**“Genetic structure and sodium channel gene mutation analysis related
to pyrethroid insecticide toxicity in sylvatic
Andean *Triatoma infestans* from Bolivia”**

Marcet Paula L.^{1*}, Santo-Orihuela Pablo^{2,3}, Messenger Louisa A.^{1,4}, Vassena Claudia V.^{2,5}

¹Centers for Diseases Control and Prevention (CDC), Division of Parasitic Diseases and Malaria;

²Centro de Investigaciones de Plagas e Insecticidas (CONICET/UNIDEF/CITEDEF), Buenos Aires,

Argentina; ³Universidad de Buenos Aires Facultad de Farmacia y Bioquímica, Química Analítica

Instrumental, Buenos Aires, Argentina; ⁴American Society for Microbiology, 1752 N Street, NW

Washington DC, 20036, United States of America; ⁵3iA, Universidad Nacional de San Martín, Buenos

Aires, Argentina.

*Corresponding author: pmarcet@cdc.gov

Pablo Santo-Orihuela psorihuela@gmail.com,

Louisa A Messenger louisa.messenger@lshtm.ac.uk

Claudia V Vassena cvassena@gmail.com,

ABSTRACT

BACKGROUND

Sylvatic populations of *Triatoma infestans* represent a challenge for vector control as these populations are not targeted by control activities and could play a key role in post-spraying house re-infestation. Improved understanding of sylvatic foci, population distribution, dispersion patterns, gene flow between sylvatic and domestic populations, as well as characterization of insecticide resistance profiles, is crucial to optimize vector control interventions.

METHODS

We analyzed the genetic relationship of five Andean populations from Bolivia from localities with distinct insecticide susceptibility profiles (sylvatic: 20 de Octubre, Illicuni, Kirus Mayu and Mataral and one domestic from Mataral). Individual multilocus genotypes based on 8 microsatellites and the DNA sequence of a fragment of the cytochrome B (cytB) gene were obtained for 92 individuals. We compared the cytB haplotypes with previously reported Andean *T. infestans* haplotypes and evaluated the directionality and possibly history of gene flow among populations. Each specimen was screened for 2 nucleotide mutations (L1014 and L9251) of the sodium channel gene (*kdr*), described for *T. infestans* and related to pyrethroid resistance.

RESULTS

Significant genetic differentiation was observed among all populations, reflecting current genetic isolation among them. However, individuals of admixed origin were detected in four populations, especially between the sylvatic and domestic populations from Mataral. Historical analysis of gene flow suggests that insecticide resistance is conferred by ancient trait(s) in *T. infestans* sylvatic populations that are capable of invading domiciles. The *kdr* mutation L1014 was identified in one individual from Mataral, while the L9251 mutation was not detected in any population. The low frequency of *kdr* mutations in these populations suggests this mechanism is unlikely to be the primary cause of the observed altered insecticide susceptibility. However, the resistance conferring mutation is present in the area and with the potential to be selected under insecticidal pressure.

CONCLUSIONS

These results emphasize the need for stronger entomological surveillance in the region, including early detection of house invasion, particularly post-spraying, monitoring for resistance to pyrethroids and the design of integrative control actions that consider both sylvatic foci around domestic settings as well as the bug dispersion dynamics.

Key words: *Triatoma infestans*, insecticide, resistance, pyrethroid, genetic structure, gene flow, *kdr*

55 INTRODUCTION

56 Chagas disease is caused by *Trypanosoma cruzi*, a protozoan parasite that infects humans as
57 well as domestic and sylvatic mammals. This disease can be transmitted by approximately 150 species
58 of hemipteran insect vectors of which *Triatoma infestans* (Klug 1834) (Hemiptera: Reduviidae,
59 Triatominae) is the most important in the southern countries of Latin America [1, 2]. The most recent
60 reports estimate that Chagas disease affects approximately 6 million people worldwide; approximately
61 70 million people are at risk of infection in 21 Latin American countries [3].

62 Domiciliary populations of *T. infestans* have been successfully controlled by spraying with
63 pyrethroid insecticides in parts of southern South America [1, 4]; this contributed to the reduction in
64 the geographical distribution of *T. infestans* from an estimated 6.28 million km² in the 1960s to
65 estimations of less than 1 million km² today [1, 5]. Moreover, several countries and regions in South
66 America, have been certified as having interrupted domestic vector transmission [6].

67 The “Gran Chaco Region”, which covers 1.3 million km² including territories in Bolivia,
68 Paraguay and Argentina, is highly endemic for Chagas disease transmission; it has proven to be
69 particularly difficult to control domestic infestation by *T. infestans* in the region. Rapid house re-
70 infestation post-spraying and insecticide resistance are among the factors associated with vector control
71 failure in parts of this region [7-11].

72 *T. infestans* has been found in sylvatic foci in several areas in the Chaco region, some
73 reportedly invading houses and are associated with *T. cruzi* transmission to humans [12-23]. Sylvatic
74 populations are not targeted by control interventions and could re-invade houses treated with
75 insecticide, especially if the population has resistant variants that can be selected under insecticidal
76 pressure. Remarkably, certain sylvatic populations from Bolivia have shown low susceptibility to
77 pyrethroid insecticides [24], and some also exhibited a lower sensitivity to fenitrothion [25].

78 Pyrethroid resistance in *T. infestans* was initially described in Argentina and Bolivia in the late
79 1990s. It later became implicated in failure of local control programs, where certain populations from
80 the Gran Chaco area required up to 1000 times the amount of insecticide to kill the same proportion of
81 susceptible individuals [26]. Increasing reports indicate that several natural populations are resistant to
82 insecticides, and significant variability in susceptibility levels has been observed among them [27-30].

83 Resistance traits are considered pre-adaptive, implying that there are a few resistant insects in each
84 natural population, which would survive under the selective pressure of insecticide exposure, selecting
85 for the resistant genetic variants that would be inherited and in turn, become more frequent in
86 subsequent generations. The dispersal rate of resistant traits will depend on pre-existing genetic
87 variation in natural populations, selection intensity, fitness cost associated to expressing the resistant
88 variant, local adaptation and levels of gene flow among populations. Such resistance dynamics have

been reported in populations of other medically-important vector species, particularly *Aedes aegypti* [31, 32]. Following this rationale, three general assumptions can be made regarding population genetic diversity and insecticide resistance:

- a) Insecticide resistant triatomine populations are expected to be less genetically diverse than a wild-type population as they have likely undergone a period of selective pressure (i.e. insecticide exposure).
- b) Sylvatic populations are expected to be more genetically diverse than domestic populations as they are in theory, ancient progenitors, that have accumulated diversification over time, in the absence of selective pressure, as they are not being targeted by vector control activities.
- c) Domestic populations are expected to have been the subject of greater selection pressures, compared to their sylvatic counterparts and therefore be much less diverse.

Elucidating population genetic dynamics through high resolution genetic markers in the context of well characterized insecticide resistance profiles could test these hypothesis and provide invaluable information for an integrated vector control approach. Establishing the role of sylvatic populations of *T. infestans* as potential sources of re-infestation and understanding the impact of insecticide resistance on triatomine population genetics is crucial to safeguard the continued effectiveness of current control efforts. Previous studies using high resolution molecular markers demonstrated that genetic differentiation is detectable up to the capture site level (local populations) if sufficient numbers of bugs are captured per site to constitute a representative population for genetic analyses comparisons [33-36].

The aim of this work was to establish and characterize the genetic profiles of four sylvatic and one domestic Andean populations of *Triatoma infestans* from Bolivia, and to interpret their genetic relationships in the context of their differing levels of insecticide susceptibility. Moreover, an evaluation of the mechanisms likely responsible for the pyrethroid resistance observed in these populations including esterase enzymes and *kdr*-mutations is discussed.

MATERIALS AND METHODS

Insect origin Collection methods were described in detail in [24]. Briefly, *T. infestans* populations were sampled from domiciliary areas in Mataral, Department of Cochabamba, Bolivia (hereby named Mataral-D) and a sylvatic area located approximately 2 km away (Mataral-S). Sylvatic bugs were also collected from sites in Cochabamba (Illicuni and 20 de Octubre) and Potosí (Kirus Mayu). Sylvatic

triatomines were sampled from rock-piles using mouse-baited sticky traps [37]. From each population a colony was established with a minimum of 10 founder individuals, and maintained at $28\pm 1^{\circ}\text{C}$, 50% RH, and a photoperiod of 12:12 (L:D) h at the Centro de Investigaciones de Plagas e Insecticidas (CIPEIN, CITEFA-CONICET), Buenos Aires, Argentina. Rearing conditions are described in detail elsewhere [38, 39]. For this study, bugs from each population from the second or third colony generations were used. Genomic DNA was obtained from three legs of each bug using the Wizard Genomic Purification Kit (Promega[®]) following the manufacturer recommendations.

The insecticide resistance profile of these populations was established by toxicological bioassays performed according to World Health Organization [40] standard protocols [24].

Multilocus microsatellite genotyping A multilocus genotype was obtained for each individual based on nine microsatellite loci, developed and optimized for *T. infestans* following the amplification conditions previously described [33, 41]. Multiplex PCR reactions were carried out with 2 pairs of loci with the same annealing temperature and different dye colors (Tinf_ms5 with Tinf_ms42 and Tinf_ms22 with Tinf_ms27). DNA fragment detection with 1bp resolution was performed with an automated DNA sequencer (ABI 3130, Applied Biosystems) and size determination was obtained with GeneMapper 4.1 (Applied Biosystems). Ninety-two colony-reared, second-third generation individuals were sampled for microsatellite genotyping (Table 1).

Microsatellite analysis. Allele number per locus and population were obtained directly after binning, and mean allele number among all loci were compared. The inbreeding coefficient **F_{is}** [42]; allelic richness **A_r** [43], gene diversity or expected heterozygosity (**H_e**) and observed heterozygosity (**H_o**) [44] were obtained with the Excel add-in MS_tools.xla for Microsoft Excel or with FSTAT2.9.3.2 [45]. Bonferroni correction was applied to determine p-values for multiple comparisons [46]. The fit to Hardy-Weinberg (HW) equilibrium expectations was evaluated using the U score test available in Genepop 4.2, under the null hypothesis of random union of gametes. Sample-size- corrected private (population-specific) allele frequency per locus (PA/L) was calculated in HP-RARE [47]. Multilocus linkage disequilibrium, estimated by the index of association (**I_A**), was calculated in MULTILOCUS 1.3b [48] and statistical significance was evaluated by comparison with a null distribution of 1000 randomizations.

Pairwise **F_{ST}** values and significance levels were calculated with ARLEQUIN 3.01 as in Weir & Cockerham [42]. Fisher exact test for population differentiation comparing genic and genotypic frequencies was implemented in Genepop 4.2. Population clustering was explored using a neighbor-

joining (NJ) tree based on pairwise distances (D_{AS} : 1 proportion of shared alleles at all loci/n). A Mantel's test for the effect of isolation by distance within populations (pairwise genetic vs. geographic distance) was performed in GenAlEx 6.5 using 10,000 random permutations [49]. A Hierarchical Analysis of Molecular Variance (AMOVA) [50] was performed at three levels of population structure: among communities, among populations and among individuals in a population, in ARLEQUIN 3.01 [51].

Population structure. Population genetic structure was examined using the Bayesian model-based approach [52-54] implemented in STRUCTURE 2.3.4. Each individual MLG in the sample is probabilistically assigned to one of K populations, or jointly to two or more populations if their genotypes might have an admixed origin. Simultaneously, the method determines the number of significant K genetic clusters within the total sample. The number of clusters evaluated ($=K$) ranged from 1 to 10. The analysis was performed using 35 replicate runs per K value, a burn-in period length of 50,000 and a run length of 50,000. The analysis model used was the admixed and correlated allele frequencies, with no prior information on the origin of the individuals. The final selection of the sample K value was based on the log-probability of the data between successive K values [55] implemented in the online version of STRUCTURE HARVESTER [56].

Mitochondrial (mtDNA) analysis. The cytB gene fragment was amplified as previously described [57] for a total of 74 specimens from the five populations (Table 1). SeqManPro Software 14.1 (DNASTar) was used to assemble and align forward (F) and reverse (R) chromatograms, and obtain the consensus sequence for each individual. MEGA 5 software [58] was used to align multiple sequences and to examine the phylogenetic relationships among sequences. Standard genetic variability (haplotype diversity, H_d and nucleotide diversity: π) and differentiation among sequences were evaluated using DnaSP version 5.1 software [59]. A Nexus matrix was constructed for haplotype network analysis in PopART using a median-joining model based on 1000 iterations with default parameters [60, 61].

To evaluate the possible directionality and history of gene flow among populations, we compared the haplotypes obtained in this work with previously reported Andean *T. infestans* haplotypes. Sequences were trimmed from 666bp to 388bp and a second haplotype network was built using the same software and criteria including 25 sequences from sylvatic Bolivian *T. infestans* previously deposited in GenBank: HQ333215-HQ333239 [15, 62].

Screening for mutations in the *kdr* gene A fragment of the DNA sequence of the sodium channel gene was amplified and screened for two nucleotide mutations previously reported in *T. infestans* associated with pyrethroid knockdown resistance (*kdr*) [63, 64]. The protocol has been previously described [65] and was used to screen 67 individuals from the five populations (Table 1).

Results

Population diversity analysis

Ninety-two individual *T. infestans* samples were grouped *a priori* into five populations based on their geographical origin (Table 1): Mataral domestic (Mat-D), Mataral sylvatic (Mat-S), Ilicuni (Ili), 20 de Octubre (20oct) and Kirus Mayu (KM) and were genotyped across eight polymorphic microsatellites. In contrast with genotyping data from *T. infestans* in Argentina and Paraguay, the locus Tims_64 did not amplify consistently across samples. In addition, locus Tims_65 did not amplify in either domestic or sylvatic Mataral populations, indicative of null alleles; therefore, neither loci were utilized in the final population analysis.

A total of 84 unique individual multilocus genotypes (MLG) were identified across populations (Table 2). Genetic diversity levels were high and consistent across most populations as evidenced by similar frequencies of unique MLGs, levels of allelic richness (3.1-5.4); distance between shared alleles (0.38-0.67); private alleles per locus (0.42-1.26) and average gene diversity (0.44-0.74) (Table 2). Only minor differences were observed between populations with larger numbers of founding individuals (e.g. 20oct and KM), indicating minimal genetic diversity was lost by culturing population members prior to genetic analysis.

Although it had the lowest number of individuals available for testing, the sylvatic population from Mataral displayed the highest parameters related to genetic diversity (average number of alleles = 5.5, $Ar = 5.37$, $He = 0.74$, $av PIC = 0.65$). Moreover, individuals from this area demonstrated the highest number of private alleles per locus (1.26) and Fis value (0.22) indicative that this may be an ancestral population with restricted intra-population gene flow (i.e. sub-structure), respectively (Table 2). The populations 20oct, KM and Mat-S presented significant deviations from HW equilibrium, and all due to significant heterozygote deficits.

Inter-population subdivision and gene flow

Both genic and genotypic Fisher tests for pairwise population differentiation were highly significant for all pairs compared. A hierarchical AMOVA, which evaluated the distribution of genetic diversity, demonstrated that the majority of variation (71.5%) was within individuals, compared to 24.1% and 4.4% among populations and among individuals within populations, respectively. A population neighbor joining (NJ) clustering analysis based on pair-wise genetic distances D_{AS} [66] is shown in Figure 1. The NJ topology was consistent with geographical origin and F_{ST} distances. The more proximate populations, both Mataral populations, were grouped together and closer to Illicuni (separated by approximately 60 km), while the more distant populations of 20oct and KM were more closely related.

Estimates of population subdivision based on allele frequency and identity (F_{ST}) indicate significant genetic differentiation between all populations, with the lowest F_{ST} values found between both populations of Mataral and Illicuni (Table 3). F_{ST} values among sylvatic populations spanning a greater geographical distance (e.g. $F_{ST}=0.27$ between Mat-S and 20Oct; 162.5 km) were largely equivalent to those observed between domestic and sylvatic triatomines sampled from a more restricted area ($F_{ST}=0.22$ between Mat-D and KM; 57 km) and between adjacent sylvatic populations ($F_{ST}=0.22$ between Mat-S and KM 102.06 km), suggesting little gene flow occurred among these populations throughout the geographic range considered. The Mantel's test conducted to evaluate the significance of the correlation between the geographical and genetic distances between population pairs, showed a slight but significant adjustment to an isolation by distance model ($R_{XY}=0.413$; $P<0.01$).

A Bayesian assignment of individual genotypes determined an optimal number of five genetic clusters to describe the dataset (Figures 2 and 3), corresponding broadly to the five populations assigned by capture site. Most individual MLGs clustered significantly with other members from the same geographic location. However, in all sylvatic populations, individuals that could be first generation migrants or of admixed origin were detected (Figure 3).

A hierarchical comparison of alternative clustering models reflects the relative genetic distance among MLGs. This analysis shows the resulting clustering pattern of MLGs from the whole sample when using different K values (i.e. considering alternative numbers of genetic clusters). The results are presented in Figure 3, which shows the closer genetic relationship between both Mataral populations and Illicuni, as the individuals from those populations cluster together when considering four or three possible genetic clusters to assign the whole dataset.

Mitochondrial genetic clustering among populations and individuals

A fragment of the cytochrome B gene (cytB) was amplified and sequenced from 74 individuals (20oct=14, Ili=15, KM=15, Mat-S=11 and Mat-D=19), and assembled into a 666 bp alignment. A total of eight haplotypes were observed, including an ambiguous (i.e. heteroplasmid) sequence detected in two individuals from KM (ht). All haplotypes were submitted and are accessible in GenBank (access numbers MH763648 to MH763654). Haplotypes were determined by 15 variable sites that were mostly synonymous except for sites 28 and 625, carried by haplotypes XVI and XLIX (Table 4).

The most common mitochondrial genotype was XVI, identified in 21 isolates from 20oct, Ili and KM, but absent from both Mataral populations (Table 5, Figure 4). Mat-D and Ili displayed the highest haplotype diversity ($H_d=0.65$ and 0.59 , respectively), and were the populations that also presented the most diverse haplotypes ($\pi=0.0056$ and 0.0044 , respectively) (Table 5). Both Mataral populations presented private unique haplotypes (XLV in Mat-D and XLVI and XLVII in Mat-S, respectively). Only one of the five haplotypes observed in Mat-S and Mat-D was shared between the two populations (XXIV).

The distribution of mitochondrial haplotypes across populations corresponded broadly to the geographical distances between populations, with KM sharing two genotypes with 20oct (XVI and XLIX); likewise Ili and both Mataral populations contained more closely-related cytB sequences. The phylogenetic tree and network analysis, clustered haplotypes into two groups based on the substitution A/G in nucleotide position 162 (Table 4). One cluster contained haplotypes XLVI and XLVII (only detected in Mat-S population) with XXIV, found in both Mataral populations and Ilicuni. The remaining haplotypes were found in all but the Mat-S population (Figure 4).

In order to interpret the results of this study in the context of known phylogeographic patterns of *T. infestans* populations in Bolivia, we built a haplotype network with previously reported haplotypes for Bolivian *T. infestans* [15, 62]. As the sequences previously reported were only 388bp long, the haplotypes we characterized were trimmed to that length for comparison. Consequently, some of our trimmed haplotypes were indistinguishable from each other within the fragment length considered, and were pooled into four haplotypes previously reported in Andean Bolivian populations: A, C, E and R [15] which are shown in red in Figure 4.

Remarkably, the haplotype network built with sequences from previous studies of Bolivian *T. infestans* (Fig. 4) showed that the haplotypes found in Mataral-S correspond to a central node (i.e. ancient), haplotype E, composed of populations from the “Central Andean area” collected in two locations in Cochabamba, Mataral [15] and Chuquisaca [67]. Another haplotype found in Mataral-S which has also been found in Mataral-D and Ilicuni (haplotype XXIV from this work), corresponded

to the previously reported haplotype C, a rare divergent haplotype from most of the others reported and considered a derivate from haplotype E [15].

In this comparison, the three haplotypes found in 20 Octubre, Kirus Mayu, Ilicuni and Mataral-D populations correspond to that described as the most abundant haplotype A, previously detected in specimens from “Northern and Central Andean areas” [15]. Finally the haplotype XLIX found in 20 Octubre, Kirus Mayu and Ilicuni correspond to haplotype R, considered a derivate of haplotype A [15]. Haplotypes A, R and C had also been previously reported in domestic populations [67].

***Kdr* screening**

Among the sixty-seven samples screened for two mutations in the *kdr* gene, related to insecticide resistance in *T. infestans*, only one individual from Mataral-S presented the resistant allele variant for the site L1014. No individual carrying the mutation L9251 was detected.

Discussion

Through high resolution genotyping of five Andean *T. infestans* populations, we provide insights into natural population genetic structures and the relationships between triatomine genetic diversity and insecticide resistance. All populations were genetically distinct, which correlates with previously described heterogeneity in insecticide susceptibility profiles [24]. The pattern of inter-population differentiation was consistent with a significant isolation-by-distance (IBD) model, with geographically closer populations being more genetically-related. This observation was supported by both mitochondrial and nuclear microsatellite markers. The genetic pattern of allelic similarity reflected possible gene flow and/or secondary contact between domestic and sylvatic Mataral populations. This observation furthers supports previous reports of dynamic gene flow between sylvatic and domestic populations in the Bolivian Andes [15, 36, 62].

Both populations from Mataral were characterized by considerably higher genetic diversity compared to the other sylvatic populations evaluated in this work, and between the Mataral populations, genetic diversity and substructure were higher for the sylvatic population. Furthermore, high levels of homozygosity and numbers of private alleles per locus suggest that Mataral-S may be ancestral to the other sylvatic populations evaluated here. This observation is independently supported by the occurrence of two mitochondrial haplotypes within this population that correspond to haplotype E, which has been previously reported as an ancient haplotype among other Andean *T. infestans* populations [15].

The populations 20oct, KM and Mat-S presented significant deviations from HW equilibrium, and all due to significant heterozygote deficits. As the samples were arbitrarily grouped per community, the significant F_{is} values are likely attributable to the population being comprised of admixed individuals from multiple, independent non-panmictic founding populations (Wahlund effect). This would also account for the higher linkage disequilibrium detected in these populations. High standard deviations associated with D_{AS} values further supports the existence of intra-population sub-structuring. Alternatively, this genetic pattern could evidence that some members have undergone long-term inbreeding, following historical isolation.

The bioassays performed in these populations demonstrated three different levels of altered susceptibility [24]. High levels of decreased susceptibility in Mataral-D and Mataral-S, moderate levels of decreased susceptibility in Kirus-Mayu and 20 de Octubre and limited altered susceptibility in Ilicuni (Table 1). The RRs obtained for these populations are not as high as values from other areas that have presented difficulties for vector control (i.e. $RR > 2000$ El Jramento, Guemes, Chaco, Argentina [26] or $RR_{50} = 541.6$, Tierras Nuevas, Tarija, Bolivia [68]). However, in all but Ilicuni, the RR values are above the level classified as resistant by PAHO guidelines $RR > 5$ (fivefold higher than reference populations) [30, 69]. Those guidelines also attribute altered susceptibility in field populations presenting $RR < 5$ to individual variability, which is not enough to cause resistance at the population level. As previously discussed [25, 70], elevated RR in a sylvatic population might result from at least three non-mutually exclusive hypotheses:

- a) existence of naturally altered (lower) susceptibility to insecticides
- b) development of resistance resulting from exposure to insecticides used in agriculture and vector control campaigns
- c) contact and exchange of genetic variants between sylvatic and domestic resistant populations

In the context of this study, the decreased susceptibility to deltamethrin observed in all populations may represent a shared ancestral trait, which has undergone further intensification in Mataral following insecticidal treatment, particularly for the domestic population, with possible geneflow between the two Mataral populations, including the dispersal of bugs with altered susceptibility to Pyrethroids. In this scenario, the intermediate RRs of 20 de Octubre and Kirus-Mayu would have been maintained by moderate selection pressure, likely the indirect effect of agricultural spraying; the difference between them could be explained by intrinsic intra-species variability, consistent with the distinct genetic patterns observed. However, the minimally altered susceptibility observed in Ilicuni would imply that this trait had been almost lost in this population in the absence of

selection. Alternatively, insecticide resistance could have been acquired independently in all populations following different histories of insecticide exposure, with population RR directly corresponding to the intensity of selection each population has undergone. Further characterization of the ecological context of the population in this area is necessary to test these hypothesis or propose alternative explanations to observed the genetic pattern.

Insecticide-based interventions are expected to impact vector population structure and levels of genetic diversity. Selective bottlenecks imposed by insecticidal pressure are likely to result in both reductions in genetic variability and population size, such that a small proportion of survivors, possessing the requisite advantageous resistance genes, survive and reproduce (i.e. diversification that would be reflected in a genetic pattern as multiple independent “island effects”). In general, insecticide resistant populations are expected to be less genetically diverse than a wild-type population, as they have likely undergone a period of selective pressure (i.e. insecticide exposure). And within this context, domestic populations are assumed to have been the subject of greater selection pressures, compared to their sylvatic counterparts and therefore be less diverse.

However, our study results challenge these traditional paradigms and may instead be explained by an alternative hypothesis for genetic diversification. In a given location, a population may undergo a selection process from insecticide exposure independent of neighboring triatomine populations [71]. A previous work comparing the genetic profiles of populations under different vector control pressure [33] described populations within a community recovering from a recent massive spraying as much more genetically structured and diverse than populations that were not under systematic vector control pressure. The authors proposed that the distinct observed genetic pattern reflected the cumulative effect of migration among adjacent populations for several generations, which generated a pattern of genetic “homogenization” even among populations located geographically far apart. In contrast, populations recovering from a pool of survivors of insecticide spraying were observed as discrete genetic entities at the individual capture site (i.e. one household). Pooling together such individuals from these “genetically discrete” populations with bugs from other sites, for a random sampling of a given area, (i.e. at the community level, pooling individuals from different houses in the same sample set), would result in a highly genetically diverse metapopulation, that is actually composed of several subpopulations, with genetic substructure and gene flow restrictions within. In this study, the elevated genetic diversity of both Mataral populations could represent the assemblage of individuals from distinct genetic units, as they have been under higher selective pressure than the sylvatic populations from 20 de Octubre, Kirus-Mayu and Illicuni, located further away from the domestic area subjected to vector control. Great genetic diversity has also been reported in several domestic non-Andean *T.*

infestans populations from the Gran-Chaco area [35, 72], of which many had been under extensive vector control pressure.

More importantly, the existence of genetically-diverse populations displaying high levels of altered insecticide resistance has potential implications for longer-term efficacy of control interventions, such that the likelihood of individuals surviving future selection events is greater when the original population is composed of more unique genetic variants. It is noteworthy that three haplotypes found in the sylvatic populations analyzed here (A, C and E) were also identified in domestic populations in the department of Cochabamba [67], indicative of household colonization by sylvatic populations in the Bolivian Andes, a pervasive and contemporary process.

Regarding the resistance mechanism(s) underlying the altered susceptibility patterns in Mataral, analyses of P450 monooxygenases and 7-CP esterases [24] showed a lack of association between enzymatic activities and susceptibility ratios, indicating that these enzymes were not responsible for modified susceptibility in these sylvatic populations. In this work, we evaluated the presence of two *kdr* mutations and found one individual from sylvatic Mataral carrying the mutation L1014 [63]. This result indicates that although not likely responsible for the mechanism altering insecticide susceptibility in these populations, the presence of the mutation in this area, represents a high risk for selection under pressure and potential development of higher resistance intensities [73]. Future work to characterize the underlying metabolic mechanisms of resistance in these populations is warranted.

Conclusions

Sylvatic populations of *Triatoma infestans* represent a challenge for vector control. These populations are not targeted by control activities and could play a key role in post-spraying house re-infestation. The results presented here suggest that there might be an ancient trait(s) that confers resistance in *T. infestans* sylvatic populations that are capable of invading domiciles. These observations emphasize the need for stronger entomological surveillance in the region, including early detection of house invasion, particularly post-spraying, monitoring for resistance to pyrethroids and the design of integrative control actions that consider both sylvatic foci around domestic settings as well as the bug dispersion dynamics.

416 **Declarations**

417 **Ethics approval and consent to participate**

418 Not applicable.

419 **Consent for publication**

420 Not applicable.

421 **Availability of data and material**

422 DNA sequences obtained in this study are available in Genebank.

423 **Competing interests**

424 The authors declare that they have no competing interests.

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

427 **Funding**

428 This study received financial support from CONICET (Consejo Nacional de Investigaciones
429 Científicas y Técnicas. Argentina PIP 411/2012) and ANPCyT (Agencia Nacional de Promoción
430 Científica y Tecnológica- Argentina). LAM is supported by an American Society for
431 Microbiology/Centers for Disease Control and Prevention Fellowship.

432 **Authors' contributions**

433 PLM, PSH, CVV conceived the idea and designed the work. PLM carried out the laboratory work.
434 PLM and LAM performed the analysis and took the lead in writing the manuscript. All authors
435 provided critical feedback for the manuscript and contributed to the conclusions of the work. All
436 authors read and approved the final version of the manuscript.

437 **Acknowledgements**

438 No applicable

439

440

441

442 **References**

- 443 1. Schofield CJ, Jannin J, Salvatella R: The future of Chagas disease control. *Trends in*
444 *Parasitology* 2006, 22:583-588.
- 445 2. Gurtler RE, Diotaiuti L, Kitron U: Commentary: Chagas Disease: 100 years since discovery
446 and lessons for the future. *International Journal of Epidemiology* 2008, 37:698-701.
- 447 3. WHO: Chagas disease in Latin America: an epidemiological update based on 2010 estimates.
448 In *Weekly Epidemiological Record (WER)* (Organization WH ed., vol. 90. pp. 33-44. Geneva,
449 Switzerland: WHO; 2015:33-44.
- 450 4. Zerba EN: Susceptibility and resistance to insecticides of Chagas disease vectors. *Medicina (B*
451 *Aires)* 1999, 59 Suppl 2:41-46.
- 452 5. Gorla D: Variables ambientales registradas por sensores remotos como indicadores de la
453 distribución geográfica de *Triatoma infestans* (Heteroptera: Reduviidae) *Ecología Austral*
454 2002, 12:117-127.
- 455 6. Coura JR, Viñas PA, Junqueira ACV: Ecoepidemiology, short history and control of Chagas
456 disease in the endemic countries and the new challenge for non-endemic countries. *Memórias*
457 *do Instituto Oswaldo Cruz* 2014, 109:856-862.
- 458 7. Cavallo MJ, Amelotti I, Gorla DE: Invasion of rural houses by wild Triatominae in the arid
459 Chaco. *Journal of Vector Ecology* 2016, 41:97-102.
- 460 8. Gurevitz JM, Ceballos LA, Gaspe MS, Alvarado-Otegui JA, Enriquez GF, Kitron U, Gurtler
461 RE: Factors Affecting Infestation by *Triatoma infestans* in a Rural Area of the Humid Chaco in
462 Argentina: A Multi-Model Inference Approach. *PLoS Negl Trop Dis* 2011, 5:e1349.
- 463 9. Gurevitz JM, Gaspe MS, Enriquez GF, Provecho YM, Kitron U, Gürtler RE: Intensified
464 Surveillance and Insecticide-based Control of the Chagas Disease Vector *Triatoma infestans* in
465 the Argentinean Chaco. *PLoS Negl Trop Dis* 2013, 7:e2158.
- 466 10. Cecere MC, Vazquez-Prokopec GM, Gürtler RE, Kitron U: Reinfestation sources for Chagas
467 disease vector, *Triatoma infestans*, Argentina. *Emerging Infectious Diseases* 2006, 12:1096-
468 1102.
- 469 11. Lardeux F, Depickere S, Duchon S, Chavez T: Insecticide resistance of *Triatoma infestans*
470 (Hemiptera, Reduviidae) vector of Chagas disease in Bolivia. *Trop Med Int Health* 2010.
- 471 12. Ceballos LA, Piccinali RV, Marcet PL, Vazquez-Prokopec GM, Cardinal MV, Schachter-
472 Broide J, Dujardin J-P, Dotson EM, Kitron U, Gürtler RE: Hidden Sylvatic Foci of the Main
473 Vector of Chagas Disease *Triatoma infestans*: Threats to the Vector Elimination Campaign?
474 *PLoS Negl Trop Dis* 2011, 5:e1365.
- 475 13. Cortez MR, Emperaire L, Piccinali RV, Gurtler RE, Torrico F, Jansen AM, Noireau F: Sylvatic
476 *Triatoma infestans* (Reduviidae, Triatominae) in the Andean valleys of Bolivia. *Acta Trop*
477 2007, 102.
- 478 14. Waleckx E, Depickere S, Salas R, Aliaga C, Monje M, Calle H, Buitrago R, Noireau F,
479 Breniere SF: New discoveries of sylvatic *Triatoma infestans* (Hemiptera: Reduviidae)
480 throughout the Bolivian Chaco. *Am J Trop Med Hyg* 2012, 86:455-458.
- 481 15. Waleckx E, Salas R, Huamán N, Buitrago R, Bosseno M-F, Aliaga C, Barnabé C, Rodriguez
482 R, Zoveda F, Monje M, et al: New insights on the Chagas disease main vector *Triatoma*
483 *infestans* (Reduviidae, Triatominae) brought by the genetic analysis of Bolivian sylvatic
484 populations. *Infection, Genetics and Evolution* 2011, 11:1045-1057.
- 485 16. Ceballos LA, Piccinali RV, Berkunsky I, kitron U, Gürtler RE: First Finding of Melanic
486 Sylvatic *Triatoma infestans* (Hemiptera: Reduviidae) Colonies in the Argentine Chaco. *J Med*
487 *Entomol* 2009, 46:1195-1202.
- 488 17. Noireau F, Bastrenta B, Catala S, Dujardin JP, Panzera F, Torres M, Perez R, Galvao C,
489 Jurberg J: Sylvatic population of *Triatoma infestans* from the Bolivian Chaco: from field
490 collection to characterization. *Mem Inst Oswaldo Cruz* 2000, 95 Suppl 1:119-122.

- 491 18. Noireau F, Cortez MG, Monteiro FA, Jansen AM, Torrico F: Can wild *Triatoma infestans* foci
492 in Bolivia jeopardize Chagas disease control efforts? *Trends Parasitol* 2005, 21:7-10.
- 493 19. Rolón M, Vega MC, Román F, Gómez A, Rojas de Arias A: First Report of Colonies of
494 Sylvatic *Triatoma infestans* (Hemiptera: Reduviidae) in the Paraguayan Chaco, Using a
495 Trained Dog. *PLoS Negl Trop Dis* 2011, 5:e1026.
- 496 20. Buitrago NLR, Bosseno MF, Waleckx E, Brémond P, Vidaurre P, Zoveda F, Brenière SF: Risk
497 of transmission of *Trypanosoma cruzi* by wild *Triatoma infestans* (Hemiptera: Reduviidae) in
498 Bolivia supported by the detection of human blood meals. *Infection, Genetics and Evolution*
499 2013, 19:141-144.
- 500 21. Buitrago R, Bosseno MF, Depickere S, Waleckx E, Salas R, Aliaga C, Barnabe C, Breniere
501 SF: Blood meal sources of wild and domestic *Triatoma infestans* (Hemiptera: Reduviidae) in
502 Bolivia: connectivity between cycles of transmission of *Trypanosoma cruzi*. *Parasit Vectors*
503 2016, 9:214.
- 504 22. Buitrago R, Waleckx E, Bosseno MF, Zoveda F, Vidaurre P, Salas R, Mamani E, Noireau F,
505 Breniere SF: First report of widespread wild populations of *Triatoma infestans* (Reduviidae,
506 Triatominae) in the valleys of La Paz, Bolivia. *Am J Trop Med Hyg* 2010, 82:574-579.
- 507 23. Breniere SF, Buitrago R, Waleckx E, Depickere S, Sosa V, Barnabe C, Gorla D: Wild
508 populations of *Triatoma infestans*: Compilation of positive sites and comparison of their
509 ecological niche with domestic population niche. *Acta Trop* 2017, 176:228-235.
- 510 24. Roca-Acevedo G, Cueto GM, Germano M, Orihuela PS, Cortez MR, Noireau F, Picollo MI,
511 Vassena C: Susceptibility of sylvatic *Triatoma infestans* from Andean valleys of Bolivia to
512 deltamethrin and fipronil. *J Med Entomol* 2011, 48:828-835.
- 513 25. Santo-Orihuela PL, Carvajal G, Picollo MI, Vassena CV: Toxicological and biochemical
514 analysis of the susceptibility of sylvatic *Triatoma infestans* from the Andean Valley of Bolivia
515 to organophosphate insecticide. *Mem Inst Oswaldo Cruz* 2013, 108:790-795.
- 516 26. Fronza G, Toloza AC, Picollo MI, Spillmann C, Mougabure-Cueto GA: Geographical
517 Variation of Deltamethrin Susceptibility of *Triatoma infestans* (Hemiptera: Reduviidae) in
518 Argentina With Emphasis on a Resistant Focus in the Gran Chaco. *Journal of Medical*
519 *Entomology* 2016, 53:880-887.
- 520 27. Germano MD, Roca Acevedo G, Mougabure Cueto GA, Toloza AC, Vassena CV, Picollo MI:
521 New findings of insecticide resistance in *Triatoma infestans* (Heteroptera: Reduviidae) from
522 the Gran Chaco. *J Med Entomol* 2010, 47:1077-1081.
- 523 28. Mougabure-Cueto G, Picollo MI: Insecticide resistance in vector Chagas disease: Evolution,
524 mechanisms and management. *Acta Trop* 2015, 149:70-85.
- 525 29. Picollo MI, Vassena C, Santo Orihuela P, Barrios S, Zaidenberg M, Zerba E: High resistance
526 to pyrethroid insecticides associated with ineffective field treatments in *Triatoma infestans*
527 (Hemiptera: Reduviidae) from Northern Argentina. *J Med Entomol* 2005, 42:637-642.
- 528 30. Pessoa GC, Vinas PA, Rosa AC, Diotaiuti L: History of insecticide resistance of Triatominae
529 vectors. *Rev Soc Bras Med Trop* 2015, 48:380-389.
- 530 31. Aguirre-Obando OA, Bona ACD, Duque L. JE, Navarro-Silva MA: Insecticide resistance and
531 genetic variability in natural populations of *Aedes* (Stegomyia) *aegypti* (Diptera: Culicidae)
532 from Colombia. *Zoologia (Curitiba)* 2015, 32:14-22.
- 533 32. Marcombe S, Paris M, Paupy C, Bringuier C, Yebakima A, Chandre F, David J-P, Corbel V,
534 Despres L: Insecticide-Driven Patterns of Genetic Variation in the Dengue Vector *Aedes*
535 *aegypti* in Martinique Island. *PLOS ONE* 2013, 8:e77857.
- 536 33. Marcet PL, Mora MS, Cutrera AP, Jones L, Gürtler RE, Kitron U, Dotson EM: Genetic
537 structure of *Triatoma infestans* populations in rural communities of Santiago Del Estero,
538 northern Argentina. *Infect Genet Evol* 2008, 8:835-846.

34. 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.
35. Perez de Rosas AR, Segura EL, Fusco O, Guinazu AL, Garcia BA: Fine-scale genetic structure in populations of the Chagas' disease vector *Triatoma infestans* (Hemiptera, Reduviidae). *Genetica* 2013.
36. Breniere SF, Salas R, Buitrago R, Bremond P, Sosa V, Bosseno MF, Waleckx E, Depickere S, Barnabe C: Wild populations of *Triatoma infestans* are highly connected to intra-peridomestic conspecific populations in the Bolivian Andes. *PLoS One* 2013, 8:e80786.
37. Noireau F, Flores R, Vargas F: Trapping sylvatic Triatominae (Reduviidae) in hollow trees. *Trans R Soc Trop Med Hyg* 1999, 93:13-14.
38. Picollo M, Wood E, Zerba E, Licastro S, Rúveda M: Laboratory test for measuring toxicity of insecticides in *Triatoma infestans*, Klug. *Acta Bioquim* 1976, 10:67-70.
39. Núñez JA, Segura EL: Rearing of Triatominae. *Chagas' Disease Vectors* 1987, 2:31-40.
40. WHO: (World Health Organization). Protocolo de evaluación de efecto insecticida sobre triatominos. *Acta Toxicológica Argentina* 1994, 2:29-32.
41. 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). *Infection, Genetics and Evolution* 2006, 6:32-37.
42. Weir BS, Cockerham CC: Estimating F-statistics for the analysis of population structure. *Evolution* 1984, 38:1358-1370.
43. 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.
44. Nei M: *Molecular Evolutionary Genetics*. New York: Columbia University Press; 1987.
45. Goudet J: FSTAT version 1.2: a computer program to calculate F-statistics *Journal of Heredity* 1995, 86:485-486.
46. Rice W: Analyzing tables of statistical tests. *Evolution* 1989, 43 223-225.
47. Kalinowski ST: HP-Rare: A Computer Program for Performing Rarefaction on Measures of Allelic Diversity. *Molecular Ecology Notes* 2005, 5:187-189.
48. Agapow P-M, Burt A: Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes* 2001, 1:101-102.
49. Peakall R, Smouse PE: GenA1Ex 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics* 2012, 28:2537-2539.
50. Excoffier L, Smouse PE, Quattro JM: Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 1992, 131:479-491.
51. Excoffier LG, Laval G, Schneider S: ARLEQUIN (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 2005, 1:47-50
52. Pritchard JK, Stephens M, Donnelly P: Inference of population structure using multilocus genotype data. *Genetics* 2000, 155:945-959.
53. Falush D, Stephens M, Pritchard JK: Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* 2007, . doi: 10.1111/j.1471-8286.2007.01758.x.
54. Hubisz MJ, Falush D, Stephens M, Pritchard JK: Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* 2009, 9:1322-1332.
55. 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.

- 588 56. Earl D, vonHoldt B: STRUCTURE HARVESTER: a website and program for visualizing
589 STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*
590 2012, 4:359-361.
- 591 57. Monteiro FA, Barrett TV, Fitzpatrick S, Cordon-Rosales C, Feliciangeli D, Beard CB:
592 Molecular phylogeography of the Amazonian Chagas disease vectors *Rhodnius prolixus* and *R.*
593 *robustus*. *Mol Ecol* 2003, 12:997-1006.
- 594 58. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S: MEGA5: Molecular
595 Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and
596 Maximum Parsimony Methods. *Molecular Biology and Evolution* 2011, 28:2731-2739.
- 597 59. Librado P, Rozas J: DnaSP v5: a software for comprehensive analysis of DNA polymorphism
598 data. *Bioinformatics* 2009, 25:1451-1452.
- 599 60. Bandelt HJ, Forster P, Rohl A: Median-joining networks for inferring intraspecific
600 phylogenies. *Mol Biol Evol* 1999, 16:37-48.
- 601 61. Leigh JW, Bryant D: popart: full-feature software for haplotype network construction. *Methods*
602 *in Ecology and Evolution* 2015, 6:1110-1116.
- 603 62. Quisberth S, Waleckx E, Monje M, Chang B, Noireau F, Brenière SF: "Andean" and "non-
604 Andean" ITS-2 and mtCytB haplotypes of *Triatoma infestans* are observed in the Gran Chaco
605 (Bolivia): Population genetics and the origin of reinfestation. *Infection, Genetics and Evolution*
606 2011, 11:1006-1014.
- 607 63. Fabro J, Sterkel M, Capriotti N, Mougabure-Cueto G, Germano M, Rivera-Pomar R, Ons S:
608 Identification of a point mutation associated with pyrethroid resistance in the para-type sodium
609 channel of *Triatoma infestans*, a vector of Chagas' disease. *Infect Genet Evol* 2012, 12:487-
610 491.
- 611 64. Capriotti N, Mougabure-Cueto G, Rivera-Pomar R, Ons S: L925I Mutation in the *Para*-Type
612 Sodium Channel Is Associated with Pyrethroid Resistance in *Triatoma infestans* from the Gran
613 Chaco Region. *PLoS Negl Trop Dis* 2014, 8:e2659.
- 614 65. Santo-Orihuela PL, Vassena CV, Carvajal G, Clark E, Menacho S, Bozo R, Gilman RH, Bern
615 C, Marcet PL: Toxicological, Enzymatic, and Molecular Assessment of the Insecticide
616 Susceptibility Profile of *Triatoma infestans* (Hemiptera: Reduviidae, Triatominae) Populations
617 From Rural Communities of Santa Cruz, Bolivia. *J Med Entomol* 2017, 54:187-195.
- 618 66. Nei M, Tajima F, Tatenio Y: Accuracy of estimated phylogenetic trees from molecular data. II.
619 Gene frequency data. *J Mol Evol* 1983, 19:153-170.
- 620 67. Giordano R, Cortez JC, Paulk S, Stevens L: Genetic diversity of *Triatoma infestans*
621 (Hemiptera: Reduviidae) in Chuquisaca, Bolivia based on the mitochondrial cytochrome b
622 gene. *Mem Inst Oswaldo Cruz* 2005, 100:753-760.
- 623 68. Germano MD, Vassena CV, Picollo MI: Autosomal inheritance of deltamethrin resistance in
624 field populations of *Triatoma infestans* (Heteroptera: Reduviidae) from Argentina. *Pest Manag*
625 *Sci* 2010, 66:705-708.
- 626 69. PAHO: Technical report: II Reunión técnica latinoamericana de monitoreo de resistencia a
627 insecticidas en triatomíneos vectores de Chagas. Panamerican Health Organization. Panamá, 11-
628 13 April; 2005.
- 629 70. Depickère S, Buitrago R, Siñani E, Baune M, Monje M, Lopez R, Waleckx E, Chavez T,
630 Brenière SF: Susceptibility and resistance to deltamethrin of wild and domestic populations of
631 *Triatoma infestans* (Reduviidae: Triatominae) in Bolivia: new discoveries. *Mem Inst Oswaldo*
632 *Cruz, Rio de Janeiro* 2012, 107:1042-1047.
- 633 71. Pessoa GC, Dias LS, Diotaiuti L: Deltamethrin pyrethroid susceptibility characterization of
634 *Triatoma sordida* Stal, 1859 (Hemiptera: Reduviidae) populations in the Northern Region of
635 Minas Gerais, Brazil. *Rev Soc Bras Med Trop* 2014, 47:426-429.

- 636 72. Piccinali RV, Marcet PL, Noireau F, Kitron U, Gürtler RE, Dotson EM: Molecular population
637 genetics and phylogeography of the Chagas disease vector *Triatoma infestans* in South
638 America. *Journal of Medical Entomology* 2009, 46 796-809.
- 639 73. Sierra I, Capriotti N, Fronza G, Mougabure-Cueto G, Ons S: Kdr mutations in *Triatoma*
640 *infestans* from the Gran Chaco are distributed in two differentiated foci: Implications for
641 pyrethroid resistance management. *Acta Tropica* 2016, 158:208-213.
- 642 74. Botstein D, White RL, Skolnick M, Davis RW: Construction of a genetic linkage map in man
643 using restriction fragment length polymorphisms. *Am J Hum Genet* 1980, 32:314-331.
- 644 75. Monteiro FA, Perez R, Panzera F, Dujardin JP, Galvao C, Rocha D, Noireau F, Schofield C,
645 Beard CB: Mitochondrial DNA variation of *Triatoma infestans* populations and its implication
646 on the specific status of *T. melanosoma*. *Mem Inst Oswaldo Cruz* 1999, 94 Suppl 1:229-238.
647
648
649

Figure legends

Figure 1: Neighbor joining (NJ) tree based on population pairwise D_{AS} distances [66] estimated with the MLGs of 8 loci and 1000 replications.

Figures 2, 3: Bayesian cluster analysis results of *T. infestans* populations from sylvatic and domestic areas in Bolivia.

2a) Mean of estimated Ln likelihood of the data ($\ln(P)$) versus K number of genetic clusters considered [52].

2b) Graphic representation of $\Delta K = \frac{\ln(L(K))}{SD(L(K))}$ [55] to evaluate the rate of change in the $\ln(P)$ versus the number of genetic clusters considered ($K=1-10$).

3) Bar plot representing the individual assignment probability of each MLG to a specific genetic cluster. Each vertical bar represents an individual and the different colors are the probability that each individual was assigned to a particular genetic cluster (K). The hierarchical approach allows for comparison of individual MLG assignment results when considering $K=3$, $K=4$ and $K=5$ genetic clusters.

Figure 4: Left: haplotype network based on a median-joining model, constructed from sequences identified in this study. The size of the nodes is proportional to the abundance of each haplotype. The colors indicate the population of origin. The distance among haplotypes is represented by the number of mutational steps between them. Right: haplotype network including 25 additional Bolivian Andean *T. infestans* haplotypes (labelled A-Y) [15, 62]. Trimmed sequences from this study corresponded to four haplotypes previously reported: A, C, E and R (shown in red).

Tables

Collection site	msat	cytB	kdr	LD50 ¹ (95% CL)	RR ¹ (95% CL)
NFS*	-	-	-	0.13 (0.11-0.15)	-
Mataral-D (Mat-D)	19	19	13	2.25 (0.28-4.80)	17.4 (11.88-25.43)
Mataral-S (Mat-S)	11	11	6	1.53 (0.53-3.34)	11.9 (9.43-14.93)
Kirus Mayu (KM)	22	23	15	0.95 (0.49-1.66)	7.4 (5.78-9.25)
Ilicuni (Ili)	18	18	16	0.25 (0.14-0.39)	1.9 (1.44-2.56)
20 de Octubre (20oct)	22	22	17	0.88 (0.08-1.84)	6.8 (4.98-9.26)

Table 1: Number of bugs per population evaluated per molecular marker: msat= microsatellites; cytB= cytochrome B; **kdr**: evaluated for 2 mutations in the sodium channel gene (L1014 and L9251). Toxicity of topically applied deltamethrin in first instars nymphs of *T. infestans* from Bolivia (¹Data previously published [24]). *Susceptible strain reared at CIPEIN. **LD50**: Lethal dose to kill 50% of the sample measured in ng/insect; **CL**= confidence limit; **RR**: resistance ratio.

Population	#MLG8/N	PA/L $\pm$ SE	av# alleles	Fis $\pm$ SE	avPIC	av A_r	He $\pm$ SE	Ho	%PL Hd	D _{AS} $\pm$ SE	I _A	I _A P-value
Mataral-D	18/19	0.55 $\pm$ 0.09	5	0.052 $\pm$ 0.03	0.57	4.51 $\pm$ 1.91	0.63 $\pm$ 0.07	0.6	50	0.55 $\pm$ 0.12	0.40	0.007
Mataral-S	10/11	1.26 $\pm$ 0.22	5.5	0.22 $\pm$ 0.04**	0.65	5.37 $\pm$ 2.07	0.74 $\pm$ 0.05	0.57	100	0.67 $\pm$ 0.19	1.08	<0.001
Kirus Mayu	20/22	0.45 $\pm$ 0.12	4.3	0.108 $\pm$ 0.03**	0.47	3.69 $\pm$ 1.83	0.52 $\pm$ 0.09	0.47	85.7	0.45 $\pm$ 0.15	0.65	<0.001
Ilicuni	18/18	0.75 $\pm$ 0.11	4.9	0.047 $\pm$ 0.05	0.58	4.34 $\pm$ 1.50	0.64 $\pm$ 0.07	0.61	62.5	0.54 $\pm$ 0.14	0.45	0.012
20 de Octubre	18/22	0.42 $\pm$ 0.13	3.6	-0.047 $\pm$ 0.07**	0.37	3.06 $\pm$ 0.96	0.44 $\pm$ 0.06	0.46	50	0.34 $\pm$ 0.17	1.44	<0.001

Table 2. Population genetic parameters for Bolivian *T. infestans* based on eight microsatellite loci. **#MGL8/N**: number of complete multilocus genotype for 8 loci / number of individuals genotyped per population; **PA/L**: private alleles per locus; **av# alleles**: average number of alleles per locus; **Fis**: inbreeding coefficient [42] **Significant Heterozygote deficit (p<0.05). ; **av PIC**: averaged Polymorphism Information Content [74]. **av A_r** : Averaged allele richness [43]; **He**: expected heterozygosity = average gene diversity [44]; **Ho**: observed heterozygosity; **% PL Hd**: percentage of polymorphic loci displaying a deficit in heterozygosity; **D_{AS}**= Pairwise distance between alleles. **I_A**: index of association that suggests multilocus linkage disequilibrium if p<0.05.

702

	20oct	Ili	KM	Mat-S	Mat-D
20oct	-	**	**	**	**
Ili	0.272 (151.76 km)	-	**	**	**
KM	0.323 (62.92 km)	0.282 (104.63 km)	-	**	**
Mat-S	0.272 (162.50 km)	0.144 (56.02 km)	0.223 (102.06 km)	-	**
Mat-D	0.278 (159.74 km)	0.163 (56.94 km)	0.216 (99.07 km)	0.145 (3.38 km)	-

703 **Table 3:** F_{ST} values (geographic distances) in a pair-wise comparisons between
704 populations (** $p < 0.005$)

705

706

707

haplotypes	variable sites (nucleotide positions)														
	28	162	279	402	450	480	526	549	564	576	588	600	625	646	636
I	G	A	A	A	A	A	C	G	T	A	T	C	G	T	T
XXIV	.	.	.	G	G	.	T	.	.	T	.	.	.	.	.
XVI	A	G	.	.	.	.	.	.	.	T	.	.	.	.	C
HT	A	G	R	.	.	R	.	.	.	Y	.	.	.	Y	C
XLIX	A	G	G	.	.	G	.	.	.	C	.	.	A	C	C
XXVII	.	G	.	.	.	.	.	.	.	T	.	.	.	.	C
XLV	.	G	.	.	.	.	.	.	.	T	.	.	.	.	C
XLVI	.	.	.	.	.	.	.	A	.	T	C	.	.	.	.
XLVII	.	.	.	.	.	.	.	A	C	T	C	.	.	Y	C

708 **Table 4:** Variable sites among the cytochrome B (cytB) haplotypes detected in sylvatic populations of Andean
709 *T. infestans* in Bolivia. Highlighted background indicates non-synonymous substitutions. R and Y are the
710 IUPAC nucleotide code indicating ambiguous base position (A/G) and (C/T) respectively. HT: heteroplasmid
711 sequence (i.e. individuals carrying more than one mitochondrial haplotype). Haplotype I was among the first
712 one described for *T. infestans* in Bolivia [75] and found widely distributed in several populations since then.
713 Green indicates the variable sites shared between haplotypes sequenced in this study and those previously
714 deposited in GenBank for Bolivian Andean *T. infestans* populations.

715

716

717

718

719

Population	XXVII	XXIV	XVI	XLV	XLVI	XLVII	XLIX	ht	# cytB seqs	Hd ± SD	π ± SD
Mat-D	7	9		3					19	0.649 ± 0.061	0.00437 ± 0.00033
Mat-S		1			9	1			11	0.455 ± 0.17	0.00241 ± 0.0011
KM	1		2				10	2	15	0.41 ± 0.154	0.00312 ± 0.0011
Ili		6	8				1		15	0.59 ± 0.077	0.00563 ± 0.00097
20oct			11				3		14	0.363 ± 0.13	0.00272 ± 0.00098

Table 5: Cytochrome B (cytB) haplotype distribution per population and diversity parameters. **Hd**= Haplotype diversity, **π** = nucleotide diversity, SD = standard deviation. **ht** = heteroplasmid