Am. J. Trop. Med. Hyg., 88(5), 2013, pp. 893–896
doi:10.4269/ajtmh.12-0451
Copyright © 2013 by The American Society of Tropical Medicine and Hygiene

Short Report: Variation in Mitochondrial NADH Dehydrogenase Subunit 5 and NADH Dehydrogenase Subunit 4 Genes in the Chagas Disease Vector *Triatoma infestans* (Hemiptera: Reduviidae)

Cintia J. Fernández, Alicia R. Pérez de Rosas, and Beatriz A. García*

Cátedra de Bioquímica y Biología Molecular, Facultad de Ciencias Médicas, INICSA (CONICET-Universidad Nacional de Córdoba), Ciudad Universitaria, Córdoba, Argentina

Abstract. Variation in mitochondrial NADH dehydrogenase subunit 5 (ND5) and NADH dehydrogenase subunit 4 (ND4) genes was surveyed in *Triatoma infestans* from 24 localities of Argentina. The DNA sequence comparisons of 2,183 basepairs of the mitochondrial genome, which include the complete sequence of ND5 (1,712 basepairs) and 401 basepairs of ND4 genes, showed 19 haplotypes determined by 48 variable sites and a nucleotide diversity value of 0.292%. Twenty-six (65%) substitutions were synonymous, and there were 14 (35%) predicted amino acid replacements in ND5. In ND4, 5 (62.5%) substitutions were synonymous and 3 (37.5%) were replacement sites. Samples from six localities studied shared one haplotype and the rest of the localities had different haplotypes. The amplified regions should be useful for population genetic studies.

Chagas disease is caused by infection with *Trypanosoma cruzi*, which is transmitted by hematophagous insects of the subfamily Triatominae (Hemiptera: Reduviidae). *Triatoma infestans* is the main vector of *T. cruzi* in the Southern Cone of Latin American countries between the latitudes 10° S and 46° S. Genetic analyses of populations of this species may provide information for development of control strategies.

Mitochondrial DNA (mtDNA) genes have been used in a number of population genetic analyses and have been recognized as particularly useful for phylogeographic studies in many species of insects.[1] However, the maternally inherited markers analyzed in *T. infestans* either exhibited low levels of variation[2–4] or have not been useful for phylogeographic inferences of the Chagas disease vector in Argentina.[5] The mitochondrial NADH dehydrogenase subunit 5 (ND5) and NADH dehydrogenase subunit 4 (ND4) are two of the most variable protein-coding genes in insects.[6–8] We tested variation of the ND5 and ND4 genes in individual *T. infestans* from 24 localities in Argentina.

The geographic origin of the specimens of *T. infestans* analyzed is shown in Figure 1 and Table 1. The specimen obtained in each locality was considered to belong to one population.

The primer pairs (Table 2) used for polymerase chain reaction (PCR) and direct sequencing were designed from conserved fragments of aligned sequences of several insects available in GeneBank. Sequences were aligned using CLUSTAL implemented in MEGA 4.[9] The PCR amplifications were conducted with genomic DNA by using a Professional Basic thermal cycler (Biometra, Goettingen, Germany). Reaction products were purified by using a QIAquick PCR Purification Kit (QIAGEN, Hilden, Germany) for direct sequencing in an ABI Prism 3130 automated DNA sequencer (Applied Biosystems, Foster City, CA).

DNA sequences were aligned using CLUSTAL in MEGA 4. Nucleotide diversity was estimated according to the estimator (θ_W) of Watterson,[10] average nucleotide diversity (π),[11] and average number of nucleotide differences (K)[12] by using DnaSp version 5.1.[13] Absolute distances and maximum composite likelihood distances between haplotypes were obtained by using MEGA 4. A haplotype network was generated on the basis of the parsimony method using the TCS program.[14]

A total of 2,365 basepairs (GeneBank accession no. KC182516) of mtDNA were amplified and sequenced from *T. infestans*. The region studied includes 55 basepairs of the glutamic acid tRNA gene (tRNA-Glu), 69 basepairs of the phenylalanine tRNA gene (tRNA-Phe), the complete ND5 gene (1712 basepairs), 65 basepairs of the histidine tRNA gene (tRNA-His), and the first 461 basepairs of the ND4 gene.

Fragments of the mtDNA region characterized above were analyzed in 24 *T. infestans*, each collected at a different site in Argentina. The length of the mtDNA fragments ranged from 2,194 to 2,365 basepairs (Table 1). The DNA sequence comparisons of 2,183 basepairs of this region showed 19 haplotypes involving 48 variable sites (Table 3). The observed total nucleotide variability was of 0.00589 and 0.292% according to θ_W and π, respectively. The average number of nucleotide differences (K) was 6.370.

The percentages of sequence variation over ND5 and ND4 gene regions are shown in Table 4. The variable nucleotide positions included 40 transitions and 8 transversions and were classified by codon positions in both genes. As expected, most variation occurred in third codon positions in ND5 and ND4 genes (65% and 62.5%, respectively), and the second codon positions showed the least variation (only 5 of the 571 second codon positions varied in ND5 and 1 of the 134 second codon positions varied in ND4); 26 (65%) of the substitutions were synonymous and 14 (35%) predicted amino acid replacements in ND5. In ND4, 5 (62.5%) were synonymous and 3 (37.5%) were replacement sites (Table 3).

Insects from Caucete, Caucete, San Juan; Santa Rosa, Valle Viejo, Catamarca; Saujil, Tinogasta, Catamarca; Saujil, Pomán, Catamarca; Vaca Huañuna, Figueroa, Santiago del Estero; and El Zapallo, General Paz, Corrientes shared haplotype H and the rest of the localities showed different haplotypes. These different haplotypes, found in the only specimen analyzed in the remaining 18 localities, suggest that this portion of the mitochondrial genome might provide a valuable tool for genetic analysis of *T. infestans* populations.

* Address correspondence to Beatriz A. García, Cátedra de Bioquímica y Biología Molecular, Facultad de Ciencias Médicas, Universidad Nacional de Córdoba, Pabellón Argentina 2do Piso, Ciudad Universitaria, 5000 Córdoba, Argentina. E-mail: bgarcia@biomed.uncor.edu

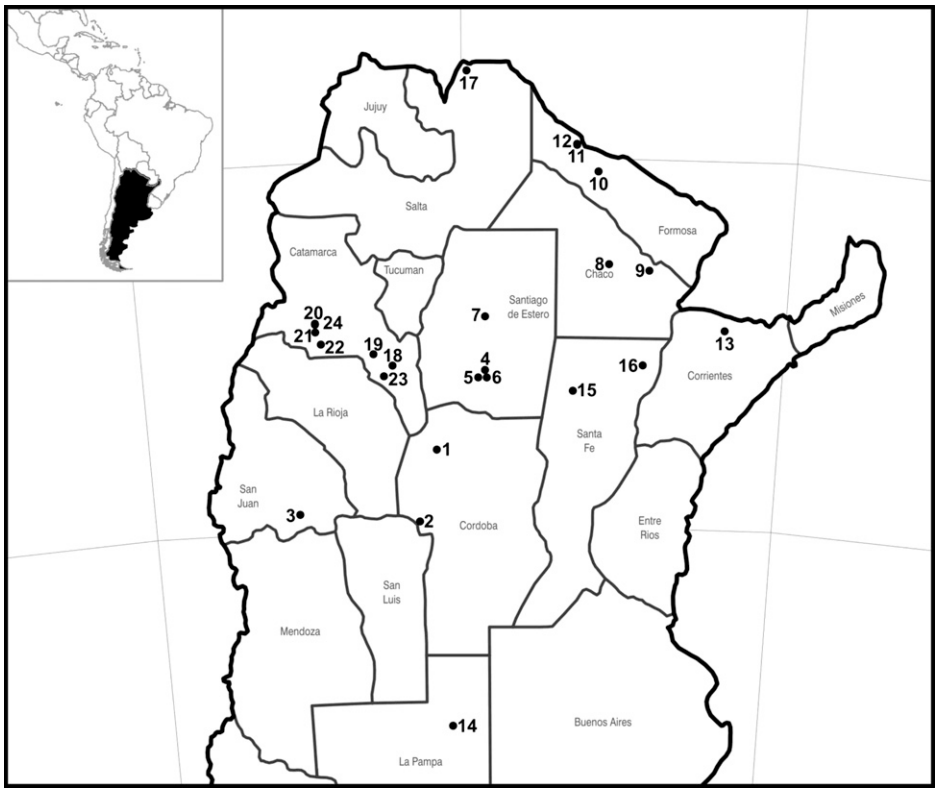

FIGURE 1. Geographic origin of *Triatoma infestans* specimens analyzed. For more details, see Table 1.

Sequences were closely related; the pair of haplotypes did not differ by more than 14 nucleotides. Pairwise divergence among *T. infestans* haplotypes ranged from 0.0005 (1 of 2,183) between haplotypes G-H, Q-H, and O-N to 0.0066 (14 of 2,183) between haplotypes A-J and B-J. Thirty nine of the 48 variable positions observed across all the haplotypes detected differed by a single substitution in one haplotype (Table 3). These sites are parsimony uninformative sites and the remaining nine nucleotide positions are informative. In this respect, it is important to point out that eight of these informative sites are from 463 to 1578 nucleotide positions of the ND5 gene.

TABLE 1

Sampling site and its code for *Triatoma infestans* from Argentina, including length of the mitochondrial DNA region analyzed, haplotypes, and corresponding GeneBank accession numbers*

Sampling site	Locality, county, province	Code	No.	Latitude	Longitude	Length (bp)	Haplotype†	GenBank accession no.
1	Chuña, Ischilin, Córdoba	CCI	1	30°28′S	64°40′W	2,343	A	KC196504
2	Sauce Arriba, San Alberto, Córdoba	CSA	1	31°54′S	65°10′W	2,362	B	KC196505
3	Caucete, Caucete, San Juan	SJC	1	31°40′S	68°16′W	2,358	H	KC196511
4	Sabagasta, Salavina, Santiago del Estero	SES	1	28°37′S	63°29′W	2,350	G	KC196510
5	Vaca Human, Salavina, Santiago del Estero	SEV	1	28°47′S	63°37′W	2,350	E	KC196508
6	Taco Totorayo, Salavina, Santiago del Estero	SET	1	28°49′S	63°26′W	2,356	F	KC196509
7	Vaca Huañuna, Figueroa, Santiago del Estero	SEH	1	27°27′S	63°27′W	2,352	H	KC196511
8	Tres Isletas, Maipú, Chaco	CHT	1	26°21′S	60°26′W	2,349	Q	KC196520
9	Siete Árboles, Gral. San Martín, Chaco	CHS	1	26°24′S	59°25′W	2,194	P	KC196519
10	Palo Santo, Bermejo, Formosa	FPS	1	24°19′S	60°55′W	2,262	M	KC196516
11	Pozo del Zuri, Bermejo, Formosa	FPZ	1	23°41′S	61°09′W	2,263	N	KC196517
12	La Esperanza, Bermejo, Formosa	FES	1	23°40′S	61°09′W	2,302	O	KC196518
13	El Zapallo, General Paz, Corrientes	COZ	1	27°41′S	57°37′W	2,349	H	KC196511
14	Santa Rosa, Capital, La Pampa	LPS	1	36°35′S	64°20′W	2,363	C	KC196506
15	El Nochero, 9 de Julio, Santa Fé	SFN	1	29°01′S	61°15′W	2,274	S	KC196522
16	Moraju, General Obligado, Santa Fé	SFM	1	28°30′S	59°30′W	2,352	R	KC196521
17	Salvador Mazza, San Martín, Salta	SSM	1	22°03′S	63°42′W	2,350	D	KC196507
18	Santa Rosa, Valle Viejo, Catamarca	CSR	1	28°28′S	65°47′W	2,364	H	KC196511
19	Saujil, Pomán, Catamarca	CSP	1	28°12′S	66°14′W	2,358	H	KC196511
20	Saujil, Tinogasta, Catamarca	CST	1	27°34′S	67°37′W	2,335	H	KC196511
21	Fiambalá, Tinogasta, Catamarca	CFI	1	27°41′S	67°37′W	2,365	L	KC196515
22	Copacabana, Tinogasta, Catamarca	CCO	1	28°12′S	67°29′W	2,351	K	KC196514
23	Huillapima, Capayán, Catamarca	CHU	1	28°44′S	65°59′W	2,345	J	KC196513
24	Medanitos, Tinogasta, Catamarca	CME	1	27°31′S	67°36′W	2,362	I	KC196512

*bp = basepairs.
†Haplotypes detected from DNA sequence comparisons of 2,183 bp of the mitochondrial DNA region analyzed.

TABLE 2

Primers used for amplification and direct sequencing of NADH dehydrogenase subunit 5 (ND5) and NADH dehydrogenase subunit 4 (ND4) genes in *Triatoma infestans*, Argentina

Primer	Sequence (5′→3′)	Fragment size (basepairs)
Ser –F	GCTAACTATCTTTTAAAGCGG	1,294
ND5-R1	GCTCARATTCCKTTTTCTTC	1,294
ND5-F1	GCAGTAACYAAAGTAGAAGAATG	1,269
ND4-R1	ATTGCTTATTCTTCDGTTGCTC	1,269
ND5-F2	AAAGGGCAAACCACACAAAG	822
ND5-R2	ATGGGTTCTGATGTTGATCG	822

TABLE 3

Haplotypes found in NADH dehydrogenase subunit 5 (ND5) and NADH dehydrogenase subunit 4 (ND4) mitochondrial DNA genes of *Triatoma infestans**

Haplotypes	ND5																																								ND4							
	24	54	114	122	150	258	312	342	399	452	463	561	575	604	651	689	758	759	761	802	864	873	875	885	907	912	942	1122	1132	1197	1203	1215	1260	1383	1466	1518	1521	1578	1682	1695	18	56	88	168	223	316	331	340
A	C	G	G	C	C	C	T	C	T	T	T	A	G	G	G	A	A	A	G	A	C	T	A	G	T	G	C	A	G	A	A	T	G	G	C	A	C	C	C	C	T	T	A	T	T	G	T	A
B	.	.	.	T	T	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	.	.	.	.	.	A	.	.	.	T	.	.	.	.	G	C	.	.	.	G
C	.	.	.	.	T	T	.	.	C	.	.	.	.	A	A	.	.	.	.	.	.	.	.	A	C	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	G
D	.	.	.	.	T	.	.	.	.	.	C	.	A	.	A	.	.	.	.	.	.	.	.	A	C	.	T	.	.	.	.	.	.	.	.	G	.	T	T	.	.	.	.	.	.	.	.	G
E	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	G
F	.	.	.	.	T	.	.	.	.	.	C	.	.	.	A	.	.	.	.	.	.	.	G	A	C	.	T	.	.	.	.	.	.	.	.	G	T	T	.	.	.	.	.	.	.	A	.	G
G	.	.	.	.	T	.	.	.	.	.	C	.	.	.	A	.	.	.	.	.	.	C	G	A	C	.	T	.	.	.	.	.	.	.	.	G	.	T	.	.	.	.	.	.	.	.	.	G
H	.	.	.	.	T	.	.	.	.	.	C	.	.	.	A	.	.	.	.	.	.	.	G	A	C	.	T	.	.	.	.	.	.	.	.	G	.	T	.	.	.	.	.	.	.	.	.	G
I	.	.	.	.	T	.	.	.	.	.	C	.	.	.	A	.	.	.	.	.	.	.	.	A	C	.	.	.	C	.	.	.	.	.	.	.	.	T	.	T	.	.	.	.	.	.	.	G
J	.	.	.	.	T	.	.	.	.	.	C	.	.	.	A	.	C	C	T	C	.	.	G	A	C	.	T	.	.	.	.	.	.	.	.	G	.	T	.	.	.	.	.	.	.	.	.	G
K	.	.	.	.	T	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	A	C	.	.	.	.	.	G	.	A	.	.	.	.	T	.	.	C	.	.	.	.	.	.	G
L	.	.	A	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	.	.	.	.	.	.	T	.	.	T	.	.	.	.	.	.	.	.	.	G
M	.	.	.	.	T	.	.	T	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	C	G
N	.	.	.	.	T	.	.	.	.	.	C	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	G	.	T	.	.	.	.	.	.	.	.	.	G
O	.	.	.	.	T	.	.	.	.	.	C	.	.	.	A	.	.	.	.	.	.	.	.	.	C	.	T	.	.	.	.	.	.	.	.	G	.	T	.	.	.	.	.	.	.	.	.	G
P	A	.	.	.	T	.	.	.	.	.	C	.	.	.	.	C	.	.	.	.	T	.	G	A	C	.	T	.	.	.	.	.	.	.	.	G	.	T	.	.	.	.	.	.	.	.	.	G
Q	.	.	.	.	T	.	.	.	.	C	C	.	.	.	A	.	.	.	.	.	.	.	G	A	C	.	T	.	.	.	.	.	.	.	.	G	.	T	.	.	.	.	.	.	.	.	.	G
R	.	.	.	.	T	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	A	C	A	.	G	.	G	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	G
S	.	A	.	.	T	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	A	C	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	A	.	.	.	.	C	G

*Variable nucleotide positions are indicated in sequential order in both genes. The sequence of the first haplotype is used as a reference. Nucleotides identical to the reference are indicated by a dot. Replacement sites for ND5 are: 24, L→F; 122, V→I; 452, K→E; 463, S→K; 604, T→M; 689, C→G; 758, C→G; 759 and 761, H→K; 802, F→C; 907, S→K; 1132, T→S; 1466, G→S; 1682, G→S. Replacement sites for ND4 are: 18, S→G; 56, H→L; 168, I→V.

TABLE 4

Percentages of variable sites in NADH dehydrogenase subunit 5 (ND5) and NADH dehydrogenase subunit 4 (ND4) genes across *Triatoma infestans* haplotypes*

Characteristic	ND4				ND5			
	1st codon position	2nd codon position	3rd codon position	Total	1st codon position	2nd codon position	3rd codon position	Total
Unvaried	98.50 (131)	99.25 (133)	96.27 (129)	98.00 (393)	98.42 (562)	99.13 (566)	95.44 (544)	97.66 (1,672)
Transitions	1.50 (2)	0.00 (0)	3.73 (5)	1.75 (7)	1.05 (6)	0.52 (3)	4.21 (24)	1.93 (33)
Transversions	0.00 (0)	0.75 (1)	0.00 (0)	0.25 (1)	0.53 (3)	0.35 (2)	0.35 (2)	0.41 (7)
Total (bp)	133	134	134	401	571	571	570	1712

*Values in parentheses are number of sites. bp = basepairs.

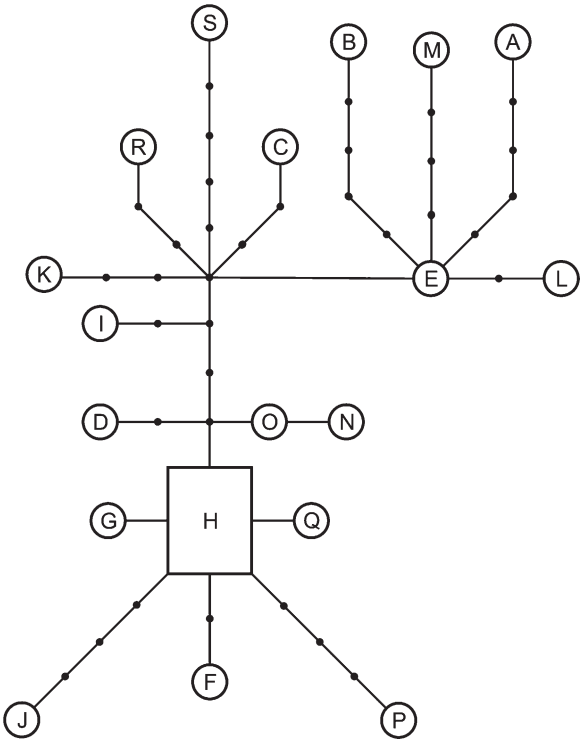

FIGURE 2. Haplotype network for mitochondrial NADH dehydrogenase subunit 5 and NADH dehydrogenase subunit 4 genes of *Triatoma infestans*. The maximum number of steps connecting parsimoniously two haplotypes is indicated. One step is indicated by lines between two haplotypes, and each additional base substitution is indicated by a solid circle. The haplotype with the highest ancestral probability (H) is displayed as a square, and the other haplotypes are displayed as circles.

The haplotype network (Figure 2) obtained with TCS software indicates that the different sequence types observed would have derived from a common ancestral haplotype (H). In contrast, the network shows that the haplotypes A and B (detected in Córdoba Province), as well as M and L (detected in Formosa and Catamarca Provinces, respectively), would have derived from haplotype E (found in Santiago del Estero Province). This haplotype is separated from H by 4 mutational events. In addition, haplotype N would have derived from haplotype O (detected in the geographically closest localities from Formosa [Pozo del Zuri, Bermejo, Formosa and La Esperanza, Bermejo, Formosa], respectively), both separated by a single mutational step.

This is the first ND5 and ND4 gene sequences analysis in the Chagas disease vector *T. infestans*. The amplified regions, fundamentally a portion of the ND5 gene, should be useful for phylogeographic studies. Because one individual was analyzed per locality, it is important to take into account that more sequences need to be examined.

Received July 26, 2012. Accepted for publication January 12, 2013.

Published online February 11, 2013.

Financial support: This study was supported by the grants from the Agencia Nacional de Promoción Científica y Tecnológica (FONCyT), the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) of Argentina, and the Secretaría de Ciencia y Tecnología de la Universidad Nacional de Córdoba. Cintia J. Fernández is a Fellow of FONCyT and Beatriz A. García is Career Investigator of CONICET.

Authors' addresses: Cintia J. Fernández, Alicia R. Pérez de Rosas, and Beatriz A. García, Cátedra de Bioquímica y Biología Molecular, Facultad de Ciencias Médicas, Universidad Nacional de Córdoba, Pabellón Argentina 2° Piso, Ciudad Universitaria, 5000 Córdoba, Argentina, E-mails: cjudithfernandez@gmail.com, arperez@biomed.fcm.edu.com.ar, and bgarcia@biomed.uncor.edu.

REFERENCES

1. Roderick GK, 1996. Geographic structure of insect populations: gene flow, phylogeography, and their uses. *Annu Rev Entomol 41:* 325–352.
2. Monteiro FA, Pérez R, Panzera F, Dujardin JP, Galvão C, Rocha D, Noireau F, Schofield C, Beard CB, 1999. Mitochondrial DNA variation of *Triatoma infestans* populations and its implication on the specific status of *T. melanosoma*. *Mem Inst Oswaldo Cruz 94:* 229–238.
3. García BA, Manfredi C, Fichera L, Segura EL, 2003. Variation in mitochondrial 12S and 16S ribosomal DNA sequences in natural populations of *Triatoma infestans* (Hemiptera: Reduviidae). *Am J Trop Med Hyg 68:* 692–694.
4. Segura EL, Torres AG, Fusco O, García BA, 2009. Mitochondrial 16S DNA variation in populations of *Triatoma infestans* from Argentina. *Med Vet Entomol 23:* 34–40.
5. Piccinali RV, Marcet PL, Noireau F, Kitron U, Gurtler RE, Dotson EM, 2009. Molecular population genetics and phylogeography of the Chagas disease vector *Triatoma infestans* in South America. *J Med Entomol 46:* 796–809.
6. Michel AP, Grushko O, Guelbeogo WM, Lobo NF, Fale Sagnon N, Costantini C, Besansky NJ, 2006. Divergence with gene flow in *Anopheles fenustus* from the Sudan Savanna of Burkina Faso, West Africa. *Genetics 173:* 1389–1395.
7. Meraner A, Brandstätter A, Thaler R, Aray B, Unterlechner M, Niederstätter H, Parson W, Zelger R, Dalla Via J, Dallinger R, 2008. Molecular phylogeny and population structure of the codling moth (*Cydia pomonella*) in central Europe: I. Ancient clade splitting revealed by mitochondrial haplotype markers. *Mol Phylogenet Evol 48:* 825–837.
8. Grisales N, Triana O, Angulo V, Jaramillo N, Parra-Henao G, Panzera F, Gómez-Palacio A, 2010. Diferenciación genética de tres poblaciones Colombianas de *Triatoma dimidiata* (Latreille, 1811) mediante análisis molecular del gen mitocondrial ND4. *Biomedica 30:* 207–214.
9. Tamura K, Dudley J, Nei M, Kumar S, 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol 24:* 1596–1599.
10. Watterson GA, 1975. On the number of segregating sites in genetical models without recombination. *Theor Popul Biol 7:* 256–276.
11. Nei M, 1987. *Molecular Evolutionary Genetics*. New York: Columbia University Press.
12. Tajima F, 1983. Evolutionary relationship of DNA sequences in finite populations. *Genetics 105:* 437–460.
13. Librado P, Rozas J, 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics 25:* 1451–1452.
14. Clement M, Posada D, Crandrall KA, 2000. TCS: a computer program to estimate gene genealogies. *Mol Ecol 9:* 1657–1659.