Population genetic structure of Meccus longipennis (Hemiptera, Reduviidae, Triatominae), vector of Chagas disease in West Mexico

Simone Frédérique Brenière, Etienne Waleckx, Ezequiel Magallón-Gastélum, Marie-France Bosseno, Xavier Hardy, Cyrille Ndo, Felipe Lozano-Kasten, Christian Barnabé, Pierre Kengne

Infection, Genetics and Evolution 12 (2012) 254–262

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

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

Population genetic structure of *Meccus longipennis* (Hemiptera, Reduviidae, Triatominae), vector of Chagas disease in West Mexico

Simone Frédérique Brenière [a,*], Etienne Waleckx [a], Ezequiel Magallón-Gastélum [b], Marie-France Bosseno [a], Xavier Hardy [a,c], Cyrille Ndo [d], Felipe Lozano-Kasten [b], Christian Barnabé [a], Pierre Kengne [a]

[a] MIVEGEC, Université de Montpellier 1 et 2, CNRS 5290, IRD 224, Maladies Infectieuses et Vecteurs: Ecologie, Génétique, Evolution et Contrôle, Institut de Recherche pour le Développement (IRD), Representation in Bolivia, Av Hernando Siles No. 5290, Esq Calle 7 Obrajes, CP 9214, La Paz, Bolivia
[b] Universidad de Guadalajara, Centro Universitario de Ciencias de la Salud, Instituto Regional de Investigación en Salud Pública, AP 2-136, Jalisco, Mexico
[c] Ecole des hautes études en santé publique (EHESP), Avenue du Professeur Léon-Bernard, CS 74312, 35043 Rennes, France
[d] Organisation de Coordination pour la lutte Contre les Endémies en Afrique Centrale (OCEAC), Laboratoire de Recherche sur le Paludisme, P.O. Box 288, Yaoundé, Cameroon

ARTICLE INFO

Article history:
Received 12 May 2011
Received in revised form 20 October 2011
Accepted 8 November 2011
Available online 28 November 2011

Keywords:
Chagas disease
Triatomine
Mexico
Population genetics
Microsatellite

ABSTRACT

The originally wild species of the *Meccus* complex are important vectors of Chagas disease in Mexico. In West Mexico, *Meccus longipennis* plays an important epidemiological role. To understand the genetic structure of the domestic and wild populations of this species, a preliminary study with five polymorphic microsatellite loci was conducted. The population genetics analysis showed high structuring between peridomestic biotopes, with breeding subunits detected in a single peridomestic structure. In the wild environment, two genetic patterns were observed according to the biotope, possible breeding subunits in large rocky formations and a larger panmictic unit in agropastoral areas, suggesting considerable dispersal of bugs in this biotope. Moreover, the discovery of two foci of wild populations at the edge of Guadalajara city raises the question of new urban areas where the phenomenon of bug incursions into households could constitute a risk of transmission of Chagas disease.

1. Introduction

In Mexico, Chagas disease is endemic in various regions where numerous species of triatomines are vectors, but a national program to control vectorial transmission of the disease does not exist yet. In West Mexico, *Meccus longipennis*, is the main species of the *Meccus* complex responsible for the transmission of *Trypanosoma cruzi*, the agent of the disease (Lozano-Kasten et al., 2008). Previous studies have shown that this primary sylvatic species can heavily colonize peridomestic areas (outdoors), and that inhabitants repeatedly found bugs in their dwellings (indoors) (Brenière et al., 2007, 2010; Espinoza-Gomez et al., 2002; Magallon-Gastelum et al., 1998, 2006; Martinez-Ibarra et al., 2008). It was observed in several parts of West and Central Mexico that the predominant species of the *Meccus* complex infesting the villages were different (e.g., *Triatoma longipennis*, *Triatoma picturata*, *Triatoma pallidipennis*); in each region, the same principal species was found in the villages studied and in their immediate sylvatic environment; this observation supports the relationship between domestic and sylvatic environments and the sylvatic origin of the infestation in dwellings (Magallon-Gastelum et al., 2001, 2004; Ramsey et al., 2005). In general, indoor infestation evaluated by conventional active research (man-hours) is low, but of the search for triatomines monitored by inhabitants over time shows that there is a regular penetration of bugs indoors, and sometimes small triatomine colonies can be found (Brenière et al., 2010; Martinez-Ibarra et al., 2008). Triatomines found in dwellings could have an outdoor origin when peridomestic structures are heavily infested, but another alternative may be the incursion of insects from the surroundings (sylvatic populations of triatomines) (Brenière et al., 2010). The movements of vectors between the sylvatic environment and human habitats are poorly known, even though this knowledge is important for the implementation of vector control strategies.

However, this topic has been addressed for some triatomine species by analyzing various genetic markers such as isoenzymes, nucleotide variation of DNA sequences, and microsatellite polymorphism. For *Triatoma infestans*, the principal vector in the Southern Cone countries of South America, a high level of structuring was observed among populations from different villages (domestic and peridomestic) and even at the micro-geographical scale between populations of different sites within the same village (Brenière et al., 1998; Quisberth et al., 2011; Pizarro et al., 2008; Piccinali et al., 2009; Perez de Rosas et al., 2007). In contrast, for *Triatoma dimidiata*, the main vector in Yucatan (Mexico), poor differentiation was observed between villages and between

* Corresponding author. Tel.: +591 2 278 29 69; fax: +591 2 278 29 44.
E-mail address: Frederique.Breniere@ird.fr (S.F. Brenière).

1567-1348/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2011.11.003

populations from different habitats (sylvatic, peridomestic and domestic) in the same village (Dumonteil et al., 2007). Similarly for *Triatoma pseudomaculata* in the northeastern state of Bahia, a species with a domiciliation trend, the results support the lack of structuring between sylvatic and peridomestic populations (Carvalho-Costa et al., 2010). For *Rhodnius prolixus* an important vector in Colombia, domestic populations appeared weakly differentiated whereas sylvatic populations are clustered apart from the domestic populations (Lopez et al., 2007). These examples show that the study of the population structure of triatomines and the relationships between populations of different biotopes (i.e., an area where environmental conditions are more or less uniform) differ between species.

In order to reach an understanding of the degree of panmixia of the *M. longipennis* triatomine populations collected indoors, outdoors, in an agropastoral area, and in other wild areas, and to assess the level of genetic differentiation between populations on different geographical scales, genetic variation at microsatellite DNA loci was examined. Briefly, according to the biotope, analyzes suggested deviation from Hardy–Weinberg expectations within populations manifested by a deficit of heterozygous, strong structuring between populations, even when they are geographically close, or gene flow between certain sylvatic populations a few hundred meters apart.

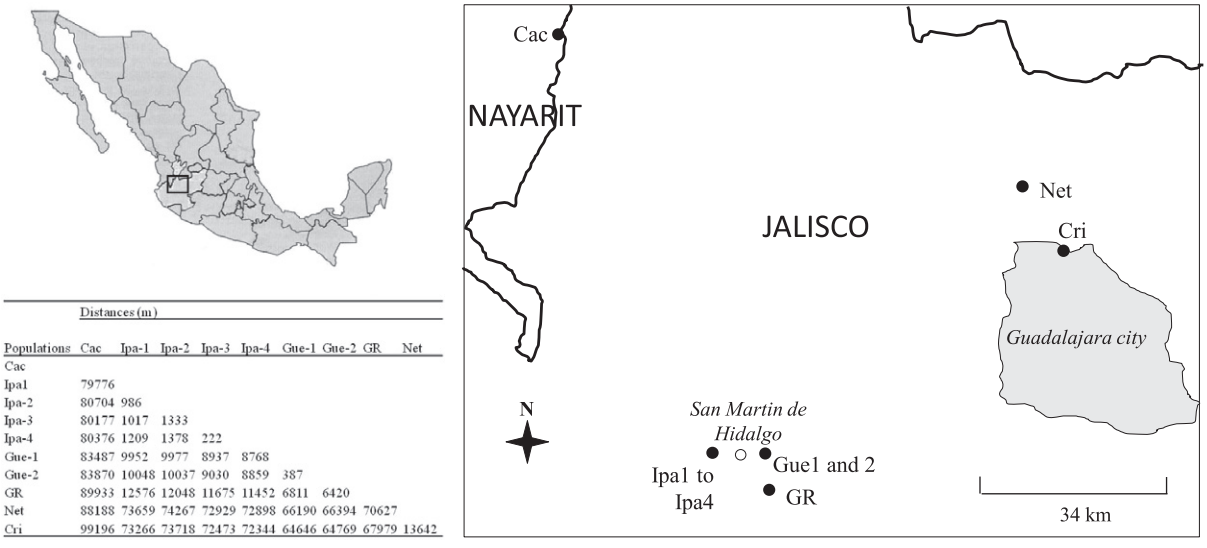
2. Material and methods

2.1. The geographic origin of M. longipennis samples

Ten samples of *M. longipennis* were collected in Western Mexico (Fig. 1, Table 1) in domestic, peridomestic, and sylvatic environments by the inhabitants themselves or by a team of professionals through active search indoors and outdoors or by using mice bait-traps in the sylvatic environments as previously described (Bosseno et al., 2009). The sample from the village of Cacalutan (Cac), whose characteristics were previously described (Brenière et al., 2010), was composed of individuals randomly taken from the full collection carried out by the inhabitants, coming from 19 different houses. The villages of Ipazoltic, Santa Cruz de Las Flores, and Los Guerrero, whose characteristics were described previously, are located in the Ameca valley close to the San Martin de Hidalgo village (Brenière et al., 2010). Samples Ipazoltic 1–4 (Ipa1–4) were collected from rock pile boundary walls separating fields of corn (*Zea mays*), Maguey tequilero (*Agave tequilana*), and grasslands around Ipazoltic by using mice bait-traps placed in hollows formed by stones approximately 100 m long; these fields are located at intervals of about 220–1380 m. The "Grandes Rocas" (GR) sample was collected in a large rocky formation at the top of a small hill in the valley, approximately 650 m from the first houses of Santa Cruz de Las Flores. The samples from Los Guerrero 1 and 2 (Gue1 and Gue2) were collected manually by professionals in two different peridomiciles in this village, in piles of bricks. These sites 387 m apart, were characterized by very large colonies of more than 200 individuals each. Two other samples were collected from two areas close to Guadalajara city. The first area was located about 10 km north of the city; the corresponding sample, called Internet (Net), was composed of bugs collected indoors (by inhabitants) and outdoors (by professionals) of an isolated inhabited house, and from a ruin on the hillside located at 100 m from the house. The second area was a kennel at the north side of the city along a geological fault that houses a residential area; the corresponding sample, called "Criadero" (Cri), consisted of bugs manually collected in the dog cages (by inhabitants and professionals) and in large boulders located near the cages (within a range of 20–50 m). This rocky formation is very similar to that where the GR sample was collected. In the laboratory, the legs of each bug were removed and preserved in 70% ethanol for DNA extraction.

2.2. Microsatellite genotyping

DNA was extracted from two bug legs using the slightly modified CTAB-chloroform method (Edwards, 1998). Briefly, legs were crushed in 200 µl of 2% CTAB (cetyl trimethyl ammonium bromide) and incubated overnight at 37 °C after the addition of 30 µl of proteinase K (20 mg/ml). The extraction was then performed as described previously (Edwards, 1998). The final DNA pellet was suspended in 20 µl of distilled water. Eight heterologous sets of primers identified for *T. dimidiata* microsatellite loci (TDMS1, TDMS3, TDMS4, TDMS6, TDMS9, TDMS11, TDMS19, TDMS22, (Anderson et al., 2002) were used in polymerase chain reaction (PCR) using *M. longipennis* DNAs. A forward primer of each set was the 5′ end labeled with one of the three fluorescent dyes,



Distances (m)

Populations	Cac	Ipa-1	Ipa-2	Ipa-3	Ipa-4	Gue-1	Gue-2	GR	Net
Cac									
Ipa1	79776								
Ipa-2	80704	986							
Ipa-3	80177	1017	1333						
Ipa-4	80376	1209	1378	222					
Gue-1	83487	9952	9977	8937	8768				
Gue-2	83870	10048	10037	9030	8859	387			
GR	89933	12576	12048	11675	11452	6811	6420		
Net	88188	73659	74267	72929	72898	66190	66394	70627	
Cri	99196	73266	73718	72473	72344	64646	64769	67979	13642

Fig. 1. Geographical origin of the 10 *M. longipennis* samples. On the left, a map of Mexico where the study area is indicated (small rectangle). On the right, the localization of the samples is indicated by black circles: Cac was from the state of Nayarit, the others were from the state of Jalisco, seven were collected near the San Martin de Hidalgo village (white circle) and two near Guadalajara city.

Table 1
Geographic origin and characteristics of *M. longipennis* samples collected in West Mexico.

State	Municipality	Nearest locality	Sample code	Latitude (N)	Longitude (W)	Sample size	Collection		
							Date	Method	Biotope
Nayarit	Ixtlán del Río	Cacalutan	Cac	21°06′98.8″	104°15′69.7″	27	February 2005–September 2006	By inhabitants	Indoor and outdoor
Jalisco	San Martin de Hidalgo	Ipazoltic	Ipa1	20°26′26.50″	103°59′36.80″	25	July 2007	By traps	Agropastoral
Jalisco	San Martin de Hidalgo	Ipazoltic	Ipa2	20°25′54.35″	103°59′35.00″	14	July 2007	By traps	Agropastoral
Jalisco	San Martin de Hidalgo	Ipazoltic	Ipa3	20°26′24.57″	103°59′1.61″	24	July 2007	By traps	Agropastoral
Jalisco	San Martin de Hidalgo	Ipazoltic	Ipa4	20°26′19.92″	103°58′55.69″	20	July 2007	By traps	Agropastoral
Jalisco	San Martin de Hidalgo	Los Guerrero	Gue1	20°26′36.05″	103°53′53.39″	9	July 2003	Active search by professionals	Outdoor
Jalisco	San Martin de Hidalgo	Los Guerrero	Gue2	20°26′23.76″	103°53′49.83″	28	July 2003	Active search by professionals	Outdoor
Jalisco	San Martin de Hidalgo	Santa Cruz de las Flores	GR	20°22′56.02″	103°53′24.47″	21	June 2006 and June 2007	By traps	Rocky formation
Jalisco	Zapopan	Tesistan-Guadalajara	Net	20°52′18.54″	103°27′13.44″	26	August 2007	Active search by professionals and by traps	Indoor, outdoor, and ruins
Jalisco	Zapopan	Guadalajara	Cri	20°46′6.14″	103°22′55.27″	16	July and August 2007	Active search by professionals and by inhabitant	Outdoor and rocky formation

FAM, NED, or VIC (Applied Biosystems, France). PCR amplifications were performed in a 25 μl solution containing 1.5 mM $MgCl_2$, 200 μM each dNTP, 2 pM of each primer (forward-labeled), 1 U of Taq polymerase (Quiagen, Courteboeuf, France), 1X Taq polymerase buffer, and 20–25 ng template DNA. Amplifications were carried out in a Perkin–Elmer 9600 thermal cycler, using the following reaction conditions: 94 °C for 5 min, 35 cycles of 94 °C for 30 s, 50 °C for 30 s, and 72 °C for 30 s, and a final step of 72 °C for 15 min. Amplified products were visualized after electrophoresis on a 2.5% agarose gel. Alleles were sized relative to GeneScan 500 LIZ internal size standards in a 3130XL Genetic Analyzer (Applied Biosystems, Villebon-sur-yvette, France) using the GeneMapper software v. 4.0. In the absence of a repeated signal (twice), the allele was considered as null allele.

2.3. Data analysis

The significance of genotype association between pairs of loci in the overall sample (linkage disequilibrium) was tested using the Fstat 2.9.3.2 software (Goudet, 2001). The *P*-value of the test was obtained as follows: genotypes at two loci were associated at random a 1500 times, and the log-likelihood *G*-test statistic was recalculated on the randomized data. The *P*-value was estimated as the proportion of statistics from randomized data sets that are larger or equal to that observed.

Analysis of genetic variability per loci, per sample, and overall consisted of general descriptive indexes. The number of alleles, unbiased expected heterozygosity (UHe), and observed heterozygosity (Ho) were estimated with Genalex 6.3 xla software (Peakall and Smouse, 2006). Allelic richness (aRich) which is a measure of the number of alleles independent of sample size, was estimated with Fstat (Goudet, 1995). To assess levels of structuring on overall samples, Wright's *F*-statistics (Wright, 1965) were estimated using Weir and Cockerham's unbiased estimators (Weir and Cockerham, 1984) in *F*stat: to explore deviations from Hardy–Weinberg (HW) expectations within samples and each locus, the F_{IS} index, which measures the heterozygote deficit or excess, was tested after 6000 random allele permutations among individuals within the samples. The 95% confidence interval of F_{IS} values per sample was estimated by bootstrapping over loci using the GENETIX v. 4.05.2 package (Belkhir et al., 2001).The level of genetic structuring was estimated through different hierarchical analyzes of molecular variance (AMOVA), with the Arlequin software v. 3.5.1.2 (Excoffier et al., 2005). The pairwise F_{ST} can be used as short-term genetic distances between samples, which were computed with Arlequin; differentiation between samples was tested by permutation procedures of multilocus genotypes between populations (1000 permutations). The genetic relationships between samples were inferred using the neighbor-joining method (NJ) (Saitou and Nei, 1987) using F_{ST} values between pairwise samples as genetic distance. This analysis was conducted in MEGA v. 4.0.2. (Tamura et al., 2007).

A Bayesian approach was implemented to determine the number of genetic clusters within the entire data set using STRUCTURE v. 2.2 (Falush and Pritchard, 2007; Pritchard et al., 2000). Several runs with a burn-in period of 50,000 iterations and 50,000 in length, with variable genetic clusters ($k = 2–10$) were carried out to identify the best assignment; this procedure places individuals in *k* clusters chosen in advance and membership coefficients are calculated for each individual in inferred clusters.

The "admixture" (individuals may have ancestry from various populations) and "no admixture" (each individual comes from one of the *k* populations) models for the ancestry of individuals were applied to the set of data with the "Allele Frequencies Correlated" option (meaning that allele frequencies in the different populations are most probably similar), the other parameters being left at their default values. The "no admixture" model was then selected because estimations of data probability expressed by the Ln *P* (D) values were always smaller for this model than for the "admixture model". The "no admixture model" it is often more powerful than the "admixture" model at determining fine structure, as may be expected the present sample between distant populations. For these runs, the LOCPRIOR option was used, which allows clustering by using sampling locations of individuals as prior information. This option is more appropriate for data sets with few markers and small sample sizes. Five independent runs for each *k* were produced to assess the consistence of the results across runs that are one of the criteria to estimate the best *k* (Pritchard et al., 2000). Moreover, the authors of the STRUCTURE program have indicated that it makes sense to focus on *k* values that capture most of the structure (by looking for the bar plot obtained for each run).

3. Results

Of the eight published sets of microsatellite primers for *T. dimidiata*, two were discarded because of inconsistent amplification with *M. longipennis*. Data from the remaining six microsatellite loci were analyzed for a total of 206 *M. longipennis* specimens. The overall amplification success of loci was 92.9% ranging from 89% (TDMS04 and TDMS11) to 98.1% (TDMS22). Out of the six loci, one was monomorphic (TDMS03). No significant association between the pairs of locus genotypes within the overall sample and within each sample was found, a result supporting the statistical independence of loci. A summary of the genetic variability per locus and sample is presented in Table 2. The mean allelic richness per locus over the samples was 3.004 (range, 1.672 for TDMS04 to 4.980 for TDMS19). Significant deviations from HW expectations due to heterozygote deficit (positive F_{IS} values that significantly differed from random predictions) were observed for all polymorphic loci except TDMS22 when all samples were considered together as a single panmictic unit, suggesting a strong Wahlund effect. When each sample was analyzed separately, five of the ten samples studied (Cac, Net, Cri, Ipa3, and GR) presented a significant deviation from HW expectations due to a heterozygote deficit when analyzed overall loci, (Table 2, Fig. 2). The high rate of null alleles in Cri and Ipa3 samples (13.3% in both samples) mainly due to the TDMS11 locus, could be partly responsible for this deficit; however, positive F_{IS} values significantly different from zero

Table 2
Summary of the genetic variability for 6 microsatellite loci in different samples of *M. longipennis*.

Locus	Sample code										Av. overall populations
	Cac	Ipa1	Ipa2	Ipa3	Ipa4	Gue1	Gue2	GR	Net	Cri	
TDMS03											
N	27	25	13	21	18	7	27	17	26	14	19.5
Nall	1	1	1	1	1	1	1	1	1	1	1
aRich	1	1	1	1	1	1	1	1	1	1	1
F_{IS}	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
UHe	0	0	0	0	0	0	0	0	0	0	0
Ho	0	0	0	0	0	0	0	0	0	0	0
TDMS04											
N	27	22	14	22	14	5	21	18	26	15	18.4
Nall	3	2	2	3	2	1	1	2	2	2	1.8
aRich	1.996	1.407	1.357	2.216	1.595	1	1	1.746	1.351	1.994	1.672
F_{IS}	**0.846***	NA	NA	**1.000***	1.000	NA	NA	1.000	1.000	**1.000***	**0.965****
UHe	0.24	0.09	0.07	0.32	0.14	0	0	0.32	0.07	0.54	0.18
Ho	0.04	0.09	0.07	0	0	0	0	0.17	0	0.27	0.06
TDMS09											
N	27	23	14	24	17	8	28	19	26	16	20.2
Nall	5	4	4	4	5	4	4	4	5	5	4.4
aRich	3.555	3.159	3.514	3.288	3.812	3.499	3.206	3.403	3.548	4.042	3.555
F_{IS}	**0.508***	0.366	−0.017	0.263	−0.052	0.462	0.418	−0.083	0.269	0.222	**0.260****
UHe	0.67	0.68	0.63	0.62	0.67	0.67	0.67	0.63	0.68	0.72	0.66
Ho	0.33	0.43	0.64	0.46	0.71	0.37	0.39	0.68	0.5	0.56	0.51
TDMS19											
N	26	23	14	21	17	8	23	18	24	14	18.8
Nall	11	12	5	8	4	4	8	7	6	7	7.2
aRich	5.55	5.204	3.689	4.468	3.171	3.615	3.567	5.134	4.418	4.409	4.98
F_{IS}	0.272	−0.072	−0.172	−0.085	−0.018	0.341	−0.146	**0.535***	0.288	0.341	**0.140****
UHe	0.84	0.77	0.68	0.66	0.52	0.74	0.57	0.82	0.76	0.75	0.71
Ho	0.61	0.83	0.79	0.71	0.53	0.5	0.65	0.39	0.54	0.5	0.61
TDMS11											
N	25	25	14	16	17	9	27	18	25	9	18.5
Nall	6	9	6	8	6	5	5	6	6	5	6.2
aRich	2.459	5.103	3.97	5.023	3.648	4.188	4.023	4.042	3.112	3.484	4.598
F_{IS}	0.186	0.177	0.517	**0.467***	0.232	0.273	**0.460***	**0.572***	0.101	0.556	**0.346****
UHe	0.29	0.83	0.72	0.81	0.68	0.75	0.75	0.64	0.44	0.48	0.64
Ho	0.24	0.68	0.36	0.44	0.53	0.56	0.41	0.28	0.4	0.22	0.41
TDMS22											
N	27	25	14	24	18	9	26	19	26	16	20.4
Nall	3	3	3	3	3	2	3	3	3	2	2.8
aRich	2.206	2.417	2.27	1.721	1.968	1.817	2.657	2.334	2.149	1.69	2.221
F_{IS}	−0.081	−0.074	−0.191	−0.051	−0.063	−0.067	−0.262	0.196	0.112	−0.071	−0.074
UHe	0.34	0.37	0.36	0.16	0.21	0.21	0.55	0.33	0.3	0.17	0.3
Ho	0.37	0.4	0.43	0.17	0.22	0.22	0.69	0.26	0.27	0.19	0.32
Overall loci											Av overall loci and populations
Av. N	26.5	23.8	13.8	21.3	16.8	7.7	25.3	18.17	25.5	14	19.3
Av. NA	4.83	5.17	3.5	4.5	3.5	2.83	3.67	3.83	3.83	3.83	3.95
Av. aRich	2.794	2.980	2.574	2.953	2.532	2.520	2.575	2.943	2.596	2.770	3.004
Av. F_{IS}	**0.335***	0.119	0.078	**0.313***	0.110	0.319	0.158	**0.392***	**0.246***	**0.441***	**0.249****
Av. UHe	0.39	0.46	0.41	0.43	0.37	0.4	0.42	0.46	0.38	0.44	0.42
Av. Ho	0.27	0.4	0.38	0.3	0.33	0.28	0.36	0.3	0.28	0.29	0.32

N = number of individuals; Nall = number of alleles; aRich = Allelic richness; Uhe and Ho = expected and observed heterozygosity; Av = average.
NA = not applicable.
**Significant deviation from Hardy–Weinberg expectations ($P \leqslant 0.002$).
*Significant deviation from Hardy–Weinberg expectations ($P \leqslant 0.05$).

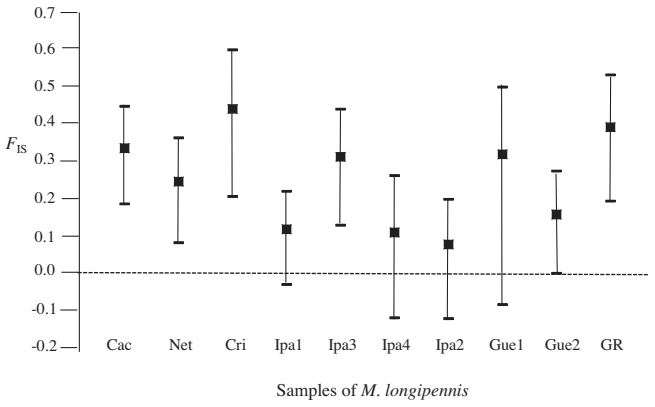


Fig. 2. F_{IS} values and their 95% confidence intervals obtained by bootstrap over loci, for the 10 *M. longipennis* samples under study and collected in West Mexico.

were also obtained after this locus was discarded (data not shown). For Cac, Net, and GR samples, only 2–5% null alleles were observed on all loci. For the other samples (Ipa1, Ipa2, Ipa4, Gue1, and Gue2), no significant deviation from HW expectations were observed. Moreover, the F_{IS} value calculated after grouping the four sylvatic samples around the locality of Ipazoltic (Ipa1–4) was not significant (0.173; $P > 0.05$). In order to better understand whether HW deviation within samples could have been induced by mistaken grouping of discrete populations of bugs, subsamples were selected according to time and space collection; Fig. 3 shows the F_{IS} values obtained for the Cac, Net, Cri, Ipa3, and GR subsamples. For Cac and GR, the subsamples corresponded to different collection times; the F_{IS} values were slightly modified, and in some cases no significant deviation from HW expectations were observed (Cac 2006, Cac April 2005, GR 2007). For the Net and Ipa3 subsamples, defined according to spatial localization, in most of the cases, no significant deviations from HW expectations were observed. Interestingly, the Net outdoor subsample (bugs collected in different sites around the house) had a significant positive F_{IS} value, but when the bugs collected in the enclosed patio behind the house where dogs are kept were separately analyzed (Net outdoor dog subsample), no significant deviation from HW expectations was observed. The first Cri subsample was composed of six bugs collected in different dog cages, and the second one was composed of 10 bugs collected in the large rocky formation located in the the breeder's field; significant deviations from HW expectations were observed for the two subsamples showing evidence of substructure. Genetic differentiation between subsamples (within each sample) was evaluated by calculating F_{ST}, but no F_{ST} between pairwise subsamples was significant.

This last result allowed us to group the subsamples and to examine the genetic structure among samples with the AMOVA procedure using all loci for distance computation and all data, for 0.12% of missing data. Most of the genetic variation was assigned to differences within samples (94.3% of the total variation, Table 3), but the low percentage of variation attributed to differences between samples (5.69%) was significant, showing some genetic differentiation among samples ($P < 10^{-4}$). To understand the genetic structure better, sample pairwise comparisons were performed

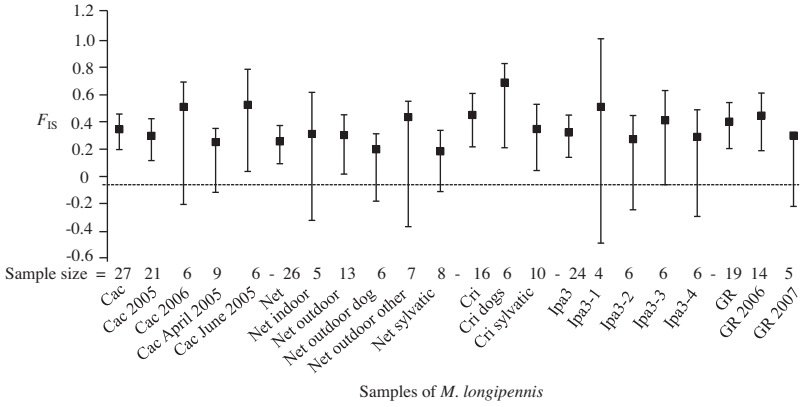


Fig. 3. F_{IS} values and their 95% confidence intervals obtained by bootstrap over loci, for subsamples of five *M. longipennis* samples that presented a heterozygote deficit. The Cac and GR subsamples were defined according to the collection time. The Net subsamples were defined according to the biotopes (Net indoor, Net outdoor, Net sylvatic, and the spatial origin in the case of additional subdivision of the Net outdoor subsample (Net outdoor dog and Net outdoor others). The four Ipa3 subsamples were defined according to the spatial origin by sequentially grouping the bugs captured by the traps placed along the rock pile boundary walls (Ipa3–1–Ipa3–4).

Table 3
Molecular variance analysis (AMOVA) for *M. longipennis* samples at different hierarchical levels.

Hierarchical level	Source of variation	d. f.	Sum of squares	Variance components	% of variation	Fixation indices	
(a) All populations	Among populations	9	32.926	0.06177 Va	5.43	F_{ST} = 0.05432	$<10^{-4}$
	Within populations	412	443.084	1.07545 Vb	94.57		
(b) Geographic groups of populations	Among groups	1	8.574	0.04332 Va	3.70	F_{CT} = 0.03696	>0.05
	Among populations within groups	7	19.464	0.04335 Vb	3.71	F_{SC} = 0.03849	$<10^{-4}$
	Within populations	353	382.301	1.08301 Vc	92.60	F_{ST} = 0.07402	$<10^{-4}$
(c) Genetic groups of populations	Among groups	1	12.687	0.04750 Va	4.03	F_{CT} = 0.04035	<0.004
	Among populations within groups	8	21.528	0.03926 Vb	3.33	F_{SC} = 0.03475	$<10^{-4}$
	Within populations	406	442.783	1.09060 Vc	92.63	F_{ST} = 0.07369	$<10^{-4}$

(a) Overall populations; (b) group 1 = two samples close to Guadalajara city, Net and Cri; group 2, 7 samples located in the municipality of San Martin de Hidalgo, Ipa1–4, Gue1–2, and GR; (c) groups defined according to the topology on the NJ tree (Fig. 3): group 1 = Cac, Net, Cri, Gue1, and GR, group 2 = Ipa1–Ipa4 and Gue2.

Table 4
F_{ST} values between *M. longipennis* samples.

Samples	Cac	Net	Cri	Ipa1	Ipa2	Ipa3	Ipa4	Gue1	Gue2
Cac	–								
Net	0.021	–							
Cri	**0.075****	**0.063****	–						
Ipa1	**0.068****	**0.066*****	**0.049****	–					
Ipa2	**0.066****	**0.081*****	**0.078****	0.018	–				
Ipa3	**0.076*****	**0.090*****	**0.086*****	−0.007	0.003	–			
Ipa4	**0.068*****	**0.089*****	**0.072****	0.014	−0.012	−0.007	–		
Gue1	0.0478	0.039	−0.038	**0.054***	0.049	0.033	**0.097*****	–	
Gue2	**0.112*****	**0.115*****	**0.106*****	**0.035****	**0.042***	**0.051****	**0.046****	**0.084****	–
GR	**0.047***	**0.079*****	**0.045***	**0.035***	0.015	0.026	0.026	0.033	**0.095*****

Significant differentiation between samples (F_{ST} values in bold).
* $P < 0.05$.
** $P < 0.01$.
*** $P < 0.001$.

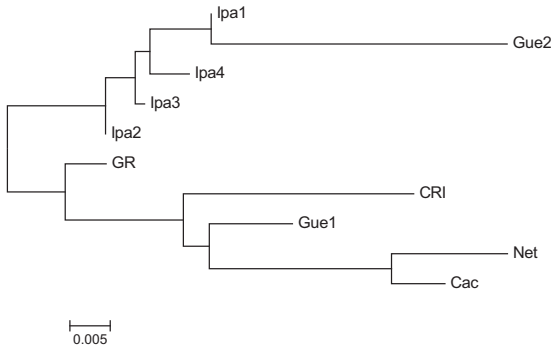


Fig. 4. Evolutionary relationships between samples of *M. longipennis* collected in West Mexico inferred by using the neighbor joining (NJ) method based on F_{ST} values as genetic distances. The tree is drawn to scale.

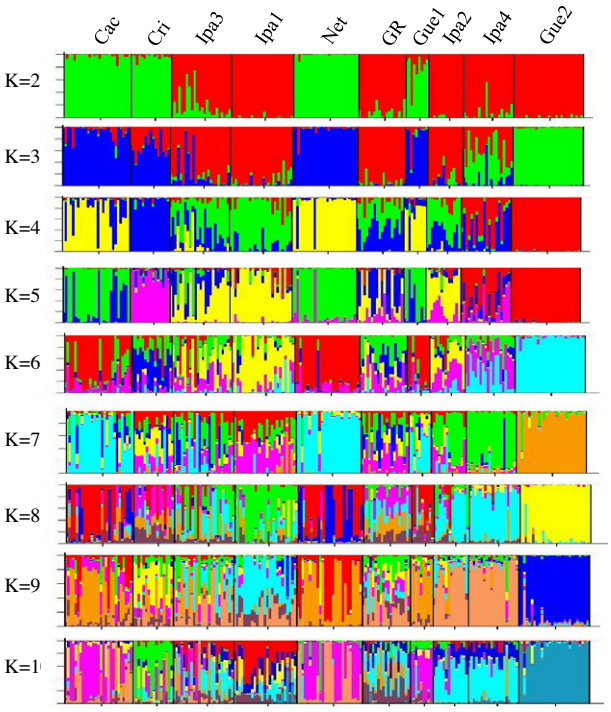


Fig. 5. Plots of ancestry estimates of *M. longipennis* individuals collected from 10 sites in West Mexico using the STRUCTURE software. The estimated membership coefficients for each individual, in each of the inferred clusters ($k = 2$–10) is plotted. Each individual in the data set is represented by a single vertical line, which is partitioned into different colored segments representing the individual membership estimate in each of the inferred clusters.

by calculating F_{ST}. Out of 45 pairwise F_{ST} comparisons, 29 (64.4%) were significant (Table 4). The Gue2, Cri, Net, and Cac samples differed significantly from most of the others. The two outdoor samples from Los Guerrero differed significantly ($F_{ST} = 0.084$, $P < 0.01$), while they were very close (~390 m). In contrast, the four samples of bugs captured in the rock pile boundary walls that separate the fields around Ipazoltic, 230–1390 m apart, were not differentiated (Table 4). In the neighbor-joining tree built from the F_{ST} values between samples (Fig. 4), the topology suggested the existence of two groups that did not fit with the geographical origin of the samples. Similarly, the correlation between geographical and genetic distances between samples was not significant suggesting that geographical distance was not the main force responsible for all the genetic differentiation observed (Mantel, $r = 0.520$, $P = 0.06$). Two additional hierarchical AMOVA were performed: the first one grouping populations according to their geographic origin, the second one grouping populations according to the observed NJ tree topology (Tables 3b and c). In both cases, the variance component distribution was higher within samples and the variability among groups was only 3.7% and 4.0% of the total variability, respectively. Interestingly, genetic differentiation between geographic groups was not significant ($P > 0.05$); however, genetic groups were corroborated by significant differentiation ($P < 0.004$).

A discrete structure over the total sample was detected by STRUCTURE (Fig. 5) for the "no admixture" model only. Estimations of data probability expressed by Ln P (D) values were consistent across runs, but only for k values of 2 and 3; hence this indicates that populations are probably subdivided into two or three clusters only. For k values greater than 3, a large proportion of individuals were assigned to several clusters, this is strong indication that an inferred structure with $k > 3$ does not fit the data. Indeed, Table 5 showed that when two or three clusters were assumed, 98.0% and 83.2% of the individuals, respectively, were assigned to only one cluster with $\geqslant$70% or higher probability. In the first hypothesis of the existence of two groups ($k = 2$), 100% of individuals of the Cac, Cri, and Net samples were significantly assigned to the same cluster (high individual probabilities) and to a lesser extent 77.7% of the Gue1 individuals, whereas all individuals from other samples were assigned to the second cluster except for Ipa3 (83.3% assigned). In the second case ($k = 3$), most individuals from Cac, Gue1, and Net were assigned to the first cluster, whereas three of the four samples around Ipazoltic and GR were assigned to the second one, and finally all individuals from Gue2 were assigned to the third one. The two remaining samples (Cri and Ipa4) presented a high percentage of individuals that could not be assigned significantly to any clusters. Surprisingly, when an

Table 5
Assignment of individual genotypes of *M. longipennis* from 10 samples to two or three inferred clusters.

Sample	No.	Number of individuals assigned to each cluster with a probability >0.70 (range of individual probability)				
		Two inferred clusters ($k = 2$)[a]		Three inferred clusters ($k = 3$)[b]		
		1	2	1	2	3
Cac	27	0	27 (0.819-1.000)	24 (0.758-1.000)	0	0
Net	26	0	26 (0.891-1.000)	25 (0.896-1.000)	0	0
Cri	16	0	16 (0.833-1.000)	4 (0.800-0.978)	0	0
Ipa1	25	25 (0.904-1.000)	0	0	0	23 (0.733-0.1.000)
Ipa2	14	14 (0.947-1.000)	0	0	0	13 (0.700-1.000)
Ipa3	24	20 (0.710-0.998)	1 (0.745)	2 (0.869-0.930)	0	19 (0.727-0.994)
Ipa4	20	20 (0.783-1.000)	0	0	6 (0.774-0.950)	1 (0.795)
Gue1	9	1 (0.811)	7 (0.737-0.948)	8 (0.832-0.987)	0	1 (0.722)
Gue2	28	28 (0.940-1.000)	0	0	28 (0.963-1.000)	0
GR	19	19 (0.706-1.000)	0	0	0	19 (0.709-1.000)

[a] Total number of individuals assigned with probability >0.70 = 204/208 = 98.0%.
[b] Total number of individuals assigned with probability >0.70 = 173/208 = 83.2%.

assignment was monitored with $k > 3$, Gue2 was the only sample for which most individuals were always assigned to a single cluster (Fig. 5).

4. Discussion

The aim of this study was to explore the genetic differentiation within and between *M. longipennis* samples collected in domestic, peridomestic, and wild environments, to assess their level of isolation at different geographic scales. These analyzes provided information on the dispersal behavior of the species, a very important issue in estimating the risks of colonization and reinfestation of dwellings from peridomestic or surrounding wild populations.

4.1. Heterozygote deficit within populations and the Wahlund effect

Five samples presented a significant heterozygote deficit: Cac, Net, Cri, Ipa3, and GR, even after discarding the locus with a high percentage of null alleles. In the Cac sample, the possible Wahlund effect could be explained by the fact that bugs had been collected for 2 years by inhabitants (possible temporal substructuring), in different houses (possible substructuring between dwellings) or because of migrant bugs coming from different places in the surrounding wild environment, as previously suggested to explain the pattern of dwelling infestation in the locality of Cacalutan (Brenière et al., 2010). For the Net population, breeding subunits (Wahlund effect) could be suspected according to their biotope; indeed eight bugs were collected in the ruin located 100 m from the dwelling (sylvatic population), five were collected indoors, and 13 others were collected outdoors in different sites around the house. F_{IS} values were not significant for the subsamples collected in the ruin and indoors, but they remained significantly positive for the subsample composed of bugs collected in different collection sites around the dwelling. This result shows that breeding subunits remain likely in peridomestic environments, and the absence of HW disequilibrium in the subsample composed of bugs all collected in an enclosed patio behind the house where dogs are kept, supports this hypothesis (Fig. 3). Similarly, in Los Guerrero, both samples collected in brick piles in two different peridomiciles were in HW equilibrium, but significantly differentiated. The subdivision of a locality into smaller panmictic units corresponding to different collection sites (indoors, chicken coops) was previously reported for *T. infestans* populations, the main vector of Chagas disease in South America (Brenière et al., 1998; Marcet et al., 2008; Piccinali et al., 2009; Pizarro et al., 2008). The reason for this pattern in the peridomestic environment could be the lack or reduced gene flow between colonies established in different structures, even close to one another, that allows genetic drift. However, this does not imply an absence of dispersion in all cases; indeed in a previous analysis, *M. longipennis* distribution in Los Guerrero showed that in 40% of positive sites, only one or two triatomines (adults or nymphs) were captured (Brenière et al., 2007). These insects are probably moving in search of food and/or a favorable microenvironment, and/or for reproduction in adults, and they could be the product of new colonies arising through founder effects.

When sylvatic samples collected in three categories of biotopes (rock pile boundary walls, large rocky formations, and a ruin) were analyzed, F_{IS} values were positive and significant in samples from large rocks suggesting a Wahlund effect (Cri sylvatic, GR, GR 2006). The existence of subunits in such a biotope was not expected because the bugs were collected at the same time, throughout the rocky formations over a distance of less than 70 m. However, these formations are composed of very large rocks that form small caves that are ideal as refuges for small mammals and triatomines. In this biotope it could be assumed that triatomine colonies are stable over very small distances; then genetic drift in each colony and the very low dispersal are sufficient to generate differences in allelic composition between colonies. On the contrary, three of the four sylvatic populations collected in rock pile boundary walls around Ipazoltic were in HW equilibrium although they were collected over a distance of approximately 100 m. The fourth one (Ipa3) had a positive and significant F_{IS} value that should be partly attributed to the high rate of null alleles (13.3%). When the four samples (Ipa1–4) were grouped, the F_{IS} value was not significant ($F_{IS} = 0.173$, $P > 0.05$). All these results show that in this biotope the panmictic unit should be larger than in large rocks, a result suggesting considerable dispersal activity of the bugs in agropastoral environments, for reasons that remain to be explored.

4.2. Genetic differentiation between samples and clustering

Significant genetic differentiations were observed between samples, independently from the geographical distance between them. For example, the Cac population, which is about 70 km away from the others, was not genetically differentiated from all samples, whereas the two populations separated by only 360 m in the village of Los Guerrero were differentiated. The hierarchical AMOVA testing differentiation between samples from San Martin de Hidalgo and those close to Guadalajara city was not significant. The presence of geographically close samples, differentiated by probable founder effects, could explain this lack of differentiation by the distance model in the present study. In the peridomestic environment, other studies on *T. infestans* a successful vector colonizing human habitat, showed that strong structuring between peridomestic units would be the rule, as observed in absence of insecticide control and after insecticide pressure and reinfestation

(Brenière et al., 1998; Marcet et al., 2008; Perez de Rosas et al., 2007). In contrast, for *T. dimidiata*, HW equilibrium was not rejected at the locality level in the Yucatan peninsula, and low genetic differentiation across localities was observed as well as between most of the samples collected at short distances in domestic, peridomestic, and sylvatic biotopes because this vector has substantial dispersal capabilities (Dumonteil et al., 2007). In this study, the Bayesian assignment procedure inferred two clusters that fit the topology of the NJ tree constructed by using FST values as genetic distance between samples. The first cluster included the samples collected in San Martin de Hidalgo except one (Gue1). Interestingly, when the inferred number of clusters was increased to $k = 10$, most of the memberships of Gue2 were always assigned to a single cluster with high probability.

These observations support the focal differentiation of some bug colonies in peridomestic structures as a probable result of founder effects and genetic drift combination that can explain the specific genetic composition of these samples. The four sylvatic samples collected around the locality of Ipazoltic were clustered together, and the absence of genetic differentiation between them suggests considerable dispersal capability (FST values not significant). In this agropastoral environment, the rock pile boundary walls are the scarce refuges for mammals, and the fauna is probably disturbed by the intense agricultural activity which prevents the stable installation of large triatomine colonies and generates dispersal of bugs. The second cluster included the sample of Cacalutan and the two collected north of Guadalajara city. This result suggests genetic affinities between populations over greater distances that can very likely be explained by the recent passive transport (e.g., humans) of bugs from one place to another.

5. Conclusion

The overall results show that the movement of triatomines is complex since there may be strong structuring on a micro-geographical scale, in domestic and sylvatic environments, but also the more distant populations could be related. Ipazoltic is particularly important because the analysis of the surrounding populations of bugs shows that they move significantly in the agropastoral area and are potentially able to reach the village where inhabitants collected a total of 29 bugs in their house or in their garden over a 6-week period (unpublished results).

Acknowledgments

We extend our gratitude to the health authorities of San Martin de Hidalgo. This study received financial support from the UNDP/WorldBank/WHO Special Program for Research and Training in Tropical Diseases (TDR) on Operational Research on Chagas disease, ID A30442 and from the "Institut de Recherche pour le Développement" (IRD) France.

References

Anderson, J.M., Lai, J.E., Dotson, E.M., Cordon-Rosales, C., Ponce, C., Norris, D.E., Beard, C.B., 2002. Identification and characterization of microsatellite markers in the Chagas disease vector *Triatoma dimidiata*. Infect. Genet. Evol. 1, 243–248.

Belkhir, K., Borsa, P., Goudet, J., Chikhi, L., Bonhomme, F., 2001. GENETIX 4.02, logiciel sous WindowsTM pour la génétique des populations. In Laboratoire Génome, Populations, Interactions CNRS UMR 5000, Université de Montpellier II (France).

Bosseno, M.-F., Barnabé, C., Sierra, M.J., Kengne, P., Guerrero, S., Lozano, F., Magallon-Gastelum, E., Brenière, S.F., 2009. Wild ecotopes and food habits of *Triatoma longipennis* infected by *Trypanosoma cruzi* lineages I and II in Mexico. Am. J. Trop. Med. Hyg. 80, 988–991.

Brenière, S.F., Bosseno, M.-F., Vargas, F., Yaksic, N., Noireau, F., Noel, S., Dujardin, J.P., Tibayrenc, M., 1998. Smallness of the panmictic unit of *Triatoma infestans* (Hemiptera: Reduviidae). J. Med. Entomol. 35, 911–917.

Brenière, S.F., Bosseno, M.-F., Magallon-Gastelum, E., Castillo Ruvalcaba, E.G., Gutierrez, M.S., Montano Luna, E.C., Basulto, J.T., Mathieu-Daudé, F., Walter, A., Lozano-Kasten, F., 2007. Peridomestic colonization of *Triatoma longipennis* (Hemiptera, Reduviidae) and *Triatoma barberi* (Hemiptera, Reduviidae) in a rural community with active transmission of *Trypanosoma cruzi* in Jalisco state, Mexico. Acta Trop. 101, 249–257.

Brenière, S.F., Bosseno, M.-F., Magallon-Gastelum, E., Soto Gutierrez, M.M., de Jesus Kasten Monges, M., Barraza Salas, J.H., Romero Paredes, J.J., Lozano-Kasten, F., 2010. Community participation and domiciliary occurrence of infected *Meccus longipennis* in two Mexican villages in Jalisco state. Am. J. Trop. Med. Hyg. 83, 382–387.

Carvalho-Costa, F.A., Dos Santos, S.M., Pires, M.Q., Lopes, C.M., Noireau, F., Pacheco, R.S., 2010. Sylvatic and peridomestic populations of *Triatoma pseudomaculata* are not significantly structured by habitat, as revealed by two genetic markers. J. Vector Ecol. 35, 295–300.

Dumonteil, E., Tripet, F., Ramirez-Sierra, M.J., Payet, V., Lanzaro, G., Menu, F., 2007. Assessment of *Triatoma dimidiata* dispersal in the Yucatan Peninsula of Mexico by morphometry and microsatellite markers. Am. J. Trop. Med. Hyg. 76, 930–937.

Edwards, J.K., 1998. Miniprep procedures for the isolation of plant DNA. Chapman & Hall, London.

Espinoza-Gomez, F., Maldonado-Rodriguez, A., Coll-Cardenas, R., Hernandez-Suarez, C.M., Fernandez-Salas, I., 2002. Presence of Triatominae (Hemiptera, Reduviidae) and risk of transmission of Chagas disease in Colima, Mexico. Mem. Inst. Oswaldo Cruz 97, 25–30.

Excoffier, L.G., Laval, G., Schneider, S., 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol. Bioinform 1, 47–50.

Falush, S.M.D., Pritchard, J.K., 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol. Ecol. Notes 7, 574–578.

Goudet, J., 1995. FSTAT (version 1.2): a computer program to calculate *F*-statistics. J. Hered. 86, 485–486.

Goudet, J., 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet (1995).

Lopez, D.C., Jaramillo, C., Guhl, F., 2007. Population structure and genetic variability of *Rhodnius prolixus* (Hemiptera: reduviidae) from different geographic areas of Colombia. Biomedica 27 (Suppl. 1), 28–39.

Lozano-Kasten, F., Magallon-Gastelum, E., Soto-Gutierrez, M., Kasten-Monges, M., Bosseno, M.-F., Brenière, S.F., 2008. Epidemiologic knowledge and current situation of Chagas disease in the state of Jalisco, Mexico. Salud Publica Mex. 50, 508–515.

Magallon-Gastelum, E., Magdaleno-Peñaloza, N.C., Kattahain-Duchateau, G., Trujillo-Contreras, F., Lozano-Kasten, F.J., Hernández-Gutiérrez, R.J., 1998. Distribución de los vectores de la enfermedad de Chagas Hemiptera: Reduviidae Triatominae en el estado de Jalisco, México. Rev Biom. 9, 151–157.

Magallon-Gastelum, E., Lozano-Kasten, F., Flores-Perez, A., Bosseno, M.-F., Brenière, S.F., 2001. Sylvatic Triatominae of the *phyllosoma* complex (Hemiptera: Reduviidae) around the community of Carrillo Puerto, Nayarit. Mexico. J. Med. Entomol. 38, 638–640.

Magallon-Gastelum, E., Lozano-Kasten, F., Bosseno, M.-F., Cardenas-Contreras, R., Ouaissi, A., Brenière, S.F., 2004. Colonization of rock pile boundary walls in fields by sylvatic triatomines (Hemiptera: Reduviidae) in Jalisco State. Mexico. J. Med. Entomol. 41, 484–488.

Magallon-Gastelum, E., Lozano-Kasten, F., Gutierrez, M.S., Flores-Perez, A., Sanchez, B., Espinoza, B., Bosseno, M.-F., Brenière, S.F., 2006. Epidemiological risk for *Trypanosoma cruzi* transmission by species of *Phyllosoma* complex in the occidental part of Mexico. Acta Trop. 97, 331–338.

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.

Martinez-Ibarra, J.A., Grant-Guillen, Y., Morales-Corona, Z.Y., Haro-Rodriguez, S., Ventura-Rodriguez, L.V., Nogueda-Torres, B., Bustos-Saldana, R., 2008. Importance of species of Triatominae (Heteroptera: Reduviidae) in risk of transmission of *Trypanosoma cruzi* in western Mexico. J. Med. Entomol. 45, 476–482.

Peakall, R., Smouse, P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6, 288–295.

Perez de Rosas, A.R., Segura, E.L., Garcia, B.A., 2007. Microsatellite analysis of genetic structure in natural *Triatoma infestans* (Hemiptera: Reduviidae) populations from Argentina: its implication in assessing the effectiveness of Chagas' disease vector control programmes. Mol. Ecol. 16, 1401–1412.

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.

Pizarro, J.C., Gilligan, L.M., Stevens, L., 2008. Microsatellites reveal a high population structure in *Triatoma infestans* from Chuquisaca, Bolivia. PLoS Negl. Trop. Dis. 2, e202.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.

Quisberth, S., Waleckx, E., Monje, M., Chang, B., Noireau, F., Brenière, S.F., 2011. "Andean" and "non-Andean" ITS-2 and mtCytB haplotypes of *Triatoma infestans* are observed in the Gran Chaco (Bolivia): population genetics and the origin of reinfestation. Infect. Genet. Evol. 11, 1006–1014.

Ramsey, J.M., Alvear, A.L., Ordonez, R., Munox, G., Garcia, A., Lopez, R., Leyva, R., 2005. Risk factors associated with house infestation by the Chagas disease vector *Triatoma pallidipennis* in Cuernavaca metropolitan area. Mexico. Med. Vet. Entomol. 19, 219–228.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.

Weir, B.S., Cockerham, C.C., 1984. Estimating *F*-statistics for the analysis of population-structure. Evolution 38, 1358–1370.

Wright, S., 1965. The interpretation of population structure by *F*-statistics with special regard to system of mating. Evolution 19, 395–420.