HHS Public Access

Author manuscript

Infect Genet Evol. Author manuscript; available in PMC 2016 January 08.

Published in final edited form as:

Infect Genet Evol. 2013 December ; 20: 352–361. doi:10.1016/j.meegid.2013.09.003.

Eco-geographical differentiation among Colombian populations of the Chagas disease vector *Triatoma dimidiata* (Hemiptera: Reduviidae)

Andrés Gómez-Palacio[a,*], **Omar Triana**[a], **Nicolás Jaramillo-O**[a], **Ellen M. Dotson**[b], and **Paula L. Marcet**[b]

[a]Grupo BCEI, Universidad de Antioquia UdeA, Calle 70 No. 52-21, Medellin, Colombia

[b]Centers for Disease Control and Prevention (CDC), Division of Parasitic Diseases and Malaria, Entomology Branch, Atlanta, GA, USA

Abstract

Triatoma dimidiata is currently the main vector of Chagas disease in Mexico, most Central American countries and several zones of Ecuador and Colombia. Although this species has been the subject of several recent phylogeographic studies, the relationship among different populations within the species remains unclear. To elucidate the population genetic structure of *T. dimidiata* in Colombia, we analyzed individuals from distinct geographical locations using the cytochrome c oxidase subunit 1 gene and 7 microsatellite loci. A clear genetic differentiation was observed among specimens from three Colombian eco-geographical regions: Inter Andean Valleys, Caribbean Plains and Sierra Nevada de Santa Marta mountain (SNSM). Additionally, evidence of genetic subdivision was found within the Caribbean Plains region as well as moderate gene flow between the populations from the Caribbean Plains and SNSM regions. The genetic differentiation found among Colombian populations correlates, albeit weakly, with an isolation-by-distance model (IBD). The genetic heterogeneity among Colombian populations correlates with the eco-epidemiological and morphological traits observed in this species across regions within the country. Such genetic and epidemiological diversity should be taken into consideration for the development of vector control strategies and entomological surveillance.

Keywords

Triatominae; Chagas disease; *Triatoma dimidiate*; Multilocus microsatellite analysis; Cytochrome c oxidase subunit 1; Population genetics

1. Introduction

Chagas disease is a parasitic disease in which the pathogenic agent, *Trypanosoma cruzi*, is transmitted by hematophagous insects of the Triatominae subfamily. *Triatoma dimidiata* is the major vector in several Central American countries as well as in regions of Ecuador and

*Corresponding author. Address: Grupo de Biología y Control de Enfermedades Infecciosas – BCEI, Laboratorio 620, Sede de Investigación Universitaria, Universidad de Antioquia, Calle 62 # 52-59, Medellin, Colombia. Tel.: +57 04 2196681. amgomezpa@gmail.com (A. Gómez-Palacio).

Colombia, where it occupies a large diversity of eco-topes and life zones (Dorn et al., 2007). Notably, *T. dimidiata* populations within Colombian seem to differ in several biological, ecological and epidemiological attributes (i.e., life cycle, infection rates) and present marked morphological differences between western and eastern populations (Esteban, 2010).

This species had been found in 47% of Colombia departments, occupying human dwellings, palm trees, bark of dead trees, rocks piles, chicken huts, huts for drying tobacco leaves and wood stoves (Grisales et al., 2010; Guhl et al., 2007). Colombian *T. dimidiata* populations from the Inter Andean Valleys eco-geographical region are frequently found colonizing dwellings and show high rates of *T. cruzi* infection (Guhl et al., 2007; Parra-Henao, personal communication). Hence, these *T. dimidiata* populations are considered more relevant as Chagas vectors than those from Northern Colombia or Caribbean regions, that are mostly sylvatic (Guhl et al., 2007; Ramírez et al., 2005).

An effective surveillance and vector control program must take into account the habitat diversity and the dispersion capacity among populations within and among eco-geographical regions, which is yet to be established for Colombian *T. dimidiata* populations. A recent study based on nucleotide sequence analysis of the mitochondrial NADH dehydrogenase gene subunit 4 (ND4), showed significant genetic differentiation and strong population structure among three localities from the departments of La Guajira, located at Northwestern of the Sierra Nevada de Santa Marta (SNSM) and Cesar, located at Caribbean Plains and Santander, in the Inter Andean Valleys region (Grisales et al., 2010). However, because the sample evaluated was rather small ($n = 40$), representing only a minimal area of the species distribution, a more comprehensive analysis is needed to elucidate the actual genetic structure of *T. dimidiata* populations throughout Colombia and to determine whether the morphological and epidemiological differences observed across regions correlate with genetic differentiation. The present study includes a larger number of individuals representing *T. dimidiata* populations across the complete distribution of the species within the country. Moreover, by analyzing information based on independent and higher resolution genetic markers, we were able to explore the degree of gene flow occurring among the different eco-geographical regions. The application of such molecular markers can help improve vector control and surveillance by identifying and characterizing genetically distinct vector populations and developing targeted intervention strategies. Quantitative measures of these aspects can be assessed by using certain parameters such as the best-fitted number of genetic clusters (K) based on allele frequencies distribution, estimation of genetic structure indexes (such as Fst, Rst or Φst) or by statistical procedures that allows the hierarchical partitioning of genetic variation among and within groups (such as AMOVA test). To estimate these parameters independent genetic markers with a significant degree of variability within a species are required.

Mitochondrial DNA (mtDNA) had been extensively used in molecular systematic studies of triatomine species across most countries in Central and South America (Lyman et al., 1999; Mas-Coma and Bargues, 2009). mtDNA loci are considered one of the more sensitive tools to infer population structure at both local and regional levels in several epidemiologically relevant triatomine species such as *T. infestans* (Monteiro et al., 1999; Piccinali et al., 2009, 2011), *Rhodnius prolixus* (Fitzpatrick et al., 2008) and *T. dimidiata* (Blandón-Naranjo et al.,

2010; Monteiro et al., 2013), among others. The cytochrome c oxidase subunit 1 (COI) gene has been used at several micro-evolutionary scales ranging from gene flow studies among sylvatic and domestic isolates in *Triatoma infestans* at a local level (Piccinali et al., 2011) to regional studies about phylogenetic relationships and population structure in North American sibling species as *Thala recurva* and *T. rubida* (Pfeiler et al., 2006), and of *Triatoma infestans* populations in South America (Piccinali et al., 2009).

Multilocus microsatellite analysis (MMA) have been developed and applied for population studies of several triatomine species over a wide range of geographic and evolutionary scales: *R. prolixus* (Fitzpatrick et al., 2009), *Rhodnius pallescens* (Gómez-Sucerquia et al., 2009), *Triatoma pseudomaculata* (Harry et al., 2008a) and *T. infestans* (Marcet et al., 2008; Pizarro et al., 2008; Pérez de Rosas et al., 2008, 2011; Richer et al., 2007). For *T. dimidiata* however, only a limited number of studies have applied MMA for genetic population studies. Low genetic differentiation and variability of *T. dimidiata* populations from several villages in the Yucatan Peninsula of Mexico (Dumonteil et al., 2007) was detected using eight previously reported microsatellite loci (Anderson et al., 2002), of which only four loci had actually provided reliable genotypic information (Dumonteil et al., 2007). Those results revealed the need to identify new microsatellite loci for *T. dimidiata* species and/or to optimize the performance of those already published. Therefore the objectives of this work were (i) to evaluate and optimize the performance of previously published microsatellite loci from different triatomine species in *T. dimidiata* individuals from different areas across the species geographical distribution, and (ii) to study the genetic structure of Colombian *T. dimidiata* populations by using both COI nucleotide sequences and microsatellites markers.

2. Materials and methods

2.1. Sample origin

The capture origin of each *T. dimidiata* from Colombia included in the analyses is detailed in Fig. 1 and Table 1. Samples were collected in 12 communities groups from seven departments that belong to three different eco-geographical regions, considered relevant for the presence of triatomine vectors in Colombia: Inter Andean Valleys, Caribbean Plains and SNSM (Fig. 1, Table 1). The sampling includes both rural and urban capture sites, encompassing a variety of eco-epidemiological attributes, which aimed to represent the ecological diversity of *T. dimidiata* populations throughout the geographical distribution of the species in Colombia.

Bug captures were carried out during 2003–2009 in collaboration with local personnel from the Ministry of Health. Sylvatic collections were performed with live-baited traps (Noireau et al., 1999). Domiciliary and peridomiciliary collections were made by the traditional time manual collection method using a dislodging spray (Gürtler et al., 1999) and by homeowners. A maximum of three insects per house were included in the sample. Captures from palm trees were obtained through palm dissection (Fitzpatrick et al., 2008), having previously obtained consent from the landowners. All specimens were identified as *T. dimidiata* according to morphological characters (Lent and Wygodzinsky (1979) and kept in 70% ethanol until processed for DNA extraction.

Genomic DNA was obtained from four legs of each insect or from thorax muscle, following an insect DNA extraction protocol (Collins et al., 1987). Additionally, in order to optimize the micro-satellite PCR conditions, high quality DNA from fresh specimens was purified using the Wizard Genomic Purification Kit (Promega®) following the manufacturer recommendations.

2.2. PCR and sequencing

A 402-bp fragment of the COI gene was PCR-amplified for each specimen, using the primers LCO1490f (5′-GGTCWMCAAATCATAAAGATATTGG-3′) and HCO2198r (5′-TAWACTTCAGGGTGWCCAAARAATCA-3′), slightly modified from the original publication (Folmer et al., 1994) to improve their performance in this species. PCR reactions were conducted in a final volume of 35 μl using 3 μl of 10 ng/μl DNA templates, 3.5 μl of 1 × PCR buffer, 4.4 μl of 2 mM dNTP, 1.4 μl of each 0.4 μM primer, 3.5 μl of 50 mM $MgCl_2$ and 1 U/μl of Taq DNA polymerase (Promega®). Amplification conditions were: 95 °C for 5 min; 35 cycles of 94 °C for 30 s, 50 °C for 30 s, and 72 °C for 60 s; 72 °C for 10 min. PCR products were sent to Macrogen Inc., Korea for purification and sequencing.

Both strands of the COI fragment were sequenced for all samples (n = 82). A multiple sequence alignment was obtained with the CLUSTAL W algorithm (Thompson et al., 1997) implemented in BioEdit v. 7.0.5 (Hall, 1999). The number of segregating sites (S), nucleotide diversity (π), number of haplotypes (h) and haplotype diversity (Hd) were estimated with DnaSP v.5.10 (Librado and Rozas, 2009). Population differentiation based on nucleotide and haplotype diversity was estimated by Kst and Hst (Hudson et al., 1992), with a permutation test of 1000 replicates and significance level of $p < 0.001$, using DnaSP v.5.10 (Librado and Rozas, 2009). A median joining (MJ) haplotype network was obtained using Network v.4.6.1.1 (http://www.fluxus-engineering.com), using default parameters (equal character weight = 10; epsilon value = 10; transversions/transitions weight = 1:1 and connection cost as a criterion).

2.3. Multilocus microsatellite analysis (MMA)

The performance of microsatellite loci previously published for triatomine species was assessed for *T. dimidiata*. Twenty seven markers were evaluated: eight from *T. dimidiata* (Anderson et al., 2002), ten from *T. infestans* (Marcet et al., 2006), six from *T. pseudomaculata* (Harry et al., 2008a) and three from *R. prolixus* (Harry et al., 2008b). Each primer set was first evaluated in 10 *T. dimidiata* insects from different locations (Belize, Costa Rica, Honduras, Nicaragua and Colombia), aiming to account for most of the genetic diversity observed within the species across its geographic distribution. Amplification conditions for each locus optimized for *T. dimidiata* are presented in Table S1. DNA fragment detection was performed using an automated DNA sequencer ABI 3130 (Applied Biosystems®). Fragment size determination with one bp resolution and allele classification was performed with GeneMapper v.3.7 (Applied Biosystems®) software.

Using the 10 loci that proven suitable for *T. dimidiata*, a multilocus genotype was obtained for each individual from Colombian populations (n = 189). Bug populations were initially defined as: “group of *T. dimidiata* captured in the same community (or group of adjoining

villages)". Only populations with >10 individuals were used for parameter estimation at the population level.

The algorithm implemented in the software Micro-Checker v.2.2.3 (Van Oosterhout et al., 2004) was used to evaluate each locus for significant deviations from HWE in the total sample and rule out the presence of null alleles, PCR artifacts or genotyping errors. The algorithm was applied with 1000 randomizations. Hardy–Weinberg equilibrium (HWE) expectations and Linkage disequilibrium (LD) between microsatellite loci pairs were evaluated at the region and community levels using Arlequin v.3.11 (Excoffier et al., 2005) and GenePop v.4.0 (Rousset, 2008). Observed (Ho) and expected (He) heterozygosis were calculated for each locus and each community using the program GenAlEx v.6.4 (Peakall and Smouse, 2006). Wright's inbreeding coefficient F_{IS} (Cockerham and Weir, 1986) and allele richness (AR) (El Mousadik and Petit, 1996) were obtained using FSTAT v.2.9.3.2 (Goudet, 1995). Bonferroni correction for multiple comparisons was applied when needed. Genetic structure of the whole dataset was assessed using hierarchical analysis of molecular variance (AMOVA) and fixation index (Fis) estimation as implemented in Arlequin v.3.11 (Excoffier et al., 2005). Statistical significance for the population structure was assessed by permutation tests with 1000 iterations.

To evaluate whether the genetic differentiation among community pairs follows an isolation-by-distance (IBD) model, the correlation between the linearized F_{ST}, $[F_{ST}/(1-F_{ST})]$ and the linear geographic distance between them, was examined with the Mantel test (Mantel, 1967) implemented in GenAlEx v.6.4 (Peakall and Smouse, 2006) and a permutation procedure (1000 permutations).

Bayesian assignment approach, as implemented in STRUCTURE v.2.3.3 (Pritchard et al., 2000), was used to determine the genetic structure of the whole sample independently of the individual capture location. Graphical representation of the results was obtained using Distruct v.1.1 (Rosenberg, 2004). The number of genetic clusters (K) was determined using the model of "admixture ancestry and correlated allele frequencies", as recommended by the authors to analyze closely related populations (Falush et al., 2003). The genetic clustering was evaluated from $K = 1$ (number of eco-geographical regions minus 2) to $K = 15$ (number of communities plus 3). Ten replicates were obtained per each K value, after 100,000 iterations and following a burn-in of 10,000 runs. Optimum K number was determined according to the Evanno et al. (2005) algorithm implemented in Structure Harvester V. 0.3 (Earl and vonHoldt, 2011). Individuals were considered assigned to aspecific cluster if the proportion of ancestry was $\geqslant$80%; otherwise the individual was considered unassigned.

3. Results

3.1. Molecular diversity, polymorphism and population differentiation

Base composition for the COI 402-bp fragment was A–T rich (61.7%), which includes a high number of variable sites (S = 165) that determined 63 haplotypes (Genbank accession codes KC489226–KC489291). The geographical distribution of these haplotypes is presented in a supplementary table (Table S2).

While all three regions presented similar haplotype diversity values (Table 2), the Caribbean Plains region presented the highest nucleotide diversity (almost twice than the other two). Hudson's statistics of comparative nucleotide (Kst = 0.329; $p < 0.001$) and haplotype diversity (Hst = 0.024; $p < 0.001$) among the three regions was significant, indicating a strong genetic differentiation among them (Table 2).

The haplotype network reflects the existence of genetic structure among the three eco-geographical regions (Fig. 2). Although only separated by a few mutational steps, most haplotypes from each region clustered together. A few haplotypes however, did not group with those from the same region (Hap49, Hap55, Hap36, Hap29 and Hap31). Furthermore, haplotype Hap37 (clustered among the SNSM haplotype group) was present in both SNSM and Caribbean Plains regions, which indicates that a small degree of gene-flow could be occurring between these two regions.

3.2. Microsatellites loci optimization

Ten out of the 27 microsatellite loci evaluated, successfully amplified *T. dimidiata* DNA from diverse geographic areas (Colombia, Costa Rica, Honduras and Nicaragua). On the other hand, locus TDMS1 did not amplify in Belize samples and locus TDMS19 did not amplify for one individual from Colombia and one from Belize (data not shown). Optimized amplification conditions, microsatellite motif, and allele size ranges are presented in Table S1.

3.3. Microsatellites loci suitability for population genetic studies within Colombia

A multilocus genotype based on 10 loci was obtained for each individual from Colombia. Linkage disequilibrium analysis considered across all Colombian capture sites and for all loci pairs, showed no significant deviations after applying Bonferroni correction. Therefore, these 10 loci are presumed to segregate independently and can be used together in population genetic analyses. The algorithm implemented in Micro-checker determined that there was no evidence of genotyping errors (i.e., allele dropouts or stuttering) affecting allele scoring at any of the markers. However, the analysis indicated that loci TDMS9 and TDMS19 showed high frequency of potential null alleles (0.4 and 0.7, respectively). This fact was also perceived through the significant excess of homozygotes detected in all three eco-geographical regions for these loci (Table S3).

No significant deviation from Hardy–Weinberg equilibrium (HWE) expectation was detected when considering the whole sample (all communities and all loci) after applying Bonferroni correction (Table S3). However significant departures from HWE for a few loci occurred in several communities (Table S3), and all deviations were associated with positive F_{IS} values, reflecting heterozygote deficit (Htd) at the community level.

Locus TDMS3 resulted monomorphic in all communities analyzed (allele = 141 bp), a fact that was previously reported in *T. dimidiata* populations from Mexico (Dumonteil et al., 2007) and in a close-related species *Meccus longipennis* (Brenière et al., 2012). On the other hand, locus TDMS9 showed no amplification in some individuals of communities 1, 2, 3, 5 and 8 (Table S3), and locus TDMS19 did not amplify on most individuals from some Inter

Andean Valleys communities (Table S3). Therefore, three loci (TDMS3, TDMS9 and TDMS19) were discarded for the final study of population differentiation analyses when the entire sample was considered.

3.4. Population structure of T. dimidiata within Colombia

The mean number of alleles per locus across the total sample was 4.8, ranging between 9 (TDMS4) to 34 (Tinf_ms42) alleles per locus (Table S1). Mean allele richness (AR) per community was 4.9, ranging from 1.2 to 6.4 (Table S4). The number of polymorphic loci per community suitably amplified ranged from 6 to 9 (Table S4). Considering all loci and all communities mean observed heterozygosis (*Ho*) ranged from 0.2 to 0.6, while expected heterozygosis (He) between 0.3 and 0.6 (Table S4). Inbreeding coefficient F_{IS} ranged from 0.10 to 0.44 (Table S4).

Pairwise comparisons among communities based on *F*st values showed significant differentiation among all communities from different regions (Table 3). Moreover, significant differentiation was observed among populations within the Caribbean Plains region (Table 3). AMOVA analysis showed the following distribution of the molecular variation of the whole sample: 22.1% was attributed to differences among eco-geographical regions, 22.5% to differences among communities, 5.8% to differences among communities within regions and 19.2% among individuals within communities (Table S5).

Similarly, population structure predictors estimated for all regions (Fct) and all communities (Fst), showed significant values, suggesting a strong genetic differentiation within Colombia (Table S5). Moreover, when the comparisons between regions was re-analyzed between pairs, low differentiation (1.4%) was observed between the Caribbean Plains and the SNSM regions, and conversely, the Inter Andean Valleys region presented the most differentiation with both other regions (27% and 32%) (Table S5).

The mantel test of IBD revealed a weak but significant correlation between linearized F_{ST} and pairwise geographic distance among populations at the community level ($r^2 = 0.57$, $p < 0.001$). (Fig. 3). Remarkably, low genetic differentiation was observed between the geographically closer population pairs from Caribbean Plains and SNSM regions. Likewise, within the Caribbean Plains region there are populations located at >200 km that showed strong genetic differentiation (Fig. 3). However, the values showed some dispersion from the lineal expectation, which explains the rather low adjustment to the IBD model.

Genetic structure analysis based on Bayesian clustering indicated that the number of genetic clusters (*K*) for the whole dataset was 4 (Fig. 4a and b). The efficiency of individual assignments to a specific genetic cluster ranged from 82% to 91%. Briefly, 91% of the individuals from communities 1–5 (Inter Andean Valleys) were assigned to cluster 1, 91% of the individuals from communities 6 and 7 (Caribbean Plains south) were assigned to cluster 2, 82% of the individuals from communities 8–10 (Caribbean Plains north) were assigned to cluster 3, and 91% of the individuals from communities 11 and 12 (SNSM) were assigned to cluster 4 (Table 4 and Fig. 4a). Remarkably, most individuals from communities 8–10 (Caribbean Plains north) not assigned to cluster 3, were significantly assigned to

clusters 2 or 4 (Table 4), suggesting the occurrence of gene flow between the communities of Caribbean Plains and SNSM regions.

4. Discussion

The analyses applied in this work, using mitochondrial and microsatellite markers were able to detect three well-defined genetic groups corresponding to the 3 Colombian eco-geographical regions analyzed. Furthermore, this high-resolution genetic approach also revealed an apparent genetic subdivision within the Caribbean Plains region.

Previously attempts to use MMA in *T. dimidiata* failed to detect any signs of genetic structure among the populations analyzed (Dumonteil et al., 2007). However, only 4 loci resulted reliable in terms of amplification success and reproducibility of results from the several microsatellite loci described for the species and assayed in that work. Moreover, among those 4 loci, one lack of polymorphism, presenting a unique allele for all populations. A significant contribution of the present work is the exhaustive testing and optimization of microsatellite loci for successful amplification in *T. dimidiata* from different geographic sources. The evaluation of the loci was designed aiming to account for the extensive natural genetic variability of the species, and thus contribute with a tool that can be applied in studies of populations for different sources. Nevertheless, as the amplification results might vary for different populations even within a region (i.e., high frequency of potential null alleles for locus TDMS19 in most of Inter Andean Valleys communities from Colombia), in order to choose the appropriate markers it is recommended to test each loci with local individuals at the time of planning a study for a particular region.

The genetic structure in *T. dimidiata* populations across Central America and Mexico has been studied in different contexts and geographical scales, using different markers, with different resolution. In Costa Rica, nucleotide comparison based on the cytochrome b (cyt b) and ITS-2 sequences failed to detect differentiation among the populations evaluated (Blandón-Naranjo et al., 2010). In the Yucatan Peninsula, despite the failure in detecting local genetic differentiation using multilocus microsatellite analyses (Dumonteil et al., 2007), significant genetic structure was encountered within the region using ITS-2 as well as cyt b markers (Bargues et al., 2008; Dorn et al., 2009; Herrera-Aguilar et al., 2009; Monteiro et al., 2013). Also within Mexico, in the Campeche state, strong genetic differentiation was detected among communities from two distinct eco-geographic regions that harbored two divergent clades (Tamay-Segovia et al., 2008). Summarizing, most studies determined that the genetically distinct population groups found within those regions, showed degrees of divergence compatible with subspecies or species levels, which imply necessary independent origins and/or genetic isolation sustained for a significant amount of time (Bargues et al., 2008; Dorn et al., 2009; Monteiro et al., 2013; Panzera et al., 2006). In contrast, within Colombia, despite the large genetic differentiation among populations from different regions, all populations belong to the same subspecies or genetic group (Bargues et al., 2008; Dorn et al., 2009; Monteiro et al., 2013; Panzera et al., 2006), implying a different micro-evolutionary process than those occurring in Central American and Mexican conspecifics. Further studies are needed to determine the extent of the genetic divergence and phylogenetic relationship among taxonomic groups within the dimidiata complex.

The genetic differentiation among the three eco-geographical regions studied here was significant with the two sets of markers applied. Both the haplotype distribution and the genetic structure index indicate high differentiation among regions, indicating that gene-flow among regions is restricted. This pattern might imply that (i) populations at the regional level differentiated from one another over time because the high degree of genetic isolation and different environmental conditions among regions or (ii) that there was an independent source of the populations at a regional level, with secondary and more recent genetic exchange among them. The haplotype network based on nucleotide differences among individual sequences showed that although only separated by a few mutational steps, most haplotypes from each region clustered together. However, there are no highly abundant haplotypes located in a central position of the network or shared among regions, which would be a pattern consistent with a common origin for all Colombian populations and a posterior diversification at the regional level. This observation favors the second hypothesis, but further inter and intra-regional analyses with a larger number of samples per region are still needed to reveal the actual degree of genetic sub-structure within each eco-geographical region.

It is worth noticing the apparent fourth genetic group evidenced by MMA, which manifests a degree of genetic subdivision of the Caribbean Plains region. The extensive geographical distance among the Caribbean Plains communities evaluated here, and the relative geographical proximity of some of these populations with those of the SNSM region is the most likely explanation for the observed pattern. Moreover, this result is in agreement with the genetic differentiation pattern at the community level, which fits, albeit weakly, an IBD model. To this date, this dataset is the largest analyzed within the country, but it is still limited in the number of populations considered, which does not represent a random geographic distribution of the sampling. Therefore, these results should be considered rather exploratory, and not the ultimate representation of *T. dimidiata* genetic pattern in Colombia, for which a larger number of population within regions have to be analyzed.

4.1. Population structure of T. dimidiata and eco-epidemiological considerations

The genetic differentiation among Colombian *T. dimidiata* populations described here is consistent with the heterogeneity reported for several aspects of this species within the country. Previous studies have shown morphological evidence of differentiation between populations of *T. dimidiata* from Western and Eastern Colombia (Esteban, 2010). Independently, genetic distinctions (based on ND4 gene results) were reported among populations of northern and central regions (Grisales et al., 2010). Moreover, epidemiological differences concerning Chagas disease transmission and prevalence were detected among Colombian regions (Guhl et al., 2007; Ramírez et al., 2005).

In Colombia *T. dimidiata* is considered the second in relevance as Chagas disease vector, after *R. prolixus*. The later is mostly domestic and does not show significant genetic structure across several eco-geographical regions (López et al., 2007). The relatively homogenous populations could be potentially eliminated effectively with the appropriate use of chemical control measurements (López et al., 2007). However, *T. dimidiata* presents sylvatic populations across its geographical distribution within the country. A diverse

genetic and epidemiological background across different eco-geographical regions requires distinct vector control and surveillance strategies to be implemented accordingly. The main eco-epidemiological attributes of *T. dimidiata* in each region, could be summarized considering the following aspects:

4.1.1. Inter andean valleys—In this region *T. dimidiata* had been considered the main Chagas disease vector in several areas, specifically in Boyacá and Santander departments (Guhl et al., 2007; Ramírez et al., 2005). *T. dimidiata* is likely to offer more problems for control in this region given its effective capability to colonize human environments (i.e., in huts for drying tobacco leaves and outside woods stoves).

Previous results based on RAPD's analysis showed no genetic differentiation among individuals from sylvatic, peridomestic and domestic habitats in the Boavita locality from the Boyacá department (Ramírez et al., 2005). The present work also failed to detect any genetic differentiation among individuals collected in sylvatic (i.e., rocks piles) and anthropogenic environments (i.e., hut for drying tobacco leaves, woods stoves and dwellings). These results suggest that within this region *T. dimidiata* from sylvatic habitats invades peridomestic and domestic structures, which supposes a high risk of reinfestation post spraying. Moreover, the influx from sylvatic habitats poses the threat of bringing new *T. cruzi* strains and re-establish the transmission cycle if ever interrupted in the domiciles. We therefore suggest continuous entomological vigilance to be established throughout the region to avoid reinfestation of treated households.

Although the genetic pattern seemed quite homogeneous among individuals from this region, a small number of putative immigrants from the Caribbean Plains regions were observed. Because of the low number of individuals analyzed, no reliable assessment of their sources could be yet established. A likely explanation for the mechanism that originated this pattern could be the passive transport of bugs among human goods. This which could pose another significant threat for the success of vector control campaigns, but a higher number of individuals need to be analyzed to further account for this observation.

4.1.2. SNSM (Northwestern zone)—The Sierra Nevada de Santa Marta region evaluated here included several localities in Northwestern-slope where *T. dimidiata* is mostly a sylvatic species, inhabiting palms trees. In this region, *T. dimidiata* individuals have only been found sporadically visiting indigenous dwellings and related with human *T. cruzi* infections (Rodríguez et al., 2009). This region shows a complex Chagas disease landscape, and no conclusive *T. cruzi* transmission cycles have yet been defined. *R. prolixus* is found domiciliated in high densities and is considered the most important Chagas disease vector in this region (Rodríguez et al., 2009). However, it is important to note that if *R. prolixus* is eliminated by adequate insecticide spraying, secondary vector species such as *T. dimidiata* and *R. pallescens* could invade the houses. Further entomological and eco-epidemiological studies are need within this region to determine the actual relevance of *T. dimidiata.* In particular, closely monitoring of its abundance during and after spraying should be applied.

4.1.3. Caribbean plains—In this region *T. dimidiata* shows and eclectic behavior and occupies variable niches, including palm trees and other sylvatic habitats. Sporadically, *T. dimidiata* adults are found in houses and peridomiciliary ecotopes and it was determined that some individuals have fed on humans (Peña et al., 2012). The genetic structure analysis of *T. dimidiata* populations from this area determined two sub-zones: southern and northern Caribbean Plains. Noteworthy is the fact that the populations analyzed in this work are not randomly distributed in space and could be considered from two different sections within the region. The northern populations (communities 8–10) are in fact geographically closer to the samples analyzed from SNSM region (11, 12), and thus the genetic similitude among them could be due to geographical proximity, reflecting a pattern of moderate gene-flow occurring between regions.

However, these results would imply that gene-flow is occurring between domiciliary bugs from northern Caribbean Plains (localities of Cesar department) and sylvatic bugs (from palm trees) from SNSM. A mechanism of passive transport of bugs between the two eco-geographical regions, might explain the colonization of indigenous dwellings in the Caribbean Plains populations.

The localities of northern Caribbean Plains evaluated here are located on the foothill of East-slope of SNSM, where the indigenous tribe Wiwa is settled. Moreover, there are other Wiwa groups inhabiting the Northwestern-slope of SNSM, and thus human movements of people between these two eco-geographic regions are quite frequent. Indigenous communities use palm tree leaves to build their roofs. Palm trees are abundant in SNSM but are scarce in the northern Caribbean localities we analyzed. A possible explanation for the observed genetic pattern is that domestic colonies in the northern Caribbean Plains houses result from bug introductions carried in palm leaves from the SNSM region.

Under this picture we consider a vector control approach based on insecticide spraying of houses could potentially control domiciliary infestations in northern Caribbean Plains, but the constant influx of bugs in palm leaves from SNSM would jeopardize the sustainability of the actions. Specific measures targeting such sources should be implemented (i.e., treating materials before building the roofs or replacing palm leaves with alternative materials and education of community members to change their house building habits).

Finally, we consider that additional genetic studies, using other nuclear or mitochondrial markers and larger population numbers at the local level, are still necessary to understand the biological and evolutionary aspects of this species at the micro and macro-evolutionary levels. These aspects should be taken into account for phylogeography and systematic studies as well as for the epidemiological implications that the genetic heterogeneity within this species in Colombia and throughout its distribution could pose.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by Estrategía de sostenibilidad 2013–2014, CODI-Universidad de Antioquia, and by COLCIENCIAS projects # 111545921460 and 111549326149. We would like to thank to researchers Gabriel Parra-Henao (Instituto Colombiano de Medicina Tropical, Universidad CES, Medellin) and Víctor Angulo (Centro de Investigaciones en Enfermedades Tropicales-CINTROP, Universidad Industrial de Santander, Bucaramanga) for providing some specimens. To Instituto Nacional de Salud de Colombia, and its entomologists staff from Boyacá, Norte de Santander and Cesar departments for their assistance in fieldwork. Oligonucleotides for microsatellite loci were synthetized by the CDC core facilities.

The findings and conclusions in this manuscript are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.meegid.2013.09.003. These data include Google maps of the most important areas described in this article.

References

Anderson JM, Lai JE, Dotson EM, Cordon-Rosales C, Ponce C, Norris DE, Beard CB. Identification and characterization of microsatellite markers in the Chagas disease vector *Triatoma dimidiata*. Infect Genet Evol. 2002; 1:243–248. [PubMed: 12798021]

Bargues MD, Klisiowicz DR, Gonzalez-Candelas F, Ramsey JM, Monroy C, Ponce C, Salazar-Schettino PM, Panzera F, Abad-Franch F, Sousa OE, Schofield CJ, Dujardin JP, Guhl F, Mas-Coma S. Phylogeography and genetic variation of *Triatoma dimidiata*, the main Chagas disease vector in Central America, and its position within the genus *Triatoma*. PLoS Negl Trop Dis. 2008; 2:e233. [PubMed: 18461141]

Blandón-Naranjo M, Zuriaga MA, Azofeifa G, Zeledón R, Bargues MD. Molecular evidence of intraspecific variability in different habitat-related populations of *Triatoma dimidiata* (Hemiptera: Reduviidae) from Costa Rica. Parasitol Res. 2010; 106:895–905. [PubMed: 20165880]

Brenière SF, Waleckx E, Magallón-Gastélum E, Bosseno MF, Hardy X, Ndo C, Lozano-Kasten F, Barnabé C, Kengne P. Population genetic structure of *Meccus longipennis* (Hemiptera, Reduviidae, Triatominae), vector of Chagas disease in West Mexico. Infect Genet Evol. 2012; 12:254–262. [PubMed: 22142488]

Cockerham CC, Weir BS. Estimation of inbreeding parameters in stratified populations. Ann Hum Genet. 1986; 50:271–281. [PubMed: 3446014]

Collins FH, Mendez MA, Rasmussen MO, Mehaffey PC, Besansky NJ, Finnerty V. A ribosomal RNA gene probe differentiates member species of the *Anopheles gambiae* complex. Am J Trop Med Hyg. 1987; 37:37–41. [PubMed: 2886070]

Dorn PL, Monroy C, Curtis A. *Triatoma dimidiata* (Latreille, 1811): a review of its diversity across its geographic range and the relationship among populations. Infect Genet Evol. 2007; 7:343–352. [PubMed: 17097928]

Dorn PL, Calderon C, Melgar S, Moguel B, Solorzano E, Dumonteil E, Rodas A, de la Rua N, Garnica R, Monroy C. Two distinct *Triatoma dimidiata* (Latreille, 1811) taxa are found in sympatry in Guatemala and Mexico. PLoS Negl Trop Dis. 2009; 3:e393. [PubMed: 19274073]

Dumonteil E, Tripet F, Ramirez-Sierra MJ, Payet V, Lanzaro G, Menu F. Assessment of *Triatoma dimidiata* dispersal in the Yucatan Peninsula of Mexico by morphometry and microsatellite markers. Am J Trop Med Hyg. 2007; 76:930–937. [PubMed: 17488918]

Earl, D.; vonHoldt, B. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour. 2011. http://dx.doi.org/10.1007/s12686-12011-19548-12687

El Mousadik A, Petit R. High level of genetic differentiation for allelic richness among populations of the argan tree (*Argania spinosa* (L.) Skeels) endemic to Morocco. Theor Appl Genet. 1996; 92:832–839. [PubMed: 24166548]

Esteban, L. MSc Thesis. Universidad Nacional de Colombia; 2010. Variabilidad morfológica entre poblaciones de Triatoma dimidiata (Latreille 1811), procedentes de cuatro departamentos de Colombia.

Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 2005; 14:2611–2620. [PubMed: 15969739]

Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online. 2005; 1:47–50. [PubMed: 19325852]

Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics. 2003; 164:1567–1587. [PubMed: 12930761]

Fitzpatrick S, Feliciangeli MD, Sanchez-Martin MJ, Monteiro FA, Miles MA. Molecular genetics reveal that silvatic *Rhodnius prolixus* do colonise rural houses. PLoS Negl Trop Dis. 2008; 2:e210. [PubMed: 18382605]

Fitzpatrick S, Watts PC, Feliciangeli MD, Miles MA, Kemp SJ. A panel of ten microsatellite loci for the Chagas disease vector *Rhodnius prolixus* (Hemiptera: Reduviidae). Infect Genet Evol. 2009; 9:206–209. [PubMed: 19061974]

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol. 1994; 3:294–299. [PubMed: 7881515]

Gómez-Sucerquia LJ, Triana-Chávez O, Jaramillo-Ocampo N. Quantification of the genetic change in the transition of *Rhodnius pallescens* Barber, 1932 (Hemiptera: Reduviidae) from field to laboratory. Mem Inst Oswaldo Cruz. 2009; 104:871–877. [PubMed: 19876559]

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

Grisales N, Triana O, Angulo V, Jaramillo N, Parra-Henao G, Panzera F, Gómez-Palacio A. Genetic differentiation of three Colombian populations of *Triatoma dimidiata* (Heteroptera: Reduviidae) by ND4 mitochondrial gene molecular analysis. Biomedica. 2010; 30:207–214. [PubMed: 20890568]

Guhl F, Aguilera G, Pinto N, Vergara D. Updated geographical distribution and ecoepidemiology of the triatomine fauna (Reduviidae: Triatominae) in Colombia. Biomedica. 2007; 27(Suppl. 1):143–162. [PubMed: 18154255]

Gürtler RE, Cecere MC, Canale DM, Castanera MB, Chuit R, Cohen JE. Monitoring house reinfestation by vectors of Chagas disease: a comparative trial of detection methods during a four-year follow-up. Acta Trop. 1999; 72:213–234. [PubMed: 10206120]

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

Harry M, Dupont L, Romaña C, Demanche C, Mercier A, Livet A, Diotaiuti L, Noireau F, Emperaire L. Microsatellite markers in *Triatoma pseudomaculata* (Hemiptera, Reduviidae, Triatominae), Chagas' disease vector in Brazil. Infect Genet Evol. 2008a; 8:672–675. [PubMed: 18571993]

Harry M, Roose CL, Vautrin D, Noireau F, Romaña CA, Solignac M. Microsatellite markers from the Chagas disease vector, *Rhodnius prolixus* (Hemiptera, Reduviidae), and their applicability to *Rhodnius* species. Infect Genet Evol. 2008b; 8:381–385. [PubMed: 18304894]

Herrera-Aguilar M, Be-Barragán LA, Ramirez-Sierra MJ, Tripet F, Dorn P, Dumonteil E. Identification of a large hybrid zone between sympatric sibling species of *Triatoma dimidiata* in the Yucatan peninsula, Mexico, and its epidemiological importance. Infect Genet Evol. 2009; 9:1345–1351. [PubMed: 19786121]

Hudson RR, Slatkin M, Maddison WP. Estimation of levels of gene flow from DNA sequence data. Genetics. 1992; 132:583–589. [PubMed: 1427045]

Lent H, Wygodzinsky P. Revision of the Triatominae (Hemiptera, Reduviidae), and their significance as vectors of Chagas' disease. Bull Am Mus Nat Hist. 1979; 163:125–520.

Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009; 25:1451–1452. [PubMed: 19346325]

López DC, Jaramillo C, Guhl F. Population structure and genetic variability of *Rhodnius prolixus* (Hemiptera: reduviidae) from different geographic areas of Colombia. Biomedica. 2007; 27(Suppl. 1):28–39. [PubMed: 18154243]

Lyman DF, Monteiro FA, Escalante AA, Cordon-Rosales C, Wesson DM, Dujardin JP, Beard CB. Mitochondrial DNA sequence variation among triatomine vectors of Chagas' disease. Am J Trop Med Hyg. 1999; 60:377–386. [PubMed: 10466963]

Mantel N. The detection of disease clustering and a generalized regression approach. Cancer Res. 1967; 27:209–220. [PubMed: 6018555]

Marcet PL, Lehmann T, Groner G, Gürtler RE, Kitron U, Dotson EM. Identification and characterization of microsatellite markers in the Chagas disease vector *Triatoma infestans* (Heteroptera: Reduviidae). Infect Genet Evol. 2006; 6:32–37. [PubMed: 16376838]

Marcet PL, Mora MS, Cutrera AP, Jones L, Gürtler RE, Kitron U, Dotson EM. Genetic structure of *Triatoma infestans* populations in rural communities of Santiago Del Estero, northern Argentina. Infect Genet Evol. 2008; 8:835–846. [PubMed: 18773972]

Mas-Coma S, Bargues MD. Populations, hybrids and the systematic concepts of species and subspecies in Chagas disease triatomine vectors inferred from nuclear ribosomal and mitochondrial DNA. Acta Trop. 2009; 110:112–136. [PubMed: 19073132]

Monteiro FA, Pérez R, Panzera F, Dujardin JP, Galvão C, Rocha D, Noireau F, Schofield C, Beard CB. Mitochondrial DNA variation of *Triatoma infestans* populations and its implication on the specific status of *T. melanosoma*. Mem Inst Oswaldo Cruz. 1999; 94(Suppl. 1):229–238. [PubMed: 10677723]

Monteiro FA, Peretolchina T, Lazoski C, Harris K, Dotson EM, Abad-Franch F, Tamayo E, Pennington PM, Monroy C, Cordon-Rosales C, Salazar-Schettino PM, Gómez-Palacio A, Grijalva MJ, Beard CB, Marcet PL. Phylogeographic pattern and extensive mitochondrial DNA divergence disclose a species complex within the Chagas disease vector *Triatoma dimidiata*. PLoS One. 2013; 8:e70974. [PubMed: 23940678]

Noireau F, Flores R, Vargas F. Trapping sylvatic Triatominae (Reduviidae) in hollow trees. Trans R Soc Trop Med Hyg. 1999; 93:13–14. [PubMed: 10492778]

Panzera F, Ferrandis I, Ramsey J, Ordòñez R, Salazar-Schettino PM, Cabrera M, Monroy MC, Bargues MD, Mas-Coma S, O'Connor JE, Angulo VM, Jaramillo N, Cordón-Rosales C, Gómez D, Pérez R. Chromosomal variation and genome size support existence of cryptic species of *Triatoma dimidiata* with different epidemiological importance as Chagas disease vectors. Trop Med Int Health. 2006; 11:1092–1103. [PubMed: 16827710]

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

Peña V, Fernández G, Gómez-Palacio A, Mejía-Jaramillo A, Cantillo O, Triana-Chávez O. High-resolution melting (HRM) of the cytochrome B gene: a powerful approach to identify blood–meal sources in Chagas disease vectors. PLoS Negl Trop Dis. 2012; 6:e1530. [PubMed: 22389739]

Pérez de Rosas AR, Segura EL, Fichera L, García BA. Macrogeographic and microgeographic genetic structure of the Chagas' disease vector *Triatoma infestans* (Hemiptera: Reduviidae) from Catamarca, Argentina. Genetica. 2008; 133:247–260. [PubMed: 17885811]

Pérez de Rosas AR, Segura EL, García BA. Molecular phylogeography of the Chagas' disease vector *Triatoma infestans* in Argentina. Heredity. 2011; 107:71–79. [PubMed: 21224874]

Pfeiler E, Bitler BG, Ramsey JM, Palacios-Cardiel C, Markow TA. Genetic variation, population structure, and phylogenetic relationships of *Triatoma rubida* and *T. recurva* (Hemiptera: Reduviidae: Triatominae) from the Sonoran Desert, insect vectors of the Chagas' disease parasite *Trypanosoma cruzi*. Mol Phylogenet Evol. 2006; 41:209–221. [PubMed: 16934496]

Piccinali RV, Marcet PL, Noireau F, Kitron U, Gürtler RE, Dotson EM. Molecular population genetics and phylogeography of the Chagas disease vector *Triatoma infestans* in South America. J Med Entomol. 2009; 46:796–809. [PubMed: 19645282]

Piccinali RV, Marcet PL, Ceballos LA, Kitron U, Gürtler RE, Dotson EM. Genetic variability, phylogenetic relationships and gene flow in *Triatoma infestans* dark morphs from the Argentinean Chaco. Infect Genet Evol. 2011; 11:895–903. [PubMed: 21352954]

Pizarro JC, Gilligan LM, Stevens L. Microsatellites reveal a high population structure in *Triatoma infestans* from Chuquisaca, Bolivia. PLoS Negl Trop Dis. 2008; 2:e202. [PubMed: 18365033]

Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000; 155:945–959. [PubMed: 10835412]

Ramírez CJ, Jaramillo CA, del Pilar Delgado M, Pinto NA, Aguilera G, Guhl F. Genetic structure of sylvatic, peridomestic and domestic populations of *Triatoma dimidiata* (Hemiptera: Reduviidae) from an endemic zone of Boyaca, Colombia. Acta Trop. 2005; 93:23–29. [PubMed: 15589794]

Richer W, Kengne P, Cortez MR, Perrineau MM, Cohuet A, Fontenille D, Noireau F. Active dispersal by wild *Triatoma infestans* in the Bolivian Andes. Trop Med Int Health. 2007; 12:759–764. [PubMed: 17550473]

Rodríguez IB, Botero A, Mejía-Jaramillo AM, Marquez EJ, Ortiz S, Solari A, Triana-Chávez O. Transmission dynamics of *Trypanosoma cruzi* determined by low-stringency single primer polymerase chain reaction and southern blot analyses in four indigenous communities of the Sierra Nevada de Santa Marta, Colombia. Am J Trop Med Hyg. 2009; 81:396–403. [PubMed: 19706903]

Rosenberg NA. Distruct: a program for the graphical display of population structure. Mol Ecol Notes. 2004; 4:137–138.

Rousset F. Genepop' 007: a complete re-implementation of the genepop software for Windows and Linux. Mol Ecol Resour. 2008; 8:103–106. [PubMed: 21585727]

Tamay-Segovia P, Alejandre-Aguilar R, Martínez F, Villalobos G, de la Serna FJ, de la Torre P, Laclette JP, Blum-Domínguez S, Espinoza B. Two *Triatoma dimidiata* clades (Chagas disease vector) associated with different habitats in southern Mexico and Central America. Am J Trop Med Hyg. 2008; 78:472–478. [PubMed: 18337346]

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The ClustalX windows interface: xexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 1997; 24:4876–4882. [PubMed: 9396791]

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes. 2004; 4:535–538.

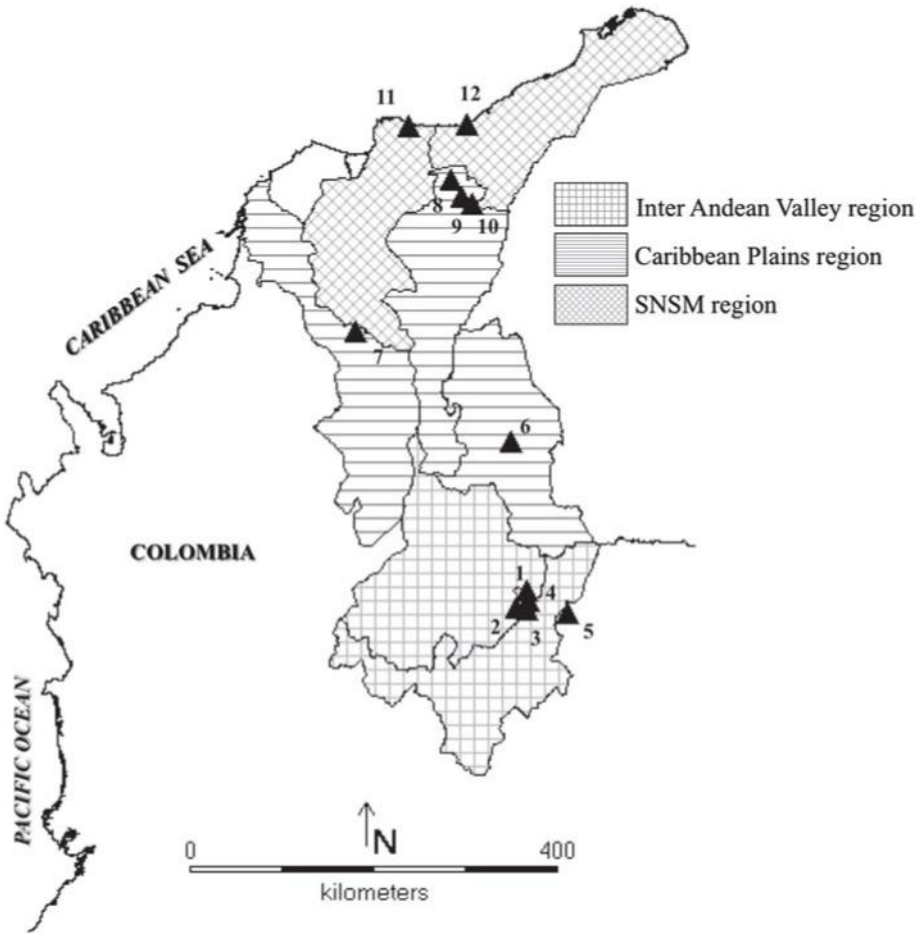

Fig. 1.

Map of Colombia showing the geographical origin of the twelve communities where the *T. dimidiata* used in the analysis were captured: (a) Inter Andean Valleys eco-geographical region: Communities 1–5; (b) Caribbean Plains region: Communities 6–10 and (c) Sierra Nevada de Santa Marta mountain (SNSM): communities 11 and 12. The number of individuals captured per site is detailed in Table 1.

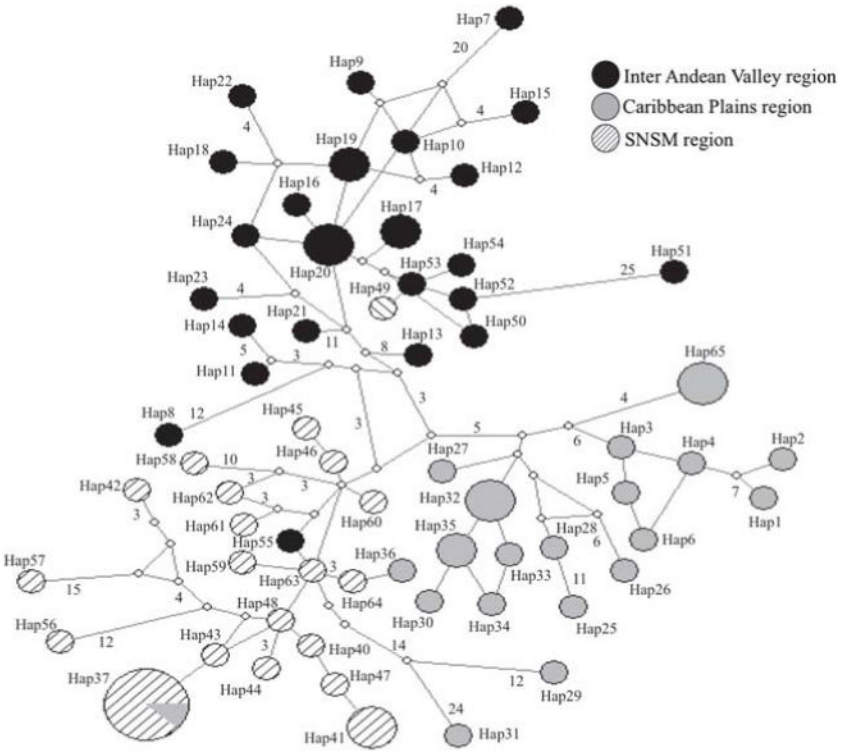

Fig. 2.

Median joining network of COI haplotypes of *T. dimidiata* from Colombia. Node size is proportional to haplotype frequencies, and hypothetical haplotypes or median vectors are indicated as solid white nodes. Numbers between correlative haplotypes represent the number of substitution between pairs (>2 mutation steps).

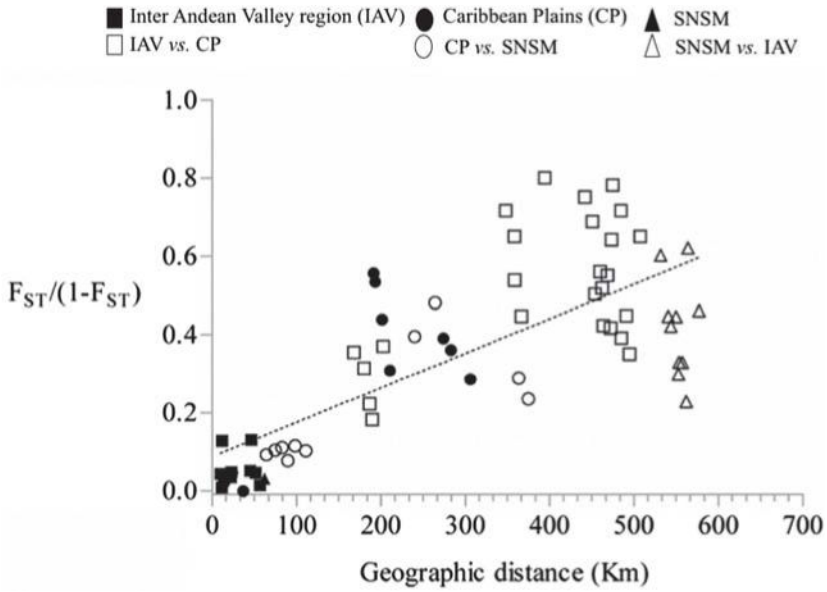

Fig. 3.

Isolation by distance test of *T. dimidiata* of Colombia. Mantel test of geographical distances measured in kilometers and pairwise genetic distances ($F_{ST}/1-F_{ST}$) of the twelve Colombian communities of *T. dimidiata* analyzed. Symbols represent pairwise F_{st} values between communities.

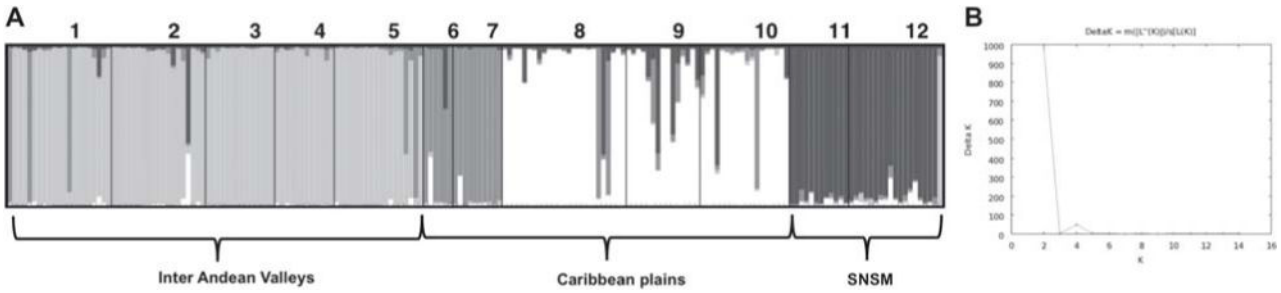

Fig. 4.

Bayesian cluster analysis of twelve *T. dimidiata* communities from three eco-geographical regions of Colombia. Bayesian inference of genetic clusters in Colombian *T. dimidiata* populations. Panel A. Bayesian arrays of individuals from each community (from community 1–12) assigned to each one of the four genetic clusters (grey scale). Each bar represents an individual and the color is the probability with which each individual was assigned to each cluster. Eco-geographical region names of communities are shown in bottom of figure. Panel B. Evannós algorithm (Evanno et al., 2005) testing the rate of change in the log probability of the likelihood associated to each one of the genetic clusters evaluated ($K = 1$ to $K = 15$).

Table 1

Origin, habitat and number of *T. dimidiata* individuals analyzed in this study.

Eco-geographical region	Map no.	Municipality (Department)	Communities	Collection sites	Number of individuals analyzed	
					COI	Microsatellites
Inter Andean Valleys	1	Capitanejo (Santander)	La Chorrera	Houses, Rocks piles and hut for drying tobacco leaves	6	21
	2	Soata (Boyacá)	La Costa	Houses, Rocks piles, Peridomicile, wood stoves and huts for drying tobacco leaves	9	19
	3	Soata (Boyacá)	El Espinal, Soataurban zone	Houses	2	14
	4	Tipacoque (Boyacá)	Tipacoque urban zone	Houses	3	12
	5	Susacon, Saltivanorte, San Mateo, El Espino and Boavita (Boyacá*)	Susaconurban zone, Baracuta, Datal and La Estancia, Huerta Vieja, Guayabal, El Espinourban zone, Espigon, Lagunillas	Houses	8	18
Caribbean plains	6	El Carmen (Norte de Santander)	San Miguel, El Carmen urban zone, Tierra azul and El Hoyo	Houses	3	6
	7	Mompos (Bolivar)	San Fernando, Margarita	Palm trees	6	10
	8	Valledupar (Cesar)	Seyminin	Houses and Chicken hut	5	25
	9	Valledupar (Cesar)	Arwamake, Bechungaka, Chemesquena	Houses, Chicken hut	5	15
	10	Valledupar (Cesar)	Sabana de Crespo, Donachui	Houses and Chicken hut	6	18
Sierra Nevada de Santa Marta Mountain	11	Santa Marta (Magdalena)	Guachaca-Cacahualito, Tarapaca and Mendihuaca	Palm trees	12	19
	12	Dibulla, San juan del Cesar (La Guajira*)	Gumake, Taminaca, Umandita, Marocazo, Ulago	Palm trees and Chicken hut	17	12
Total					82	189

*Insects from adjoining communities were combined and analyzed as one population.

Table 2

(a) Summary of genetic diversity indices and (b) pairwise comparison based on haplotype (*H*st above diagonal) and nucleotide diversities (*K*st below the diagonal) of the COI (402 bp) gene fragment in *T. dimidiata* communities from three eco-geographical regions in Colombia.

Eco-geographical region	***N***	***n***	**π (±SD)**	***h***	***Hd* (±SD)**
COI gene diversity					
Inter Andean Valleys	5	28	0.029 (0.005)	24	0.987 (0.014)
Caribbean Plains	5	25	0.041 (0.007)	20	0.977 (0.019)
SNSM	2	29	0.020 (0.004)	20	0.924 (0.041)
Total	12	82	0.044 (0.003)	63	0.985 (0.007)
Nucleotide-based (Kst)/Haplotype-based (Hst) differentiation statistics					
	Inter Andean Valleys		Caribbean Plains		SNSM
Inter Andean Valleys			0.009**		0.023***
Caribbean Plains	0.238***				0.017***
SNSM	0.334***		0.388***		

Notations: *N*: number of communities; *n*: number of COI sequences; π (±SD): nucleotide diversity (standard deviation); h: number of haplotypes and *Hd* (±SD): Haplotype diversity (standard deviation); ns: not significant.

** $p < 0.01$.

*** $p < 0.001$.

Table 3

Pairwise Fst values (Cockerham and Weir, 1986) between communities (below diagonal) and pairwise lineal geographical distance between localities in kilometers (above diagonal).

Community ID	Inter Andean Valleys					Caribbean plains					SNSM	
	1	**2**	**3**	**4**	**5**	**6***	**7**	**8**	**9**	**10**	**11**	**12**
1	–	22	52	12	23	168	348	450	442	473	531	540
2	0.034	–	45	10	12	190	367	472	464	495	553	562
3	0.011	0.036	–	46	56	202	394	484	475	507	564	577
4	0.078	0.052	0.092	–	14	180	358	462	454	485	543	552
5	0.052	0.025	0.031	0.009	–	187	359	468	460	491	550	557
6*	–	–	–	–	–	–	211	283	274	306	364	375
7	**0.474**	**0.356**	**0.522**	**0.413**	**0.392**	**–**	**–**	**191**	**193**	**201**	**264**	240
8	**0.491**	**0.384**	**0.518**	**0.486**	**0.458**	**–**	**0.365**	**–**	**14**	**24**	**83**	98
9	**0.493**	**0.371**	**0.517**	**0.459**	**0.441**	**–**	**0.291**	**0.037**	**–**	**37**	**90**	111
10	**0.472**	**0.344**	**0.489**	**0.417**	**0.404**	**–**	**0.266**	**0.081**	**0.003**	**–**	**64**	74
11	**0.450**	**0.329**	**0.468**	**0.412**	**0.399**	**–**	**0.331**	**0.144**	**0.080**	**0.133**	**–**	62
12	**0.382**	**0.257**	**0.405**	**0.338**	**0.333**	**–**	**0.269**	**0.132**	**0.096**	**0.136**	**0.039**	–

Values in bold remain significant following sequential Bonferroni correction.

* Not estimated due to low sample size.

Infect Genet Evol. Author manuscript; available in PMC 2016 January 08.

Table 4

Assignment (with probability ⩾0.80) of individual genotypes of Colombian *T. dimidiata* from twelve communities to each one of the four genetic clusters defined by Bayesian Analysis.

Community ID	*n*	cluster1	cluster2	cluster3	cluster4	Total individuals assigned	Eco-geographical region
1	21	17	3			20	Inter Andean valleys
2	19	17			1	18	
3	14	14				14	
4	12	12				12	
5	18	16	2			18	
6	6		5			5	Caribbean plains
7	10		10			10	
8	25		2	21	1	24	
9	15		1	12	2	15	
10	18		1	15	1	17	
11	12				11	11	Sierra Nevada de Santa Marta
12	19	1		1	17	19	