Infection, Genetics and Evolution 11 (2011) 167–174

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Statistical phylogeography of Chagas disease vector *Triatoma infestans*: Testing biogeographic hypotheses of dispersal

Fernando Torres-Pérez [a,b,c,*], Mariana Acuna-Retamar [b,d], Joseph A. Cook [a], Antonella Bacigalupo [d], Alejandro García [e], Pedro E. Cattan [d]

[a] Biology Department and Museum of Southwestern Biology, University of New Mexico, Albuquerque, NM 87131, USA
[b] Department of Pathology, University of New Mexico Health Sciences Center, Albuquerque, NM 87131, USA
[c] Facultad de Ciencias de la Educación, Universidad Central de Chile, Santiago, Chile
[d] Facultad de Ciencias Veterinarias y Pecuarias, Universidad de Chile, Santiago, Chile
[e] Facultad de Medicina Occidente, Universidad de Chile, Santiago, Chile

ARTICLE INFO

Article history:
Received 14 March 2010
Received in revised form 11 August 2010
Accepted 13 September 2010
Available online 8 October 2010

Keywords:
Chagas disease
Chile
Spatial genetics
Triatominae
Cytochrome oxidase I

ABSTRACT

Chagas disease is one of the most important vector-borne diseases in Latin America. The disease, caused by the flagellate protozoan *Trypanosoma cruzi*, is commonly transmitted to humans by *Triatoma infestans* in South America. Using mitochondrial DNA sequences, we assessed alternative biogeographic scenarios of dispersal of *T. infestans* using coalescence simulations. We also assessed phylogeographic structure and spatial genetics of *T. infestans* in Chile. Two major routes of dispersal in southern South America were supported including a dual-origin of *T. infestans* in Chile. Phylogeographic analyses identified two primary clades with Chilean haplotypes partitioned into either a northern cluster with Peruvian and Bolivian haplotypes or a north-central cluster with Argentinean and Uruguayan haplotypes. The north-central clade is further divided into two subgroups. Domestic and sylvatic *T. infestans* in central Chile were not segregated in the phylogeographic reconstruction. Spatial genetic analyses show higher distances in northern Chile, congruent with the presence of two divergent lineages of *T. infestans*. Phylogenetic evidence does not unequivocally support the hypothesized Bolivian origin of *T. infestans*, so we discuss alternative scenarios.

1. Introduction

The evolutionary and biogeographic history of zoonotic pathogens and associated vectors can provide critical insight into the management and control of endemic diseases (Lloyd-Smith et al., 2009). Chagas disease is one of the most important vector-borne diseases in Latin America with over 8 million people infected (Gurtler et al., 2008). The disease is caused by the flagellate protozoan *Trypanosoma cruzi*, which is commonly transmitted to humans in South America by *Triatoma infestans* (Reduviidae, Triatominae). Epidemiological importance of Chagas disease led to the launch of a united strategy (the southern cone initiative for control of Chagas disease) to combat the disease mainly through elimination of the main vector, *T. infestans*. This effort corresponded to a decrease in incidence of new infections by 70% and effective vector control campaigns (Schofield and Dias, 1999; Schofield et al., 2006; Moncayo and Silveira, 2009). Control programs in Chile have eliminated *T. infestans* from large domestic areas; however, new sylvatic foci have been reported in central Chile (Bacigalupo et al., 2006, 2010) highlighting the necessity of vector studies to understand the origin of these populations. This reduviid vector occurs in Chile along a latitudinal range (18–34°S; Fig. 1) that spans multiple biomes, including northern desert, semi-arid desert, and central Mediterranean areas (Schofield et al., 1982; Bargues et al., 2006).

Because of the broad distribution of *T. infestans*, knowledge of its origin and colonization history may provide key knowledge into the spread and geographic limits of the associated disease. *T. infestans* was historically considered an exclusively domestic and peridomestic insect, with the only known sylvatic populations confirmed in the Andean valleys of Cochabamba and Sucre in Bolivia (Torrico, 1946; Usinger et al., 1966; Dujardin et al., 1987, 1998). The Bolivian highlands were hypothesized as the center of origin for *T. infestans* and from there it spread to other South American countries (Usinger et al., 1966; Dujardin et al., 1998; Panzera et al., 2004; Giordano et al., 2005; Bargues et al., 2006). Biogeographic hypothesis of dispersal of *T. infestans* are thought to involve two routes initially (Bargues et al., 2006): one was through the Andean regions of Bolivia and Perú, while the second route led

* Corresponding author at: Pontificia Universidad Católica de Chile, Center for Advanced Studies in Ecology and Biodiversity, Alameda 340, Santiago, Chile.
E-mail addresses: ftorresp@gmail.com, ftorresp@unm.edu (F. Torres-Pérez).

1567-1348/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2010.09.005

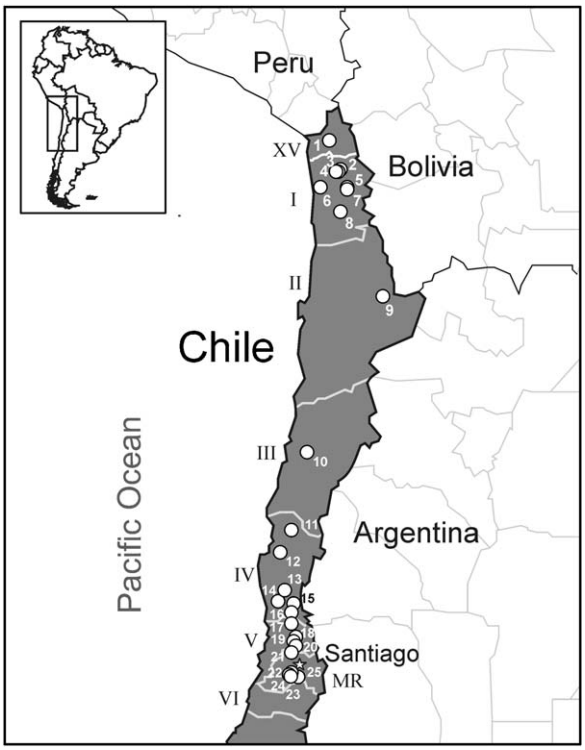

Fig. 1. Map of the sampled localities (see numbers in Table 1) of *T. infestans* in Chile. Map generated using Planiglobe (http://www.planiglobe.com). Administrative regions in Chile are shown as Roman numerals (I, XV–VI); MR: metropolitan region.

to the lowlands of Argentina, Paraguay, Uruguay, Brazil, and Chile. These primary routes were supported by independent nuclear and mitochondrial studies that defined two major lineages of *T. infestans*, the Andean and non-Andean lineages (Monteiro et al., 1999; Panzera et al., 2004; Bargues et al., 2006; Piccinali et al., 2009). Using the traditional view in formulating hypotheses about the events, populations of *T. infestans* in Chile may be the result of colonizers originating from each of the two primary lineages (Bargues et al., 2006).

Statistical phylogeographic approaches provide the opportunity to test alternative models of historical processes (Knowles and Maddison, 2002). By using a coalescent framework (Kingman, 1982; Richards et al., 2007), data simulated under explicit models can be compared to observed data (Knowles, 2004, 2009). In this study, we used mitochondrial sequence variation to assess four *a priori* biogeographic hypotheses of dispersal of *T. infestans* in South America, and to determine the origin of Chilean populations. Previous evidence reported that *T. infestans* in central Chile clustered within both the Andean and non-Andean lineages, suggesting a dual colonization route (Bargues et al., 2006). Hence, the first hypothesis (named dual-origin of Chilean samples with Bolivia ancestral – DOBA; Fig. 2A) posits that *T. infestans* from Bolivia are ancestral and from there populations followed two independent dispersal routes: one to Perú and then northern Chile, and a second route to Argentina, Uruguay, and north-central Chile. Samples from Chile were segregated either in the north (including exclusively those from Regions I and XV; see Fig. 1), or north-central (Regions II to Metropolitan) based on strong genetic subdivision in a second Chagas disease vector of the genus *Mepraia* (Calleros et al., 2010), and under the assumption that both vectors experienced similar processes of colonization and population subdivision. The second hypothesis (dual origin of Chilean samples with Bolivia non-ancestral – DOBnA; Fig. 2B) also posits two major dispersal events but Bolivian samples are not considered ancestral. In this case, the northern cluster would include Bolivia, Perú and northern Chile. This hypothesis arises from the maximum parsimony topology obtained by Piccinali et al. (2009) which recovered two major clades: Argentina plus Uruguay and Bolivia plus Perú. Although these two hypotheses describe a pattern of

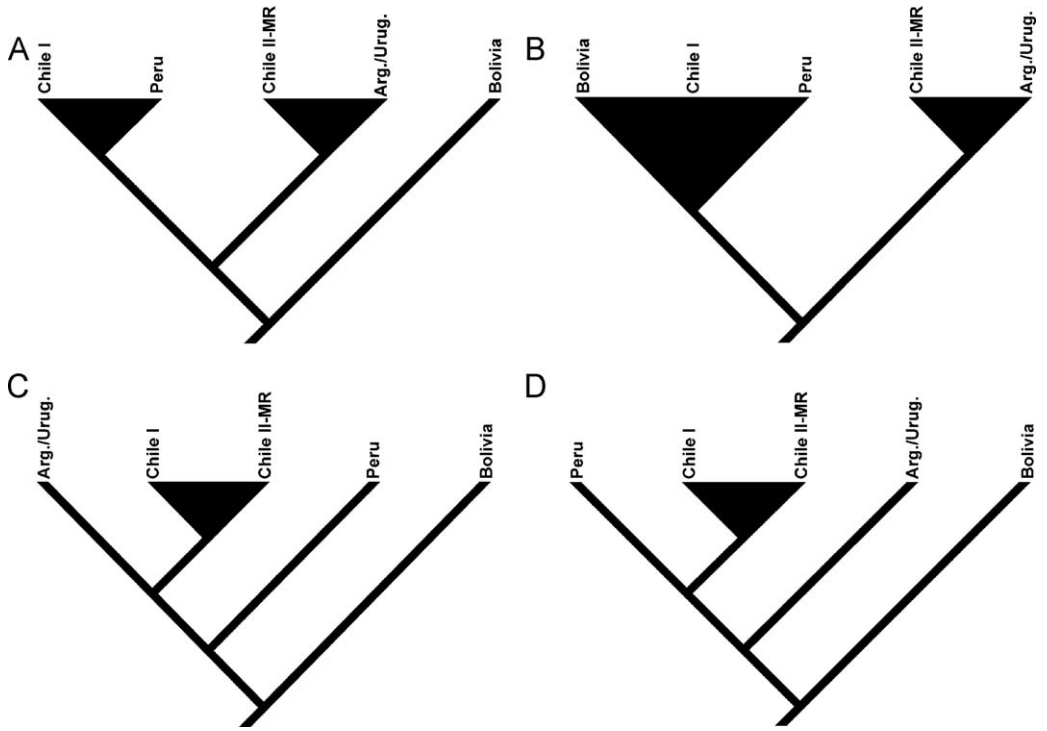

Fig. 2. Illustration representing the two dual-dispersal hypotheses (A and B) and two single dispersal hypotheses (C and D) of *T. infestans* in South America tested using coalescent simulations in Mesquite v.2.72. Chile I, samples from Region I and XV in Chile (see Fig. 1); Chile II-MR, samples from Region II to metropolitan region in Chile; Arg./Urug., samples from Argentina and Uruguay.

dual-colonization of *T. infestans* in Chile, alternative historical scenarios that represent a range of possible explanations may be generated. Two alternative (simpler sensu Knowles, 2004) hypotheses were also tested assuming a single dispersal route of *T. infestans* into Chile: a single northern colonization from Perú with Bolivia as ancestral – SN (Fig. 2C), and a single southern colonization from Argentina, also with Bolivia as the most ancestral group – SS (Fig. 2D). Finally, our study also aims to estimate the date of the most recent common ancestor (MRCA) of all *T. infestans* in South America.

2. Materials and methods

2.1. Sampling and sequencing

A total of 80 *T. infestans* were collected in Chile ranging from 18°S to 34°S (Fig. 1; Table 1). Samples represent a collecting effort of more than 3 years in localities representative of the distribution of *T. infestans*, combined with samples provided by the Chilean Ministry of Health under the program for control of Chagas disease in Chile. Kissing bugs were collected as previously described (Marcet et al., 2008; Bacigalupo et al., 2010). Genomic DNA from legs was extracted using QIAamp Mini Kit (QIAGEN™). A fragment of 661-nt of the mitochondrial DNA cytochrome oxidase I (COI) was amplified using polymerase chain reaction (PCR) and the primers COI-121 5′-AGTTATAATTGGAGGCTTCGGTAAC-3′ and COI-942 5′-AATACAAATCCGAAGGCTCATAGA-3′. The following conditions were used to amplify the COI gene: 95 °C denaturation for 30 s, 51 °C annealing for 30 s, and 72 °C extension for 1 min 33 cycles. Amplification reactions were carried out in 20 μl reaction mixture containing 1–50 ng of template DNA, 10× buffer PCR, 25 mM $MgCl_2$, 10 mM of dNTPs mix, 10 μM of each primers and 5 U/ml AmpliTaq Gold polymerase. PCR-amplified products were checked by electrophoresis on a 1% agarose gel in 1× TBE running buffer. Double-stranded PCR products were purified using QIAquick PCR purification kit (Qiagen Inc., Valencia, CA, USA). Sequencing was conducted in the DNA Research Facility (University of New Mexico, USA; http://hsc.unm.edu/som/Pathology/DNA/) through cycle sequencing on an ABI Prism 3100 automated sequencer (Perkin Elmer, Norwalk, CT, USA), using the same primers as employed for PCR amplification, but diluted at 1 μM. Sequences were edited using the BioEdit Sequence Alignment Editor (Hall, 1999), and aligned using Clustal W implemented in BioEdit. Sequences are deposited in GenBank (accession numbers HM439795–HM439874). Thirty-six COI haplotypes from *T. infestans* from Perú, Bolivia, Uruguay, and Argentina were downloaded from GenBank (EF451005-EF451040), and a matrix of 235 sequences was re-built using the information provided in Piccinali et al. (2009). The complete matrix used for analyses includes 315 COI sequences.

2.2. Phylogeographic structure and historical biogeography

The Median Joining method (Bandelt et al., 1999) implemented in Network 4.2.0.1 software (http://www.fluxus-engineering.com/sharenet.htm) was used to assess intraspecific relationships of all samples (315 sequences) from Chile, Perú, Bolivia, Argentina and Uruguay.

We conducted coalescent simulations in Mesquite version 2.72 (Maddison and Maddison, 2009) under the four *a priori* biogeographic hypotheses of dispersal for *T. infestans*. Coalescent simulations were used to evaluate the fit between the historical hypotheses (Fig. 2A–D) and empirical data. Simulations are performed under conditions that mirror the empirical data (e.g. the amount of data and mutational model) following established procedures (Knowles, 2001; Knowles and Maddison, 2002; Richards et al., 2007). Briefly, the analysis involves first generating a large number of genealogies simulated by a neutral coalescent process under each biogeographic hypothesis. Sequence data are then simulated on these genealogies, summary statistics calculated for each genealogy, and a null distribution of the summary statistic (e.g. number of deep coalescents) is generated. If the value of the summary statistic estimated from empirical data falls within 95% of the null distribution, then that model is not rejected as a possible biogeographic scenario. We first estimated the effective population size (N_e) for *T. infestans* using the Bayesian Skyline Plot analysis (Drummond et al., 2005) implemented in the program BEAST (Drummond and Rambaut, 2007), and using an average of

Table 1
Sampling localities, sample size and coordinates of *T. infestans* of Chile using mitochondrial DNA cytochrome oxidase I. Haplotypes are depicted in Fig. 3. Category of sites: D, domestic and/or peridomestic; S, sylvatic.

Administrative region in Chile	Locality	Category of sites	Sample size	Haplotype	Map number	Latitude	Longitude
XV-Arica y Parinacota	Livilcar	D	3	B	1	−18.482214	−69.732971
I-Tarapacá	Apamilco	D	1	B	2	−19.308713	−69.391819
I-Tarapacá	Camiña	D	1	B	3	−19.332202	−69.455223
I-Tarapacá	Jasjara	D	1	C	4	−19.345484	−69.498825
I-Tarapacá	Huaviña	D	1	B	5	−19.784796	−69.221420
I-Tarapacá	Huara	D	4	A, B, C	6	−19.808054	−69.971581
I-Tarapacá	Limaxiña	D	1	B	7	−19.862310	−69.211121
I-Tarapacá	Matilla	D	1	A	8	−20.509998	−69.367676
II-Antofagasta	San Pedro de Atacama	D	5	D	9	−22.919820	−68.209991
III-Atacama	Copiapó	D	1	D	10	−27.371158	−70.327606
IV-Coquimbo	La Serena	D	1	F	11	−29.578234	−70.795898
IV-Coquimbo	Andacollo	D	1	F	12	−30.232374	−71.082230
IV-Coquimbo	Combarbalá	D	1	E	13	−31.309062	−70.982666
IV-Coquimbo	Illapel	D	1	E	14	−31.594913	−71.117249
IV-Coquimbo	Coirón	D	1	E	15	−31.676758	−70.735474
IV-Coquimbo	Chillepín	D	2	E	16	−31.886887	−70.773926
V-Valparaíso	Petorca	D	3	E, F	17	−32.220934	−70.835724
V-Valparaíso	Putaendo	S	5	F	18	−32.615510	−70.682330
V-Valparaíso	San Felipe	D	1	G	19	−32.729530	−70.735474
V-Valparaíso	Villa Alegre	D	1	G	20	−32.864016	−70.667496
MR-Metropolitan	Til Til	D	8	D, F, G	21	−33.075204	−70.774367
MR-Metropolitan	Calera de Tango/Talagante	D	1	G	22	−33.641171	−70.818538
MR-Metropolitan	Calera de Tango	S, D	14	F, G	23	−33.652975	−70.783947
MR-Metropolitan	Talagante	S, D	20	G	24	−33.657919	−70.814696
MR-Metropolitan	Pirque	D	1	G	25	−33.730190	−70.573929

both 1 (Rojas et al., 2007) and 2 generations per year (Gorla and Schofield, 1989). We then simulated DNA sequences ($N = 1000$) with the same parameters as the empirical data constraining the coalescence of lineages within the topology of our four *a priori* biogeographic hypotheses. Simulations were performed using the HKY + I model of sequence evolution and the following parameters were extracted from jModelTest version 0.1.1 (Posada, 2008): $\pi A = 0.2856$, $\pi C = 0.2386$, $\pi G = 0.1732$, $\pi T = 0.3025$, proportion of invariant sites = 0.8155, and transition/transversion ratio = 31.2987. We then calculated the number of deep gene coalescents (ndC), and built a null distribution of ndC values. Using our reconstructed Bayesian gene tree, we fit each of the *a priori* hypotheses and calculated the observed ndC value, which was then compared with the null distribution from the simulations.

2.3. Spatial genetic analyses

Spatial analysis of molecular variance was performed in SAMOVA v.1.0 (Dupanloup et al., 2002). This method uses a simulated annealing approach to identify groups of populations (K), which are geographically homogeneous and maximally differentiated by maximizing Fct (the proportion of the total genetic variance due to differences among groups of populations). Fct values were calculated by running the program sequentially (500 random initial conditions), and forcing the data into k groups (where $k = 2–18$). Population subdivision was also estimated using the fixation index (Fst) between all groups generated in Arlequin 3.1 (Excoffier et al., 2005).

An interpolation-based graphical method was employed to generate a three-dimensional genetic landscape shape (GLS) within the program Alleles in Space (Miller, 2005). This analysis provides a visual perspective of the spatial distribution of recent genetic structure over landscapes, with peaks in areas where genetic distances between individuals are high, and valleys where genetic distances between individuals are low (Miller et al., 2006). Georeference coordinates (Universal Transverse Mercator system) were provided for each individual, and analyzed for the COI sequences.

2.4. Molecular diversity and dates of the most recent common ancestor (MRCA)

We evaluated variability of COI sequences (661-nt) in the three groups resulting from SAMOVA analyses using DnaSp 5.1 (Librado and Rozas, 2009). Neutrality tests assume that populations are in mutation-drift and migration-drift equilibrium, so we assessed population equilibrium in *T. infestans* by performing Tajima's *D*-test (Tajima, 1989), Fu's *Fs*-test (Fu, 1997), and R_2-test (Ramos-Onsins and Rozas, 2002), testing the significance of the statistics from 5000 simulated samples. Significantly negative values in neutrality tests reflect an excess of low frequency (Tajima's *D*) mutations in a population relative to what is expected under a standard neutral model, which is consistent with either directional or purifying selection (i.e. they are not in mutation-drift equilibrium) or an increase in population size (i.e. they are not in migration-drift equilibrium). Fu's *Fs* statistics measures departure from neutrality by detecting excess or deficiency of alleles, consistent with recent population expansion or genetic hitchhiking (negative values), and recent population bottlenecks or selection (positive values). Ramos-Onsis R_2-test is based on the difference between the number of singleton mutations and the average number of nucleotide differences, with lower values of the statistics expected under a population growth event. Population equilibrium tests were performed using DnaSp 5 (Librado and Rozas, 2009).

We used the Bayesian Markov Chain Monte Carlo (BMCMC) method available in BEAST v1.5.3 (Drummond and Rambaut, 2007) to estimate dates of the most recent common ancestor (MRCA) for *T. infestans*. Analyses were performed for the groups resulting from SAMOVA, the Andean and non-Andean lineages independently, and all data pooled using 218 (of the 315) randomly selected COI samples. Two separate MCMC analyses were run for 8×10^8 generations (10% burn-in) with parameters sampled every 80,000 steps. TRACER v1.5 (Rambaut and Drummond, 2009) was used to determine convergence, measure the effective sample size of each parameter (all resulting effective sample sizes exceeded 200), and calculate the mean and 95% highest posterior density interval (HPD). Analyses were performed comparing uniform rates across branches (strict clock) and uncorrelated relaxed clock assumptions. We used the strict clock assumption after comparing the marginal posterior distribution to prior distribution of the uncorrelated log normal distribution (standard deviation) using the Savage–Dickey ratio estimator of the Bayes factor (Suchard et al., 2001). We used as a prior a fixed mean substitution rate of 0.016 sub/site/my calculated for COI and COII (Arensburger et al., 2004).

3. Results

We used mitochondrial COI sequences (661-nt) to assess spatial structure and alternative phylogeographic hypotheses of dispersal of *T. infestans* in Chile. A total of 315 sequences of *T. infestans* from Argentina, Bolivia, Chile, Perú, and Uruguay produced 38 haplotypes. The unrooted phylogeographic network (Fig. 3) showed three major groups of haplotypes. Haplotypes A, B, and C representing Chilean localities from the extreme North (Regions I and XV, see Fig. 1) were closer to haplotypes found in Bolivia and Perú. The predominant haplotype (HapB) occurred in these three countries. A second group consisted of haplotypes from Argentina and Chile, with one haplotype (HapF) shared between both countries. The third and most diversified group included samples from Chile, Argentina, and Uruguay where haplotype D showed the highest frequency. Haplotypes G and E occurred in these three countries. In Chile, haplotype G was restricted to the central regions. Both the second and third groups included samples from central and northern Chile ranging from regions II to Metropolitan (see Fig. 1). Haplotypes from domestic and sylvatic *T. infestans* in central Chile were not segregated in the phylogeographic network.

The number of deep coalescent events (nDC) for the Bayesian topology of our observed data differed across the population trees

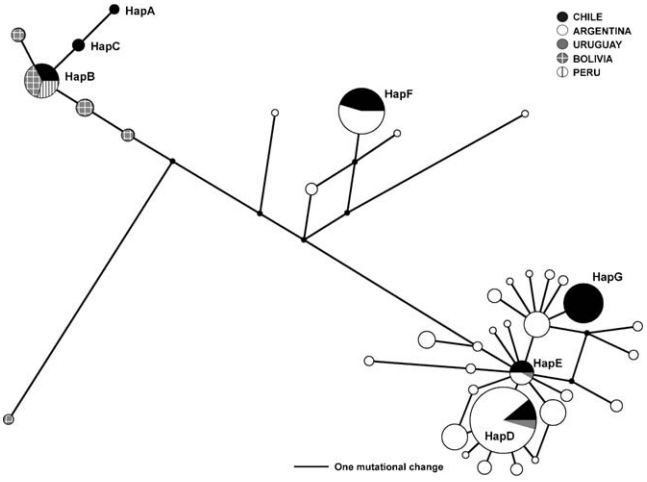

Fig. 3. Unrooted network of South American *T. infestans* using cytochrome oxidase I sequences with haplotypes depicted according to countries. Size of circles represents the number of individuals per haplotypes. Haplotypes found in Chile are labeled from HapA to HapG (see Table 1).

Table 2
Tests using the program Mesquite of the number of deep coalescents (ndC) for four biogeographic hypotheses related to the colonization of Chile by *T. infestans*. DOBA, dual-origin of Chilean kissing bugs with Bolivia ancestral; DOBnA, dual-origin of Chilean kissing bugs with Bolivia non-ancestral; SN, single northern colonization for Chilean kissing bugs; SS, single southern colonization. Ne_1 and Ne_2 are from the statistic (ndC) using an effective population size of 300,000 and 150,000, respectively.

Hypothesis	Observed ndC	Mean $ndC_{(Ne_1)}$	95% C.I.	*P*-value (Ne_1)	Mean $ndC_{(Ne_2)}$	95% C.I.	*P*-value (Ne_2)
DOBA	67	64.59	47–82	0.04	47.28	26–67	0.045
DOBnA	55	58.68	38–83	0.31	43.17	26–61	0.137
SN	89	66.94	44–87	0.02	49.78	29–72	0.001
SS	105	71.12	48–94	<0.0001	51.76	30–72	<0.0001

Table 3
Fixation index (Fst) among *T. infestans* sampling groups using mitochondrial DNA cytochrome oxidase I (*N* = sample size). Numbers in parenthesis represent *P*-values (±standard deviation) after 10,000 permutations.

Group	Chile I	Chile II-MR	Argentina	Bolivia
Chile I-XV (*N* = 13)	–			
Chile II-MR (*N* = 67)	0.354 (0.000 ± 0.000)	–		
Argentina (*N* = 199)	0.262 (0.000 ± 0.000)	0.165 (0.000 ± 0.000)	–	
Bolivia (*N* = 24)	0.102 (0.023 ± 0.002)	0.280 (0.000 ± 0.000)	0.196 (0.000 ± 0.000)	–
Perú (*N* = 7)	0.149 (0.138 ± 0.003)	0.499 (0.000 ± 0.000)	0.401 (0.000 ± 0.000)	0.239 (0.011 ± 0.001)

of each of the hypothesized biogeographic scenarios: 67 for the DOBA hypothesis, 55 for the DOBnA hypothesis, 89 for the SN hypothesis, and 105 for the SS hypothesis. Effective population size was estimated at 300,000 (Ne_1) and 150,000 (Ne_2) assuming 1 and 2 generations of *T. infestans* per year, respectively. Regardless of Ne, coalescent simulations rejected both of the single-colonization hypotheses (Table 2). Analyses failed to reject the hypothesis of dual-origin of Chilean samples with Bolivia non-ancestral (Fig. 2B). The hypothesis of a dual-origin of Chilean samples with Bolivia ancestral (Fig. 2A) was not rejected using a Ne = 300,000, but was marginally rejected using Ne = 150,000 (Table 2).

Spatial analysis of molecular variance was performed to assess substructure within *T. infestans* in South America. F_{CT} values ranged from 0.140 to 0.6369, with the group structure maximized at $k = 4$. Collection sites from Argentina, Uruguay, and Regions II to Metropolitan in Chile formed the first group (congruent with the previously described non-Andean lineage), samples from Bolivia formed the second group, samples from Perú the third group, and samples from Regions I and XV in Chile the fourth group. Fst-values among *T. infestans* (Table 3) ranged between 0.102 (Bolivia vs. Chile I-XV) to 0.499 (Perú vs. Chile II-RM).

The genetic landscape shape interpolation analysis (Fig. 4) showed that the spatial distribution of haplotypes across the north–south gradient in Chile was not uniform. The lowest pairwise genetic distances among *T. infestans* populations were found in the extreme north (18°29′S to 20°30′S), and the highest values between populations 8 (Matilla) and 9 (San Pedro de Atacama; Fig. 1). South of San Pedro de Atacama (22°55′S), some peaks and valleys (areas of high and low genetic distance between individuals) were found, with higher genetic heterogeneity in the South (Valparaiso and Metropolitan Regions; 32° to 33°S). Grid size and distance weighting parameters did not affect landscape shape.

Estimates of genetic variation (Table 4) varied across groups. For Bolivia, haplotype diversity (Hd) was 0.779 ± 0.05 and nucleotide diversity (π) 0.0034 ± 0.0011. Chile II-XV showed lower Hd (0.590 ± 0.122) and π (0.0012 ± 0.0003). Chile II-RM, Argentina and Uruguay samples showed the highest Hd (0.839 ± 0.016) and π (0.0061 ± 0.0003) values. Historical processes such as demographic

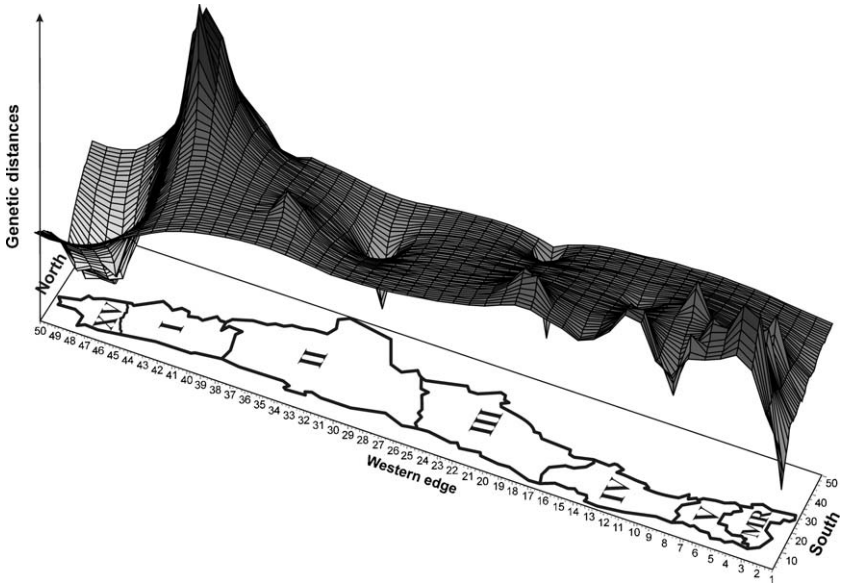

Fig. 4. A graphical interpolation-based representation of the genetic structure was made using a 50 × 50 grid and a distance weighting parameter of 0.5 for *T. infestans* in Chile. *X*- and *Y*-axes represent geographic coordinates, while surface heights along the *Z*-axis indicate genetic distances. Peaks (black) and valleys (white) are indicative of areas with high or low pairwise genetic distance between individuals, respectively.

Table 4
Descriptive statistics of genetic variation and neutrality tests of *T. infestans* sequences. *N*: Number of individuals; *S*: number of segregating sites; Hd: haplotype diversity; π: nucleotide diversity; SD: standard deviation. R_2: Ramos-Onsins and Rozas' neutrality test. Significance (*P-values*) of the statistics is shown in parenthesis.

Lineage/group	*N*	*S*	Hd ± SD	Π ± SD	Fu's-*Fs*	R_2	Tajima's *D*
Bolivia	24	10	0.779 ± 0.050	0.0034 ± 0.0011	1.3641 (*P* = 0.775)	0.1134 (*P* = 0.371)	−0.5105 (*P* = 0.350)
Chile I-XV	13	2	0.590 ± 0.122	0.0012 ± 0.0003	0.3131 (*P* = 0.619)	0.1987 (*P* = 0.739)	0.6554 (*P* = 0.772)
Chile II-MR, Argentina, Uruguay	271	37	0.839 ± 0.016	0.0061 ± 0.0003	−7.5635 (*P* = 0.085)	0.0542 (*P* = 0.174)	−0.9115 (*P* = 0.194)
Chile II-MR	67	11	0.677 ± 0.034	0.0071 ± 0.0005	9.9856 (*P* = 0.997)	0.2122 (*P* = 0.998)	2.8443 (*P* = 0.998)

expansion and contraction were evaluated under the assumption of neutrality using COI sequences of *T. infestans*. Neutrality tests were non-significant after 5000 permutations in all groups analyzed, rejecting the expansion hypothesis (Table 4).

BMCMC analyses using a strict molecular clock estimated MRCA of *T. infestans* at 0.89 million years ago (Mya) (95% HPD = 0.281–2.277 Mya); the divergence of the lineage Bolivia, Perú, and Chile north (Andean lineage) was estimated at 0.388 Mya (95% HPD = 0.086–1.015 Mya) and the lineage Argentina, Uruguay, and Chile central north (non-Andean lineage) at 0.588 Mya (95% HPD = 0.148–1.529 Mya). MRCA of samples from Bolivia were estimated at 0.31 Mya (95% HPD = 0.053–0.815 Mya).

4. Discussion

Coalescent simulations support the hypothesis that two major routes of dispersal of *T. infestans* in South America are represented by two independent lineages (Bargues et al., 2006), a finding congruent with the Andean and non-Andean groups (Panzera et al., 2004). Both lineages can be distinguished using mitochondrial and nuclear markers (Monteiro et al., 1999; Giordano et al., 2005; Bargues et al., 2006; Piccinali et al., 2009), the number of C-banded autosomes, C-banding on the X chromosome, and DNA content (Panzera et al., 2004). These two primary lineages of *T. infestans* colonized Chile across the Andes from both the north and southeast. The northern (Andean) lineage is found in Chile exclusively in Regions I and XV (around 18° to 20°30′S), and clusters with samples from Bolivia and Perú; while the non-Andean lineage includes samples from north-central Chile ranging from Regions II to Metropolitan (ca. 22°55′ to 33°43′S), and clusters with samples from Argentina and Uruguay. Other studies have shown that samples from the Bolivian Chaco, Brazil, and Paraguay are placed within the non-Andean group (Panzera et al., 2004). Higher variability, shared haplotypes between Chile and Argentina, and a haplotype network that lacks geographic structure in the non-Andean lineage may reflect multiple colonization events from Argentina or simply retention of ancestral polymorphisms (Maddison and Knowles, 2006). Sylvatic populations of *T. infestans* in Chile (Bacigalupo et al., 2006, 2010) were included within the non-Andean lineage and hence may be the result of colonization from Argentina. Whether sylvatic populations were the result of natural colonization of Chile, or they arrived first as domestic populations associated with human movement that then moved into the wild, remains to be investigated.

Analyses of spatial genetic structure (Fst, genetic landscape shape) showed higher divergence in the area between Regions I/XV and Region II (around 21° to 22°S) which is probably associated with the early split of the two major lineages that then independently reached northern Chile. In another kissing bug (genus *Mepraia*), separate species (Frias et al., 1998) are found in either Region I (*Mepraia gajardoi*) or Regions III to Metropolitan (*M. spinolai*) (Calleros et al., 2010). Samples of *Mepraia* from Region II showed a mixed phylogeographic pattern which may reflect either mitochondrial introgression or incomplete lineage sorting (Calleros et al., 2010). We do not have evidence suggesting that this area harbors intermediate forms (e.g. hybrids) between the two *T. infestans* lineages, but more intensive sampling and multiple loci should be used to further investigate this dynamic.

The geographic origin of *T. infestans* has been hypothesized to be the Andean valleys in Bolivia, either in Cochabamba (Usinger et al., 1966; Panzera et al., 2004; Bargues et al., 2006) or the area of Sucre-Vallegrande-Potosí (Dujardin et al., 1998; Giordano et al., 2005). The main evidence supporting this hypothesis was that these areas were thought to support the only known sylvatic populations, while reports of the species elsewhere surfaced only in domestic/peridomestic areas. The subsequent discovery of sylvatic populations elsewhere challenged this view (Noireau et al., 2005; Bacigalupo et al., 2006, 2010; Ceballos et al., 2009; Noireau, 2009; Buitrago et al., 2010). Although molecular studies have generally suggested a Bolivian Andean origin of *T. infestans*, those studies remain inconclusive. For example, phylogeographic studies recovered two reciprocally monophyletic groups of *T. infestans* (Monteiro et al., 1999; Bargues et al., 2006; Piccinali et al., 2009), and Bolivian populations were not basal (as expected for an ancestral group). Older alleles have a greater probability of being interior haplotypes in intraspecific gene genealogies (Posada and Crandall, 2001). Our phylogeographic haplotype network showed that samples from Bolivia were peripheral, suggesting a low probability that they represent the location of ancestral populations. Although one of our simulations (assuming a higher effective population size) did not reject the hypothesis that samples from Bolivia represent the ancestral group, results tend to favor the hypothesis that samples from Bolivia clustered with Perú and northern Chile. Alternative hypotheses may explain these results: (1) The origin of *T. infestans* was not Cochabamba, but instead elsewhere in Bolivia, as previously postulated (Dujardin et al., 1998; Giordano et al., 2005). Dujardin et al. (1998) postulated Sucre (and nearby areas) as the geographical origin of the species. Giordano et al. (2005) found a haplotype from Chuquisaca that may be the common ancestor of all *T. infestans*, but their limited sampling from only two other countries (Argentina and Brazil) precluded robust discrimination between ancestral and derived haplotypes. We included *T. infestans* from Cochabamba (14 domestic and 10 sylvatic) (Piccinali et al., 2009) but not Chuquisaca or Sucre, so ancestral haplotypes may not have been sampled. (2) Bolivia is not the center of origin of *T. infestans*. Given that molecular phylogeographic studies have not recovered any particular group as ancestral (Bargues et al., 2006; Piccinali et al., 2009), there exists the possibility that the most ancient haplotypes might have originated in an unsampled area outside of Bolivia. For example, Piccinali et al. (2009) found that a single haplotype from Santiago del Estero (Argentina) diverged earlier than the remaining samples from South America (Piccinali et al., 2009). Within the non-Andean lineage, higher variability in C-banding patterns of populations of the Argentinean Chaco led Panzera et al. (2004) to postulate this area as the primary focus of dispersal into non-Andean regions. Ceballos et al. (2009) recovered a basal haplotype from sylvatic *T. infestans* sampled in the Argentinean Chaco, but only a single peridomestic sample from Bolivia was included in that set of analyses. Our study is consistent with the placement of the MRCA of the non-Andean lineage in the Pleistocene, earlier than the Andean lineage (Bargues et al., 2006), suggesting that older haplotypes may be found outside of Bolivia. Hence, under the assumption that ancient haplotypes are not extinct, alternative models about the origin of *T. infestans* should be

considered and more fully explored in the future with expanded geographic sampling. For example, intermediate chromosomal forms between the Andean and non-Andean groups has been detected near the southeastern border of Bolivia (Departamento de Tarija) and in northwestern Argentina (Provincia de Salta) (Panzera et al., 2007).

T. infestans is found mostly in domestic and peridomestic environments, an adaptation that possibly arose during pre-Columbian times (Schofield, 1988; Panzera et al., 2004). This feature coupled with poor dispersal ability (Schofield, 1992) led to the hypothesis that passive dispersal associated with human migrations was the main mechanism of geographic expansion of *T. infestans* throughout South America (Schofield, 1988; Panzera et al., 2004; Noireau, 2009; Cortez et al., 2010). Low population genetic variation was postulated as a product of a recent rapid spread of the species (Dujardin et al., 1998; Monteiro et al., 1999; Marcilla et al., 2001; Segura et al., 2009). However, neutrality tests show that *T. infestans* has not experienced recent demographic expansion. Using a mitochondrial marker we estimated the divergence of the Andean and non-Andean lineages around 0.388–0.588 My, which globally agrees with dates previously reported based on nuclear markers (Bargues et al., 2006). These dates widely pre-date human colonization of the American continent (Dillehay and Collins, 1988; Dixon, 2001), and suggest that the earliest dispersal events for *T. infestans* in South America were not human-mediated. However, as domestication evolved, humans likely played an important role in the passive dispersal of *T. infestans* beginning with Andean tribes (5000 BC) and later during pre-Columbian and recent times (Usinger et al., 1966; Aufderheide et al., 2004; Cortez et al., 2010). Contemporary population structure has been dramatically influenced by human–vector interactions as reflected in the complex biogeographic patterns for *T. infestans* at both local and regional scales.

The phylogeographic split observed for *T. infestans* in northern Chile raises new questions regarding the possibility of other differences between these two independent lineages that will require further investigation. Furthermore, an understanding of the dynamic relationship between domestic and sylvatic populations is also key to the design and implementation of effective preventive and control strategies.

Acknowledgments

We thank SEREMI de Salud Región Metropolitana, IV and V, and Instituto de Salud Pública de Chile for providing valuable samples. Financial support was provided by Fondecyt 1070960. The manuscript was significantly improved by two anonymous reviews and the Editor's commentaries.

References

Arensburger, P., Buckley, T.R., Simon, C., Moulds, M., Holsinger, K.E., 2004. Biogeography and phylogeny of the New Zealand cicada genera (Hemiptera: Cicadidae) based on nuclear and mitochondrial DNA data. J. Biogeogr. 31, 557–569.

Aufderheide, A.C., Salo, W., Madden, M., Streitz, J., Buikstra, J., Guhl, F., Arriaza, B., Renier, C., Wittmers Jr., L.E., Fornaciari, G., Allison, M., 2004. A 9,000-year record of Chagas' disease. Proc. Natl. Acad. Sci. U.S.A. 101, 2034–2039.

Bacigalupo, A., Segura, J.A., Garcia, A., Hidalgo, J., Galuppo, S., Cattan, P.E., 2006. First finding of Chagas disease vectors associated with wild bushes in the Metropolitan Region of Chile. Rev. Med. Chil. 134, 1230–1236.

Bacigalupo, A., Torres-Perez, F., Segovia, V., García, A., Correa, J.P., Moreno, L., Arroyo, P., Cattan, P.E., 2010. Sylvatic foci of the Chagas disease vector Triatoma infestans in Chile: description of a new focus and challenges for control programs. Mem. Inst. Oswaldo Cruz 105, 633–641.

Bandelt, H.J., Forster, P., Rohl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.

Bargues, M.D., Kiisiowicz, D.R., Panzera, F., Noireau, F., Marcilla, A., Perez, R., Rojas, M.G., O'Connor, J.E., Gonzalez-Candelas, F., Galvao, C., Jurberg, J., Carcavallo, R.U., Dujarding, J.P., Mas-Coma, S., 2006. Origin and phylogeography of the Chagas disease main vector *Triatoma infestans* based on nuclear rDNA sequences and genome size. Infect. Genet. Evol. 6, 46–62.

Buitrago, R., Waleckx, E., Bosseno, M.F., Zoveda, F., Vidaurre, P., Salas, R., Mamani, E., Noireau, F., Breniere, S.F., 2010. First report of widespread wild populations of *Triatoma infestans* (Reduviidae, Triatominae) in the valleys of La Paz, Bolivia. Am. J. Trop. Med. Hyg. 82, 574–579.

Calleros, L., Panzera, F., Bargues, M.D., Monteiro, F.A., Klisiowicz, D.R., Zuriaga, M.A., Mas-Coma, S., Pérez, R., 2010. Systematics of *Mepraia* (Hemiptera-Reduviidae): cytogenetic and molecular variation. Infect. Genet. Evol. 10, 221–228.

Ceballos, L.A., Piccinali, R.V., Berkunsky, I., Kitron, U., Gurtler, R.E., 2009. First finding of melanic sylvatic *Triatoma infestans* (Hemiptera: Reduviidae) colonies in the Argentine Chaco. J. Med. Entomol. 46, 1195–1202.

Cortez, M.R., Monteiro, F.A., Noireau, F., 2010. New insights on the spread of *Triatoma infestans* from Bolivia-Implications for Chagas disease emergence in the Southern Cone. Infect. Genet. Evol. 10, 350–353.

Dillehay, T.D., Collins, M.B., 1988. Early cultural evidence from Monte Verde in Chile. Nature 332, 150–152.

Dixon, E.J., 2001. Human colonization of the Americas: timing, technology and process. Q. Sci. Rev. 20, 277–299.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

Drummond, A.J., Rambaut, A., Shapiro, B., Pybus, O.G., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol. Biol. Evol. 22, 1185–1192.

Dujardin, J.P., Schofield, C.J., Tibayrenc, M., 1998. Population structure of Andean *Triatoma infestans*: allozyme frequencies and their epidemiological relevance. Med. Vet. Entomol. 12, 20–29.

Dujardin, J.P., Tibayrenc, M., Venegas, E., Maldonado, L., Desjeux, P., Ayala, F.J., 1987. Isozyme evidence of lack of speciation between wild and domestic *Triatoma infestans* (Heteroptera: Reduviidae) in Bolivia. J. Med. Entomol. 24, 40–45.

Dupanloup, I., Schneider, S., Excoffier, L., 2002. A simulated annealing approach to define the genetic structure of populations. Mol. Ecol. 11, 2571–2581.

Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol. Bioinform. Online 1081, 39–56.

Frias, D.A., Henry, A.A., Gonzalez, C.R., 1998. *Mepraia gajardoi:* a new species of Triatominae (Hemiptera: Reduviidae) from Chile and its comparison with *Mepraia spinolai*. Rev. Chil. Hist. Nat. 71, 177–188.

Fu, Y.X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147, 915–925.

Giordano, R., Cortez, J.C.P., Paulk, S., Stevens, L., 2005. Genetic diversity of *Triatoma infestans* (Hemiptera: Reduviidae) in Chuquisaca, Bolivia based on the mitochondrial cytochrome b gene. Mem. Inst. Oswaldo Cruz 100, 753–760.

Gorla, D.E., Schofield, C.J., 1989. Population dynamics of *Triatoma infestans* under natural climatic conditions in the Argentine Chaco. Med. Vet. Entomol. 3, 179–194.

Gurtler, R.E., Diotaiuti, L., Kitron, U., 2008. Commentary: Chagas disease: 100 years since discovery and lessons for the future. Int. J. Epidemiol. 37, 698–701.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Kingman, J.F., 1982. The coalescent. Stoch. Process. Appl. 13, 235–248.

Knowles, L.L., 2001. Did the pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshopprers. Mol. Ecol. 10, 691–701.

Knowles, L.L., 2004. The burgeoning field of statistical phylogeography. J. Evol. Biol. 17, 1–10.

Knowles, L.L., 2009. Statistical phylogeography. Annu. Rev. Ecol. Evol. Syst..

Knowles, L.L., Maddison, W.P., 2002. Statistical phylogeography. Mol. Ecol. 11, 2623–2635.

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

Lloyd-Smith, J.O., George, D., Pepin, K.M., Pitzer, V.E., Pulliam, J.R., Dobson, A.P., Hudson, P.J., Grenfell, B.T., 2009. Epidemic dynamics at the human–animal interface. Science 326, 1362–1367.

Maddison, W.P., Knowles, L.L., 2006. Inferring phylogeny despite incomplete lineage sorting. Syst. Biol. 55, 21–30.

Maddison, W.P., Maddison, D.R., 2009. Mesquite: A Modular System for Evolutionary Analysis. , http://mesquiteproject.org.

Marcet, P.L., Mora, M.S., Cutrera, A.P., Jones, L., Gurtler, R.E., Kitron, U., Dotson, E.M., 2008. Genetic structure of *Triatoma infestans* populations in rural communities of Santiago del Estero, northern Argentina. Infect. Genet. Evol. 8, 835–846.

Marcilla, A., Bargues, M.D., Ramsey, J.M., Magallon-Gastelum, E., Salazar-Schettino, P.M., Abad-Franch, F., Dujardin, J.P., Schofield, C.J., Mas-Coma, S., 2001. The ITS-2 of the nuclear rDNA as a molecular marker for populations, species, and phylogenetic relationships in Triatominae (Hemiptera: Reduviidae), vectors of Chagas disease. Mol. Phylogenet. Evol. 18, 136–142.

Miller, M.P., 2005. Alleles In Space (AIS): computer software for the joint analysis of interindividual spatial and genetic information. J. Hered. 96, 722–724.

Miller, M.P., Bellinger, M.R., Forsman, E.D., Haig, S.M., 2006. Effects of historical climate change, habitat connectivity, and vicariance on genetic structure and diversity across the range of the red tree vole (*Phenacomys longicaudus*) in the Pacific Northwestern United States. Mol. Ecol. 15, 145–159.

Moncayo, A., Silveira, A.C., 2009. Current epidemiological trends for Chagas disease in Latin America and future challenges in epidemiology, surveillance and health policy. Mem. Inst. Oswaldo Cruz 104, 17–30.

Monteiro, F.A., Perez, R., Panzera, F., Dujardin, J.P., Galvao, C., Rocha, D., Noireau, F., Schofield, C., Beard, C.B., 1999. Mitochondrial DNA variation of *Triatoma infestans* populations and its implication on the specific status of *T. melanosoma*. Mem. Inst. Oswaldo Cruz 94, 229–238.

Noireau, F., 2009. Wild *Triatoma infestans*, a potential threat that needs to be monitored. Mem. Inst. Oswaldo Cruz 104, 60–64.

Noireau, F., Cortez, M.G.R., Monteiro, F.A., Jansen, A.M., Torrico, F., 2005. Can wild *Triatoma infestans* foci in Bolivia jeopardize Chagas disease control efforts? Trends Parasitol. 21, 7–10.

Panzera, F., Dujardin, J.P., Nicolini, P., Caraccio, M.N., Rose, V., Tellez, T., Bermudez, H., Bargues, M.D., Mas-Coma, S., O'Connor, J.E., Perez, R., 2004. Genomic changes of Chagas disease vector, South America. Emerg. Infect. Dis. 10, 438–446.

Panzera, F., Pérez, R., Lucero, C., Ferrandis, I., Ferreiro, M.J., Calleros, L., Romero, V., 2007. Cambios genómicos en la subfamilia Triatominae, con énfasis en *Triatoma infestans*. In: Rojas, M. (Ed.), Triatominos de Bolivia y la enfermedad de Chagas. Ministerio de Salud y Deportes. Programa Nacional de Chagas, Bolivia, pp. 149–172.

Piccinali, R.V., Marcet, P.L., Noireau, F., Kitron, U., Gurtler, R.E., Dotson, E.M., 2009. Molecular population genetics and phylogeography of the chagas disease vector *Triatoma infestans* in South America. J. Med. Entomol. 46, 796–809.

Posada, D., 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256.

Posada, D., Crandall, K.A., 2001. Intraspecific gene genealogies: trees grafting into networks. Trends Ecol. Evol. 16, 37–45.

Rambaut, A., Drummond, A.J., 2009. Tracer v1.5. , Available from:http://beast.bio.ed.ac.uk/Tracer.

Ramos-Onsins, S.E., Rozas, J., 2002. Statistical properties of new neutrality tests against population growth. Mol. Biol. Evol. 19, 2092–2100.

Richards, C.L., Carstens, B.C., Knowles, L.L., 2007. Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. J. Biogeogr. 34, 1833–1845.

Rojas, M., Emperaire, L., Piccinali, R.V., Gurtler, R.E., Torrico, F., Jansen, A.M., Noireau, F., 2007. Sylvatic *Triatoma infestans* (Reduviidae, Triatominae) in the Andean valleys of Bolivia. Acta Trop. 102, 47–54.

Schofield, C.J., 1988. Biosystematics of the Triatominae. In: Service, M.W. (Ed.), Biosystematic of Haematophagous Insects. Clarendon Press, Oxford, pp. 284–312.

Schofield, C.J., 1992. Eradication of *Triatoma infestans*: a new regional programme for southern Latin America. Ann. Soc. Belge Med. Trop. 72 (Suppl. 1), 69–70.

Schofield, C.J., Apt, W., Miles, M.A., 1982. The ecology of Chagas disease in Chile. Ecol. Dis. 1, 117–129.

Schofield, C.J., Dias, J.C., 1999. The Southern Cone Initiative against Chagas disease. Adv. Parasitol. 42, 1–27.

Schofield, C.J., Jannin, J., Salvatella, R., 2006. The future of Chagas disease control. Trends Parasitol. 22, 583–588.

Segura, E.L., Torres, A.G., Fusco, O., Garcia, B.A., 2009. Mitochondrial 16S DNA variation in populations of *Triatoma infestans* from Argentina. Med. Vet. Entomol. 23, 34–40.

Suchard, M.A., Weiss, R.E., Sinsheimer, J.S., 2001. Bayesian selection of continuous-time Markov chain evolutionary models. Mol. Biol. Evol. 18, 1001–1013.

Tajima, F., 1989. The effect of change in population size on DNA polymorphism. Genetics 123, 597–601.

Torrico, R.A., 1946. Hallazgo de *Eratyrus mucronatus*, infestacion natural de vinchucas de cerro y *Eutriatoma sordida* en Cochabamba. Ann. Lab. Central 1, 19–23.

Usinger, R.L., Wygodzinsky, P., Ryckman, R.E., 1966. The biosystematics of Triatominae. Annu. Rev. Entomol. 11, 309–330.