Gene 543 (2014) 118–124

Contents lists available at ScienceDirect

Gene

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

cDNA isolation and characterization of two vitellogenin genes in the Chagas' disease vector *Triatoma infestans* (Hemiptera, Reduviidae)

María J. Blariza [a], Néstor W. Soria [b], Adrián G. Torres [a], Carla G. Grosso [a], Beatriz A. García [a,*]

[a] INICSA, CONICET and Cátedra de Bioquímica y Biología Molecular, Facultad de Ciencias Médicas, Universidad Nacional de Córdoba, Córdoba, Argentina
[b] Laboratorio de Análisis Clínicos Especializados (LACE), Córdoba, Argentina

ARTICLE INFO

Article history:
Received 10 December 2013
Received in revised form 7 March 2014
Accepted 27 March 2014
Available online 28 March 2014

Keywords:
Triatoma infestans
Chagas' disease vector
Vitellogenin genes

ABSTRACT

Two vitellogenin genes (*Vg1* and *Vg2*) were identified in the Chagas' disease vector *Triatoma infestans*. The putative coding sequence corresponding to *Vg2* was found to be 5553 bp long, encoding 1851 amino acids in a single open reading frame. The comparative analysis of the deduced amino acid sequences from *Vg1* and *Vg2* cDNA fragments of *T. infestans* revealed 58.94% of identity with 76.43% of homology. The phylogenetic tree based on the complete Vg amino acid sequences of hemimetabolous insects unambiguously supported two clusters, one consisting of Vg sequences from dictyopteran and the other containing Vg sequences of hemipteran. The *Vg1* and *Vg2* mRNAs were detected in fat bodies and ovaries of adult females with the highest levels of both *Vg* transcripts in the first tissue. Quantitative PCR showed low expression of *Vg2* in head and muscle of adult females, while the *Vg1* transcript was not present in these organs. Neither *Vg1* nor *Vg2* was expressed in fifth instar nymph fat bodies or in adult male fat bodies, heads, and muscles.

1. Introduction

Chagas' disease (American trypanosomiasis) is the fourth cause of economic losses through illness in Latin America, with about 8 million persons infected and around 109 million living in endemic areas (Rassi et al., 2010). The disease is produced by infection with *Trypanosoma cruzi*, parasite transmitted by hematophagous insects of the subfamily Triatominae (Hemiptera, Reduviidae). The members of this subfamily are hemimetabolous insects with incomplete metamorphosis; the adults of both sexes and the five nymphal instars are bloodsuckers and at least one full blood meal is necessary for each molt. The subfamily Triatominae includes more than 140 species, most of which are actual or potential vectors of *T. cruzi*. Considering the different habitats described for the members of Triatominae, from exclusively sylvatic species to those well adapted to human dwellings, only a few, those with a high degree of adaptation to the domestic environment, have been recognized as effective vectors of trypanosomiasis in humans. Among them, *Triatoma infestans* is the most important and widespread vector of Chagas' disease in South America, where it has been the target of control programs as part of the Southern Cone Initiative. The interruption of transmission of this parasitic and infectious disease consists of vector control by insecticide treatment of infested dwellings. However, though Uruguay, Brazil, and Chile have been declared free of Chagas' disease transmission by *T. infestans* (Lorca et al., 2001; Silveira and Vinhaes, 1999; World Health Organization, 1994), high levels of *T. infestans* reinfestation after spraying were observed in Argentina, Bolivia, and Paraguay (Gürtler et al., 2007). Moreover, vector control failure caused by resistance to pyrethroid insecticides has been detected (Germano et al., 2010; Picollo et al., 2005; Santo Orihuela et al., 2008; Toloza et al., 2008; Vassena et al., 2000). Analyses of fundamental genes in these insect vectors as those related to their reproduction are of considerable importance and should be undertaken.

The majority of blood-feeding insects, including triatomines, ingest a large amount of blood in a single meal. The digested blood contains a large amount of amino acids and lipids that can be used to produce molecules such as vitellogenins (Vgs). Vg is the precursor protein of egg yolk vitellin (Vn) that provides energy reserves in oviparous vertebrates and invertebrates. In most insects, Vgs are synthesized exclusively in the fat body, while in others, the ovary is a complementary vitellogenic organ (Belles, 1998, 2005; Giorgi et al., 2005; Melo et al., 2000). In female insects, the main function of the fat body is to produce massive amounts of yolk protein precursors, which are secreted into the hemolymph and then sequestered by competent oocytes via receptor-mediated endocytosis (Snigirevskaya and Raikhel, 2005). During these processes, Vgs and Vns are modified by glycosylation, lipidation, phosphorylation and proteolytic cleavage (Dhadialla and Raikhel, 1990). Following uptake by oocytes, Vgs are stored in cytoplasmic granules as Vns providing raw materials for development of the embryos (Kunkel and

Abbreviations: bp, base pair(s); cDNA, DNA complementary to RNA; RNA, ribonucleic acid; mRNA, messenger RNA; DNase, deoxyribonuclease; dNTP, deoxyribonucleoside triphosphate; kb, kilobase(s) or 1000 bp; kDa, kilodalton(s); Oligo, oligodeoxyribonucleotide; *Vg*, vitellogenin gene; *Vgs*, vitellogenin genes; Vg, vitellogenin protein; Vgs, vitellogenin proteins; Vn, vitellin; Vns, vitellins.

* Corresponding author at: 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 address: bgarcia@biomed.uncor.edu (B.A. García).

http://dx.doi.org/10.1016/j.gene.2014.03.054
0378-1119/

Nordin, 1985). Synthesis of Vg occurs in smaller amounts in males of some insect species (Engelmann, 1979; Piulachs et al., 2003; Trenczek and Engels, 1986; Valle, 1993).

Vg cDNAs have been isolated from several insect species (Tufail and Takeda, 2008). Insect Vgs are phospholipoglycoproteins encoded by mRNAs of 6–7 kb that are translated as primary products of ~200 kDa. These primary products have been characterized at a molecular level for Hemimetabola and Holometabola species (Sappington et al., 2002; Tufail et al., 2004). The amino acid sequence, structure, and composition of Vgs are sufficiently conserved between insects and other oviparous animals indicating an origin from a common ancestral protein (Sappington et al., 2002). The insect *Vgs* constitute a multigene family, with two or three members reported from a number of species (Hughes, 2010).

Here we report for the first time the nucleotide sequence and deduced amino acid sequence for a *Vg* in a triatomine species. We have identified two *Vg* genes (*Vg1* and *Vg2*) and have explored their expression patterns at the mRNA level in both sexes, different tissues, and two developmental stages.

2. Material and methods

2.1. Insects and tissue collection

T. infestans was reared at 28 ± 1 °C at a relative humidity of 60–70% with a 6-hour light/18-hour dark cycle and fed once every two weeks after molt on restrained chickens. For the experiments, fifth instar male and female nymphs were sexed by the differences described by Espinola (1966) and separated before feeding. Females and males were maintained segregated after emergence until they were able to have a blood meal (day 7 post-ecdysis). After feeding, females were placed together with males in individual containers (each couple in one container). Mated females were checked by observation of the spermatophore. For the *Vg* cDNA sequence identification, adult female fat bodies were extracted after 3–6 days of the blood feeding. Each sample was a pool of tissue from five adult female specimens. The expression analysis of the *Vg* genes was carried out using adult specimens and fifth instar nymphs after 4–5 days of feeding. Thoracic muscles, heads and fat bodies were collected from adult females and males. Additionally, ovaries from adult females and fat bodies from female nymphs of the fifth instar were also obtained. Each sample was a pool of tissue from three adult specimens and five fifth instar nymphs. The tissues were dissected under aseptic conditions and stored in liquid nitrogen until used for RNA extraction.

2.2. Isolation of total RNA and cDNA synthesis

Total RNA was extracted from pools of insect tissues using TRizol reagent according to the manufacturer's specifications (Invitrogen, Carlsbad, CA) and eluted in 20 μL of nuclease-free water. For quantitative PCR (qPCR) total RNA was isolated from pools of insect tissues using MasterPure RNA Purification Kit (Epicentre, Madison, WI) following the manufacturer's protocol and eluted in 33 μL of nuclease-free water. This kit includes a DNase treatment in order to eliminate potential genomic DNA contamination. The RNA concentration was determined by absorption at 260 nm.

First-strand cDNA synthesis was performed with 1 μL of Oligo-dT_{20} (50 μM) (Invitrogen), 1 μg of total RNA, and 400 U of SuperScript III RT (reverse transcriptase, Invitrogen) in a 20 μL reaction volume incubated at 55 °C for 1 h.

2.3. Amplification and sequencing of vitellogenin cDNA

Rapid amplification of 3′ cDNA end (3′-RACE) was performed using the GeneRacer commercial kit (Invitrogen). 1 μg of total RNA was reverse transcribed with GeneRacer Oligo-dT and Superscript III RT. A forward primer designed by Lee et al. (2000), based on the conserved amino acid sequence GL/ICG present near the C-terminus, and a GeneRacer 3′ primer (homologous to GeneRacer Oligo-dT Primer) (Table 1) were used to amplify the first-strand cDNA and to obtain the 3′ cDNA encoding for Vg. Only mRNA with a polyA tail was reverse transcribed and amplified using the polymerase chain reaction (PCR). After the PCR product was electrophoresed, a band corresponding to the expected size of approximately 500 bp was excised from the agarose gel and purified using QIAquick PCR Purification Kit (QIAGEN, Hilden, Germany). The PCR product was then sequenced in an ABI 3130XL automated DNA sequencer (Applied Biosystems, Foster City, CA). The comparative analysis of the cDNA fragment of *T. infestans* with cDNA sequences of the *Vg* gene of other insect species revealed that it was the 3′ end of this gene. Following initial identification of the *T. infestans Vg* gene, a conventional PCR was performed using a specific primer designed from the 5′ end of the 3′ cDNA partial sequence of *T. infestans Vg* gene and a degenerate or unspecific primer (Table 1) to amplify a longer fragment of *Vg* cDNA. To obtain sequences closest to

Table 1
Sequences of all PCR primers used in this study.

Name	Sequences	Function
PCR degenerate primers (for conserved cDNA isolation)		
PF1D	5′-GTNTTYGARCCNTAYYTNGARGGN-3′	Sense
RipF 2318	5′-CTTCAAMRNATGGCAGA-3′	Sense
C1	5′-CCBRTDGCCADDGGGAATCC-3′	Antisense
Unspecific primers (for RT-PCR)		
RhoF4766	5′-CATCAAGCCACAACTTTC-3′	Sense
RhoF4820	5′-GAAGGTGCTTCATATCAAAAGG-3′	Sense
FA5	5′-ACTCTCCATCGAATGTGCCGC-3′	Sense
FA12	5′-ACTTGGAACTGCAGGTAACTTGG-3′	Sense
RA1	5′-AGAGCACAGCAGTGCCTGTCCGGG-3′	Antisense
RA5	5′-AGACTGAATATGGTTATTCAGGCC-3′	Antisense
RA9	5′-CAACCGGATCTTCAGAAGTAGTTGG-3′	Antisense
RA12	5′-CCCAAGTTACCTGCAGTTCCA-3′	Antisense
Specific primers (for RT-PCR)		
PR1	5′-TCGGGATTCGCTCCTTTTCTTGT-3′	Antisense
PF2	5′-TGATGTAGAAGCCGGGCGTAACA-3′	Sense
PF3	5′-AAGTTTTAGATTTCTCTAAGGAATCAA-3′	Sense
PF2	5′-GTACAAATTTATGCCCTTCCCAACGG-3′	Sense
PR3	5′-CCGTTGGGAAGGGCATA-3′	Antisense
PF4	5′-AAGCCACAACTTTCGACAACAG-3′	Sense
PF5	5′-CCAGTAAATGTGGGTAAATGCTGG-3′	Sense
PR4	5′-GCATGCCAGCATTTACCCA-3′	Antisense
PF6iso	5′-CTGCCCGGGCTTTTACT-3′	Sense
PFA	5′-ATTTGTCGTTCAGATCAGCCCC-3′	Sense
PRA	5′-TAAACCAAGCGTTTGCCGGT-3′	Antisense
PFB3′	5′-TAAGCGAAGAAGAAGCTGGACG-3′	Sense
TIR7	5′-TATATCTGCCCATGAGATGGC-3′	Antisense
TIR9	5′-CCGTGAATGCAGTGAGAACGC-3′	Antisense
TIF8	5′-CAGAGTCTACCCCAGACTTGCC-3′	Sense
TIF9	5′-CAGAGAAACTGCACACAACCG-3′	Sense
Cloning		
M13F	5′-GTAAAACGACGGCCAG-3′	Sense
M13R	5′-CAGGAAACAGCTATGAC-3′	Antisense
RACE primers (for full cDNA isolation)		
I-6	5′-GGTATTTGCGG-3′ (Lee et al., 2000)	3′-RACE of Vg-1
PF5	5′-CCAGTAAATGTGGGTAAATGCTGG-3′	3′-RACE of Vg-2
Race5′	5′-TAGACTTAGAAATTAATACGACTCACTA TAGGCGCGCCACCG-3′	5′-RACE of Vg-2
GeneRacer3′	5′-GCTGTCAACGATACGCTACGTAACG-3′	3′-RACE of Vg-1
GeneRacer3′	5′-GCTGTCAACGATACGCTACGTAACG-3′	3′-RACE of Vg-2
TIR8	5′-CGTGATGGTTACTAGAGCAGGTCC-3′	5′-RACE of Vg-2
qPCR primers (for real time PCR)		
QAF5	5′-TATAATCAAGGATCCCTTAGAATTTGC-3′	*Vg-1, forward*
QAF2	5′-CGCATGAGAAGCTCTCTAACCA-3′	*Vg-2, forward*
QβactinaF	5′-CCCCTTTCAGTGAGGATCTTCA-3′	Internal control
QAR5	5′-TTGGCTGGTCCTTCACAAGTT-3′	*Vg-1, reverse*
QAR2	5′-ACGCGATGATTAGTGCATCCT-3′	*Vg-2, reverse*
QβactinaR	5′-CGCCATCCTTCGATTGGA-3′	Internal control

the 5′ end of the gene, subsequent PCRs were carried out using unspecific or degenerate primers and specific primers designed from the new sequences obtained (Table 1). All PCR amplifications were carried out in a Thermocycler (MyCycler; Bio-Rad, Hercules, CA) in 15 μL of a solution containing 1 μL of first-strand cDNA (template), 1–4 μM of each degenerated primer, 0.7 μM of each specific primer, 10 mM Tris–HCl (pH 9), 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, and 0.6 U of Amplitaq DNA polymerase (Amersham Biosciences, Piscataway, NJ). The degenerate and unspecific primers were designed from the conserved regions of sequence of the Vg from other hemipteran insects: *Riptortus clavatus* (AAB72001), *Graptopsaltria nigrofuscata* (BAA85987), and *Plautia stali* (Vg1: BAA88075, Vg2: BAA88076, Vg3: BAA88077). Thermal profiles consisted of an initial denaturation step at 94 °C for 5 min, followed by 40 cycles at 94 °C for 60 s, 50 °C for 1 min, and 72 °C for 60 s, with a final extension step of 10 min at 72 °C. Reaction products were visualized after electrophoresis on 1.5% agarose gels (Tris-acetate EDTA buffer, pH 8) containing 0.5 μg/mL of ethidium bromide. The bands corresponding to the expected size were purified and sequenced as explained above. Some of the resulting PCR products were cloned into the pCR4-TOPO TA cloning vector (Invitrogen) before being sequenced. The comparative analysis of the cDNA fragments obtained with cDNA sequences of other insects' *Vg* gene showed that they were part of this gene.

In order to complete the 5′ end of the gene, the 5′-RACE procedure was conducted using the commercial kit ExactSTART Eukaryotic mRNA 5′–&3′-RACE (Epicentre) according to the manufacturer's instruction. 5 μg of total RNA was treated with heat-labile alkaline phosphatase which converts RNA with a 5′ monophosphate or a 5′ triphosphate to RNA with 5′-hydroxyl ends to ensure that they will not participate in RACE amplification process; 5′-capped RNAs are not affected by the phosphatase treatment. The phosphatase-treated sample was treated with tobacco acid pyrophosphatase to remove the 5′ cap structure from intact full-length mRNA. Using T4 RNA ligase, the 5′ RACE RNA oligonucleotide was ligated to the 5′ region of the RNAs. The RNA oligonucleotide provides a known priming site. Reverse transcription was conducted using the reverse gene-specific primer TIR8 (Table 1) and MMLV reverse transcriptase. To amplify the first-strand cDNA of the 5′ end, the reverse gene-specific primer TIR8 and the primer Race5′ (homologous to the 5′ RACE RNA oligonucleotide) were used (Table 1). The amplification conditions used were an initial denaturation at 94 °C for 5 min, followed by 35 cycles at 94 °C for 60 s, 62 °C for 1 min, and 72 °C for 3 min, and finally an incubation at 72 °C for 10 min. The PCR product was separated by electrophoresis, and a band corresponding to the expected size of approximately 1000 bp was purified and sequenced. The comparative analysis of this *T. infestans* fragment with cDNA sequences of the *Vg* gene of other insects confirmed that it was the 5′ end of the gene.

2.4. cDNA sequence analysis

The *Vg* cDNA sequences of *T. infestans* were compared with those of other insects deposited in GenBank using the "BLAST-N" or "BLAST-X" tools available on the National Center for Biotechnology Information (NCBI) website. The *T. infestans* amino acid Vg sequence was deduced from the corresponding cDNA using the translation tool from the ExPASy Proteomics website (http://www.expasy.org/tools/dna.html). The putative phosphorylation sites were detected by using the NetPhos 2.0 Server (http://www.cbs.dtu.dk/services/NetPhos/).

2.5. Sequence comparison and phylogenetic relationship

Complete amino acid sequences were aligned using the Clustal W program in MEGA version 5.1 (Tamura et al., 2011). The Vg sequences used for comparison were from the following species (GenBank accession number in parentheses): *P. stali* (Vg1: BAA88075.1, Vg2: BAA88076.1, Vg3: BAA88077.1), *R. clavatus* (AAB72001.1), *T. infestans* (Vg2: KF915267), *Lethocerus deyrollei* (BAG12118.1), *Homalodisca coagulata* (DQ118408.1), *Nilaparvata lugens* (AEL22916.1), *G. nigrofuscata* (BAA85987.1), *Periplaneta americana* (Vg1: BAA86656.1, Vg2: BAB32673.1), *Blattella germanica* (CAA06379.2), and *Leucophaea maderae* (Vg1: AB052940.1, Vg2: AB194976.1). The phylogeny of the sequences was constructed using neighbor-joining methods (NJ) with MEGA version 5.1 software (Tamura et al., 2011) and *Tenebrio molitor* (AAU20328.2), *Apis mellifera* (AJ517411.1), *Aedes aegypti* (AAA18221.1), and *Bombyx mori* (BAA06397.1) as outgroups for phylogenetic analyses. In all the analyses the branch support was assessed by bootstrapping based on 1000 replicates.

2.6. Quantitative PCR (qPCR)

To observe yolk protein gene expression, the transcript levels of *Vg1* and *Vg2* in a variety of tissues, sexes and development stages of *T. infestans* were measured by qPCR. Gene specific primers (Table 1) and Taqman probes were designed according to the corresponding cDNAs using Primer Express program (Applied Biosystems). Reverse transcription polymerase chain reaction (RT-PCR) analysis was used to verify that the PCR products showed a single band and the expected sizes. The bands corresponding to the expected size were cloned into the pCR4-TOPO TA cloning vector (Invitrogen) and sequenced to confirm the identity of the amplified products. Based on the results of RT-PCR, relative qPCR was performed to further investigate changes in *Vg* expression. Quantitative PCR was carried out using Mx3005P qPCR System with Brilliant qPCR Core Reagent Kit (STRATAGENE, La Jolla, CA). Using temperature gradients, we have identified 58 °C as an annealing temperature appropriate for all primer sets. This has allowed us to analyze three genes in one real-time PCR plate. PCR conditions were 40 cycles of 10 min at 95 °C, 15 s at 95 °C, and 60 s at 58 °C. The relative copy number of *Vg* mRNA was calculated according to $2^{-\Delta\Delta CT}$ (Livak and Schmittgen, 2001). The threshold cycle value difference $^{\Delta}$CT between *Vg* mRNA and *β-actin* RNA of each reaction was used to normalize the level of total RNA. Three replicates were performed for each reaction to account for intra-experiment variation. The assay was repeated three times with separately extracted total RNA samples. One and Two-way ANOVA with Bonferroni post test were performed using GraphPad Prism version 5.00 for Windows (GraphPad Software, San Diego California USA). Statistical significance was assumed at $P \leq 0.05$.

3. Results and discussion

3.1. Identification of vitellogenin genes

The vitellogenin (*Vg*) cDNA of *T. infestans* was obtained using a primer designed by Lee et al. (2000) against the GL/ICG motif, which is conserved among all insect Vgs, and an adaptor primer at 3′ region. The sequence obtained had 466 bp, corresponding to 3′ end of *Vg* gene with a termination codon (TAA) and a polyadenylation signal (AAAT TAA) 54 nucleotides downstream from the termination codon. Subsequent reverse transcription PCR (RT-PCR) amplifications using specific primers designed from this first fragment and unspecific or degenerate primers designed from conserved regions of *Vg* of other hemipteran species, allowed one to obtain 3′ cDNA partial sequences of two *Vg* genes for *T. infestans* (*Vg1* and *Vg2*). The partial cDNA sequence of the *Vg1* gene, obtained from the first 466 bp isolated fragment, comprises 861 bp (GenBank accession number KF915268) with 794 bp encoding for 264 amino acids and 67 bp 3′-untranslated region (3′-UTR). The procedure of rapid amplification of cDNA 5′ end (RACE) resulted in the cDNA sequence of the *Vg2* gene (GenBank accession number KF915267). In total, 5639 bp of cDNA corresponding to *Vg2* of *T. infestans* was sequenced and shown to comprise an open reading frame (ORF) of 5565 nucleotides that encoded a protein of 1855

amino acids with a predicted MW of approximately 208.69 kDa (Fig. 1). The cDNA sequence includes 74 bp 3′-UTR and the stop codon TGA.

The comparative analysis of *Vg1* and *Vg2* cDNAs allowed examining a total of 843 bp for both genes. Within this *Vg* cDNA fragment, 69.99% of the nucleotide sequence was identical and the rest (30.01%) corresponded to 253 variable sites. The variable nucleotide positions included 101 transitions (39.92%), 102 transversions (40.32%), and 50 indels (19.76%, insertions and deletions). The alignment of the deduced amino acid sequences from *Vg1* and *Vg2* cDNA fragments of *T. infestans* revealed 58.94% of identical residues with 76.43% of homology (Fig. 2). Multiple *Vg* genes/cDNAs have been observed in many insect species including *Locusta migratoria*, *L. maderae*, *R. clavatus*, *P. stali*, *P. americana*, and *A. aegypti* (Tufail and Takeda, 2008).

3.2. Sequence analysis of Vg2

The National Center for Biotechnology Information (NCBI) Conserved Domain Database (CDD) search predicted three conserved regions for Vg2 (Fig. 1). Two Vitellogenin-N domains, also known as the lipoprotein amino-terminal regions implicated in lipid transporter activity (Smolenaars et al., 2007), were found between the amino acid positions 6–323 and 408–750, respectively. A third domain, the von Willebrand factor type D (VWD) domain, was positioned near the carboxy-terminus of the protein from amino acids 1507–1692. In addition, the sequence included one RXXR consensus cleavage site, two polyserine regions, and the conserved motifs GL/ICG and DGXR (Fig. 1). As already described for the majority of the insects: a) the

```
PQKLGWKNGNLYRYEVRGRTVVGLHQVADQFAGVLVKGKLTVQPKSDEILNLQVHLEQAE   60
VHTNLSSGWSTYIEDHDLNYKKLPISDTPFELKLKNGVVDKMYVDKNLPTWEVNVLKAIA  120
SQLQVDTQAENLKKSRVNQVPVEGQTIGVYKTVEDTVTGECETIYDISPLPQYVLQSRPE  180
LVPLPNLLGAHGQVIDIVKTKNYSNCEQRMAYHFGLTGLTDWEPASNQMGSFLSRSSVTR  240
VIVTGTLTSYTIQSSVTTNKIVLAPHLYNNQKGIVASRMNLTLEAVTSSSGSPQPVPNPR  300
TVKNLVYEYNAATTEENAQSHLMGRYIQGDEPTKVHKHSSSSESSSSSSEESLYTLMMRS  360
RRSIGTRRVENYSESSSSESNSSDDSDSSTGSSEDFWQPKPTLNEAFNSPFLPYFIGYLG  420
NSIQASSKVNGVSVVYKLAKEIGEEVLDGNSITGGNTLSKFTVLVNVIHTMNTTQLLEAT  480
EKLYYPYSKVVQTTEGDHPLYQAWVVFRDAVAEVGTGPALVTITHWILNHKITGAEAAHV  540
VGVLPNSAHYPTTEYMNTFFELLKNEEVQKQRYLNTSVLTAFTDLVRKAQVNRETAHNRY  600
PVHSFGDLTPKTDRVVSEQYVTYLETQLKKAVEQGDSPNIQVYIRALGNTAHPKILGVYE  660
PYLEGHEGVTDFQRLTIVASLDKITRVYPRLARSVLFKIYQNSGETPEVRVAAVMQLMKT  720
NPPAQILQRMAQHTNYDHSRQVNAAVKSAIENAARLHGPTSYQLAQNAKSAVHLLTTQDF  780
GFHQSYNFLHDYVVKEEELMYTTYLAYIQGEDSVVPSSAFYTLWRGLGGYKRQPFQASWM  840
TSHAGSLVDLLFDQFTEVDQSGKPKHGHGSEQGEGDWTLEKVGKFLNFETEYLDQVEGNV  900
LVSAFGVKNFFTFDNHTIEYLPEYVKYLTQQLRNGHHFNYNKLYNKHTLQFGFPTATGLP  960
FFYTFSTPTRTHFAGEVKFGSHPDLAGEQGGDKILIPKTADFTGLFKLMYTTKTEGVLGF 1020
TTPYNHKTFVADFVKNFHFGLPVKATVEVDLEHKKFQTKFQPWGENRNEKFVEYSTVAYT 1080
SVHDLLELEPIAEGDNTEEIHVRPVQRFETTFGEADTGFAFDVKVESEQKFVDWATFYNG 1140
ASKYDFMYALLFPFHDGTINNNNFTVTYNTHKTTAKYVKFFFTYDDLDVKTGGTGTHISH 1200
NKHRVVSGTGDESKDNDAVPSSTTPASQQRQKEFLQRAVAGIKDSYAQVFDLGVQFEGQT 1260
NAEYVATAAYSRSYVHPNRTFLFYFGKQPAKVTGHTKPYQVALQVNTEFPTVPVTDFEKA 1320
LHADPTSHISAKLYFGENVYSGGKVYVQGTFKQSEGRREYVEHHPVSALCRRQMEEGNYL 1380
LPACRNATASANFLDHYKFAVKYEGVPEVVKYHVYKTYTIARYFANHYVSEDFVTPSGKE 1440
GYLDFGVKFTHDLKAVNVSIETPVLTSEFNNVYFPEWFTPFVVVHPVYDFDDRFGQHFFR 1500
AQHFPTCVFDKHQATTFDNRTYPVNVGKCWHALVHSVPTTEFGYTGQSTYYDYDDFTVLV 1560
REGTANQKELLILLGKDVVQVLPTGGSEIVGKVVVNGQTADFSKESVAQFKYKDGEVQVQ 1620
VYALPTGEVRLLFPGTARAFTPGLEAFTPGLEVVYDGTHVKLQASNYYRGRVRGLCGTFD 1680
GEYVTDFTTPRNCVFKNPSEFAASYAVVDETCQGPAKELHQRVLNLPCYKQTVLLGDVVS 1740
EEEAGRFQPRIRTSQSNLLRGIGGQQTEERCSKLCSKVIEQSGKTCFSLHPQMTCTSKCK 1800
ATSKLEKRVEFHCVPNSAATKHWVEMIKKGANPDFSQKPANEWFKVNVPERCVPN*      1855
```

Fig. 1. Deduced amino acid sequence of the *Triatoma infestans* vitellogenin (Vg2) protein. Amino acids are numbered on the right. The three conserved regions predicted by the National Center for Biotechnology Information Conserved Domain Search are indicated: a) the two vitellogenin-N domains are shown with boxes at the amino-terminus, and b) the von Willebrand factor is shown by a dark-shaded box at the carboxy-terminus. The regions rich in polyserine tracks are underlined. The consensus RXXR sequence motif for possible cleavage site is indicated in bold and underlined. The putative phosphorylated serine (S), threonine (T), and tyrosine (Y) residues (predicted using the NetPhos 2.0 program) are shown with light-shaded frames. The DGXR and GL/ICG motifs are shown with boxes. The asterisk indicates the stop codon.

```
TiVg1   ------------------LSFRSAPGPLEGGKVKVNGQVLDFSKESITQVKNKDGDVLVQ  42
TiVg2   REGTANQKELLILLGKDVVQVLPTGGSEIVGKVVVNGQTADFSKESVAQFKYKDGEVQVQ  1620

TiVg1   IYALPNGVLRFSVPE-------------TGTEVIYDGIRILLQINNNYRGKVRGLCGTFD  89
TiVg2   VYALPTGEVRLLFPGTARAFTPGLEAFTPGLEVVYDGTRVKLQASNYYRGRVRGLCGTFD  1680

TiVg1   GEYVTDFTTPKNCIIKDPLEFAATYAIVDGTCEGPAKELNQRAKISPCYEKIVMLGDVIS  149
TiVg2   GEYVTDFTTPRNCVFKNPSEFAASYAVVDETCQGPAKELHQRVLNLPCYKQTVLLGDVVS  1740

TiVg1   DVEAGRNKPRMRSSLTKMLQGGSRKHSEVGCTNHRVEIVEQPGKTCFSLRPLVSCSSKCK  209
TiVg2   EEEAGRFQPRIRTSQSNLLRGIGGQQTEERCSKLCSKVIEQSGKTCFSLHPQMTCTSKCK  1800

TiVg1   PTKKIEKKIDFYCVQNSSASKHWVQMTRKGANPDFSQKPANAWFNINVPERCISI*       264
TiVg2   ATSKLEKRVEFHCVPNSAATKHWVEMIKKGANPDFSQKPANEWFKVNVPERCVPN*       1855
```

Fig. 2. Comparison of the carboxy-terminal part of Vg1 and Vg2. The alignment was performed using Clustal W program. The DGXR and GL/ICG motifs are shown with light-shaded frames. Common residues are indicated with dark-shaded frames. The asterisk indicates the stop codon.

tetra-residue RXXR has been found near the amino-terminal region and is flanked by polyserine domains; and b) the DGXR motif is located few residues upstream of the GL/ICG motif (Tufail and Takeda, 2008). Additionally, the 121 potential phosphorylation residues (serine: S, threonine: T, tyrosine: Y) found in the Vg2 sequence suggest that this molecule is likely to be highly phosphorylated. It has been reported that dephosphorylation of Vg reduces its uptake by oocytes, suggesting that phosphorylated residues may contribute to the interaction between Vg and its receptor on the oocyte surface (Dhadialla et al., 1992; Miller et al., 1982).

The deduced amino acid sequence of Vg2 from *T. infestans* was compared with the corresponding sequences of other hemipteran Vgs. Sequence alignment revealed that *T. infestans* Vg2 was most similar to Vg from *L. deyrollei* (47% identity), followed by *H. coagulata* (43%), Vg1, Vg2, and Vg3 of *P. stali* (from 39% to 37%), *N. lugens* (35%), *G. nigrofuscata* (35%), and *R. clavatus* (32%).

3.3. Phylogenetic relationship with Vgs of other hemimetabolous insects

The evolutionary relationship of 14 Vgs derived from 10 hemimetabolous insects was evaluated after aligning the complete amino acid sequences and conducting a phylogenetic analysis using neighbor-joining methods (NJ) with MEGA version 5.1 and representative species of the orders Lepidoptera, Diptera, Coleoptera and Hymenoptera (*B. mori*, *A. aegypti*, *T. molitor*, and *A. mellifera*, respectively) as outgroups (Fig. 3). The phylogenetic tree based on the entire Vg sequences of

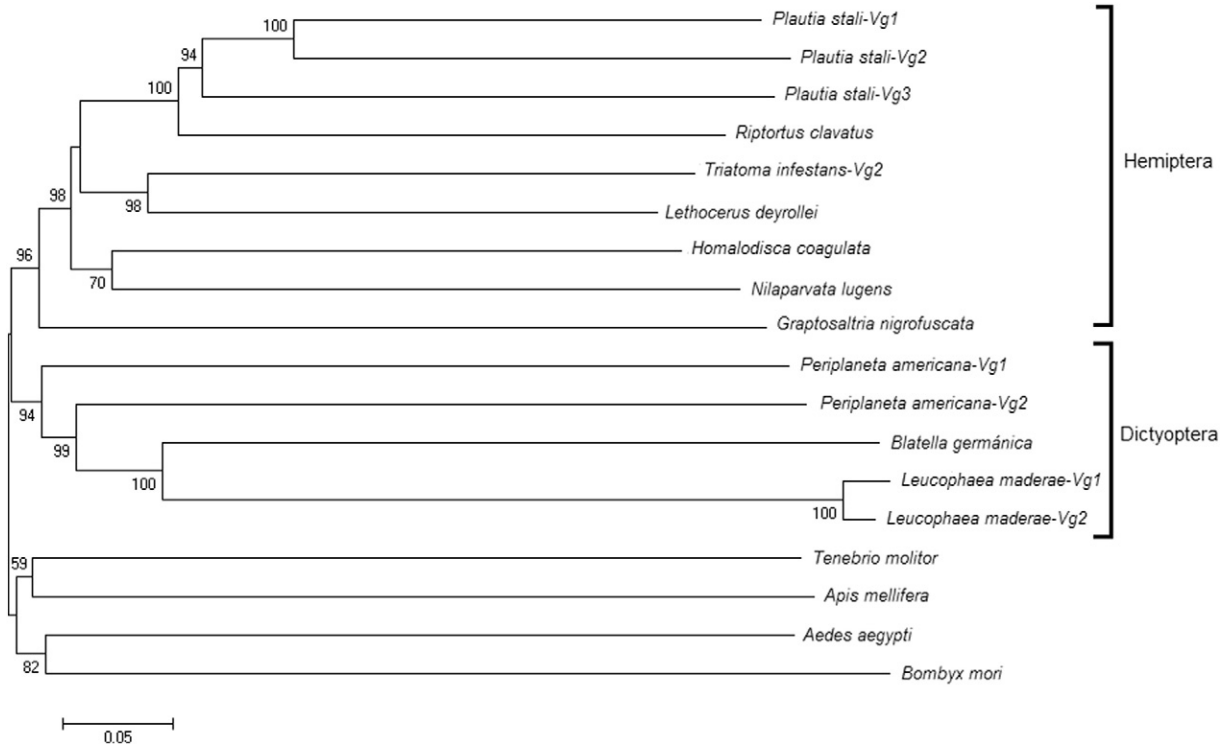

Fig. 3. Neighbor-joining bootstrap consensus tree based on 18 Vg sequences, 14 from hemimetabolous insects and four from insects that were used as outgroups for phylogenetic analyses (*Tenebrio molitor*, *Apis mellifera*, *Aedes aegypti*, and *Bombyx mori*). A distance analysis of amino acid sequences was performed using the Clustal W program and used as the input for a neighbor-joining tree construction program (MEGA 5.1). Numbers at the nodes are bootstrap values as percentages of 1000 replicates. Only values >50% are reported.

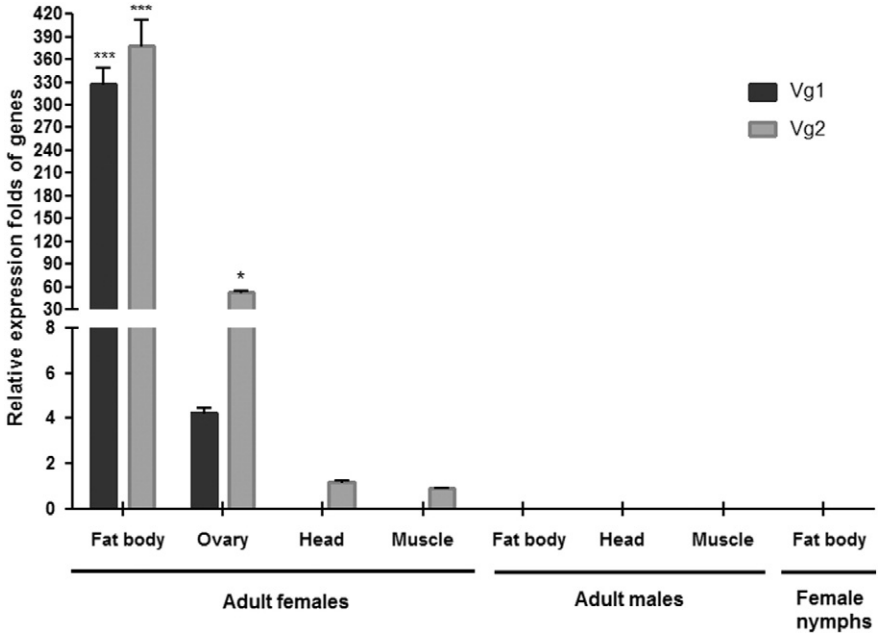

Fig. 4. Expression profiles of *Vg1* and *Vg2* in *Triatoma infestans*. Abundance of *Vg1* and *Vg2* transcripts in a panel of cDNA derived from different tissues, development stages, and two sexes were measured by qPCR using the specific primers for *Vg1*, *Vg2*, and *β-actin* (used as internal control), respectively. The error bars represent the standard deviation of mean. Statistical significance: * $P < 0.05$, *** $P < 0.001$.

hemimetabolous insects unambiguously supported two clusters, these being clearer than in previous phylogenetic analyses (Tufail and Takeda, 2008; Tufail et al., 2010). Dictyopteran and hemipteran Vg sequences appeared in separate clusters, with 94% and 96% bootstrap values, respectively. Within Hemiptera, *T. infestans* clustered with *L. deyrollei* with strong support (98% bootstrap value) and *R. clavatus* was also closely related to *P. stali* (100% bootstrap value). *G. nigrofuscata*, which in previous studies was clustered with cockroaches or was related to hemipteran species with weak bootstrap support (Tufail and Takeda, 2008; Tufail et al., 2010), occurs here in a basal position of the hemipteran cluster with high bootstrap support (96%).

3.4. Vg gene expression

Expression at mRNA level of *T. infestans Vg1* and *Vg2* genes was explored in both sexes, different tissues, and two developmental stages. Quantitative PCR showed that both *Vg1* and *Vg2* were expressed only in fat bodies and ovaries of adult females (Fig. 4). The level of the *Vg1* and *Vg2* transcripts was similar in pools of adult female fat bodies and the *Vg2* transcript was predominant in pools of adult female ovaries. The highest levels of *Vg* transcripts were detected in adult female fat bodies. This result is in agreement with other studies which have shown that in insects the female fat body is the major site of Vg biosynthesis (Tufail and Takeda, 2008). Additionally, the presence of *Vg* transcripts in ovaries of *T. infestans* is consistent with studies carried out in other insects showing that the fat body is not the only vitellogenic tissue, since ovarian follicular epithelium also produces yolk proteins (Belles, 1998, 2005; Giorgi et al., 2005). In this respect, studies in another triatomine, *Rhodnius prolixus*, have shown that the Vg of the hemolymph is clearly synthesized by fat bodies (Valle et al., 1993), but additional synthesis of Vg by the follicle cells of the ovary is also observed (Melo et al., 2000). Interestingly, we found mild *Vg2*, but not *Vg1*, expression in head and muscle of adult females. Neither *Vg1* nor *Vg2* was expressed in fifth instar nymph fat bodies or in adult male fat bodies, heads, and muscles. Contrary to this result, it has been described that synthesis of Vg occurs in males of some species of insects (Engelmann, 1979; Piulachs et al., 2003; Trenczek and Engels, 1986; Valle, 1993), including male adults of *R. prolixus* (Valle et al., 1993).

4. Conclusions

This study is the first report of genes involved in the reproduction in triatomines. We have identified two *Vg* genes (*Vg1* and *Vg2*) in *T. infestans*. The phylogenetic analysis revealed that the amino acid sequence of Vg from this species is related with strong support to Vg sequences of other hemipteran species. We detected *Vg* expression, at mRNA level, only in adult females. *Vg1* and *Vg2* mRNA transcripts were found in fat body and ovary and, to a much less extent, *Vg2* mRNA was detected in head and muscle. This work has contributed sequence data as well as information about expression of *Vg1* and *Vg2* genes in a vector of Chagas' disease.

Conflict of interest

The authors confirm that this article content has no conflicts of interest.

Acknowledgments

We thank the Centro de Referencia de Vectores, Servicio Nacional de Chagas de Córdoba, Argentina, for providing insects used in our studies. This research was supported by grants from Agencia Nacional de Promoción Científica y Tecnológica (FONCyT) of Argentina (PICT-01265), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) of Argentina (PIP-11220090100437), and Secretaría de Ciencia y Tecnología, Universidad Nacional de Córdoba, Argentina (162/12-124/13). M. J. Blariza and C. G. Grosso are fellows of CONICET. Beatriz A. Garcia is a Career Investigator of CONICET.

References

Belles, X., 1998. Endocrine effectors in insect vitellogenesis. In: Coast, G.M., Webster, S.G. (Eds.), Resent Advances in Arthropod Endocrinology. Cambridge University Press, Cambridge, pp. 71–90.

Belles, X., 2005. Vitellogenesis direct by juvenile hormone. In: Raikhel, A.S., Sapington, T. W. (Eds.), Progress in Vitellogenesis. Reproductive Biology of Invertebrates, vol. XII, Part B. Science Publishers, Inc., Enfield, USA-Plymouth, pp. 157–197.
Dhadialla, T.S., Raikhel, A.S., 1990. Biosynthesis of mosquito vitellogenin. Journal of Biological Chemistry 265, 9924–9933.
Dhadialla, T.S., Hays, A.R., Raikel, A.S., 1992. Characterization of the solubilized mosquito vitellogenin receptor. Insect Biochemistry and Molecular Biology 22, 803–816.
Engelmann, F., 1979. Insect vitellogenin: identification, biosynthesis, and the role in vitellogenesis. Advances in Insect Physiology 14, 49–108.
Espinola, N.H., 1966. Note on sex differences in immature forms of Triatominae (Hemiptera, Reduviidae). Revista Brasileira de Biologia 26, 263–267.
Germano, M.D., Roca Acevedo, G., Mougabure Cueto, G.A., Toloza, A.C., Vassena, C.V., Picollo, M.I., 2010. New findings of insecticide resistance in *Triatoma infestans* (Heteroptera: Reduviidae) from the Gran Chaco. Journal of Medical Entomology 47, 1077–1081.
Giorgi, F., Snigirevskaya, E.S., Raikel, A.S., 2005. The cell biology of yolk protein precursor synthesis and secretion. In: Raikel, A.S., Sappington, T.W. (Eds.), Progress in Vitellogenesis. Reproductive Biology of Invertebrates, vol. XII, Part B. Science Publishers, Inc., Enfield, USA-Plymouth, pp. 33–68.
Gürtler, R.E., Kitron, U., Cecere, M.C., Segura, E.L., Cohen, J., 2007. Sustainable vector control and management of Chagas disease in the Gran Chaco, Argentina. Proceedings of the National Academy of Sciences of the United States of America 104, 16194–16199.
Hughes, A.L., 2010. Evolutionary conservation of amino acid composition in paralogous insect vitellogenins. Gene 467, 35–40.
Kunkel, J.G., Nordin, J.H., 1985. Yolk protein. In: Kerkut, G.A., Gilbert, L.I. (Eds.), Comprehensive Insect Physiology, Biochemistry and Pharmacology, vol. 1. Pergamon Press, Oxford, pp. 83–111.
Lee, J.M., Hatakeyama, M., Oishi, K., 2000. A simple and rapid method for cloning insect vitellogenin cDNAs. Insect Biochemistry and Molecular Biology 30, 189–194.
Livak, K.J., Schmittgen, T.D., 2001. Análisis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta}$CT method. Methods 25, 402–408.
Lorca, M., García, A., Contreras, M.C., Schenone, H., Rojas, A., 2001. Evaluation of a *Triatoma infestans* elimination program by the decrease of *Trypanosoma cruzi* infection frequency in children younger than 10 years, Chile, 1991–1998. American Journal of Tropical Medicine and Hygiene 65, 861–864.
Melo, A.C.A., Valle, D., Machado, E.A., Salerno, A.P., Paiva-Silva, G.O., Cunha, E., Silva, N.L., de Souza, W., Masuda, H., 2000. Synthesis of vitellogenin by the follicle cells of *Rhodnius prolixus*. Insect Biochemistry and Molecular Biology 30, 549–557.
Miller, M.S., Benore, P.M., White, H.B., 1982. Dephosphorylation of chicken riboflavin-binding protein and phosvitin decreases their uptake by oocytes. Journal of Biological Chemistry 257, 6818–6824.
Picollo, M.I., Vassena, C., Santo Orihuela, P., Barrios, S., Zaidemberg, M., Zerba, E., 2005. High resistance to pyrethroid insecticides associated to the ineffectiveness of field treatments in *Triatoma infestans* (Hemiptera, Reduviidae) from the north of Argentina. Journal of Medical Entomology 42, 637–642.
Piulachs, M.D., Guidugli, K.R., Barchuk, A.R., Cruz, J., Simoes, Z.L.P., Belles, X., 2003. The vitellogenin of the honey bee *Apis mellifera*: structural analysis of the cDNA and expression studies. Insect Biochemistry and Molecular Biology 33, 49–465.
Rassi Jr., A., Rassi, A., Marin-Neto, J.A., 2010. Chagas disease. Lancet 375, 1388–1402.
Santo Orihuela, P.L., Vassena, C.V., Zerba, E.N., Picollo, M.I., 2008. Relative contribution of monooxygenase and esterase to pyrethroid resistance in *Triatoma infestans* (Hemiptera: Reduviidae) from Argentina and Bolivia. Journal of Medical Entomology 45, 298–306.
Sappington, T., Oishi, K., Raikhel, A.S., 2002. Structural characteristics of insect vitellogenins. In: Raikhel, A.S., Sappington, T.W. (Eds.), Progress in Vitellogenesis. Reproductive Biology of Invertebrates, vol. XII, Part A. Science Publishers, Inc., Enfield, USA-Plymouth, pp. 69–101.
Silveira, A., Vinhaes, M., 1999. Elimination of vector-borne transmission of Chagas disease. Memórias do Instituto Oswaldo Cruz 94, 405–411.
Smolenaars, M.M.W., Madsen, O., Rodenburg, K.W., Van der Horst, D.J., 2007. Molecular diversity and evolution of the large lipid transfer protein superfamily. Journal of Lipid Research 48, 489–502.
Snigirevskaya, E.S., Raikhel, A.S., 2005. Receptor-mediated endocytosis of yolk proteins in insect oocytes. In: Raikhel, A.S., Sappington, T.W. (Eds.), Progress in Vitellogenesis. Reproductive Biology of Invertebrates, vol. XII, Part B. Science Publishers, Inc., Enfield, USA-Plymouth, pp. 199–228.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28, 2731–2739.
Toloza, A.C., Germano, M., Cueto, G.M., Vassena, C., Zerba, E., Picollo, M.I., 2008. Differential patterns of insecticide resistance in eggs and first instars of *Triatoma infestans* (Hemiptera: Reduviidae) from Argentina and Bolivia. Journal of Medical Entomology 45, 421–426.
Trenczek, T.A., Engels, W., 1986. Occurring of vitellogenin in drone honeybees. International Journal of Invertebrate Reproduction and Development 10, 307–311.
Tufail, M., Takeda, M., 2008. Molecular characteristics of insect vitellogenins. Journal of Insect Physiology 54, 1447–1458.
Tufail, M., Raikhel, A.S., Takeda, M., 2004. Biosynthesis and processing of insect vitellogenins. In: Raikhel, A.S., Sappington, T.W. (Eds.), Progress in Vitellogenesis. Reproductive Biology of Invertebrates, vol. XII, Part B. Science Publishers, Inc., Enfield, USA-Plymouth, pp. 1–32.
Tufail, M., Naeemullah, M., Elmogy, M., Sharma, P.N., Takeda, M., Nakamura, C., 2010. Molecular cloning, transcriptional regulation, and differential expression profiling of vitellogenin in two wing-morphs of the Brown planthopper, *Nilaparvata lugens* Stal (Hemiptera: Delphacidae). Insect Molecular Biology 2–12.
Valle, D., 1993. Vitellogenesis in insects and other groups—a review. Memórias do Instituto Oswaldo Cruz 88, 1–26.
Valle, D., Kun, J., Linss, J., de Souza García, E., Goldenberg, S., 1993. cDNA Cloning and expression of *Rhodnius prolixus* vitellogenin. Insect Biochemistry and Molecular Biology 23, 457–465.
Vassena, C.V., Picollo, M.I., Zerba, E.N., 2000. Insecticide resistance in Brazilian *Triatoma infestans* and Venezuelan *Rhodnius prolixus*. Medical and Veterinary Entomology 14, 51–55.
World Health Organization, 1994. Chagas disease. Elimination of transmission. Weekly Epidemiological Record 69, 38–40.