The cys-loop ligand-gated ion channel gene superfamilies of the cockroaches *Blattella germanica* and *Periplaneta americana*

Running Title: Cockroach ligand-gated ion channel genes

Andrew K Jones[1], Delphine Goven[2], Josy-Anne Froger[2], Alexandre Bantz[2] and Valerie Raymond[2]

[1]Department of Biological and Medical Sciences, Faculty of Health and Life Sciences, Oxford Brookes University, Oxford, OX3 0BP, UK

[2]Laboratoire « Signalisation Fonctionnelle des Canaux Ioniques et Récepteurs » (SiFCIR), UPRES-EA2647 USC INRAE 1330, SFR 4207 QUASAV, UFR Sciences, Université d'Angers, 2 Bld Lavoisier, 49045 Angers Cedex 01, France

Corresponding author: Andrew Jones; Department of Biological and Medical Sciences, Faculty of Health and Life Sciences, Oxford Brookes University, Oxford, OX3 0BP, UK; +44 (0)1865 483602; a.jones@brookes.ac.uk

Email addresses: Alexandre Bantz, alexandre.bantz@univ-angers.fr; Josy-Anne Froger, josy-anne.froger@univ-angers.fr; Delphine Goven, delphine.goven@univ-angers.fr; Valerie Raymond, valerie.raymond@univ-angers.fr

This manuscript is intended to be included in the Pest Management Science special issue based on the Fall 2019 ACS AGRO symposia honouring Vincent Salgado for the International Award and Tom Sparks for the Kenneth A. Spencer Award.

Abstract

BACKGROUND: Cockroaches are serious urban pests that can transfer disease-causing microorganisms as well as trigger allergic reactions and asthma. They are commonly managed by pesticides that act on cys-loop ligand-gated ion channels (cysLGIC). To provide further information that will enhance our understanding of how insecticides act on their molecular targets in cockroaches, we used genome and reverse transcriptase PCR data to characterise the cysLGIC gene superfamilies from *Blattella germanica* and *Periplaneta americana*.

RESULTS: The *B. germanica* and *P. americana* cysLGIC superfamilies consist of 30 and 32 subunit-encoding genes, respectively, which are the largest insect cysLGIC superfamilies characterized to date. As with other insects, the cockroaches possess ion channels predicted to be gated by acetylcholine, gamma-aminobutyric acid, glutamate and histamine, as well as orthologues of the Drosophila pH-sensitive chloride channel (pHCl), CG8916 and CG12344. The large cysLGIC superfamilies of cockroaches are a result of an expanded number of divergent nicotinic acetylcholine receptor subunits, with *B. germanica* and *P. americana* respectively possessing eight and ten subunit genes. Diversity of the cockroach cysLGICs is also broadened by alternative splicing and RNA A-to-I editing. Unusually, both cockroach species possess a second glutamate-gated chloride channel as well as another CG8916 subunit.

CONCLUSION: These findings on *B. germanica* and *P. americana* enhance our understanding of the evolution of the insect cysLGIC superfamily and provide a useful basis for the study of their function, the detection and management of insecticide resistance, and for the development of improved pesticides with greater specificity towards these major pests.

Keywords: cockroach, *Blattella germanica*, *Periplaneta americana*, cys-loop ligand gated ion channel, insecticide target, nicotinic acetylcholine receptor

1 INTRODUCTION

We are delighted to be contributing a paper as part of a special issue honouring Vincent Salgado and Tom Sparks. In working on insecticides we have naturally come across both Tom's and Vincent's studies and the resulting wealth of information investigating the actions of various compounds. In the process of developing our research careers, 'Salgado' and 'Sparks' have struck us as big names in the field, whose work is influential, inspirational and aspirational in helping us to push forwards in our investigations. We hope that the following study, characterizing molecular targets of insecticides at the genomic and transcript level, will complement well the more functional aspect of Tom's and Vincent's work.

In insects, members of the cys-loop ligand-gated ion channel (cysLGIC) superfamily mediate both fast excitatory and inhibitory synaptic transmission in the nervous system and may also be performing non-neuronal roles.[1] The superfamily includes cation-permeable nicotinic acetylcholine receptors (nAChRs),[2] γ-aminobutyric acid (GABA)-gated ion channels[3] as well as chloride channels gated by glutamate (GluCls),[4] histamine[5, 6] or zinc.[7] Insect cysLGICs are of considerable interest as they are targets of effective pesticides.[8] For example, nAChRs are targets of neonicotinoids and spinosyns,[9, 10] GABA receptors are targets of fiproles and isoxazolines[11] whilst avermectins act mainly on GluCls.[4]

Cockroaches are serious urban pests that can transfer disease-causing microorganisms as well as trigger allergic reactions and asthma.[12] Management of cockroaches is heavily reliant on the use of insecticides, including the neonicotinoids imidacloprid and dinotefuran, fipronil and abamectin, in bait formulations.[13] There have been reports of cockroaches showing resistance to insecticidal bait containing fipronil, abamectin and imidacloprid[14-16] highlighting the need for improved knowledge of resistance mechanisms and detection as well as the development of novel control agents. The use of cockroach neurons have resulted in many instructive studies investigating the actions of insecticides on cysLGICs in the insect nervous system.[17-23] To date, several cockroach

cysLGIC subunit sequences have been reported. These include the nAChR subunits α1, α2, α3, α4, α6, α7, α8 and β1 from *Periplaneta americana*.[22, 24, 25] Also, fragments of *Rdl* from *Blattella germanica* were analyzed in order to investigate the basis of cyclodiene resistance.[26] Characterizing the full complement of cockroach cysLGIC subunits represents a critical step in identifying key components of the cockroach nervous system as well pinpointing particular molecular targets underlying responses to insecticides. Here we have used genome sequences of *B. germanica*[27] and *P. americana*[12] as well as reverse transcriptase PCR to provide descriptions of their complete cysLGIC gene superfamilies.

2 MATERIALS AND METHODS

2.1 Identification of cysLGIC subunits from *B. germanica* and *P. americana*

To identify putative cockroach cysLGIC subunits, we used tBLASTn[28] to screen the *B. germanica* genome assembly or transcript sequences (available at the United States Department of Agriculture https://i5k.nal.usda.gov/webapp/blast/) and the *P. americana* genome or transcript sequences (available at the NCBI National Center for Biotechnology Information https://blast.ncbi.nlm.nih.gov/Blast.cgi) with every member of the *Acyrthosiphon pisum* or *Tribolium castaneum* cysLGIC superfamilies.[29, 30] Candidate cockroach cysLGIC subunits were identified based on their considerable sequence homology with previously characterized subunits particularly in the N-terminal ligand-binding domain and the four transmembrane regions. In some cases, the cockroach cysLGIC sequences were corrected based on protein sequence alignments or reverse-transcriptase PCR sequence data (see supporting information Tables 1A and 1B for accession numbers of the cockroach cysLGIC subunit sequences).

2.2 Insects

Adult male cockroach *P. americana* used for RT-PCR experiments were maintained under standard laboratory conditions (28 °C with 12 h light / 12 h dark photo-cycle, food and water *ad libitum*). These cockroaches come from a laboratory-maintained strain that is susceptible to insecticides.

2.3 Reverse transcription polymerase chain reaction

To study cysLGIC subunits expressed in the *P. americana* central nervous system and confirm coding sequences identified in the cockroach genome, RT-PCR was carried out on the terminal abdominal ganglion (TAG), which was removed from the nerve cord of cockroaches as previously described in Benzidane *et al*.[22] The Nucleospin RNA kit (Macherey Nagel, Düren, Germany) was used to extract total RNA from TAG following the manufacturer's instructions and 500 ng of sodium acetate/ethanol purified RNA was reverse transcribed using the RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, Waltham, MA, USA). The PCR amplification reactions were performed with KOD hot start polymerase (Novagen) and specific cysLGIC primer sets (see Supporting information Table 2) using the following thermocycling conditions: 2 min at 95°C then 35 cycles of 20 sec at 95°C, 10 sec at optimal annealing temperature of primers (between 50 to 65°C) and 20 sec/kb at 70°C before being held at 10°C. The PCR products were finally purified using NucleoSpin Gel and PCR clean-up kit (Macherey Nagel, Düren, Germany) then directly sequenced (Eurofins).

2.4 Sequence analysis

Sequence of PCR products were translated to their respective protein sequences then the multiple protein sequence alignments were constructed with Clustal X2[31] using the slow-accurate mode with a gap-opening penalty of 10 and a gap-extension penalty of 0.1 and applying the Gonnet 250 protein weight matrix. The protein alignments were viewed using GeneDoc (http://www.nrbsc.org/gfx/genedoc/index.html), which was also used to calculate identity and similarity values between subunit sequences. Signal-peptide cleavage sites were predicted using the SignalP 3.0 server[32] and membrane-spanning regions were predicted using the TMpred program

(http://www.ch.embnet.org/software/TMPRED_form.html). The PROSITE database[33] was used to identify potential N-glycosylation and phosphorylation sites.

Phylogenetic trees were constructed with MEGAX software[34] using the Maximum Likelihood method and Dayhoff matrix based model. The tree with the highest log likelihood after 1000 bootstrap replications is shown.

3 RESULTS

3.1 The *B. germanica* and *P. americana* cysLGIC superfamilies consist of 30 and 32 subunit members, respectively

Using tBLASTn, 30 and 32 cysLGIC subunit genes were identified in the genomes of *B. germanica* and *P. americana*, respectively. These are notably large insect cysLGIC superfamilies considering that the previously identified superfamilies from genomes of *Drosophila melanogaster*,[35] *Apis mellifera*,[36] *Tribolium castaneum*,[30] *Nasonia vitripennis*,[37] *Acyrthosiphon pisum*,[29] *Aedes aegypti*[38] and *Aethina tumida*[39] consist of 21-26 genes.[1] An alignment of *P. americana* protein sequences shows that the cockroach subunits possess features characteristic to members of the cysLGIC superfamily[40] (Figures 1 and 2). These include: (i) an extracellular N-terminal domain containing distinct regions (loops A–F) that form the ligand-binding site;[41] (ii) the dicysteine loop (cys-loop) consisting of two disulphide bond-forming cysteines separated by 13 amino-acid residues; (iii) four transmembrane regions (TM1–4); (iv) a highly variable intracellular loop between TM3 and TM4. As with other cysLGIC subunits, the cockroach sequences also possess potential N-glycosylation sites in the extracellular N-terminal domain, which can affect receptor maturation, channel desensitization and conductance,[42, 43] and putative phosphorylation sites in the TM3–TM4 intracellular loop, which regulates several aspects of receptor function such as desensitization, aggregation and conformation changes.[44-46]

A comparison of sequence identities between *P. americana* and *T. castaneum* cysLGIC subunits (Tables 1 and 2), as well as the use of phylogenetic trees (Figures 3, 4 and 5), indicate orthologous relationships between the cockroach, beetle and aphid (*A. pisum*) subunits. In order to maintain consistent nomenclature and to facilitate comparisons between different insect species, cockroach subunits were named after their tribolium counterparts.[30] For example, the *B. germanica* or *P. americana* orthologues of *T. castaneum* nAChRα1, RDL and CG12344 were designated Bgerα1, BgerRDL, Bger12344 or Pameα1, PameRDL, Pame12344, respectively. Phylogenetic analysis indicates that an early duplication event lead to one branch evolving into the nAChR subunits whilst the other branch generated the remaining subunits (Figure 3).

3.2 Cockroach nicotinic acetylcholine receptor subunits

In the genomes of *B. germanica* and *P. americana*, 17 and 19 candidate nAChR-subunit encoding genes were identified, respectively (Figures 1, 3 and 4). This is larger than other insect nAChR gene families described, such as from *D. melanogaster* (10 subunits),[47] *Anopheles gambiae* (10),[48] *A. mellifera* (11),[49] *A. tumida* (12),[39] *A. pisum* (11),[50] *Bombyx mori* (12),[51] *T. castaneum* (12),[30] *Ae. aegypti* (14)[38] and *N. vitripennis* (16).[37] Both *B. germanica* and *P. americana* possess typical core groups of nAChR subunits that are highly conserved in different insect species,[2] these being α1-α8 and β1 (Figure 4). As is the case for Dα1, Dα2, Dα3, Dα4, Dβ2 and their orthologues in other insect species,[2] the corresponding cockroach subunits (α1–4 and α8) have an insertion in loop F (Figure 1), which may contribute to interactions with the neonicotinoid, imidacloprid.[52] The Bgerα1, Bgerα2 and Bgerα8 genes are located within 686 kb of each other, although not as tightly clustered together as the corresponding genes in the *D. melanogaster* (Dα1, Dα2 and Dβ2) and *An. gambiae* (Agamα1, Agamα2 and Agamα8) genomes, where they are found within 220 kb.[48] It is unknown whether the orthologous genes in *P. americana* are similarly arranged as its genome is still at scaffold assembly stage and the scaffolds are not of sufficient length to determine whether they are in proximity to each other. The Pameα7 and Pameβ1 subunits, however, as well as Bgerα7 and Bgerβ1, were found

within 420 kb of each other. Clustering of the equivalent subunits have also been seen in the genomes of *A. mellifera*, *An. gambiae* and *T. castaneum*.[30, 48, 49] RT-PCR of *P. americana* cDNA generated amplification products of all the core group subunits, indicating that these subunits are transcribed in cockroaches.

B. germanica and *P. americana* possess eight and ten divergent nAChR subunits, respectively. This is notably larger than in other genomes such as those of *Ae. aegypti*, *An. gambiae*, *A. mellifera*, *A. pisum*, *A. tumida, B. mori*, *D. melanogaster*, *T. castaneum* and *N. vitripennis*, which possess between 1 and 7 divergent subunit genes.[30, 37-39, 47-51] Of the divergent nAChR subunits present in *B. germanica*, two are of the α subtype and six are β whereas there are one α and nine β in the *P. americana* genome (Figures 3 and 4). As with one of the tribolium divergent subunits, Tcasα9, Bgerα10 possesses an atypical FxCC amino-acid motif (Supporting information Table 1), instead of the highly conserved YxCC found in loop C, which may result in unusual ligand-binding properties.[53] In addition, the cockroach divergent subunits, with the exception of α9, lack the GEK motif characteristic of nAChR subunits, which precedes TM2 (Figure 1), and has an important role in ion permeation and selectivity.[54] This raises the possibility that these subunits may form nAChRs with distinct ion channel characteristics. As is common for insect divergent nAChR subunits, several cockroach subunit genes are found clustered together in the genome indicating recent gene duplication events. Thus, Bgerα10 and Bgerβ7 are found within 63 kb of each whilst Pameβ7 and Pameβ9 are located within 53 kb. Also, Bgerβ3, Bgerβ4, Bgerβ5 and Bgerβ6 are all found within 50 kb whereas an extra duplication event appears to have expanded this gene cluster in *P. americana* where Pameβ3, Pameβ4, Pameβ5, Pameβ6 and Pameβ8 are all located within 50 kb. RT-PCR amplification products were seen for the of *P. americana* divergent nAChR subunits (except for β7, β9 and β10), indicating that several of the cockroach divergent nAChRs are expressed.

Bgerα4, Pameα4, Bgerα6 and Pameα6 have alternatively spliced exons, which are common to insect α4 and α6 subunits.[2] Thus, Bgerα4 and Pameα4 possesses two alternatives for exon 4

(denoted exon4 and exon4')[55] (Figure 6A), which introduces variation in LoopE and the cys-loop. Similar to α6 in *T. castaneum* and other species,[30] Bgerα6 and Pameα6 have two alternatives for exon 3 and three alternatives for exon 8 (Figure 6B). Bgerα7 and Pameα7 also have alternative splicing for exons 7 and 8 (Figure 6C), which, so far, has only been reported for the aphid, *A. pisum*.[50] Here, there are two alternatives for exon 7 in the aphid and cockroach genomes, which introduces variation in Loop C and the first transmembrane domain. Exon 7b of Bgerα7 and Pameα7 differ by one amino acid in the region between Loop C and TM1, with a valine being present in *P. americana* and a tyrosine in *B. germanica*. Bgerα7 and Pameα7 also have four alternatives for exon 8 (Figure 6C) whereas *A. pisum* has three, at a site equivalent to splicing of exon 8 and in the α6 subunit (Figure 6B), which introduces variation in or near the second transmembrane domain.

In analyzing sequences of *P. americana* nAChR subunits amplified by RT-PCR, potential RNA A-to-I editing in α6 was detected as indicated by adenosine in the genomic DNA sequence being replaced by guanosine in the cDNA sequence (Figure 6D). This results in the recoding of two amino acid residues within Loop E, effectively removing a putative N-glycosylation site, which is frequently affected by RNA editing in α6 of other species such as *T. castaneum* (Figure 6D).[2, 30]

3.3 Cockroach GABA receptor subunits

The genomes of *B. germanica* and *P. americana* contain orthologues of the known *D. melanogaster* GABA-gated ion channel subunits, RDL and the two GABA receptor-like subunits, GRD and LCCH3 (Figures 3 and 5, Table 2).[3, 56] As it is the case for other species, cockroach RDL possesses a PAR sequence before TM2 (Figure 2) which is important for forming anion channels,[54] whereas cockroach GRD and LCCH3 lack this sequence, consistent with drosophila GRD and LCCH3 forming cation-permeable channels when expressed in oocytes of *Xenopus laevis*.[57]

The *Rdl* genes of *B. germanica* and *P. americana* have alternative splicing, similar to that of many other species such as *D. melanogaster*,[58] *T. casteneum* ,[30] *A. mellifera*[36] and *A. pisum*,[29] where there are two alternatives for exons 3 and 6 (Figure 7A). Exon 6a of Bger RDL and Pame RDL differ by

one amino acid, in Loop F, with a threonine being present in *P. americana* and an isoleucine in *B. germanica*. In addition, two alternatives for exon 9 in Bger RDL were detected, which introduces variation at the C-terminal end of the protein, including TM4 (Figure 7A). Alternative splicing of this exon has so far only been reported for the small brown planthopper, *Laodelphax striatellus*.[59] Exon 9 of *Rdl* in the genome and transcriptome of *P. americana* was not observed to have alternatively spliced variants. Analysis of RNA-seq data identified two variants of *P. americana Rdl* (variant 1 and 2) with different length intracellular loops between TM3 and TM4 (Figure 7B) arising from differential use of gt splice donor sites. Variant 2 has an insertion of 21 amino acid residues, which disrupts a putative protein kinase C phosphorylation site. Insertion of amino acids at this site has been seen in RDL of diverse species including *A. mellifera*[60] and the miridbug, *Cyrtorhinus lividipennis*,[61] although the actual sequence inserted is species-specific.

3.4 Cockroach GluCls and HisCls

B. germanica and *P. americana* possess two HisCl genes, HisCl1 and HisCl2 (Figures 3 and 5, Table2), which are highly conserved in diverse insect species such as *D. melanogaster*,[5, 6] *A. mellifera*,[36] *T. castaneum*,[30] *A. pisum*[29] and the butterfly *Papilio Xuthus*.[62] *B. germanica* and *P. americana* also possess a GluCl gene, which again, is highly conserved amongst insects.[4, 37, 63] This includes alternative splicing where two alternatives for exon 3 has been observed for *B. germanica* and *P. americana* GluCls (Figure 7C). When compared to the three splice variants identified in *T. castaneum* GluCl,[30] *B. germanica* and *P. americana* GluCl exon3a was identical to tribolium exon3a whilst the second cockroach variant (here denoted exon 3b) was most similar to tribolium exon 3c (Figure 7B). Exon 3b of Bger GluCl and Pame GluCl differ by one amino acid with an isoleucine being present in *P. americana* and a tyrosine in *B. germanica*.

As with *A. pisum*,[29] a second putative GluCl gene was found in the genomes of *B. germanica* and *P. americana*, which were tentatively denoted as Bger GluCl2 or Pame GluCl2, respectively, since when compared to tribolium cysLGIC subunits, they shared highest sequence identity with Tcas

GluCl at 33% (Table 2). This level of identity did not change when considering only the N-terminal extracellular domains without the signal peptide. Phylogenetic analysis shows that the cockroach GluCl2 subunits do not appear to be closely related to GluCls in other species (Figure 5) although Bger GluCl2 and Pame GluCl2 do cluster with Apisum GluCl2 sharing a considerably high amino acid identity of 81%. Analysis of RT-PCR data of *P. americana* cDNA confirmed that GluCl2 is transcribed.

3.5 Cockroach pHCls and other cysLGICs

The *B. germanica* and *P. americana* genomes possess an orthologue of the pH-sensitive chloride channel (Figure 5 and Table 2), which was first identified in *D. melanogaster*.[64] The cockroach genomes also clearly possess orthologues of drosophila CG8916 and CG12344 (Figure 5 and Table 2), the latter of which may be involved in mediating responses to glycine.[65] *B. germanica* and *P. americana* genomes possess two subunits (CLGC1 and CLGC2, Figure 5) belonging to the Insect Group 1 of cysLGIC subunits.[30] Initially identified in *D. melanogaster* as CG6927, CG7589 and CG11340,[66] members of this group generally show unclear orthologous relationships between species (Table 2) and are amongst the least studied insect cysLGICs.[1] However recently, it has been shown that drosophila CG11340, referred to as *hodor*, encodes for a zinc-gated chloride channel.[7] As with the cockroaches, *A. pisum* has two subunits in this group[29] whereas *T. castaneum* has three (Figure 5)[30] and *A. mellifera* has one.[36]

Interestingly, an additional cysLGIC subunit gene was found in the genomes of *B. germanica* and *P. americana*. Phylogenetic analysis suggests that this subunit is most closely related to GRD or 8916 in other species (Figure 5) sharing comparable levels of identity and similarity with tribolium GRD and 8916 (Table 2). Evolutionary distance estimates suggests that it is more closely related to tribolium 8916 (Supporting information Table 3), therefore this subunit was tentatively denoted as Bger 8916_2 or Pame 8916_2.

4 DISCUSSION

4.1 Cockroaches have unusually large cysLGIC gene superfamilies

Cockroaches are pests of urban environments and are commonly managed by pesticides such as fipronil, imidacloprid, dinotefuran and abamectin,[13] which target members of the cysLGIC superfamily.[4, 9, 11] Here, we have used the available genome sequences of *B. germanica* and *P. americana*[12, 27] to characterize the complete cysLGIC gene superfamilies from these major pests. Consisting of 30 and 32 subunit genes, *B. germanica* and *P. americana*, respectively, possess the largest insect cysLGIC superfamilies so far characterized. Transcriptome analysis of complete and partial transcripts has suggested that the locust, *Locusta migratoria manilensis*, may have an even more extensive cysLGIC gene superfamily with potentially 67 subunit-encoding genes.[67] However, it appears that several sequences identified are of the same subunit and many of the 21 putative GluCls belong to other ion channel families such as AMPA, kainate or NMDA receptors. The complement and extent of the locust cysLGIC superfamily, therefore, remains to be clarified. Expansion in the divergent nAChRs mostly accounts for the large cockroach cysLGIC superfamilies. Insects usually possess up to three divergent subunits with *N. vitripennis* being a notable exception in that it has seven subunits although one of these is a pseudogene.[37] *B. germanica* and *P. americana* have the largest sets of divergent nAChR subunits characterized to date, respectively possessing eight and ten subunits (Figures 3 and 4). RT-PCR analysis has shown that most of these are transcribed indicating that they may be functional. The functional roles of divergent nAChRs and the impact of species-specific complements of these nAChR subunits remains to be elucidated.[68] For instance, are the two extra nAChR subunits found in *P. americana* increasing the functional/pharmacological repertoire of receptor function when compared to that of *B. germanica*? It has been shown that the divergent nAChR subunit, β3, from *L. migratoria* is part of a high affinity binding site for imidacloprid[69] highlighting that these subunits may be involved in species-specific receptor subtypes targeted by certain insecticides. It would be of interest to see whether divergent nAChR subunits form part of insecticide targets in cockroaches as well. Also, it would be informative to determine whether expression of divergent nAChR subunits are altered as part of an adaptation

mechanism for cockroaches to manage exposure to insecticides, such as with α1 and α2 subunits when *P. americana* was exposed to imidacloprid.[22]

In addition to possessing distinct groups of divergent nAChR subunits, another difference observed in the cysLGIC superfamilies of *B. germanica* and *P. americana* was alternative splicing of *Rdl*. Two putative alternatives for exon 9 of *Rdl* was detected in the genome of *B. germanica* but not of *P. americana* (Figure 7A). Such alternative splicing of this exon has only been reported once before for *L. striatellus* where four splice variants affected potency of the agonists GABA and β-alanine.[59] Perhaps with the extra exon 9 influencing the action of GABA, the functional capabilities of the GABA receptor is broadened in *B. germanica*.

4.2 Cockroach and aphid cysLGIC gene superfamilies share unusual features

Recently, a report concluded that, as with *A. pisum*, *P. americana* lacks an α5 subunit.[24] We can confirm here that *P. americana*, as well as *B. germanica*, do indeed possess a putative α5 subunit, which cluster together with *T. castaneum* α5 (Figure 4). However, some cockroach cysLGICs do share features with those particular to *A. pisum*. For instance, the unusual alternative splicing of exons 7 and 8 detected in the *A. pisum* α7 nAChR subunit[50] has also been observed for *B. germanica* and *P. americana* α7 theoretically giving rise to eight different variants (Figure 6C). Expression of homomeric *P. americana* α7 has been achieved in *Xenopus laevis* oocytes with limited success.[24] The splice variant presented in that study corresponds to exon7a+exon8a (Figure 6C). It would be of interest to see whether the expression of other α7 splice variants, or combinations of them, result in more robust expression.

Another unusual feature that the cockroach genome shares with that of *A. pisum* is the presence of a second putative GluCl subunit, designated as GluCl2 (Figure 5).[29] It remains to be determined whether GluCl2 functions as a glutamate-gated chloride channel. However, in this regard, it is interesting to note that electrophysiological studies on neurones from *P. americana* demonstrated the co-existence of two pharmacologically distinct glutamate-gated chloride

channels,[19, 70] perhaps as a result of GluCl and GluCl2 acting as independent glutamate receptors. If this is the case, GluCl2, with its low sequence identity to GluCls commonly found in other insects (Table 2), may be a potential target for the design of novel insecticides with higher specificity towards cockroaches and aphids.

Analysis of the *A. pisum* genome revealed the presence of *Rdl2*, *GluCl2* and *pHCl2* genes and the absence of *CG8916*, *LCCH3* and *GRD* orthologues (Figure 5).[29] This led to the suggestion that genes arising from duplication of *Rdl*, *GluCl* and *pHCl* in evolutionary ancient insects would eventually evolve into *CG8916*, *LCCH3* and *GRD* in more recent species.[29] The cockroach cysLGIC gene superfamilies may provide further insights into this evolutionary process considering it shares unusual features with the aphid yet possess orthologues of *8916*, *LCCH3* and *GRD* (Figure 5). Thus, it is tempting to speculate that, for instance, duplication of *GluCl* produced *GluCl2* providing new genetic material for mutation and functional drift resulting in either *CG8916*, *LCCH3* or *GRD*.[71] In cockroaches, *GluCl2* is retained but in more evolved species this gene duplicate is lost. Phylogenetic analysis shows that cockroach and *A. pisum* GluCl2 do not clearly cluster with other subunits, including CG8916, LCCH3 or GRD (Figure 5). However, when this analysis is repeated with sequences consisting only of the N-terminal extracellular domains without the signal peptide, GluCl2 clusters with LCCH3 (Supporting information Figure 1), indicating that duplication of *GluCl* early on in insect evolution may have given rise to the *LCCH3* gene. In accord with this, evolutionary distance estimates show that PameGluCl2 is more closely related to tribolium LCCH3 than to TcasGRD or Tcas8916 (Supporting information Table 3). Analysis of complete cysLGIC gene superfamilies from other species representing different stages of insect evolution should shed further light on the evolutionary fate of *Rdl2*, *GluCl2* and *pHCl2* identified in *A. pisum*.

4.3 Conclusions

The characterization of *B. germanica* and *P. americana* cysLGICs provides further insights into the evolution of this gene superfamily in insects. Also, identification of cockroach cysLGIC subunit genes

will provide an important basis for future studies of these ion channels as well as for the rational design of insecticides that control cockroaches, ideally with less effect on non-target organisms. Such studies can include the heterologous expression of certain subunits to determine their functional/pharmacological properties, which will also provide powerful tools enabling screening for improved insecticides that show higher selectivity towards particular cockroach receptors.[72] In addition, RNA interference, which has been successfully used to study several core nAChR subunits in *P. americana*,[73] may also be applied to the newly identified divergent nAChRs or other cysLGIC subunits to assess their role in neuronal function and response to insecticides.

ACKNOWLEDGEMENTS

This work was supported by INRAE, Inserm, la Région Pays de la Loire and the RFI Objectif Végétal.

CONFLICT OF INTEREST DECLARATION

The authors declare no conflicts of interest.

REFERENCES

1 Jones AK, Genomics, cys-loop ligand-gated ion channels and new targets for the control of insect pests and vectors. *Curr Opin Insect Sci* **12**:30:1-7 (2018).

2. Jones AK and Sattelle DB, Diversity of insect nicotinic acetylcholine receptor subunits. *Adv Exp Med Biol* **683**:25-43 (2010).

3. Buckingham SD, Biggin PC, Sattelle BM, Brown LA and Sattelle DB, Insect GABA receptors: splicing, editing, and targeting by antiparasitics and insecticides. *Mol Pharmacol* **68**:942-951 (2005).

4. Wolstenholme AJ, Glutamate-gated chloride channels. *J Biol Chem* **287**:40232-40238 (2012).

5. Gisselmann G, Pusch H, Hovemann BT and Hatt H, Two cDNAs coding for histamine-gated ion channels in *D. melanogaster*. *Nat Neurosci* **5**:11-12 (2002).

6. Zheng Y, Hirschberg B, Yuan J, Wang AP, Hunt DC, Ludmerer SW, Schmatz DM and Cully DF, Identification of two novel *Drosophila melanogaster* histamine-gated chloride channel subunits expressed in the eye. *J Biol Chem* **277**:2000-2005 (2002).

7. Redhai S, Pilgrim C, Gaspar P, Giesen L, Lopes T, Riabinina O, Grenier T, Milona A, Chanana B, Swadling JB, Wang Y-F, Dahalan F, Yuan M, Wilsch-Brauninger M, Lin W-H, Dennison N, Capriotti P, Lawniczak MKN, Baines RA, Warnecke T, Windbichler N, Leulier F, Bellono NW and Miguel-Aliaga I,

366 An intestinal zinc sensor regulates food intake and developmental growth. *Nature* **580**:263-268
367 (2020).
368 8. Sparks TC, Crossthwaite AJ, Nauen R, Banba S, Cordova D, Earley F, Ebbinghaus-Kintscher U,
369 Fujioka S, Hirao A, Karmon D, Kennedy R, Nakao T, Popham HJR, Salgado V, Watson GB, Wedel BJ
370 and Wessels FJ, Insecticides, biologics and nematicides: Updates to IRAC's mode of action
371 classification - a tool for resistance management. *Pestic Biochem Physiol* **167**:104587 (2020).
372 9. Ihara M, Buckingham SD, Matsuda K and Sattelle DB, Modes of Action, Resistance and
373 Toxicity of Insecticides Targeting Nicotinic Acetylcholine Receptors. *Curr Med Chem* **24**:2925-2934
374 (2017).
375 10. Lorsbach BA, Sparks TC, Cicchillo RM, Garizi NV, Hahn DR and Meyer KG, Natural products: a
376 strategic lead generation approach in crop protection discovery. *Pest Manag Sci* **75**:2301-2309
377 (2019).
378 11. Buckingham SD, Ihara M, Sattelle DB and Matsuda K, Mechanisms of Action, Resistance and
379 Toxicity of Insecticides Targeting GABA Receptors. *Curr Med Chem* **24**:2935-2945 (2017).
380 doi:10.2174/0929867324666170613075736
381 12. Li S, Zhu S, Jia Q, Yuan D, Ren C, Li K, Liu S, Cui Y, Zhao H, Cao Y, Fang G, Li D, Zhao X, Zhang J,
382 Yue Q, Fan Y, Yu X, Feng Q and Zhan S, The genomic and functional landscapes of developmental
383 plasticity in the American cockroach. *Nat Commun* **9**:1008 (2018).
384 13. Zhu F, Lavine L, O'Neal S, Lavine M, Foss C and Walsh D, Insecticide Resistance and
385 Management Strategies in Urban Ecosystems. *Insects* **7**(1) (2016).
386 14. Wang C, Scharf ME and Bennett GW, Genetic basis for resistance to gel baits, fipronil, and
387 sugar-based attractants in German cockroaches (Dictyoptera: blattellidae). *J Econ Entomol* **99**:1761-
388 1767 (2006).
389 15. Ko AE, Bieman DN, Schal C and Silverman J, Insecticide resistance and diminished secondary
390 kill performance of bait formulations against German cockroaches (Dictyoptera: Blattellidae). *Pest*
391 *Manag Sci* **72**:1778-1784 (2016).
392 16. Hu IH, Chen SM, Lee CY and Neoh KB, Insecticide Resistance, and Its Effects on Bait
393 Performance in Field-Collected German Cockroaches (Blattodea: Ectobiidae) From Taiwan. *J Econ*
394 *Entomol* **113**:1389-1398 (2020).
395 17. Peterson RT, Nass R, Boyd WA, Freedman JH, Dong K and Narahashi T, Use of non-
396 mammalian alternative models for neurotoxicological study. *Neurotoxicology* **29**:546-555 (2008).
397 18. Nomura K, Kawase A, Ozoe Y and Banba S, Further characterization of distinct high-affinity
398 binding sites for dinotefuran in the abdominal nerve cord of the American cockroach *Periplaneta*
399 *americana* (Blattodea). *Pestic Biochem Physiol* **165**:104554 (2020).

19. Zhao X, Salgado VL, Yeh JZ and Narahashi T, Kinetic and pharmacological characterization of desensitizing and non-desensitizing glutamate-gated chloride channels in cockroach neurons. *Neurotoxicology* **25**:967-980 (2004).
20. Taillebois E, Cartereau A, Jones AK and Thany SH, Neonicotinoid insecticides mode of action on insect nicotinic acetylcholine receptors using binding studies. *Pestic Biochem Physiol* **151**:59-66 (2018).
21. Onozaki Y, Horikoshi R, Ohno I, Kitsuda S, Durkin KA, Suzuki T, Asahara C, Hiroki N, Komabashiri R, Shimizu R, Furutani S, Ihara M, Matsuda K, Mitomi M, Kagabu S, Uomoto K and Tomizawa M, Flupyrimin: A Novel Insecticide Acting at the Nicotinic Acetylcholine Receptors. *J Agric Food Chem* **65**:7865-7873 (2017).
22. Benzidane Y, Goven D, Abd-Ella AA, Deshayes C, Lapied B and Raymond V, Subchronic exposure to sublethal dose of imidacloprid changes electrophysiological properties and expression pattern of nicotinic acetylcholine receptor subtypes in insect neurosecretory cells. *Neurotoxicology* **62**:239-247 (2017).
23. Salgado VL, Antagonist pharmacology of desensitizing and non-desensitizing nicotinic acetylcholine receptors in cockroach neurons. *Neurotoxicology* **56**:188-195 (2016).
24. Cartereau A, Taillebois E, Selvam B, Martin C, Graton J, Le Questel J-V and Thany SH, Cloning and Expression of Cockroach α7 Nicotinic Acetylcholine Receptor Subunit. *Front Physiol* **11**:418 (2020).
25. Taillebois E and Thany SH, Characterization of nicotine acetylcholine receptor subunits in the cockroach *Periplaneta americana* mushroom bodies reveals a strong expression of β1 subunit: involvement in nicotine-induced currents. *Arch Insect Biochem Physiol* **93**:40-54 (2016).
26. Kaku K and Matsumura F, Identification of the site of mutation within the M2 region of the GABA receptor of the cyclodiene-resistant German cockroach. *Comp Biochem Physiol C Pharmacol Toxicol Endocrinol* **108**:367-76 (1994).
27. Harrison MC, Jongepier E, Robertson HM, Arning N, Bitard-Feildel T, Chao H, Childers CP, Dinh H, Doddapaneni H, Dugan S, Gowin J, Greiner C, Han Y, Hu H, Hughes DST, Huylmans A-K, Kemena C, Kremer LPM, L Lee S, Lopez-Ezquerra A, Mallet L, Monroy-Kuhn JM, Moser A, Murali SC, Muzny DM, Otani S, Piulachs M-D, Poelchau M, Qu J, Schaub F, Wada-Katsumata A, Worley KC, Xie Q, Ylla G, Poulsen M, Gibbs RA, Schal C, Richards S, Belles X, Korb J and Bornberg-Bauer E, Hemimetabolous genomes reveal molecular basis of termite eusociality. *Nat Ecol Evol* **2**:557-566 (2018).
28. Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ, Basic local alignment search tool. *J Mol Biol* **215**:403-410 (1990).

29. Del Villar SG and Jones AK, Cloning and functional characterisation of the duplicated RDL subunits from the pea aphid, *Acyrthosiphon pisum*. *Int J Mol Sci* **19**(8) (2018).
30. Jones AK and Sattelle DB, The cys-loop ligand-gated ion channel gene superfamily of the red flour beetle, *Tribolium castaneum*. *BMC Genomics* **8**:327 (2007).
31. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F and Higgins DG, The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**:4876-4882 (1997).
32. Bendtsen JD, Nielsen H, von Heijne G and Brunak S, Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* **340**:783-795 (2004).
33. Hulo N, Bairoch A, Bulliard V, Cerutti L, De Castro E, Langendijk-Genevaux PS, Pagni M and Sigrist CJ, The PROSITE database. *Nucleic Acids Res* **34**(Database issue):D227-230 (2006).
34. Kumar S, Stecher G, Li M, Knyaz C and Tamura K, MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* **35**:1547-1549 (2018).
35. Dent JA, Evidence for a diverse Cys-loop ligand-gated ion channel superfamily in early bilateria. *J Mol Evol* **62**:523-535 (2006).
36. Jones AK and Sattelle DB, The cys-loop ligand-gated ion channel superfamily of the honeybee, *Apis mellifera*. *Invert Neurosci* **6**:123-132 (2006).
37. Jones AK, Bera AN, Lees K and Sattelle DB, The cys-loop ligand-gated ion channel gene superfamily of the parasitoid wasp, *Nasonia vitripennis*. *Heredity (Edinb)* **104**:247-259 (2010).
38. Matthews BJ, Dudchenko O, Kingan SB, Koren S, Antoshechkin I, Crawford JE, Glassford WJ, Herre M, Redmond SN, Rose NH, Weedall GD, Wu Y, Batra SS, Brito-Sierra CA, Buckingham SD, Campbell CL, Chan S, Cox E, Evans BR, Fansiri T, Filipović I, Fontaine A, Gloria-Soria A, Hall R, Joardar VS, Jones AK, Kay RGG, Kodali VK, Lee J, Lycett GJ, Mitchell SN, Muehling J, Murphy MR, Omer AD, Partridge FA, Peluso P, Aiden AP, Ramasamy V, Rašić G, Roy S, Saavedra-Rodriguez K, Sharan S, Sharma A, Smith ML, Turner J, Weakley AM, Zhao Z, Akbari OS, Black WC 4th, Cao H, Darby AC, Hill CA, Johnston JS, Murphy TD, Raikhel AS, Sattelle DB, Sharakhov IV, White BJ, Zhao L, Aiden EL, Mann RS, Lambrechts L, Powell JR, Sharakhova MV, Tu Z, Robertson HM, McBride CS, Hastie AR, Korlach J, Neafsey DE, Phillippy AM and Vosshall LB, Improved reference genome of *Aedes aegypti* informs arbovirus vector control. *Nature* **563**:501-507 (2018).
39. Rinkevich FD and Bourgeois L, In silico identification and assessment of insecticide target sites in the genome of the small hive beetle, *Aethina tumida*. *BMC Genomics* **21**:154 (2020).
40. Sine SM and Engel AG, Recent advances in Cys-loop receptor structure and function. *Nature* **440**:448-455 (2006).

41. Corringer PJ, Le Novère N and Changeux JP, Nicotinic receptors at the amino acid level. *Annu Rev Pharmacol Toxicol* **40**:431-458 (2000).
42. Gehle VM, Walcott EC, Nishizaki T and Sumikawa K, N-glycosylation at the conserved sites ensures the expression of properly folded functional ACh receptors. *Brain Res Mol Brain Res* **45**:219-229 (1997).
43. Nishizaki T, N-glycosylation sites on the nicotinic ACh receptor subunits regulate receptor channel desensitization and conductance. *Brain Res Mol Brain Res* **114**:172-176 (2003).
44. Hopfield JF, Tank DW, Greengard P and Huganir RL, Functional modulation of the nicotinic acetylcholine receptor by tyrosine phosphorylation. *Nature* **336**:677-680 (1988).
45. Borges LS and Ferns M, Agrin-induced phosphorylation of the acetylcholine receptor regulates cytoskeletal anchoring and clustering. *J Cell Biol* **153**:1-12 (2001).
46. Talwar S and Lynch JW, Phosphorylation mediated structural and functional changes in pentameric ligand-gated ion channels: implications for drug discovery. *Int J Biochem Cell Biol* **53**:218-223 (2014).
47. Sattelle DB, Jones AK, Sattelle BM, Matsuda K, Reenan R and Biggin PC, Edit, cut and paste in the nicotinic acetylcholine receptor gene family of *Drosophila melanogaster*. *Bioessays* **27**:366-376 (2005).
48. Jones AK, Grauso M and Sattelle DB, The nicotinic acetylcholine receptor gene family of the malaria mosquito, *Anopheles gambiae*. *Genomics* **85**:176-187 (2005). doi:10.1016/j.ygeno.2004.09.001
49. Jones AK, Raymond-Delpech V, Thany SH, Gauthier M and Sattelle DB, The nicotinic acetylcholine receptor gene family of the honey bee, *Apis mellifera*. *Genome Res* **16**:1422-1430 (2006).
50. Dale RP, Jones AK, Tamborindeguy C, Davies TGE, Amey JS, Williamson S, Wolstenholme A, Field LM, Williamson MS, Walsh TK and Sattelle DB, Identification of ion channel genes in the *Acyrthosiphon pisum* genome. *Insect Mol Biol* **2**:141-153 (2010).
51. Shao YM, Dong K and Zhang CX, The nicotinic acetylcholine receptor gene family of the silkworm, *Bombyx mori*. *BMC Genomics* **8**:324 (2007).
52. Shimomura M, Yokota M, Matsuda K, Sattelle DB and Komai K, Roles of loop C and the loop B-C interval of the nicotinic receptor alpha subunit in its selective interactions with imidacloprid in insects. *Neurosci Lett* **363**:195-198 (2004).
53. Galzi JL, Bertrand D, Devillers-Thiéry A, Revah F, Bertrand S and Changeux JP, Functional significance of aromatic amino acids from three peptide loops of the alpha 7 neuronal nicotinic receptor site investigated by site-directed mutagenesis. *FEBS Lett* **294**:198-202 (1991).

501 54. Jensen ML, Schousboe A and Ahring PK, Charge selectivity of the Cys-loop family of ligand-
502 gated ion channels. *J Neurochem* **92**:217-225 (2005).
503 55. Lansdell SJ and Millar NS, Cloning and heterologous expression of Dalpha4, a Drosophila
504 neuronal nicotinic acetylcholine receptor subunit: identification of an alternative exon influencing
505 the efficiency of subunit assembly. *Neuropharmacology* **39**:2604-2614 (2000).
506 56. Henry C, Cens T, Charnet P, Cohen-Solal C, Collet C, van-Dijk J, Guiramand J, de Jésus-Ferreira
507 M-C, Menard C, Mokrane N, Roussel J, Thibault J-B, Vignes M and Rousset M, Heterogeneous
508 expression of GABA receptor-like subunits LCCH3 and GRD reveals functional diversity of GABA
509 receptors in the honeybee *Apis mellifera*. *Br J Pharmacol* May 2020;doi:10.1111/bph.15135
510 57. Gisselmann G, Plonka J, Pusch H and Hatt H, *Drosophila melanogaster* GRD and LCCH3
511 subunits form heteromultimeric GABA-gated cation channels. *Br J Pharmacol* **142**:409-413 (2004).
512 58. ffrench-Constant RH and Rocheleau TA, Drosophila gamma-aminobutyric acid receptor gene
513 Rdl shows extensive alternative splicing. *J Neurochem* **60**:2323-2326 (1993).
514 59. Jiang J, Huang LX, Chen F, Sheng C-W, Huang Q-T, Han Z-J and Zhao C-Q, Novel alternative
515 splicing of GABA receptor RDL exon 9 from *Laodelphax striatellus* modulates agonist potency. *Insect*
516 *Sci* Apr 2020;doi:10.1111/1744-7917.12789
517 60. Taylor-Wells J, Hawkins J, Colombo C, Bermudez I and Jones AK, Cloning and functional
518 expression of intracellular loop variants of the honey bee (*Apis mellifera*) RDL GABA receptor.
519 *Neurotoxicology* **60**:207-213 (2017).
520 61. Jiang F, Zhang Y, Sun H, Meng X, Bao H, Fang J and Liu Z, Identification of polymorphisms in
521 *Cyrtorhinus lividipennis* RDL subunit contributing to fipronil sensitivity. *Pestic Biochem Physiol*
522 **117**:62-67 (2015).
523 62. Akashi HD, Chen PJ, Akiyama T, Terai Y, Wakakuwa M, Takayama Y, Tominaga M and
524 Arikawa K, Physiological responses of ionotropic histamine receptors, PxHCLA and PxHCLB, to
525 neurotransmitter candidates in a butterfly, *Papilio xuthus*. *J Exp Biol* **221**(Pt 21) (2018)
526 63. Wang JD, Chen LF, Lin DJ, Zhang J-S, Zhao J-H, Xiao D, Wang R, Wang R and Gao S-J,
527 Molecular cloning, characterization and functional analysis of GluCl from the oriental armyworm,
528 *Mythimna separata* Walker. *Pestic Biochem Physiol* **156**:56-62 (2019).
529 doi:10.1016/j.pestbp.2019.02.004
530 64. Schnizler K, Saeger B, Pfeffer C, Gerbaulet A, Ebbinghaus-Kintscher U, Methfessel C, Franken
531 E-M, Raming K, Wetzel CH, Saras A, Pusch H, Hatt H and Gisselmann G, A novel chloride channel in
532 *Drosophila melanogaster* is inhibited by protons. *J Biol Chem* **280**:16254-16262 (2005).

65. Frenkel L, Muraro NI, Beltrán González AN, Marcora MS, Bernabó G, Hermann-Luibl C, Romero JI, Helfrich-Förster C, Castaño EM, Marino-Busjle C, Calvo DJ and Ceriani MF, Organization of Circadian Behavior Relies on Glycinergic Transmission. *Cell Rep* **19**:72-85 (2017).

66. Littleton JT and Ganetzky B, Ion channels and synaptic organization: analysis of the Drosophila genome. *Neuron* **26**:35-43 (2000).

67. Wang X, Meng X, Liu C, Gao H, Zhang Y and Liu Z, Cys-loop ligand-gated ion channel gene discovery in the *Locusta migratoria manilensis* through the neuron transcriptome. *Gene* **561**:276-282 (2015).

68. Jones AK, Buckingham SD, Brown LA and Sattelle DB, Alternative splicing of the *Anopheles gambiae* nicotinic acetylcholine receptor, Agamalphabeta9, generates both alpha and beta subunits. *Invert Neurosci* **9**:77-84 (2009).

69. Bao H, Liu Y, Zhang Y and Liu Z, Two distinctive β subunits are separately involved in two binding sites of imidacloprid with different affinities in *Locusta migratoria manilensis*. *Pestic Biochem Physiol* **140**:36-41 (2017).

70. Raymond V, Sattelle DB and Lapied B, Co-existence in DUM neurones of two GluCl channels that differ in their picrotoxin sensitivity. *Neuroreport* **11**:2695-2701 (2000).

71. Magadum S, Banerjee U, Murugan P, Gangapur D and Ravikesavan R, Gene duplication as a major force in evolution. *J Genet* **92**:155-161 (2013).

72. Buckingham SD, Pym L and Sattelle DB, Oocytes as an expression system for studying receptor/channel targets of drugs and pesticides. *Methods Mol Biol* **322**:331-345 (2006).

73. Sun H, Liu Y, Li J, Cang X, Bao H and Liu Z, The potential subunits involved in two subtypes of α-Bgt-resistant nAChRs in cockroach dorsal unpaired median (DUM) neurons. *Insect Biochem Mol Biol* **81**:32-40 (2017).

TABLES

Table 1. Percentage identity/similarity between *P. americana* and *T. castaneum* nAChR subunit protein sequences.

Subunit	Tcasα1	Tcasα2	Tcasα3	Tcasα4	Tcasα5	Tcasα6	Tcasα7	Tcasα8	Tcasβ1	Tcasα9	Tcasα10	Tcasα11
Pameα1	**75/83**	52/67	55/68	52/63	27/44	33/48	33/50	53/66	38/55	15/32	18/35	53/68
Pameα2	52/67	**79/86**	50/66	47/62	26/46	34/52	35/51	49/65	37/54	14/32	19/36	50/66
Pameα3	54/67	49/64	**84/90**	67/76	27/44	33/49	32/48	53/65	39/54	14/31	19/34	54/65
Pameα4	51/62	46/61	67/76	**85/89**	26/43	33/49	32/49	51/65	38/54	14/31	18/33	52/66
Pameα5	29/49	27/48	28/46	27/46	**65/75**	34/54	31/50	28/48	30/50	16/34	22/42	28/50
Pameα6	32/49	33/50	32/48	32/40	33/52	**82/90**	65/76	33/50	32/52	16/35	22/41	33/52
Pameα7	34/50	35/51	34/50	33/49	31/51	67/78	**77/83**	34/50	32/51	15/32	21/40	35/52
Pameα8	54/69	50/67	56/68	54/67	27/45	34/52	34/51	**70/79**	38/56	14/34	19/36	72/81
Pameβ1	38/54	38/54	39/55	38/54	27/45	32/52	32/51	38/54	**84/88**	14/32	19/36	39/56
Pameα9	20/36	19/36	20/36	20/35	21/42	22/41	20/38	20/38	20/38	20/39	**50/69**	20/39
Pameβ2	11/25	11/24	11/25	11/24	10/26	10/26	10/25	12/27	10/25	13/28	16/31	12/26
Pameβ3	13/29	13/29	12/28	13/28	16/32	16/32	13/30	13/31	13/30	23/45	18/37	13/31
Pameβ4	13/30	13/30	14/29	13/28	16/32	16/32	14/30	14/31	13/28	23/45	18/37	14/31
Pameβ5	13/30	14/30	14/29	14/30	15/34	14/31	14/29	14/32	13/30	24/49	20/37	14/32
Pameβ6	13/29	14/30	14/28	14/28	16/34	14/31	13/30	15/33	14/31	27/51	20/38	14/32
Pameβ7	13/29	12/26	13/27	12/28	17/34	15/29	14/29	14/32	13/29	24/44	20/38	14/31
Pameβ8	12/29	12/29	12/27	12/27	13/32	13/30	13/27	12/32	13/30	20/46	19/36	12/31
Pameβ9	14/29	12/28	13/28	12/27	18/35	14/30	13/28	14/31	14/29	25/44	17/39	14/31
Pameβ10	14/30	14/31	14/30	14/30	19/37	18/34	17/33	16/36	15/32	28/50	18/40	16/35

Proposed orthologues are shown in bold.

Table 2. Percentage identity/similarity between *P. americana* and *T. castaneum* non-nAChR subunit protein sequences.

Subunit	TcasRDL	TcasGRD	TcasLCCH3	Tcas8916	TcasGluCl	TcasHisCl1	TcasHisCl2	TcaspHCl	Tcas12344	Tcas6927	Tcas7589	Tcas11340
PameRDL	**83/86**	28/42	32/50	25/41	28/42	23/40	24/39	19/38	23/40	18/33	17/33	17/34
PameGRD	29/42	**67/74**	28/44	37/51	23/37	21/36	20/35	17/31	19/33	18/30	16/29	15/31
PameLCCH3	32/49	26/42	**73/83**	26/42	25/42	23/39	24/38	17/35	21/38	18/37	18/35	17/35
Pame8916	24/37	36/49	25/39	**64/75**	20/34	19/31	19/33	16/29	17/30	16/29	15/29	15/30
Pame8916_2	21/35	25/35	20/33	24/37	19/33	16/30	16/27	15/28	15/28	14/26	13/26	14/28
PameGluCl	28/43	22/38	24/43	21/35	**83/90**	27/44	26/44	26/46	21/39	21/38	20/35	21/39
PameGluCl2	25/41	21/38	26/42	20/36	33/50	27/45	28/46	26/44	22/39	22/39	21/36	22/40
PameHisCl1	22/39	20/36	23/40	19/33	28/45	**81/87**	53/67	20/41	28/49	18/34	17/34	19/38
PameHisCl2	23/38	21/37	24/39	19/33	27/46	54/69	**79/85**	20/38	28/50	17/34	17/34	18/38
PamepHCl	19/37	16/30	18/35	16/30	26/46	21/39	20/39	**82/87**	18/36	20/35	18/32	20/37
Pame12344	22/39	20/35	22/38	18/33	22/40	27/49	28/47	20/38	**56/74**	19/36	18/34	20/38
PameCLGC1	20/37	18/31	20/38	16/32	23/40	19/35	20/36	20/36	19/35	34/51	34/51	30/51
PameCLGC2	19/37	18/33	20/38	17/33	19/38	15/34	17/35	20/36	16/35	34/53	32/50	29/48

Proposed orthologues are shown in bold.

FIGURES

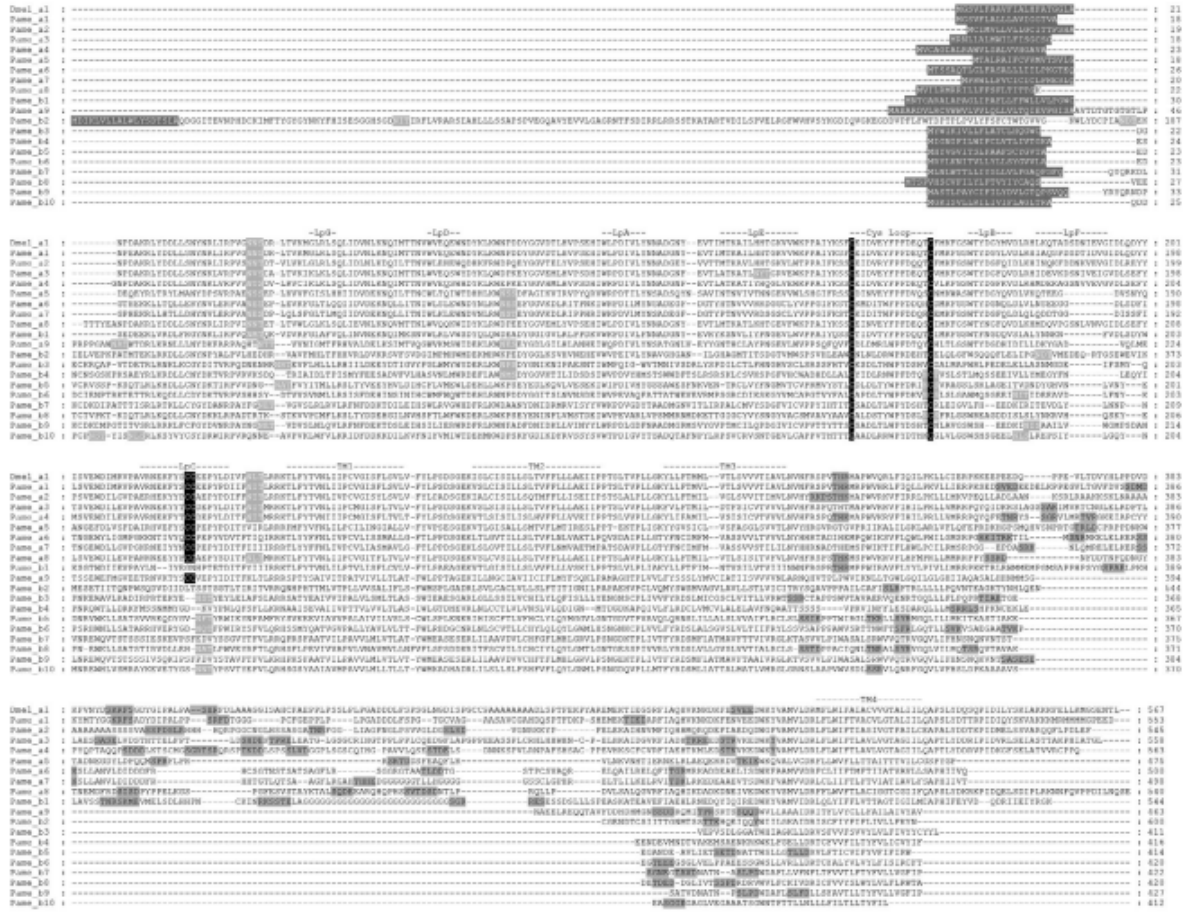

Figure 1. Protein sequence alignment of *P. americana* nAChR subunits. Dα1 of *D. melanogaster* (Accession Number CAA30172) is included for comparison. N-terminal signal leader peptides are highlighted by dark grey shading with white text and putative N-glycosylation sites in the N-terminal extracellular domain are shown in light grey shading with white text. Loops implicated in ligand binding (LpA-F) as well as transmembrane regions (TM) are indicated. The two cysteines forming the cys-loop as well as the two adjacent cysteines defining α subunits in loop C are highlighted in black shading. Casein kinase II, protein kinase C, tyrosine kinase and cAMP- and cGMP-dependent protein kinase phosphorylation sites in the intracellular loop between TM3 and TM4 are highlighted by grey shading.

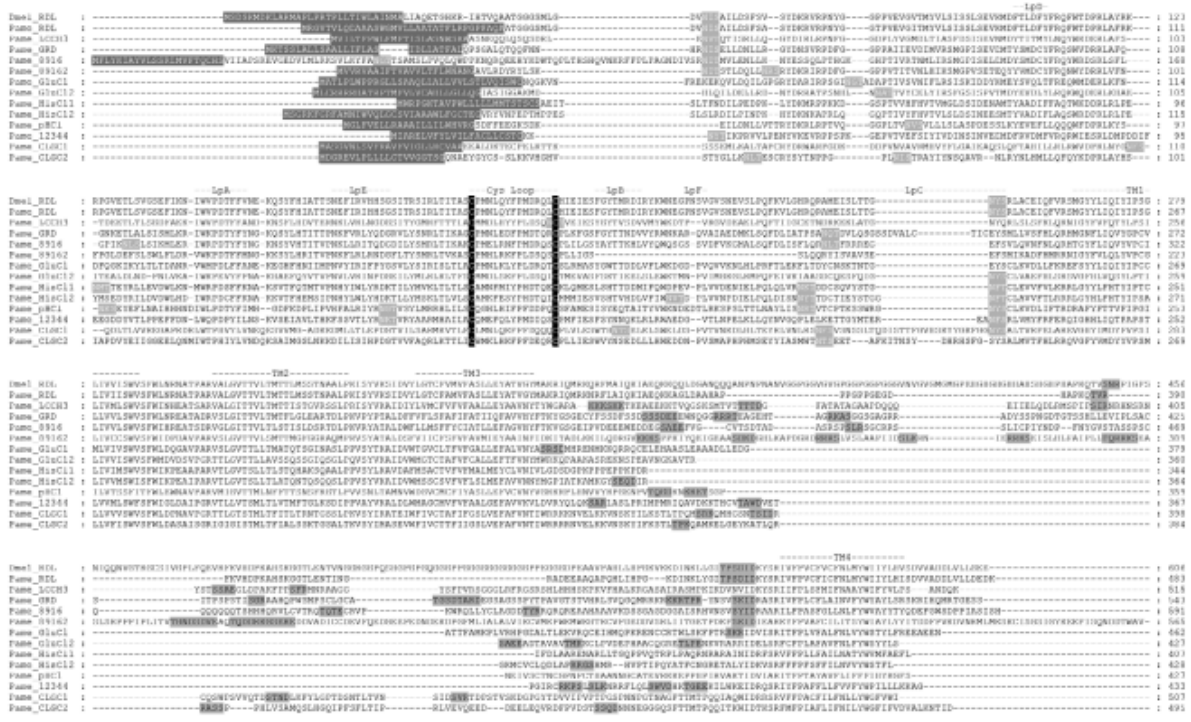

Figure 2. Protein sequence alignment of *P. americana* cysLGIC subunits other than nAChRs. RDL of *D. melanogaster* (Accession Number AAA28556) is included for comparison. N-terminal signal leader peptides are highlighted by dark grey shading with white text and putative N-glycosylation sites in the N-terminal extracellular domain are shown in light grey shading with white text. Loops implicated in ligand binding (LpA-F) as well as transmembrane regions (TM) are indicated. The two cysteines forming the cys-loop are highlighted in black shading. Casein kinase II, protein kinase C and cAMP- and cGMP-dependent protein kinase phosphorylation sites in the intracellular loop between TM3 and TM4 are highlighted by dark grey shading.

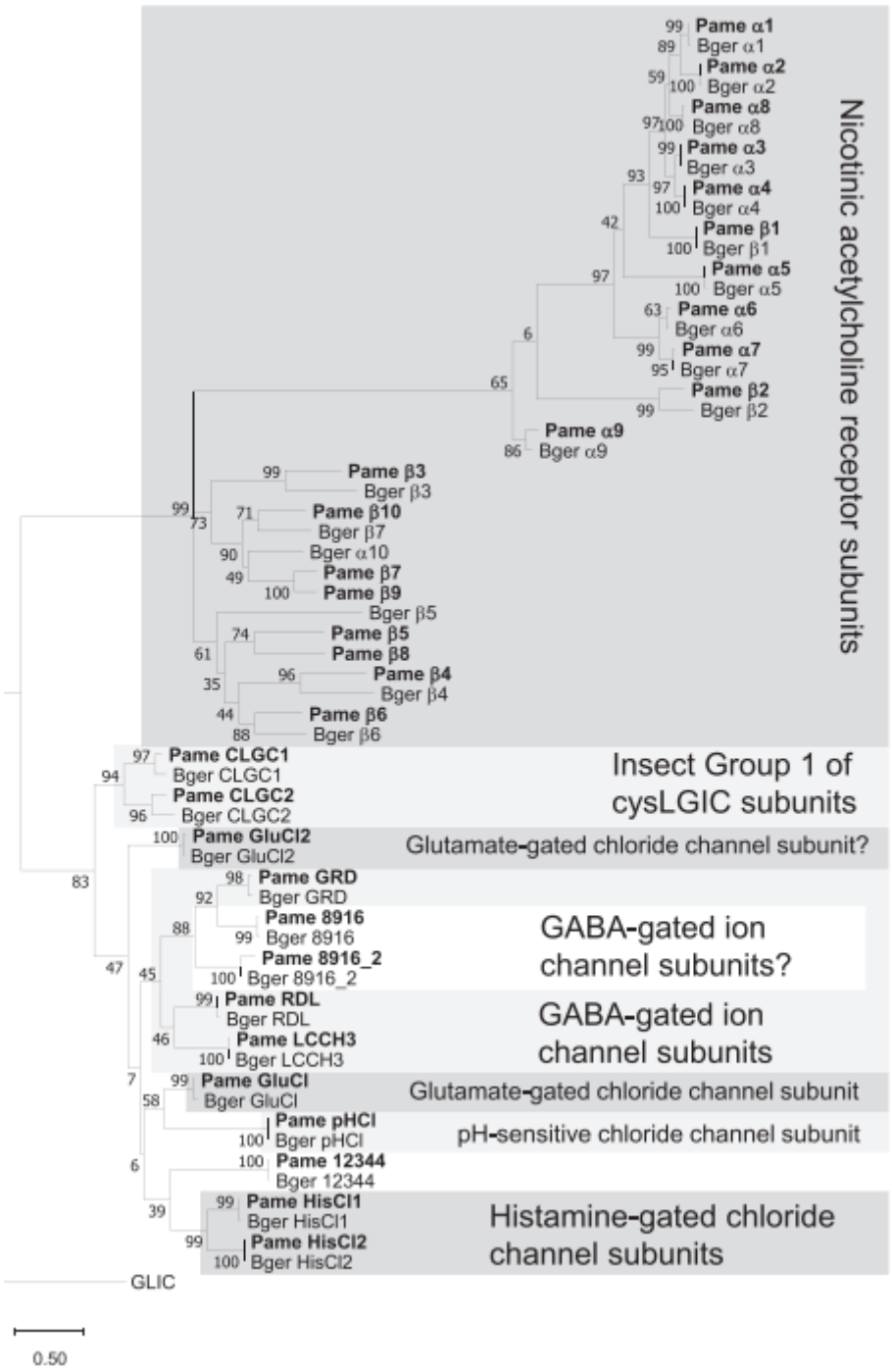

Figure 3. Tree showing relationships of the complete complement of *B. germanica* and *P. americana* cysLGIC subunit protein sequences. *P. americana* subunits are highlighted in bold. Numbers next to branches signify bootstrapping 1000 times represented as a percentage of trees in which the associated taxa clustered together. GLIC (Accession number 2XQ3_A), from *Gloebacter violaceus*, a bacterial ancestor of cysLGICs, was used as an outgroup. The tree is drawn to scale and the scale bar represents substitutions per site.

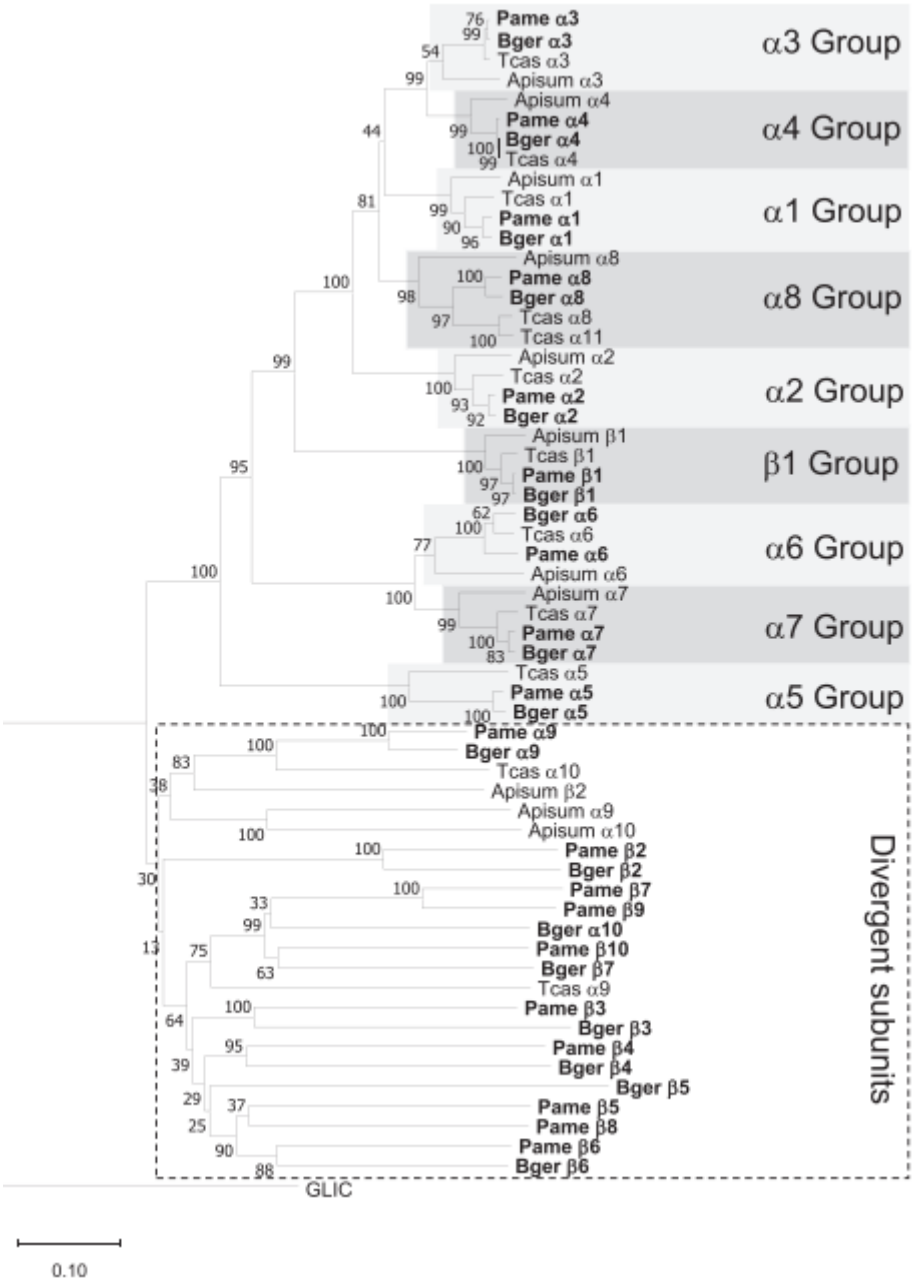

Figure 4. Tree showing relationships of *B. germanica*, *P. americana*, *A. pisum* and *T. castaneum* nAChR subunit protein sequences. Cockroach subunits are highlighted in bold. Numbers next to branches signify bootstrapping 1000 times represented as a percentage of trees in which the associated taxa clustered together. GLIC (Accession number 2XQ3_A), from *Gloebacter violaceus*, a bacterial ancestor of cysLGICs, was used as an outgroup. The tree is drawn to scale and the scale bar represents substitutions per site.

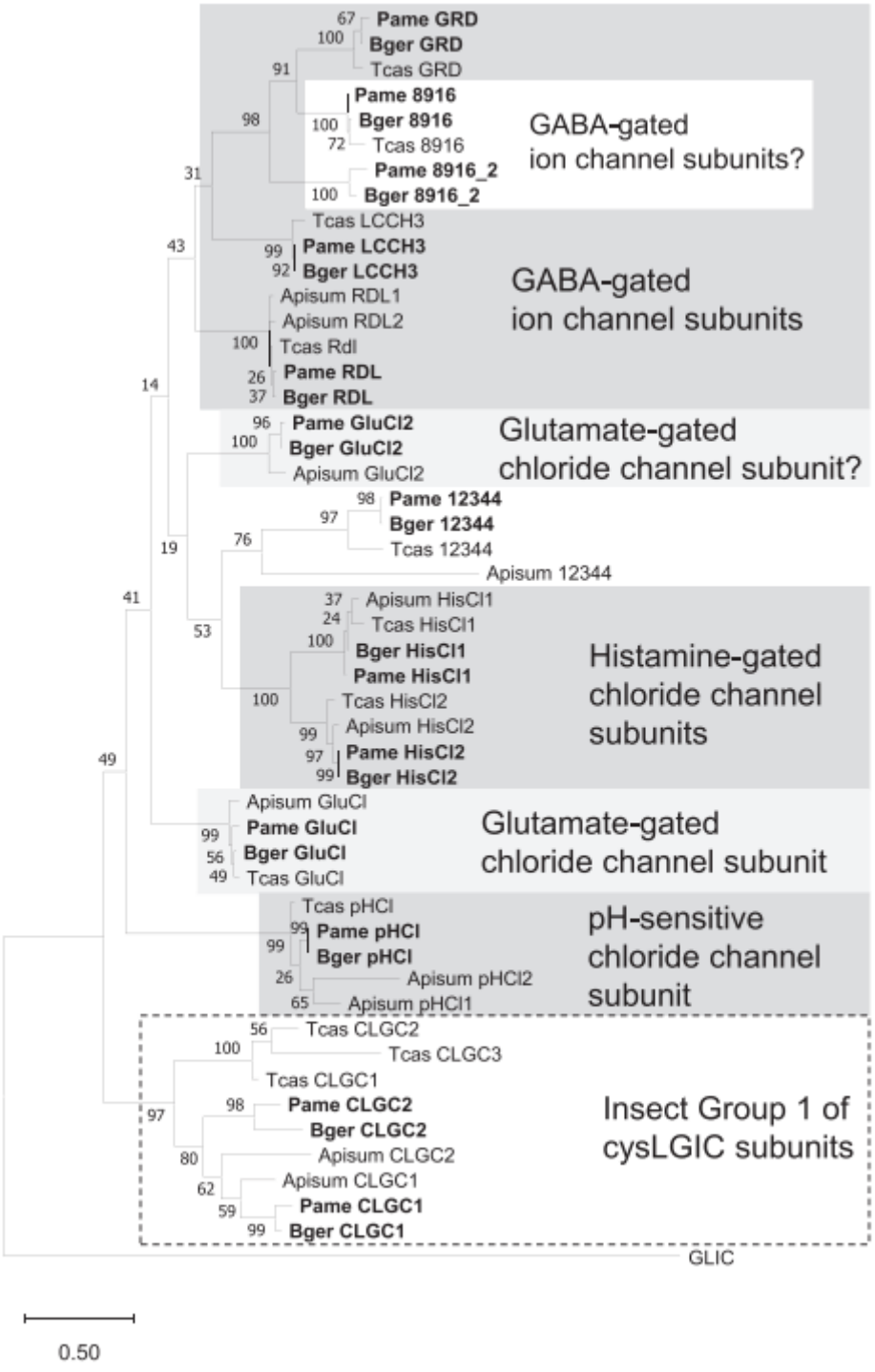

Figure 5. Tree showing relationships of *B. germanica*, *P. americana*, *A. pisum* and *T. castaneum* non-AChR subunit protein sequences. Cockroach subunits are highlighted in bold. Numbers next to branches signify bootstrapping 1000 times represented as a percentage of trees in which the associated taxa clustered together. GLIC (Accession number 2XQ3_A), from *Gloebacter violaceus*, a

bacterial ancestor of cysLGICs, was used as an outgroup. The tree is drawn to scale and the scale bar represents substitutions per site.

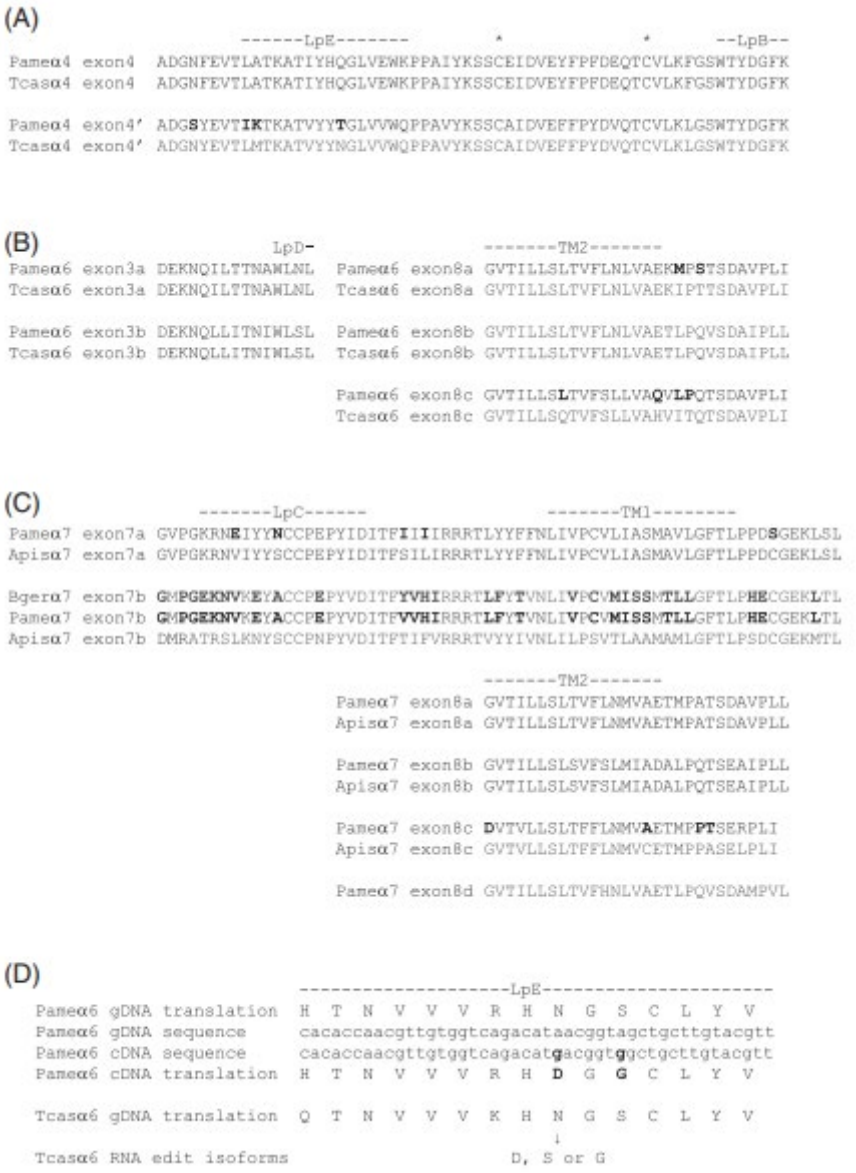

Figure 6. Alternative splicing of exons and RNA A-to-I editing in *B. germanica* and *P. americana* nicotinic acetylcholine receptor subunits. Equivalent alternate exons of *P. americana* and *T. castaneum* cysLGIC subunits are aligned. Unless shown, *B. germanica* exon sequences are identical to those of *P. americana*. (A) Exon 4 splice variants in Pameα4 and Tcasα4. The cysteine residues forming the cys-loop are marked by asterisks. (B) Splice variants of exons 3 and 8 in Pameα6 and Tcasα6. The second transmembrane region (TM2) is indicated. (C) Splice variants of exons 7 and 8 in

Bgerα7, Pameα7 and *A. pisum* α7. (D) RNA A-to-I editing of Pameα6 detected by RT-PCR. Genomic DNA (gDNA, sequence available at NCBI) as well as cDNA sequences encoding for LoopE are shown. Residues altered by putative A-to-I editing are highlighted in bold. The amino acid residue affected by RNA editing in *T. castaneum* α6 is also shown. Throughout the figure, cockroach residues that differ from those of the orthologous *T. castaneum* or *A. pisum* exon are highlighted in bold and loops B to E, which contribute to ligand binding, are indicated.

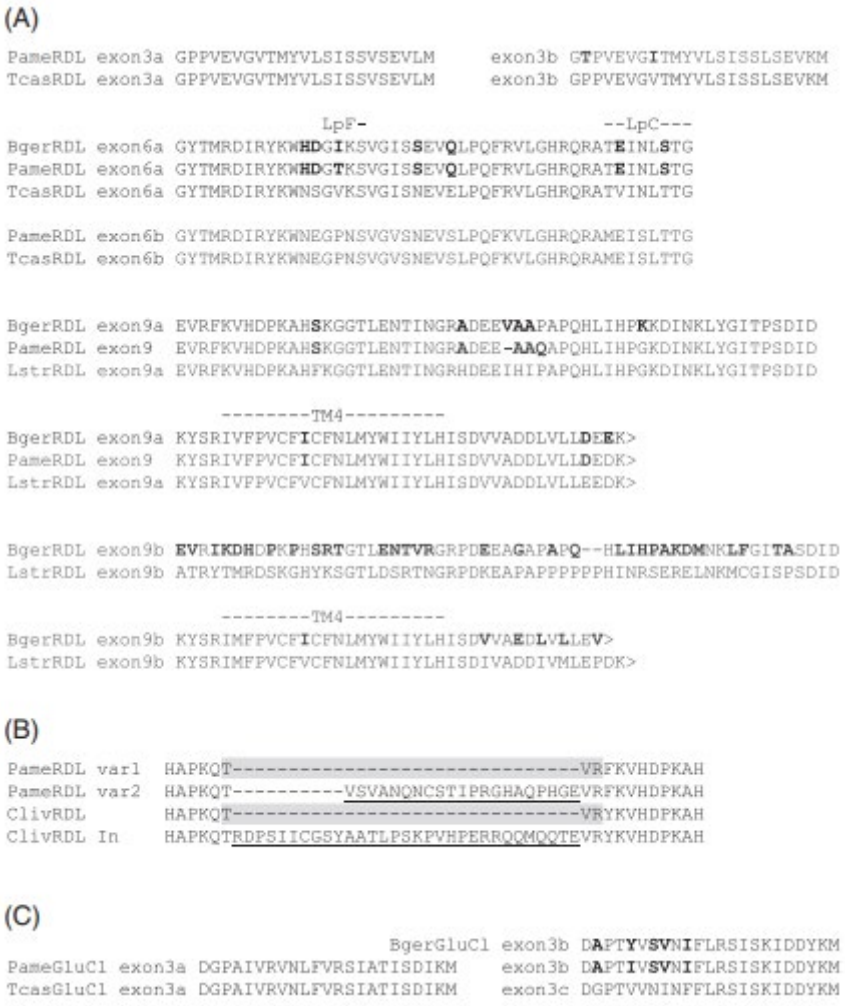

Figure 7. Alternative splicing of exons in *B. germanica* and *P. americana* RDL and GluCl subunits. (A) Splice variants of exons 3 and 6 of *B. germanica*, *P. americana* and *T. castaneum* RDLs as well as variants of exon 9 of *B. germanica* and *Laodelphax striatellus* RDLs. The fourth transmembrane region (TM4) is indicated and '>' denotes a stop codon. (B) Variants of *P. americana* and *C. lividipennis* RDLs with intracellular loops between TM3 and TM4 of varying lengths arising from

differential use of splice sites. Potential protein kinase C phosphorylation sites are highlighted in grey shading. (C) Splice variants of exon 3 of *B. germanica*, *P. americana* and *T. castaneum* GluCls. Throughout the figure, cockroach residues that differ from those of the orthologous *T. castaneum* or *L. striatellus* exon are highlighted in bold and loops C to F, which contribute to ligand binding, are indicated.