	*Meccus pallidipennis* (Insect)	Mexico	This work
CGH3	TCI	*Meccus longipennis* (Insect)	Mexico	This work
CGH4	TCI	*Meccus pallidipennis* (Insect)	Mexico	This work
CGH6	TCI	*Meccus pallidipennis* (Insect)	Mexico	This work
KR1	TCI	*Triatoma picturata* (Insect)	Mexico	This work
NINOA	TCI	*Homo sapiens*	Mexico	This work
NINOA-CONGENITA	TCI	*Mus musculus* (Mammal)	Mexico	This work
INC-5	TCI	*Homo sapiens*	Mexico	This work

(*Continued*)

https://doi.org/10.1017/S0031182019000428 Published online by Cambridge University Press

Table 1. (*Continued.*)

Reference strain	*Trypanosoma* Isolate/Species	Biological origin	Geographic origin	GenBank accession number
LL051-P23RO	TCI	*Canis familiaris* (Mammal)	Argentina	This work
LL017-PoRO	TCI	*Triatoma infestans* (Insect)	Argentina	This work
TEV55cl1	TCI	*Triatoma infestans* (Insect)	Argentina	This work
LL027-21R1	TCI	*Triatoma infestans* (Insect)	Argentina	This work
PalDa30-Po1RO	TCI	*Didelphis albiventris* (Mammal)	Argentina	This work
PalDa31-Po1RO	TCI	*Didelphis albiventris* (Mammal)	Argentina	This work
Mutum	TCI	*Panstrongylus megistus* (Insect)	Brazil	This work
Alvany	TCI	*Panstrongylus megistus* (Insect)	Brazil	This work
AQ1-7	TCI	*Triatoma sordida* (Insect)	Brazil	This work
1527	TCI	*Homo sapiens*	Brazil	This work
1240	TCI	*Homo sapiens*	Brazil	This work

North–Central America and subsequently become widespread throughout South America (Zumaya-Estrada *et al*., 2012).

Nevertheless, despite intensive efforts to elucidate the remarkable genetic diversity of *T. cruzi* DTUs, there are some gaps related to the understanding of the origin, relationships, and ecological and epidemiological relevance of the DTUs [reviewed in (Messenger *et al*., 2015; Brenière *et al*., 2016)]. There are a few reasons for these knowledge gaps. First, there is scarce information about the parasite's geographical distribution and ecotopes, including the vast variety of sylvatic hosts that remain unsampled, especially at local and microgeographic scales. Second, use of different markers (nuclear and mitochondrial) has led to different genealogical histories and thus to distinct inter- and intra-DTU division proposals, ranging among geography, transmission cycles, or vertebrate/invertebrate hosts [e.g. (Herrera *et al*., 2007, 2009; Llewellyn *et al*., 2009*c*; Ramírez *et al*., 2013)]. In addition, it is widely recognized that the majority of gene trees may be incongruent with the true underlying species tree, real speciation events, and within species population structure (Fujita *et al*., 2012; Leaché *et al*., 2009). Third, despite the current technological advances for collecting genomic datasets that include high-resolution markers and approaches [e.g. nuclear Multilocus Sequence Typing (MLST), mitochondrial MLST, and Multilocus Microsatellite Typing (MLMT)], phylogenetic methods to describe *T. cruzi* diversity and DTU relationships are mainly based on clustering algorithms and/or concatenation [e.g. (Llewellyn *et al*., 2009*a*; Flores-López and Machado, 2011; Yeo *et al*., 2011; Diosque *et al*., 2014)]. Although concatenation is a heuristic strategy that provides phylogenies with high resolution, it assumes that the evolutionary history of each gene tree is identical to the species tree (Ogilvie *et al*., 2017). While it is very true that incongruence is frequently observed between individual loci in *T. cruzi*, most recent publications using MLST, first utilize software (e.g. MLSTest), to detect statistically significant incongruence. Concatenation decisions are then based upon the absence of significant evidence for independent evolutionary histories of particular loci. This fact must be acknowledged, that this is one way to circumvent these limitations. However, novel methodologies should be explored in order to fill the gaps and underpin a better understanding of the evolutionary history of the DTUs and in particular TcI.

To help fill some of the knowledge gaps described above and avoid gene tree/species conflicts, we implemented a multispecies coalescent (MSC) approach to infer intra-TcI relationships and determine lineage boundaries testing the phylogenetic position of TcI_{DOM}, and included previously reported isolates and new isolates from Mexico, Brazil, and Argentina. We conducted the first evaluation (Using MSC) of the current position of TcI_{DOM} as an independent genotype of the TcI complex. Moreover, using a set of different independent loci, including housekeeping genes, we assessed for possible incongruence in gene genealogies and examined intra-lineage diversity through genetic diversity and network analyses.

Materials and methods

Trypanosome culture, sequencing, and molecular data collection

T. cruzi isolates (Tables 1 and 2) were cryopreserved in liquid nitrogen and maintained in LIT medium, pH 7.4, and supplemented with 10% fetal bovine serum (v/v) for DNA preparation. Genomic DNA of cultured trypanosomes was extracted from pellets of approximately 10^6 parasites using the GeneJet kit (Thermo Scientific®), according to the manufacturer's instructions. DNA amplification of the desired region was done in a thermocycler (PTC-100 MJ Research®) using a 50-μL PCR mix as follows: 100 ng of DNA, 1 U of Phusion High-Fidelity DNA Polymerase (Thermo Scientific®) with 5 μL of Phusion HF Buffer, and 1 mM of each dNTP. PCR conditions were subjected to an initial denaturation temperature of 95 °C/10 min followed by 40 cycles (denaturation at 95 °C for 1 min; annealing temperature for 1 min as indicated in Supplementary file 1; elongation at 72 °C for 1 min). Amplicons were precipitated in 70% ethanol and suspended in water. Sequencing was done by the dideoxy-terminal method in an automatic sequencer (AB3730, Applied Biosystems® Genetic Analyzer) by both strands at ACTGene Molecular Analysis (Brazil).

Resulting sequences with expected sizes of each marker (Table 3) were employed to generate multiple sequence alignments using Muscle v.3.8 (Edgar, 2004) with default settings, and manually edited in GeneDoc v.2.6.01 (Nicholas *et al*., 1997). Based on previously reported TcI sequences from GenBank, we built different sets of alignments for phylogenetic and population genetic analyses. GenBank accession numbers, hosts, and geographical origin for the samples are listed in Tables 1 and 2 and include the new sequences derived from this study.

Isolates and sequences included in the study

This study mainly examined TcI complex genetic diversity, relationships based on previously reported sequences in GenBank,

https://doi.org/10.1017/S0031182019000428 Published online by Cambridge University Press

Table 2. SL-IR sequences and isolates used in this study

Reference strain	*Trypanosoma* Isolate/Species	Biological origin	Geographic origin	GenBank accession number
TryCC1089	*T.c. marinkellei*	*Artibeus planirostris* (Mammal)	Brazil (MS)	EU867797
TryCC1093	*T.c. marinkellei*	*Artibeus planirostris* (Mammal)	Brazil (MS)	EU867798
P3	*T. dionisii*	*Pipistrellus pipistrellus* (Mammal)	England	AJ250744
TryCC454	*T. dionisii*	*Desmodus rotundus* (Mammal)	Brazil (MS)	EU867796
MG	TCI (Ia) TCI_{DOM}	*Homo sapiens*	Colombia (Arauca)	EU626722
DA	TCI (Ia) TCI_{DOM}	*Homo sapiens*	Colombia (Boyacá)	EU626723
NV	TCI (Ia) TCI_{DOM}	*Homo sapiens*	Colombia (Tolima)	EU626724
JV	TCI (Ib) TCI_{DOM}	*Homo sapiens*	Colombia (Cesar)	EU626725
SEV	TCI (Ia) TCI_{DOM}	*Homo sapiens*	Colombia (Boyacá)	EU626727
DYR	TCI (Ia) TCI_{DOM}	*Homo sapiens*	Colombia (Boyacá)	EU626728
D12	TCI (Id)	*Didelphis marsupialis* (Mammal)	Colombia (Tolima)	EU626729
H10	TCI (Ib)	*Canis familiaris* (Mammal)	Colombia (Boyacá)	EU626730
Gal61	TCI (Id)	*Rattus rattus* (Mammal)	Colombia (Sucre)	EU626731
Rp523	TCI (Id)	*Rhodnius prolixus* (Insect)	Colombia (Casanare)	EU626732
N5P14	TCI (Id)	*Rhodnius prolixus* (Insect)	Colombia (Casanare)	EU626733
Necocli	TCI (Ib)	*Rhodnius pallescens* (Insect)	Colombia (Antioquia)	EU626735
Coy11	TCI (Id)	*Rhodnius colombiensis* (Insect)	Colombia (Tolima)	EU626736
Mg9	TCI (Ib) TCI_{DOM}	*Triatoma dimidiata* (Insect)	Colombia (Antioquia)	EU626738
Tev91cl5	TCI (Ia)	*Triatoma infestans* (Insect)	Argentina (Chaco)	FJ713402
Cas16	TCI (Ib)	*Rhodnius prolixus* (Insect)	Colombia (Casanare)	FJ713362
Flop2	TCI (Ia)	*Didelphis virginiana* (Mammal)	USA (Florida)	GU179077
V81	TCI (Id)	*Triatoma infestans* (Insect)	Paraguay (Gran Chaco)	GQ398818
V195	TCI (Id)	*Triatoma infestans* (Insect)	Paraguay (Eastern)	GQ398820
PalDa	TCI (Id)	*Didelphis albiventris* (Mammal)	Argentina (Chaco)	GQ398811
PALV2-2cl5	TCI (Ie)	*Triatoma infestans* (Insect)	Argentina (Chaco)	GQ398812
PALV1Cl1	TCI (Ie)	*Triatoma infestans* (Insect)	Argentina (Chaco)	FJ713383
TALAVERDE	TCI (Ie)	*Triatoma infestans* (Insect)	Argentina (La Rioja)	GQ398816
SpGuayacan	TCI (Ie)	*Mepraia spinolai*(Insect)	Chile (IV Region)	GU903141
Sp130cl6	TCI (Ie)	*Mepraia gajardoi* (Insect)	Chile (Arica)	GU903139
Col108	TCI (Ie)	*Mepraia spinolai* (Insect)	Chile (Met. Region)	GU903125
LGNcl7	TCI (Ie)	*Homo sapiens*	Chile (Met. Region)	GU903130
AS	TCI (Ie)	*Homo sapiens*	Bolivia	FJ713356
13379cl7	TCI (Ie)	*Homo sapiens*	Bolivia	GU903124
CGC	TCI	*Homo sapiens*	Colombia (Caquetá)	AM259467
SN6C	TCI	*Rhodnius prolixus* (Insect)	Colombia (Magdalena)	AM259471
PALC	TCI	*Rhodnius prolixus* (Insect)	Colombia (Casanare)	AM259473
EFC	TCI_{DOM}	*Triatoma dimidiata* (Insect)	Colombia (Boyacá)	AM259474
CGH1	TCI	*Meccus longipennis* (Insect)	Mexico	This work
CGH2	TCI	*Meccus pallidipennis* (Insect)	Mexico	This work
CGH3	TCI	*Meccus longipennis* (Insect)	Mexico	This work
CGH4	TCI	*Meccus pallidipennis* (Insect)	Mexico	This work
CGH6	TCI	*Meccus pallidipennis* (Insect)	Mexico	This work
KR1	TCI	*Triatoma picturata* (Insect)	Mexico	This work
NINOA	TCI	*Homo sapiens*	Mexico	This work
NINOA-CONGENITA	TCI	*Mus musculus* (Mammal)	Mexico	This work
INC-5	TCI	*Homo sapiens*	Mexico	This work

(*Continued*)

https://doi.org/10.1017/S0031182019000428 Published online by Cambridge University Press

Table 2. (*Continued.*)

Reference strain	*Trypanosoma* Isolate/Species	Biological origin	Geographic origin	GenBank accession number
LL051-P23RO	TCI	*Canis familiaris* (Mammal)	Argentina	This work
LL017-PoRO	TCI	*Triatoma infestans* (Insect)	Argentina	This work
TEV55cl1	TCI	*Triatoma infestans* (Insect)	Argentina	This work
LL027-21R1	TCI	*Triatoma infestans* (Insect)	Argentina	This work
PalDa30-Po1RO	TCI	*Didelphis albiventris* (Mammal)	Argentina	This work
PalDa31-Po1RO	TCI	*Didelphis albiventris* (Mammal)	Argentina	This work
Mutum	TCI	*Panstrongylus megistus* (Insect)	Brazil	This work
Alvany	TCI	*Panstrongylus megistus* (Insect)	Brazil	This work
AQ1-7	TCI	*Triatoma sordida* (Insect)	Brazil	This work
1527	TCI	*Homo sapiens*	Brazil	This work
1240	TCI	*Homo sapiens*	Brazil	This work

Table 3. Molecular data obtained in this study

Locus	Number of sequences	Length (bp)	Variable sites	Parsimony informative sites	DNA substitution model[a]
Cytb	20	610	138	111	*121343*
SL-IR	20	262	33	19	*121134*
TcSC5D	15	734	5	5	*123141*
TcMK	15	615	25	24	*121323*
NTR	15	696	63	47	*111111*
Total	–	2917	264	206	–

[a]Nucleotide substitutions models were selected under a Bayesian framework in bModelTest (see (47) for details).

and by applying recent advances in Bayesian phylogenetic species tree estimation (MSC), we also examined genetic variation patterns in 20 new isolates from Mexico [9], Brazil [5], and Argentina [6] (Tables 1 and 2). We sequenced three nuclear regions [*TcSC5D*, a putative lathosterol/episterol oxidase (Cosentino and Agüero, 2012); *TcMK*, mevalonate kinase (Cosentino and Agüero, 2012); and *SL-IR*, spliced leader intergenic region of the miniexon gene (Burgos *et al.*, 2007)] and two mitochondrial genes [*Cytb*, cytochrome b (Messenger *et al.*, 2012); *NTR*, Nitroreductase (Hall *et al.*, 2012)]. We selected these housekeeping genes because of their lack of availability in parasite databases and their importance in the maintenance of parasite cellular functions.

Because of the lack of availability of TcMK, TcSC5D, and NTR genes in GenBank, only parasite isolates from our work were considered when evaluating these genes, except for those from Brazil. MS, SP: Brazilian States, Mato Grosso do Sul and São Paulo, respectively.

Parasite genetic diversity and network analysis

We phased TcMK, NTR, Cytb, TcSC5D, and SL-IR loci in DNAsp v.5.10.01 (Librado and Rozas, 2009) with a threshold of 0.9 and 300 iterations, although several exploratory analyses yielded identical results. General sequence diversity statistics for each locus were calculated in DNAsp v.5.10.01, including number of haplotypes, h; haplotype diversity, Hd; number of polymorphic sites, S; and nucleotide diversity, π.

To infer possible associations within and between TcI isolates from our data and those previously reported in GenBank (Tables 1 and 2; Fig. 1), resulting haplotypes of Cytb and SL-IR were employed to construct haplotype networks by the median-joining network method in popART v.1.7 (Leigh and Bryant, 2015) and were selected if ε values ranged from 0.5 to 1. For both analyses, sites with gaps were excluded, including the microsatellite motif in SL-IR, as previously suggested (Tomasini *et al.*, 2011).

Intra-lineage relationships: gene and species tree estimation

The main goal of coalescent-based methods is to identify independently evolving lineages (Fujita *et al.*, 2012). Unlike concatenated phylogenies, which assume that gene trees match for all loci in species trees (a process often called reciprocal monophyly), the MSC approach uses multilocus sequence data in a Bayesian framework and accounts for gene tree discordance by modeling coalescent stochasticity in considered populations (Yang and Rannala, 2010). Consequently, evolutionary lineages can be identified and species trees can be precisely estimated, even in the absence of monophyly or when the phylogenetic signal present in the loci is weak, especially because of recent divergence processes (Degnan and Rosenberg, 2009).

Thus, we estimated gene and species tree topologies using a Bayesian MSC approach using StarBEAST2 (Ogilvie *et al.*, 2017) and implemented in BEAST v.2.4.7 (Bouckaert *et al.*, 2014). This multilocus method co-estimates the gene trees embedded in a shared species tree and does not require that each gene alignment/sample has the same number of sequences, and only requires that each sequence in each gene alignment is mapped to the appropriate species (Heled and Drummond, 2009; Ogilvie *et al.*, 2017). StarBEAST2 enables faster species tree inference and more accurate estimates of substitution rates compared with previous versions of StarBEAST and likelihood-based

https://doi.org/10.1017/S0031182019000428 Published online by Cambridge University Press

Fig. 1. Geographical origin of TcI isolates employed in this study across the endemic distribution of Chagas disease in the Americas. TcI, TcI_{DOM}, and TcBat are indicated by color. The coordinates of each isolates were used to build a georeferenced map of isolates location. The map was built on ArcGIS10.3 using Esri Colombia PublicadorSIG layer (http://www.arcgis.com/home/item.html?id=b051fbef7fba406fbb8e62b90925f365#overview).

methods that include concatenation (Ogilvie *et al.*, 2017). Phylogenetic comparative methods as MSC that incorporate intraspecific variability are relatively new and, so far, not especially widely used in empirical studies. MSC is also helpful in depicting and explaining the genetic diversity signals and useful to infer intraspecific relationships.

For the analysis, TcI sequences of each gene alignment were previously mapped as TcI, TcBat, or TcI_{DOM} based on genotyping assignment and earlier studies (Tables 1 and 2). *Trypanosoma cruzi marinkellei* and *T. dionisii* were used as outgroups. Models of DNA evolution were determined using bModelTest (Bouckaert and Drummond, 2017) (Table 3) using the

https://doi.org/10.1017/S0031182019000428 Published online by Cambridge University Press

Table 4. Genetic diversity measures delimited by country

	n	*S*	*H*	Hd (s.d.)	π (s.d.)
Cytb					
Argentina	6	8	5	0.80 (0.09)	0.004 (0.00098)
Brazil	5	64	6	0.9 (0.062)	0.035 (0.011)
Mexico	9	79	10	0.85 (0.077)	0.034 (0.011)
SL-IR					
Argentina	6	10	6	0.88 (0.06)	0.016 (0.0017)
Brazil	5	25	8	0.93 (0.077)	0.026 (0.007)
Mexico	9	3	4	0.7 (0.084)	0.0037 (0.00071)
TcSC5D					
Argentina	6	5	8	0.894 (0.078)	0.003 (0.00038)
Brazil	–	–	–	–	–
Mexico	9	0	1	–	–
TcMK					
Argentina	6	9	7	0.924 (0.0022)	0.005 (0.00067)
Brazil	–	–	–	–	–
Mexico	9	17	12	0.935 (0.041)	0.011 (0.0011)
NTR					
Argentina	6	39	12	1 (0.034)	0.017 (0.003)
Brazil	–	–	–	–	–
Mexico	9	27	14	0.97 (0.03)	0.0097 (0.0017)

n: number of sequences analyzed; *S*: number of polymorphic sites; *h*: number of haplotypes; Hd: haplotype diversity; π: nucleotide diversity; s.d.: standard deviation.

transition/transversion split setting. We used BModelAnalyser (Bouckaert *et al.*, 2014) to visualize bModelTest log output, after discarding 20% as burn-in. Because divergence time estimates were not goals of this study, we used an uninformative strict clock prior (1/×) for each gene to simplify the model and help the analysis converge. Additionally, we ran two independent MCMC runs of 2.2×10^7 generations each, with a sample frequency of 1000 and using a Yule tree prior with constant population sizes. Convergence of the chains were checked in Tracer v.1.6 (Rambaut *et al.*, 2018) (effective sample sizes >200). The two replicated analyses were combined in LogCombiner v.2.2 (Bouckaert *et al.*, 2014) after discarding the first 10% of trees as burn-in. Final trees were summarized in a maximum clade credibility tree using TreeAnnotator v.2.4.7 (Bouckaert *et al.*, 2014) and edited in iTol v.3 (Letunic and Bork, 2016).

Results

Molecular data collection and parasite isolates

In this study, we sequenced 20 parasitic isolates for five molecular markers, all of which were previously genotyped as a TcI DTU (Gómez-Hernández *et al.*, 2011; Lauthier *et al.*, 2012) (Tables 1 and 2). The resulting sequences were deposited in Genbank under the accession numbers (MH549646-MH549700). The five loci employed varied in length from 262 to 734 bp, and represented a total of 2917 bp, and included 264 and 206 variable and parsimony-informative sites, respectively (Table 3). It is worth noting that it was not possible to amplify Brazilian isolates for genes other than SL-IR and Cytb because some loci were recalcitrant to PCR amplification, possibly because of mutations in some regions of the primers (Supplementary file 1). Moreover, we found evidence of an indel (12-bp long) in the NTR gene from the Mexican INC-5 isolate (Supplementary file 2).

Parasite genetic diversity and network analysis

Overall, based on our dataset, we found evidence of high genetic diversity among TcI strains isolated from Mexico, Brazil, and Argentina (Table 4), but relatively low π estimates in contrast with high Hd indicates a certain amount of similarity between considered strains. Interestingly, Mexican isolates sampled for TcSC5D showed a unique haplotype and no evidence of polymorphic sites (Table 4). Despite the results obtained, genetic diversity estimates were not delimited by country or geography because each gene yielded different estimates [Cytb: Argentina < Brazil < Mexico; SL-IR: Mexico < Argentina < Brazil; TcSC5D: Mexico < Argentina; TcMK: Argentina < Mexico; NTR: Argentina < Mexico].

To infer associations between shared Cytb and SL-IR alleles and haplotypes, we included a representative set of sequences from North–Central America, Northern South America, and Southern South America that belong to the endemic CD distribution (Tables 1 and 2; Fig. 1). Cytb and SL-IR haplotype networks showed no evidence of genetic structure based on geography or host (Fig. 2A and B). Thus, our Cytb haplotypes shared alleles between Mexico, Guatemala, Panama, Colombia, Venezuela, Brazil, and Argentina (H5; Supplementary file 3) and other related haplotypes with short mutational steps (Fig. 2A). However, other haplotypes from our Mexican isolates (*KR1*, *CHG1*, *CGH2*, *CGH3*, and *CGH4*) and Brazilian isolates AQ1–*7* were strongly isolated and separated by several mutational steps (Fig. 2A). Moreover, Cytb had quite homogeneous and divergent haplotypes between TcI_{DOM} and TcBat (Fig. 2A). Similarly, the SL-IR haplotype network (Fig. 2B) shared mixed haplotypes between Mexico, Colombia, the United States, and Argentina, and grouped closed to Paraguayan haplotypes (*H8–H9*; Supplementary file 4; Bolivia and Argentina had mixed haplotypes (*H34*; Table 2; Supplementary file 4).

https://doi.org/10.1017/S0031182019000428 Published online by Cambridge University Press

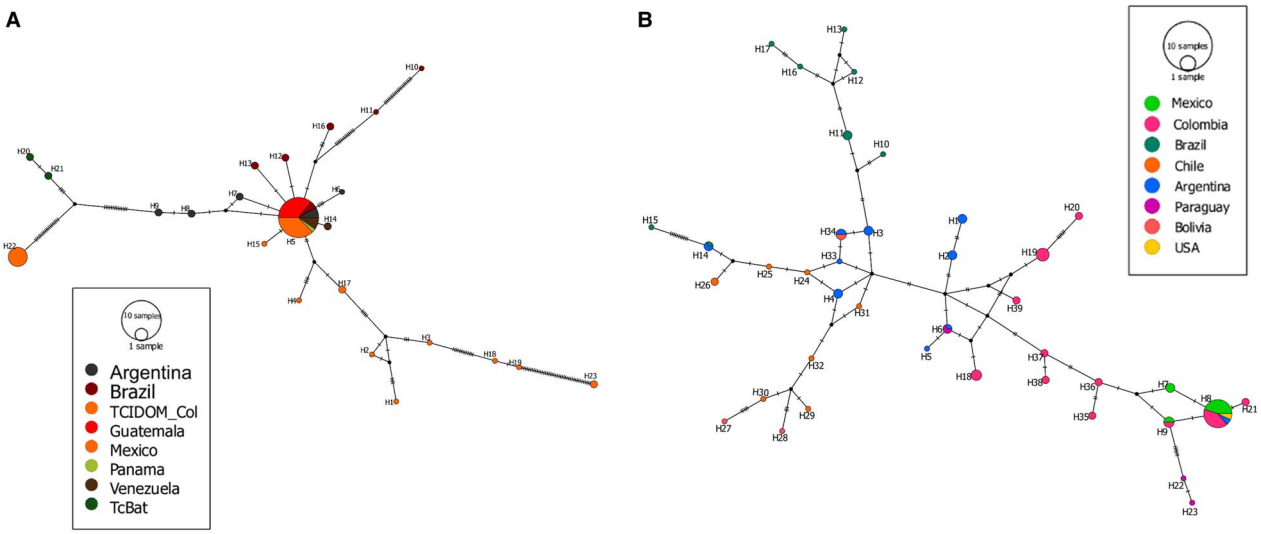

Fig. 2. haplotypes inferred by the median-joining network method in popART v.1.7 (A). Cytb network; (B). SL-IR network. Size of the circles corresponds to the frequency of isolates per haplotype, and length of the vertical lines that connect the networks represents the number of mutations. Black dots indicate the median vectors, which include inferred ancestral nodes.

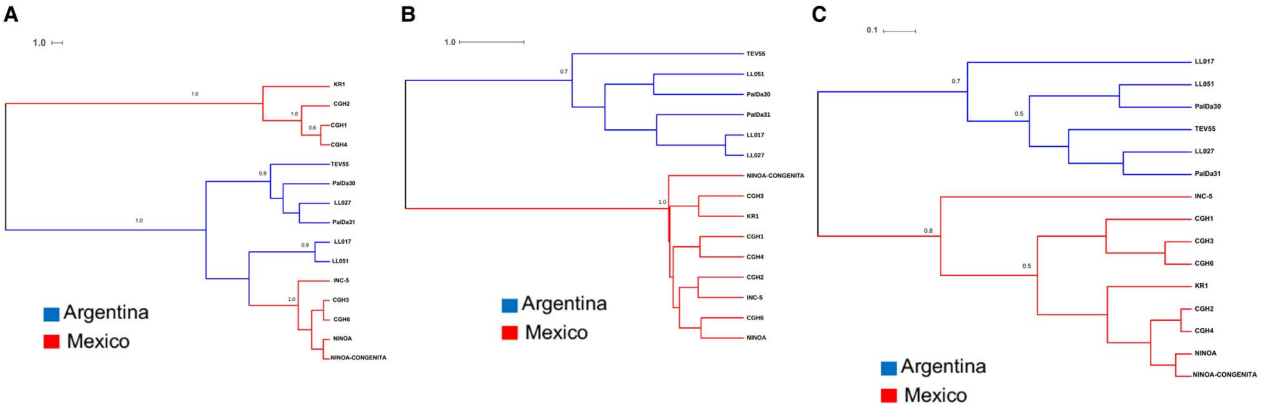

Fig. 3. Maximum Clade Credibility (MCC) Tree of TcI isolates inferred by StarBeast2. (A). Tree constructed using the TcMK (Mevalonate Kinase) sequences; (B). Tree constructed using the TcSC5D (putative lathosterol/episterol oxidase) sequences; (C). Tree constructed using the NTR (Nitroreductase) sequences.

Gene trees and species tree discordance: phylogenetic position of TcI$_{DOM}$

Our results support high phylogenetic incongruence between gene trees and embedded species tree topologies (Figs 3–6). Additionally, we found evidence of incongruence between our parasite isolates and gene trees (Figs 3–5). Consequently, the TcMK gene tree (Fig. 3A) showed also no strict association with either geography or host, and isolates were grouped into two clades: the first monophyletic clade included isolates *KR1*, *CGH2*, *CGH1*, and *CGH4* from Mexico, and the second clade included Argentinean and Mexican isolates (*CGH3*, *NINOA*, and *NINOA CONGENITA*) (Fig. 3A). In contrast, TcSC5D and NTR gene trees showed two clades that were clearly separated by country, but with different tree topologies and branch lengths (Fig 3B and C).

Based on both the SL-IR and Cytb gene trees (Tables 1 and 2), TcI$_{DOM}$ from Colombia and Venezuela were not monophyletic (Figs 4 and 5). However, Cytb gene clearly differentiated TcI$_{DOM}$ Colombia from other TcI isolates, whereas TcI$_{DOM}$ Venezuela clustered with Mexican and Central American isolates (Fig. 4). Something similar was observed for SL-IR and TcI$_{DOM}$ Colombia (Fig. 5). Regardless, the gene trees did not reject TcI$_{DOM}$ as a discrete and emergent TcI genotype. The species tree recovered strong posterior probability values for TcI and TcI$_{DOM}$ (Fig. 6).

Discussion

A comprehensive understanding of *T. cruzi* epidemiology across its geographical range based on genetic diversity and phylogeny is essential for elucidating the parasite's evolution and natural history, which in turn provides important insights for further diagnosis, treatment, prevention and control efforts. Our results here support the emergence of TcI$_{DOM}$ (Ramírez *et al.*, 2012; Ramírez *et al.*, 2013) as an independent and discrete genotype of the TcI complex at least in Colombia and Venezuela (Figs 2–6), and indicate that TcI natural history is indeed complex and dynamic across geography and transmission cycles (Figs 3–5). This result was also recently suggested by network analysis at the country level (Gómez-Palacio *et al.*, 2016). Thus, the high phylogenetic incongruence among gene trees strongly indicates that gene flow occur through ecotopes (Figs 3–5). Consequently, different biological processes such as Intra Lineage Sorting (ILS) could be shaping the current parasite genetic structure, and this process is probably more widespread than previously thought (Ramírez and Llewellyn, 2014).

https://doi.org/10.1017/S0031182019000428 Published online by Cambridge University Press

Fig. 4. Cytb Maximum Clade Credibility (MCC) Tree of TcI isolates inferred by StarBeast2. Numeric values correspond to posterior probability ranging from 0 to 1.

Recently, Zumaya-Estrada *et al.* proposed that TcI_{DOM} originated in North/Central America before moving southwards and may be as ancient as humans in South America based on microsatellite data and concatenated clustering analyses (Zumaya-Estrada *et al.*, 2012); these results are also consistent with those of Ramírez *et al.* (2012). However, methods that assume reciprocal monophyly, such as concatenation, may be biased by subjectivity (Hey, 2012); such methods are also bad estimators of real branch lengths in tree topologies and worse estimators of divergence times (Ogilvie *et al.*, 2017). Our method herein employed for the first time in *T. cruzi,* supports the independence of TcI_{DOM}; SL-IR and Cytb gene tree topologies illustrate that TcI_{DOM} from Venezuela and Colombia are divergent (Figs 4 and 5), which could be accounted for by ILS and introgression. Indeed, our results indicate that, although TcI_{DOM} is an emergent TcI genotype. Thus, a parsimonious explanation regarding these patterns in TcI_{DOM} could be 'divergence with gene flow,' which has been demonstrated for a vast array of taxa [reviewed in (Shapiro *et al.*, 2016)]. In addition, human activity in the Anthropocene continuously changed the landscape, and potentially facilitated the dispersion of several parasitic strains among sylvatic and domestic cycles [e.g. (Lima *et al.*, 2014; Poveda *et al.*, 2017)]. However, one limitation of our study was the low number of loci and samples studied. Future studies should incorporate more loci and samples covering all the CD endemic range.

Moreover, the fact that TcI_{DOM} is an emergent and discrete genotype of TcI adapted to domestic cycles with its own divergent evolutionary history (Fig. 6), implies different biological properties in this genotype. Recently, experimental infections in murine models (ICR-CD1/NIH and Balb-c mice) have demonstrated important differences in terms of parasitemia and tissue tropism as well as histopathological damage between Colombian TcI_{DOM} strains and those sylvatic ones (TcI), in which TcI_{DOM} seems to be less virulent (Cruz *et al.*, 2015; León *et al.*, 2016). These experimental findings have also corroborated by evidence of histotropism between TcI_{DOM} and sylvatic-like TcI strains in human patients with cardiomyopathy (Burgos *et al.*, 2010), supporting the current genetic subdivision.

Consequently, despite international consensus regarding *Trypanosoma cruzi* lineages or DTUs (TcI–VI) (Zingales *et al.*, 2012), and recent designation of TcBat as the seventh DTU (Marcili *et al.*, 2009; Lima *et al.*, 2015), we did not think there was an optimal consensus method for identifying the differences and relationships within and between these major lineages for the described reasons above [reviewed in (Brenière *et al.*, 2016)]. Additionally, division proposals are frequently contradictory [e.g. (Barnabé *et al.*, 2016; Brisse *et al.*, 1998; de Freitas *et al.*, 2006; Flores-Lopez and Machado, 2011; Lewis *et al.*, 2011; Tomasini and Diosque, 2015)]. This issue is even more contentious at the intra-TcI level, where classifications have fluctuated among hosts and transmission cycles based on concatenated methods and single gene trees (Herrera *et al.*, 2007, 2009; Cura *et al.*, 2010; Ramírez *et al.*, 2011) until geography using multilocus data Llewellyn *et al.*, 2009*b*). Until recently, based on nuclear and mitochondrial concatenated MLST, classification of TcI has reduced to TcI_{DOM} and sylvatic isolates, which indicates TcI_{DOM} is an independently evolving lineage in the TcI complex (Ramírez *et al.*, 2012; Ramírez *et al.*, 2013).

https://doi.org/10.1017/S0031182019000428 Published online by Cambridge University Press

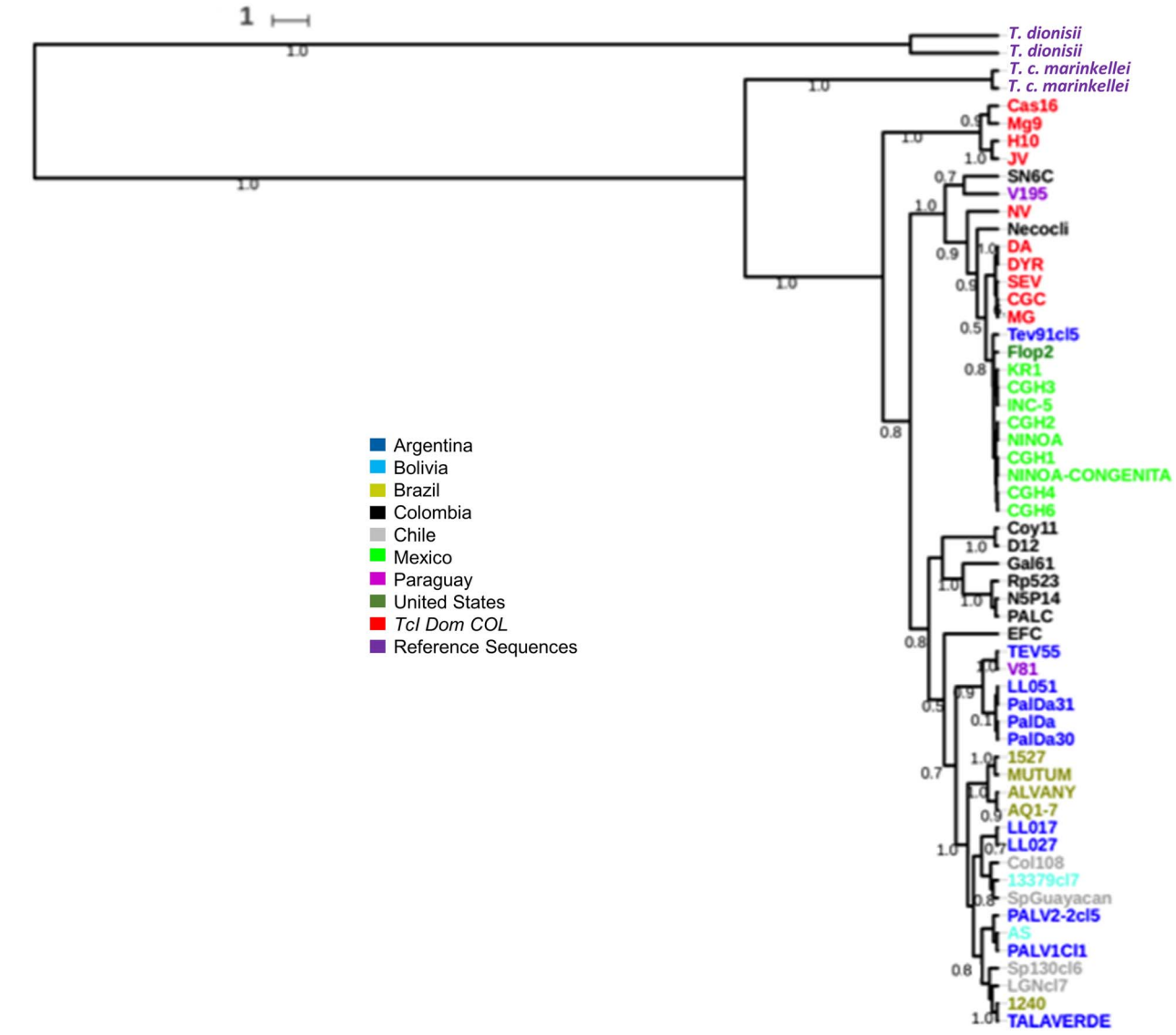

Fig. 5. SL- IR Maximum Clade Credibility (MCC) Tree of TcI isolates inferred by StarBeast2. Numeric values correspond to posterior probability ranging from 0 to 1.

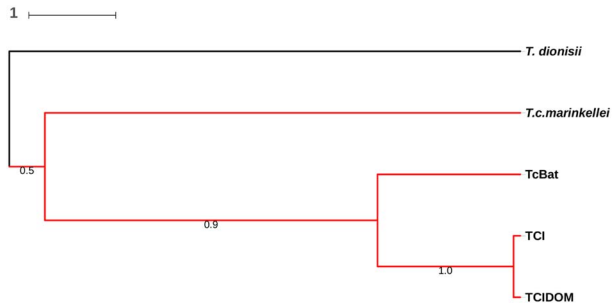

Fig. 6. Maximum Clade Credibility (MCC) species tree recovered from StarBeast2 analysis, illustrating Intra-TcI relationships and using *T. dionisii* and *T. c. marinkellei* as outgroups. Numeric values correspond to posterior probability ranging from 0 to 1.

A couple of studies have employed species delimitation approaches to clarify *T. cruzi* DTU relationships (Tomasini and Diosque, 2015) and examine the diversity of bat trypanosomes by Poisson tree processes (Cottontail *et al.*, 2014). By employing current advances in Bayesian MSC (Ogilvie *et al.*, 2017), our study represents the first comprehensive attempt to understand intra-lineage diversity and relationships in the TcI complex. Nevertheless, distinguishing between ILS and introgression is often challenging, and both are frequently confounded (Joly *et al.*, 2009). Here, we suggest that gene tree discordance may be influenced by these two processes. Consequently, it is important to note that StartBeast2 (Ogilvie *et al.*, 2017) only accounts for ILS and not migration models. Moreover, in the face of complex *T. cruzi* population dynamics, future work that focuses on trypanosome biology and evolution must take into account lineage delimitation, conduct additional hypothesis testing, and further investigate if the patterns reflect introgression or ILS.

Based on random samples from Northern/Central America to South America and applying a Bayesian MSC approach, our results support the emergence of TcI_{DOM} as an independently evolving genotype of the TcI complex at least in Colombia and Venezuela. In addition, we determined that neither geography nor hosts explain the gene tree discordance and genetic clustering among TcI strains (Introgression events). Other processes such as gene flow, divergence with gene flow, and ILS could shape the current parasite genetic structure and evolution.

However, the systematics, natural history, and epidemiology of *T. cruzi* are not fully understood. Here, we propose that combining MLST approaches with species coalescent-based methods for lineage delimitation of *T. cruzi* may be useful to prevent subjectivity and misidentification of parasite genetic structure that could be expanded to the other DTUs. Concatenation must be carefully

https://doi.org/10.1017/S0031182019000428 Published online by Cambridge University Press

used and taken into account in the light of the evaluation of hypotheses about the relationships between lineages and species.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0031182019000428.

Author ORCIDs. Juan David Ramírez González, 0000-0002-1344-9312.

Acknowledgments. We thank Mallory Eckstut, PhD, from Edanz Group (http://www.edanzediting.com/ac) for technical editing of an early version of this paper.

Financial support. This work was funded by DIRECCIÓN DE INVESTIGACIÓN E INNOVACIÓN from Universidad del Rosario.

Conflict of interest. None.

Ethical standards. Not applicable.

References

Barnabé C, Mobarec HI, Jurado MR, Cortez JA and Brenière SF (2016) Reconsideration of the seven discrete typing units within the species Trypanosoma cruzi, a new proposal of three reliable mitochondrial clades. *Infect Genet Evol.* **39**, 176–186.

Bouckaert R and Drummond A (2017) Bmodeltest: bayesian phylogenetic site model averaging and model comparison. *BMC Evolutionary Biology* **17**, 42.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C, Xie D, … Drummond A (2014) BEAST 2: a software platform for bayesian evolutionary analysis. *PLoS Computational Biology* **10**, e1003537.

Brenière S, Waleckx E and Barnabé C (2016) Over six thousand *Trypanosoma cruzi* strains classified into Discrete Typing Units (DTUs): attempt at an inventory. *PLOS Neglected Tropical Diseases* **10**, e0004792.

Brisse S, Barnabé C, Bañuls A, Sidibé I, Noël S and Tibayrenc M (1998) A phylogenetic analysis of the *Trypanosoma cruzi* genome project CL Brener reference strain by multilocus enzyme electrophoresis and multiprimer random amplified polymorphic DNA fingerprinting. *Molecular and Biochemical Parasitology* **92**, 253–263.

Brisse S, Henriksson J, Barnabé C, Douzery E, Berkvens D, Serrano M, … Tibayrenc M (2003) Evidence for genetic exchange and hybridization in *Trypanosoma cruzi* based on nucleotide sequences and molecular karyotype. *Infection, Genetics and Evolution* **2**, 173–183.

Burgos J, Altcheh J, Bisio M, Duffy T, Valadares H, Seidenstein M, … Schijman A (2007) Direct molecular profiling of minicircle signatures and lineages of *Trypanosoma cruzi* bloodstream populations causing congenital Chagas disease. *International Journal for Parasitology* **37**, 1319–1327.

Burgos J, Diez M, Vigliano C, Bisio M, Risso M, Duffy T, … Schijman A (2010) Molecular identification of *Trypanosoma cruzi* discrete typing units in end-stage chronic chagas heart disease and reactivation after heart transplantation. *Clinical Infectious Diseases* **51**, 485–495.

Cosentino R and Agüero F (2012) A simple strain typing assay for *Trypanosoma cruzi*: discrimination of major evolutionary lineages from a single amplification product. *PLoS Neglected Tropical Diseases* **6**, e1777.

Cottontail V, Kalko E, Cottontail I, Wellinghausen N, Tschapka M, Perkins S and Pinto C (2014) High local diversity of *Trypanosoma* in a common bat species, and implications for the biogeography and taxonomy of the *T. cruzi* clade. *PLoS ONE* **9**, e108603.

Cruz L, Vivas A, Montilla M, Hernández C, Flórez C, Parra E and Ramírez JD (2015) Comparative study of the biological properties of *Trypanosoma cruzi* I genotypes in a murine experimental model. *Infection, Genetics and Evolution* **29**, 110–117.

Cura C, Mejía-Jaramillo A, Duffy T, Burgos J, Rodriguero M, Cardinal M, … Gürtler R (2010) *Trypanosoma cruzi* I genotypes in different geographical regions and transmission cycles based on a microsatellite motif of the intergenic spacer of spliced-leader genes. *International Journal for Parasitology* **40**, 1599–1607.

de Freitas J, Augusto-Pinto L, Pimenta J, Bastos-Rodrigues L, Gonçalves V, Teixeira S, … Pena S (2006) Ancestral genomes, sex, and the population structure of *Trypanosoma cruzi*. *PLoS Pathogens* **2**, e24.

Degnan JH and Rosenberg NA (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution* **24**, 332–340.

De Meeûs T, McCoy KD, Prugnolle F, Chevillon C, Durand P, Hurtrez-Bousses S and Renaud FJI and Genetics and Evolution. (2007) Population genetics and molecular epidemiology or how to "débusquer la bête". **7**, 308–332.

Diosque P, Tomasini N, Lauthier J, Messenger L, Monje Rumi M, Ragone P, … Yeo M (2014) Optimized Multilocus Sequence Typing (MLST) scheme for *Trypanosoma cruzi*. *PLoS Neglected Tropical Diseases* **8**, e3117.

Edgar R (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792–1797.

Flores-Lopez CA and Machado CA (2011) Analyses of 32 loci clarify phylogenetic relationships among *Trypanosoma cruzi* lineages and support a single hybridization prior to human contact. *PLoS Neglected Tropical Diseases* **5**, e1272.

Fujita M, Leaché A, Burbrink F, McGuire J and Moritz C (2012) Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution* **27**, 480–488.

Gómez-Hernández C, Rezende-Oliveira K, Nascentes G, Batista L, Kappel H, Martinez-Ibarra J, … Ramírez L (2011) Molecular characterization of *Trypanosoma cruzi* Mexican strains and their behavior in the mouse experimental model. *Revista da Sociedade Brasileira de Medicina Tropical* **44**, 684–690.

Gómez-Palacio A, Lopera J, Rojas W, Bedoya G, Cantillo-Barraza O, Marín-Suarez J, … Mejía-Jaramillo A (2016) Multilocus analysis indicates that *Trypanosoma cruzi* I genetic substructure associated with sylvatic and domestic cycles is not an attribute conserved throughout Colombia. *Infection, Genetics and Evolution* **38**, 35–43.

Hall B, Meredith E and Wilkinson S (2012) Targeting the substrate preference of a type I nitroreductase to develop antitrypanosomal quinone-based prodrugs. *Antimicrobial Agents and Chemotherapy* **56**, 5821–5830.

Heled J and Drummond A (2009) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* **27**, 570–580.

Herrera C, Bargues M, Fajardo A, Montilla M, Triana O, Vallejo G and Guhl F (2007) Identifying four *Trypanosoma cruzi* I isolate haplotypes from different geographic regions in Colombia. *Infection, Genetics and Evolution* **7**, 535–539.

Herrera C, Guhl F, Falla A, Fajardo A, Montilla M, Adolfo Vallejo G and Bargues M (2009) Genetic variability and phylogenetic relationships within *Trypanosoma cruzi* I isolated in Colombia based on miniexon gene sequences. *Journal of Parasitology Research* 1–9, pii: 897364.

Hey J and Pinho C (2012) Population genetics and objectivity in species diagnosis. *Evolution* **66**, 1413–1429.

Jansen A, Xavier SCC and Roque A (2017) Ecological aspects of *Trypanosoma cruzi*. American Trypanosomiasis Chagas Disease: One Hundred Years of Research: Elsevier.

Joly S, McLenachan PA and Lockhart P (2009) A statistical approach for distinguishing hybridization and incomplete lineage sorting. *The American Naturalist* **174**, E54–E70.

Justi SA and Galvão C (2017) The evolutionary origin of diversity in chagas disease vectors. *Trends in Parasitology* **33**, 42–52.

Lauthier J, Tomasini N, Barnabé C, Rumi M, D'Amato A, Ragone P, … Diosque P (2012) Candidate targets for multilocus sequence typing of *Trypanosoma cruzi*: validation using parasite stocks from the Chaco Region and a set of reference strains. *Infection, Genetics and Evolution* **12**, 350–358.

Leaché AD, Koo MS, Spencer CL, Papenfuss TJ, Fisher RN and McGuire JA (2009) Quantifying ecological, morphological, and genetic variation to delimit species in the coast horned lizard species complex (Phrynosoma). *PNAS* **106**, 12418–12423.

Leigh J and Bryant D (2015) Popart: full-feature software for haplotype network construction. *Methods in Ecology and Evolution* **6**, 1110–1116.

Letunic I and Bork P (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research* **4**, 1.

Lewis MD, Llewellyn MS, Yeo M, Acosta N, Gaunt MW and Miles MA (2011) Recent, independent and anthropogenic origins of *Trypanosoma cruzi* hybrids. *PLoS Neglected Tropical Diseases* **5**, e1363.

León CM, Hernández C, Montilla M and Ramírez JD (2015) Retrospective distribution of *Trypanosoma cruzi* I genotypes in Colombia. *Memórias do Instituto Oswaldo Cruz* **110**, 387–393.

León C, Montilla M, Vanegas R, Castillo M, Parra E and Ramírez JD (2016) Murine models susceptibility to distinct *Trypanosoma cruzi* I genotypes infection. *Parasitology* **144**, 512–519.

Librado P and Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics (Oxford, England)* **25**, 1451–1452.

Lima V, Jansen A, Messenger L, Miles M and Llewellyn M (2014) Wild *Trypanosoma cruzi* I genetic diversity in Brazil suggests admixture and disturbance in parasite populations from the Atlantic Forest region. *Parasites and Vectors* **7**, 263.

Lima L, Espinosa-Álvarez O, Ortiz PA, Trejo-Varón JA, Carranza JC, Pinto CM, Serrano MG, Buck GA, Camargo EP, Teixeira MM (2015) Genetic diversity of Trypanosoma cruzi in bats, and multilocus phylogenetic and phylogeographical analyses supporting Tcbat as an independent DTU (discrete typing unit). *Acta Trop* **151**, 166–77.

Llewellyn M, Lewis M, Acosta N, Yeo M, Carrasco H, Segovia M, … Gaunt M (2009*a*) *Trypanosoma cruzi* IIc: phylogenetic and phylogeographic insights from sequence and microsatellite analysis and potential impact on emergent chagas disease. *PLoS Neglected Tropical Diseases* **3**, e510.

Llewellyn MS, Miles MA, Carrasco HJ, Lewis MD, Yeo M, Vargas J, Torrico F, Diosque P, Valente V, Valente SA and Gaunt MW (2009*b*) Genome-Scale multilocus microsatellite typing of *Trypanosoma cruzi* discrete typing unit I reveals phylogeographic structure and specific genotypes linked to human infection. *PLoS Pathogens* **5**, e1000410.

Llewellyn MS, Miles MA, Carrasco HJ, L MD, Yeo M, Vargas J, … Gaunt MW (2009*c*) Genome-scale multilocus microsatellite typing of *Trypanosoma cruzi* discrete typing unit I reveals phylogeographic structure and specific genotypes linked to human infection. *PLoS Pathogens* **5**, e1000410.

Marcili A, Lima L, Cavazzana M, Junqueira AC, Veludo HH, Maia Da Silva F, … Teixeira MM (2009) A new genotype of *Trypanosoma cruzi* associated with bats evidenced by phylogenetic analyses using SSU rDNA, cytochrome b and Histone H2B genes and genotyping based on ITS1 rDNA. *Parasitology* **136**, 641–655.

Messenger L, Llewellyn M, Bhattacharyya T, Franzén O, Lewis M, Ramírez J, … Miles M (2012) Multiple mitochondrial introgression events and heteroplasmy in *Trypanosoma cruzi* revealed by maxicircle MLST and next generation sequencing. *PLoS Neglected Tropical Diseases* **6**, e1584.

Nicholas KB, Nicholas HBJ and Deerfield DWI (1997) Genedoc: analysis and visualization of genetic variation. *EMBNEW News* **4**.

Ogilvie H, Bouckaert R and Drummond A (2017) StarBEAST2 brings faster species tree inference and accurate estimates of substitution rates. *Molecular Biology and Evolution* **34**, 2101–2114.

Poveda C, Higuera A, Urbano P and Ramírez J (2017) Ecology of *Trypanosoma cruzi* I genotypes across Rhodnius prolixus captured in Attalea butyracea palms. *Infection, Genetics and Evolution* **49**, 146–150.

Rambaut A, Drummond AJ, Xie D, Baele G and Suchard MA (2018) Tracer v1.6. Disponible en. Available at http://tree.bio.ed.ac.uk/software/tracer/.

Ramírez J and Llewellyn M (2014) Reproductive clonality in protozoan pathogens-truth or artefact? *Molecular Ecology* **23**, 4195–4202.

Ramírez J and Hernández C (2018) *Trypanosoma cruzi* I: Towards the need of genetic subdivision?, Part II. *Acta Tropica, In Press, Corrected Proof*.

Ramírez J, Duque M and Guhl F (2011) Phylogenetic reconstruction based on Cytochrome b (Cytb) gene sequences reveals distinct genotypes within Colombian *Trypanosoma cruzi* I populations. *Acta Tropica* **119**, 61–65.

Ramírez J, Guhl F, Messenger L, Lewis M, Montilla M, Cucunuba Z, … Llewellyn M (2012) Contemporary cryptic sexuality in *Trypanosoma cruzi*. *Molecular Ecology* **21**, 4216–4226.

Ramírez J, Tapia-Calle G and Guhl F (2013) Genetic structure of *Trypanosoma cruzi* in Colombia revealed by a High-throughput Nuclear Multilocus Sequence Typing (nMLST) approach. *BMC Genetics* **14**, 96.

Ramírez JD, Hernández C, Montilla M, Zambrano P, Flórez AC, Parra E and Cucunubá ZM (2014) First report of human *Trypanosoma cruzi* infection attributed to TcBat genotype. *Zoonosis and Public Health* **61**, 477–479.

Shapiro BJ, Leducq JB and Mallet J (2016) What Is Speciation? *PLoS Genet.* **12**, e1005860.

Tomasini N and Diosque P (2015) Evolution of *Trypanosoma cruzi*: clarifying hybridisations, mitochondrial introgressions and phylogenetic relationships between major lineages. *Memórias do Instituto Oswaldo Cruz* **110**, 403–413.

Tomasini N, Lauthier J, Rumi M, Ragone P, D'Amato A, Brandan C, … Diosque P (2011) Interest and limitations of Spliced Leader Intergenic Region sequences for analyzing *Trypanosoma cruzi* I phylogenetic diversity in the Argentinean Chaco. *Infection, Genetics and Evolution* **11**, 300–307.

Westenberger S, Barnabé C, Campbell DA and Sturm NR (2005) Two hybridization events define the population structure of *Trypanosoma cruzi*. *Genetics* **171**, 527–543.

(WHO), WHO (2016) Global Burden of Disease Estimates for 2000–2015. In. Geneva.

Yang Z and Rannala BJPOTNAOS (2010) Bayesian species delimitation using multilocus sequence data. 200913022.

Yeo M, Mauricio I, Messenger L, Lewis M, Llewellyn M, Acosta N, … Miles M (2011) Multilocus Sequence Typing (MLST) for lineage assignment and high resolution diversity studies in *Trypanosoma cruzi*. *PLoS Neglected Tropical Diseases* **5**, e1049.

Zingales B, Miles M, Campbell D, Tibayrenc M, Macedo A, Teixeira M, … Sturm N (2012) The revised *Trypanosoma cruzi* subspecific nomenclature: rationale, epidemiological relevance and research applications. *Infection, Genetics and Evolution* **12**, 240–253.

Zumaya-Estrada F, Messenger L, Lopez-Ordonez T, Lewis M, Flores-Lopez C, Martínez-Ibarra A, … Llewellyn M (2012) North American import? Charting the origins of an enigmatic *Trypanosoma cruzi* domestic genotype. *Parasites & Vectors* **5**, 226.

https://doi.org/10.1017/S0031182019000428 Published online by Cambridge University Press