Identification and comparative analysis of accessory gland proteins in Orthoptera

W. Evan Braswell, José A. Andrés, Luana S. Maroja, Richard G. Harrison, Daniel J. Howard, and Willie J. Swanson

Abstract: Accessory reproductive gland proteins (Acps) in *Drosophila* evolve quickly and appear to play an important role in ensuring the fertilization success of males. Moreover, Acps are thought to be involved in establishing barriers to fertilization between closely related species. While accessory glands are known to occur in the males of many insect groups, the proteins that are passed on to females by males during mating have not been well characterized outside of *Drosophila*. To gain a better understanding of these proteins, we characterized ESTs from the accessory glands of two cricket species, *Allonemobius fasciatus* and *Gryllus firmus*. Using an expressed sequence tag (EST) approach, followed by bioinformatic and evolutionary analyses, we found that many proteins are secreted and, therefore, available for transfer to the female during mating. Further, we found that most ESTs are novel, showing little sequence similarity between taxa. Evolutionary analyses suggest that cricket proteins are subject to diversifying selection and indicate that *Allonemobius* is much less polymorphic than *Gryllus*. Despite rapid nucleotide sequence divergence, there appears to be functional conservation of protein classes among *Drosophila* and cricket taxa.

Key words: Allonemobius, *Gryllus*, rapid evolution, reproductive isolation, seminal fluid.

Résumé : Les protéines des glandes reproductives accessoires chez le genre *Drosophila* évoluent rapidement et semblent jouer un rôle important dans le succès des mâles en matière de fécondation. De plus, ces protéines sont soupçonnées d'une implication dans l'établissement de barrières à la fécondation entre espèces proches. Tandis que des glandes accessoires sont présentes chez les mâles de nombreux groupes d'insectes, les protéines transmises des mâles aux femelles lors de l'accouplement n'ont pas été bien caractérisées à l'extérieur du genre *Drosophila*. Afin de mieux connaître ces protéines, les auteurs ont caractérisé des EST des glandes accessoires chez deux espèces de grillons, *Allonemobius fasciatus* et *Gryllus firmus*. À l'aide d'une approche EST, suivie d'analyses bioinformatiques et évolutives, les auteurs ont trouvé que plusieurs protéines étaient sécrétées et ainsi disponibles pour transfert à la femelle lors de l'accouplement. De plus, la plupart des EST étaient inédits puisqu'ils montraient peu de similarité entre taxons. Des analyses évolutives suggèrent que les protéines de grillons sont sujettes à une sélection divergente et indiquent que l'*Allonemobius* est beaucoup moins polymorphe que le *Gryllus*. En dépit d'une rapide divergence de la séquence nucléotidique, il semble y avoir une conservation fonctionnelle des classes de protéines chez le genre *Drosophila* et les grillons.

Mots clés : Allonemobius, *Gryllus*, évolution rapide, isolement reproductif, fluide séminal.

[Traduit par la Rédaction]

Introduction

In *Drosophila*, the male accessory reproductive glands are the source of a variety of secreted proteins (accessory gland proteins, Acps) that function to improve a male's probability of paternity. As understanding of the identity and function of these proteins has increased, their importance in modulating reproductive processes has become more evident (Wolfner 1997, 2002; Gillott 2003). Acps appear to act on every aspect of fertilization in *Drosophila*, from the initial transfer of sperm (Grimnes et al. 1986) to oviposition (Herndon and Wolfner 1995). They may even alter the fertilization probabilities of competing sperm from males mated to the same female (Clark et al. 1995). As a result, evolutionary changes in these secretions, which may be driven by sexual selection and sexual conflict, can have important implications for fertilization compatibility between males and females and the evolution of reproductive isolation.

Thus far, detailed genetic studies of Acps have been limited to *Drosophila* (Gillott 2003). Although some work has been done on grasshoppers, mosquitoes, moths, and beetles, most of this work has been based on protein extracts of whole accessory glands (Gillott 2003). As a result, we know very little about genetic variation of Acps found in other insects and the relationship of these Acps to those found in *Drosophila*. Specifically, we do not know if Acps

Received 11 February 2006. Accepted 7 May 2006. Published on the NRC Research Press Web site at http://genome.nrc.ca on 6 October 2006.

Corresponding Editor: A. Civetta.

W.E. Braswell[1] and D.J. Howard. Department of Biology, New Mexico State University, Las Cruces, NM, USA.
J.A. Andrés, L.S. Maroja, and R.G. Harrison. Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY, USA.
W.J. Swanson. Department of Genome Sciences, University of Washington, Seattle, WA, USA.

[1]Corresponding author (e-mail: wbraswel@nmsu.edu).

 doi:10.1139/G06-061 © 2006 NRC Canada

are conserved across divergent taxa or if they function in similar ways.

Here, we present the first characterization of accessory gland ESTs in 2 species of Orthoptera—the striped ground cricket *Allonemobius fasciatus* and the field cricket *Gryllus firmus*. *Allonemobius* and *Gryllus* represent distinct lineages within Orthoptera, with divergence time between the lineages possibly exceeding 200 million years. However, each of these species is one of a pair of hybridizing sister species, for which barriers to gene exchange have been well characterized. Given what we know about the dynamics of the mating systems in these crickets, it is possible that Acps play a major role in determining paternity. As in most cricket species, females in both genera orient toward male calling songs and, upon contact, are subsequently courted by males. During courtship, the final decision to mate is controlled by the female, as she must mount the male for copulation (Alexander and Otte 1967; Mays 1971; Sadowski et al. 2002). Moreover, females of these 2 species are highly promiscuous, mating repeatedly and with multiple males (Gregory and Howard 1996; Wagner et al. 2001; Tregenza and Wedell 2002). Thus, males have little pre-insemination control over paternity—they cannot force matings nor can they ensure monogamy. As a result, Acps may represent the only means by which males can assert influence over paternity.

This work represents the first in a series of papers aimed at addressing the evolution of cricket Acps and their importance in reproductive isolation. We describe the kinds of proteins found in cricket accessory glands, compare these proteins with those found in *Drosophila*, and present initial efforts towards assessing factors that have influenced the evolution of these proteins.

Materials and methods

Dissections, tissue, and RNA preparation

Male reproductive accessory glands were dissected from each of 10 and 8 adult male *Allonemobius fasciatus* and *Gryllus firmus*, respectively, and immediately placed in 10 mL of 4 mol/L guanidinium thiocyanate solution (MacDonald et al. 1987). Tissues were disrupted and homogenized by hand using a glass tissue grinder, and centrifuged to remove soluble debris. Total RNA was isolated by ultracentrifugation of the homogenate through a pad of 5.7 mol/L cesium chloride in 4 mol/L EDTA (MacDonald et al., 1987). The pellets of total RNA were isolated and resuspended to assess the quantity, purity, and integrity by ultraviolet spectrophotometry and denaturing gel electrophoresis (Sambrook et al. 2001). mRNA was purified using oligo(dT) binding (Oligotex mRNA Spin Column, Qiagen, Valencia, Calif.).

Library construction

Complementary DNA (cDNA) was constructed from approximately 5 μg of mRNA using the Superscript cDNA synthesis system (Invitrogen, Carlsbad, Calif.). Briefly, first-strand cDNA was produced using Superscript II and an oligo(dT)–*Not*I primer adapter. This cDNA was made double stranded through nick-translated replacement of the mRNA with cDNA using *Escherichia coli* polymerase I, RNase H, and DNA ligase. The blunt-ended cDNAs were ligated to *Sal*I adapters, which upon digestion with *Not*I and *Sal*I ensured an asymmetrical cDNA that would efficiently ligate to the vector and thus produce a directional library. Adapter-ligated and digested cDNAs were size fractionated by column chromatography to remove digested adapters. Finally, size-fractionated cDNA was ligated to the pCMV-SPORT6 vector and introduced to ElectroMax DH10B cells (Invitrogen) by electroporation. Recombinant colonies were grown on ampicillin-selective agar Luria–Bertani (LB) medium, and individual colonies were arrayed by manual picking and transferred to liquid ampicillin-selective LB medium in 96-well plates for independent growth.

Differential hybridization

Because we were interested in genes coding for proteins specific to the male accessory reproductive glands, we probed the libraries with female cDNA (cDNA from whole-body extracts of *Allonemobius* and *Gryllus* females, respectively) to identify and eliminate "housekeeping" genes. To do this we used a plate replicator to spot the library (colonies grown independently in liquid medium) onto nylon membranes placed on ampicillin-selective agar LB medium. This allowed the colonies to grow through the Hybond XL membranes (Amersham Biosciences, Piscataway, N.J.) and retain the 96-well organization. Following growth of colonies, membranes were removed from the media, cells were lysed, and DNA was bonded to the membrane by baking at 80 °C for 2 h.

Oligo(dT)-primed, first-strand, female, whole-body cDNA was prepared using Superscript II reverse transcriptase (Invitrogen) to incorporate ^{32}P-labeled dCTP and then denatured at 65 °C for 30 min in 0.3 mol/L NaOH. Membranes were pre-hybridized for 1 h at 65 °C in 5× standard saline phosphate (SSPE), 5× Denhardt's (0.02% polyvinylpyrrolidone, 0.02% Ficoll, 0.02% bovine serum albumin), and 0.5% sodium dodecyl sulphate (SDS). Denatured salmon sperm DNA was added to the pre-hybridization mixture at 0.2 mg/mL to block non-specific binding. ^{32}P-labeled probe was added to the pre-hybridizing membranes and incubated at 65 °C for 14 h. Following hybridization, membranes were washed with decreasing concentrations (2×, 1×, and 0.1×) of SSPE with 0.1% SDS at 65 °C for 10 min.

Ten independent colonies were initially analyzed to determine the quality of the library. Plasmid DNA was purified and the cDNA inserts were digested from the plasmid using *Eco*RI and *Xho*I restriction enzymes to estimate size on an agarose gel and then sequenced. This initial analysis demonstrated that 1 EST was highly over-represented. To avoid repeated sequencing we designed PCR primers (5′-GCTAATGCCAACGCTCTCTC-3′ and 5′-CGTTGGACAAGCTACGACCT-3′) to amplify this sequence from each species, and labeled it with ^{32}P using a random-priming DNA-labeling kit (DECA prime II, Ambion, Austin, Tex.). This labeled product was then used to probe the library of each species (as described earlier), thereby identifying clones to avoid when sequencing.

Plasmid isolation and sequencing

Putative male-specific clones (i.e., those not hybridizing with female probes) were transferred to deep 96-well plates with ampicillin-selective medium. Plasmid DNA was puri-

© 2006 NRC Canada

fied from these cells using the Qiaprep 96-well miniprep system (Qiagen). All cDNAs were single-pass sequenced from the 5′ end using the ABI Prism Big Dye Terminator Cycle Sequencing Ready Reaction kit and a vector-specific primer, SP6 (5′-ATTTAGGTGACACTATGA-3′). Sequencing reactions were analyzed using an ABI 3100 automatic DNA sequencer (PE Applied Biosystems, Foster City, Calif.).

EST clustering and gene identity searches

Sequencing chromatographs were visualized and organized with Sequencher (Gene Codes, Ann Arbor, Mich.), or SeqMan (DNASTAR, Madison, Wis.). In addition, we used these programs to identify and trim vector and poor sequence. We did not set a lower bound on sequence length because accessory gland products are known to contain small but important proteins (Wolfner 1997; Swanson et al. 2001; Gillott 2003). Similar sequences were then clustered into contigs using these programs with significant manual examination. Consensus (for contigs) and singleton sequences were then converted to FASTA format and all unique gene products (hereafter unigenes) were submitted for BLAST searches against National Center for Biotechnology Information (NCBI) databases. We searched NCBI non-redundant databases with both nucleotide sequences (BLASTn) and translated sequences (BLASTx). Additionally, we searched *Gryllus* ESTs with *Allonemobius* ESTs (and vice versa) using local tBLASTx. Finally, we searched each of these datasets against other EST datasets deposited at NCBI (excluding mouse and human ESTs).

Functional inference

To begin exploring functionality of the ESTs, we determined the gene ontology of those ESTs with significant similarity to any *Drosophila* protein. Identifiers (CG numbers) for *Drosophila* proteins that showed significant similarity to cricket ESTs (i.e., $\log_{10}(\exp) \geq 3$ in a BLASTx search against the NCBI *Drosophila* non-redundant protein database) were submitted to Panther 5.0 (http://panther.appliedbiosystems.com/; Mi et al. 2005) to acquire their gene ontology terms. Further, we used GeneMark.SPL (http://opal.biology.gatech.edu/GeneMark/; Borodovsky and McIninch 1993) for open reading frame prediction. We predicted secretory signal sequences with SignalP (http://www.cbs.dtu.dk/services/SignalP/; Bendtsen et al. 2004) and transmembrane regions with TMHMM (http://www.cbs.dtu.dk/services/TMHMM/; Krogh et al.2001).

Polymorphism assay

Because each cDNA library was constructed using pooled mRNA from 8–10 crickets, we evaluated our sequences for polymorphisms. Unigenes composed of multiple EST sequences (i.e., contigs) were explored to assess single nucleotide polymorphism (SNPs) and insertion–deletion polymorphism (indels). Owing to the wide range in the number of ESTs per contig (2–286 for *Allonemobius* and 2–52 for *Gryllus*), we did not limit the "depth" of the alignment for inferring polymorphism (i.e., all contigs were analyzed). We classified polymorphism (SNPs and indels) by visually inspecting sequencing chromatograms. Aligned sites were considered polymorphic if at least 1 EST had a high-quality base call that differed from high-quality base calls in other ESTs in the alignment.

Evolutionary analyses

To assess the evolutionary forces involved in protein divergence between *Allonemobius* and *Gryllus*, we determined the number of radical versus conservative amino acid substitutions. Specifically, for those unigenes showing significant similarity between *Allonemobius* and *Gryllus*, we aligned predicted amino acid sequences (from GeneMark above) to acquire codon alignments. Non-synonymous nucleotide substitutions were then analyzed using the approach of Zhang (2000), and included amino acid classifications based on charge, polarity, or polarity and volume. More powerful codon-based estimates of selection (i.e., $\frac{dN}{dS}$) were not used because the divergence time of these 2 genera is quite large and such analyses prove unreliable in the face of saturation at synonymous sites.

Three-dimensional protein structural prediction

To assess the impact of amino acid substitutions on protein function, we inferred 3-dimensional protein structure. For each unigene found within both *Allonemobius* and *Gryllus* libraries, we used BLASTp against the PDB database. We found a significant match for a unigene to a hydrolase. This structure was then used as a homology model to determine if any of the differences between the 2 cricket species occur in functionally important regions of the protein.

Results

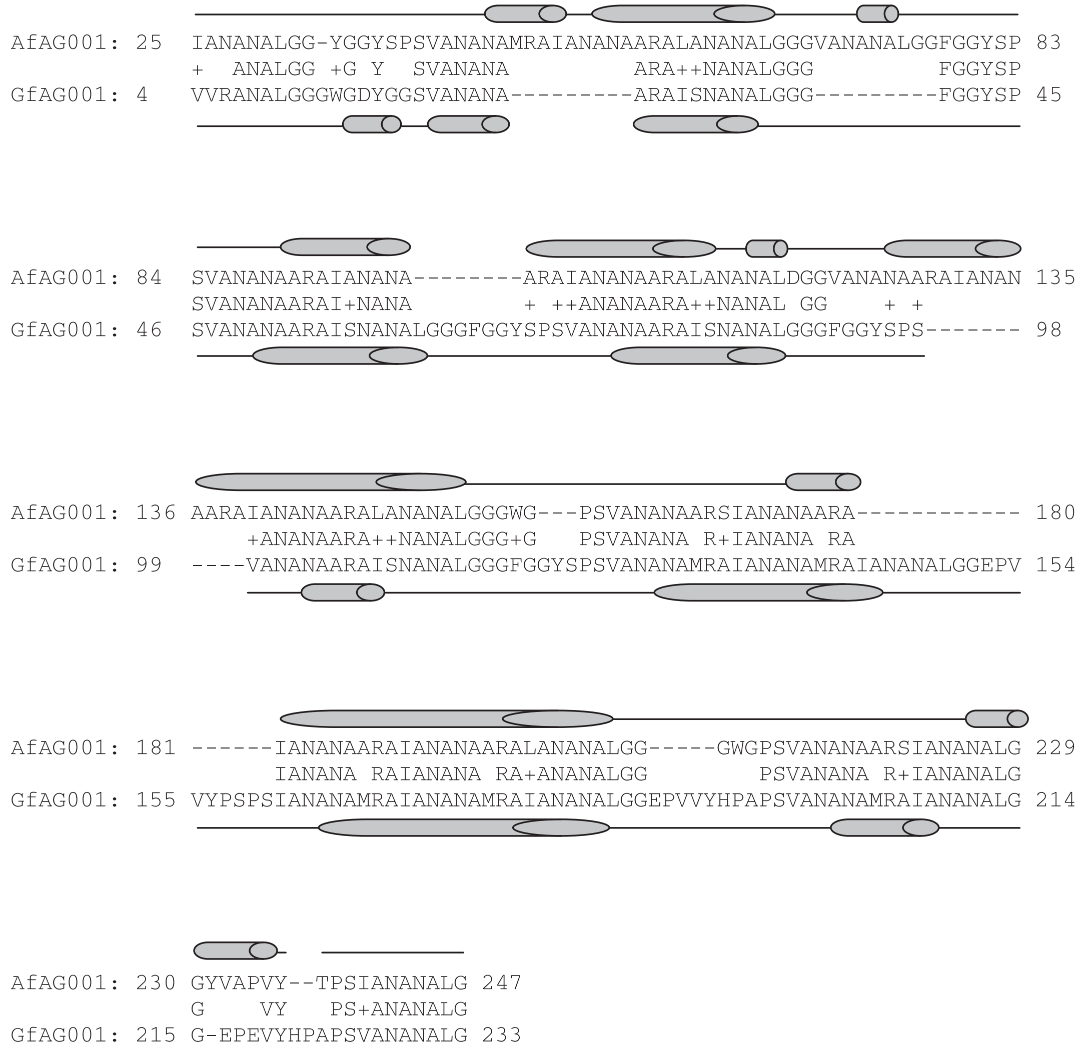
An overview of the libraries

Initial analysis of the cDNA libraries revealed good quality libraries with high transformation rates and large cDNA inserts (average insert size was about 1 kb). To enrich for male-expressed ESTs, these libraries were subsequently probed with female cDNA, and female-expressed ESTs were excluded from further analysis. During the initial analysis we discovered 1 EST with exceptionally high expression, accounting for approximately 60% of all clones in each library. This EST (AfAG001 and GfAG001 in *Allonemobius* and *Gryllus*, respectively) was sequenced from each species and aligned (Fig. 1). It is a highly repetitive sequence making accurate alignment difficult, but showing clear homology within the repetitive structure. Further, protein secondary structural predictions by PSIPRED (McGuffin et al. 2000), indicated similar patterns of helical and coiled regions (Fig. 1). However, no significant similarity was found in GenBank.

Following screening to remove female ESTs and the highly expressed repetitive sequence, clones were randomly picked and sequenced. From the sequenced clones, we found 183 *Allonemobius* and 247 *Gryllus* unigenes composed of either single ESTs or overlapping contigs (Table 1). Regardless of which database was searched, relatively few cricket unigenes showed significant similarity to previously identified sequences (using a liberal level of significance, $E < 1 \times 10^{-3}$). Indeed, a large proportion of each library was composed of novel sequences; 59% and 45% of all unigenes in *Allonemobius* and *Gryllus*, respectively, showed no significant similarity to sequences in any database searched. Of those unigenes

© 2006 NRC Canada

Fig. 1. Amino acid alignment and predicted secondary structure of the most commonly encountered EST in both *Allonemobius* and *Gryllus* accessory reproductive gland libraries. The sequences account for over 60% of the clones in each library. Lines represent coils and shaded tubes represent helices.



showing significant similarity to database sequences, most showed similarity to sequences in the *Drosophila* genome (Table 1). Only a few additional “hits” were acquired by searching non-redundant or EST databases. Specifically, for *Allonemobius*, 6 and 27 new “hits” were acquired by searching the NCBI non-redundant database with BLASTx and BLASTn, respectively, and 31 by searching the NCBI EST database. Similarly, for *Gryllus*, 20 and 41 additional hits were found in the NCBI non-redundant database using BLASTx and BLASTn, respectively, and 47 new hits occurred in the NCBI EST database.

Comparisons between cricket libraries revealed that 27 sequences displayed significant similarity between *Allonemobius* and *Gryllus* (reciprocal best hits; Table 2). Sixteen of the 27 *Allonemobius* unigenes returned no significant matches in a BLASTx search of the NCBI non-redundant database. On the other hand, only 10 of the 27 *Gryllus* unigenes could not be matched with sequences in this database. Similarly, 18 of the 27 *Allonemobius* unigenes had no matches in a BLASTx search of *Drosophila* sequences, whereas 15 of 27 *Gryllus* unigenes had no matches. Thus, even those sequences that appear to be conserved between *Allonemobius* and *Gryllus* demonstrate relatively low levels of similarity to previously described sequences.

Secretory signals

SignalP predicted secretory signals in 16.4% and 20.2% of *Allonemobius* and *Gryllus* unigenes, respectively (Table 1). The presence of a secretory signal is an indication that the proteins encoded by these genes are destined for secretion and possible transfer to the female during mating. However, since we have not identified the full-length sequence for all transcripts, these estimates represent the minimum number of secreted proteins. As expected, fewer unigenes from either species show evidence of transmembrane domains as predicted by TMHMM (Table 1).

Gene ontology

The majority of sequences from *Allonemobius* and *Gryllus* could not be assigned a gene ontology because they did not

© 2006 NRC Canada

Table 1. Summary of accessory gland EST libraries.

	Allonemobius	*Gryllus*
No. of ESTs	704	657
No. of unigenes	183	247
No. of clusters	50	64
No. of singletons	133	183
No. of significant BLASTx hits (%)		
Non-redundant	48 (26)	118 (48)
Drosophila	42 (23)	98 (40)
No. of significant BLASTn hits (%)		
Non-redundant	41 (22)	74 (30)
Drosophila	14 (7.7)	33 (28)
EST	55 (30)	80 (32)
No. with predicted signal sequence (%)	30 (16.4)	50 (20.2)
No. with predicted transmembrane domains (%)	22 (12)	9 (4)
No. with predicted anchor sequence (%)	2 (2)	1 (0.4)

display significant similarity to sequences found in *Drosophila.* However, the 42 and 98 sequences in *Allonemobius* and *Gryllus*, respectively, which were significantly similar to sequences in the *Drosophila* database were assigned gene ontologies broadly similar to gene ontologies of *Drosophila* accessory gland ESTs (Swanson et al. 2001).

The most common biological process associated with unigenes of the 2 cricket species is protein metabolism and modification (Fig. 2). This biological process is also associated with many *Drosophila* Acps, although proteins with no classified biological process are the most abundant constituents of *Drosophila* accessory glands. In addition, proteins involved with nucleic acid metabolism, developmental processes, signal transduction, and intracellular protein traffic, were relatively common in the accessory glands of all 3 species. Despite the overall similarity, some clear differences existed among the 3 groups. *Gryllus* has a low percentage, relative to *Drosophila*, of proteins involved in lipid, fatty acid, and steroid metabolism; no proteins involved in this process have yet been identified in *Allonemobius*. Cell-adhesion proteins are more common in *Drosophila* and *Allonemobius* than in *Gryllus*; cell structure and motility proteins are relatively common in both cricket species, but are rare in *Drosophila*; and carbohydrate metabolism proteins have not been found in *Allonemobius*, but are present in *Drosophila* and are relatively abundant in *Gryllus*. Finally, relative to *Allonemobius* and *Drosophila*, *Gryllus* has a low percentage of proteins involved in immunity and defense (Fig. 2).

The molecular functions assigned to the cricket unigenes were also quite similar to those associated with *Drosophila* accessory gland ESTs, although differences in molecular function were more marked than differences with regard to biological process. Many of the unigenes of all 3 species demonstrate nucleic acid binding activity or have an unclassified molecular function. Proteases and cytoskeletal proteins are more highly represented in the *Gryllus* and *Allonemobius* libraries than in *Drosophila. Allonemobius* and *Gryllus* showed reduced levels of proteins functioning as hydrolases relative to *Drosophila*. *Allonemobius* harbored more proteins functioning as chaperones, but no proteins, at least thus far, acting as transporters, transferases, transcription factors, or regulatory molecules; *Gryllus* was similar to *Drosophila* in all of these functions. It is also noteworthy that the *Gryllus* accessory gland library contains proteins with a wider array of molecular functions than the *Allonemobius* accessory gland library (Fig. 3).

Polymorphisms

Although most unigenes in each library were represented by single ESTs, at least 25% of the unigenes were represented by contigs of multiple ESTs. Because each library was constructed from multiple individuals, each contig may represent up to 16 or 20 unique alleles in *Gryllus* and *Allonemobius*, respectively.

SNPs were the most common form of polymorphism within our EST libraries. Most SNPs (73% and 82% in *Allonemobius* and *Gryllus*, respectively) occurred in the predicted coding regions as opposed to the up- or down-stream untranslated regions. This is not surprising, as greater than 70% of the sequenced bases in each library fall within predicted coding regions.

Gryllus appears to contain more variation than *Allonemobius*. Within predicted coding regions, *Gryllus* has twice as many SNPs per contig as *Allonemobius* (0.8 and 2.2 in *Allonemobius* and *Gryllus*, respectively). Additionally, there are 12 indels in *Gryllus* and only 3 in *Allonemobius*. In contrast, variation within untranslated regions was equal (mean of 0.15 SNPs per contig for both *Allonemobius* and *Gryllus*).

Transitions were the most common form of polymorphism comprising 58% and 53% of all SNPs in *Allonemobius* and *Gryllus*, respectively. Within codons, more SNPs occurred at the third position in both *Allonemobius* (49%) and *Gryllus* (59%) and, as expected, most of these (95% in each species) were synonymous substitutions. In total, 51% and 47% of SNPs were nonsynonymous in *Allonemobius* and *Gryllus*, respectively. In comparison, 52% and 36% of Acp SNPs in *D. simulans* and *D. melanogaster*, respectively, are nonsynonymous (Begun et al. 2000).

In an attempt to gain some insight into the effect of nonsynonymous variation, we classified the resulting amino acid substitutions as radical or conservative on the basis of charge, polarity, or polarity and volume (Fig. 4). Approxi-

© 2006 NRC Canada

Table 2. Sequence similarity search results for cricket unigenes showing significant similarity between cricket species.

			BLASTn (non-redundant database)		
Allonemobius	*Gryllus*	*P*	Q: *Allonemobius*	*P*	Q: *Gryllus*
AfAG021	GfAG213	5×10^{-39}	*Anopheles* gi:31206650	9×10^{-8}	*Gallus* gi: 46237943
AfAG044	GfAG145	5×10^{-21}	No matches	n.a.	No matches
AfAG086	GfAG188	1×10^{-19}	No matches	n.a.	*Mus* TPA regulated locus gi: 23337012
AfAG110	GfAG273	1×10^{-5}	No matches	n.a.	*Rattus* neu differentiation factor gi: 408384
AfAG163	GfAG130	4×10^{-17}	No matches	n.a.	gi:161850
AfAG165	GfAG074	2×10^{-63}	*Carabus* ribosomal protein gi:50344445	1×10^{-18}	*Biphyllus* ribosomal protein gi:50344449
AfAG170	GfAG158	2×10^{-90}	*Pteronemobius* 18S rRNA gi:27463976	3×10^{-93}	*Acheta* 18S rRNA gi:1199972
AfAG178	GfAG072	3×10^{-71}	No matches	n.a.	*Apis* QM protein gi:48107652
AfAG186	GfAG220	7×10^{-57}	*Allonemobius* 16S rRNA gi:27463887	0	*Gryllus* 16S rRNA gi:11141444
AfAG190	GfAG042	4×10^{-11}	No matches	n.a.	gi:21114787
AfAG198	GfAG085	7×10^{-5}	No matches	n.a.	No matches
AfAG209	GfAG098	0	*Trichosurus B*-actin gi:3320891	0	*Monopterus B*-actin gi:33526988
AfAG210	GfAG228	8×10^{-7}	No matches	n.a.	No matches
AfAG211	GfAG014	1×10^{-35}	*Oreophoetes* 28S rRNA gi: 27923540	4×10^{-33}	*Acheta* 28S rRNA gi: 2190320
AfAG218	GfAG001	2×10^{-22}	No matches	n.a.	No matches
AfAG221	GfAG179	4×10^{-27}	*Blattella* 18S rRNA gi: 8272565	1×10^{-18}	*Tenebrio* rRNA gi: 32527555
AfAG227	GfAG236	5×10^{-31}	gi:27503343	2×10^{-10}	gi:18253042
AfAG223	GfAG235	2×10^{-35}	*Brachiastoma* AmphiP2 gi:2113804	3×10^{-7}	*Ceratitus* ribosomal protein P2 gi:4239710
AfAG232	GfAG131	2×10^{-8}	No matches	n.a.	No matches
AfAG233	GfAG029	1×10^{-9}	No matches	n.a.	gi:2190322
AfAG237	GfAG101	3×10^{-11}	*Haemaphysalis* serin protease 2 gi:3970892	7×10^{-4}	No matches
AfAG247	GfAG129	3×10^{-4}	No matches	n.a.	No matches
AfAG248	GfAG053	0	*Allonemobius* 18S ribosomal RNA gi:27463970	0	*Gryllus* 18S rRNA gi: 27463972
AfAG253	GfAG076	7×10^{-17}	No matches	n.a.	No matches
AfAG254	GfAG084	1×10^{-13}	No matches	n.a.	No matches
AfAG258	GfAG063	2×10^{-12}	*Homo* putative tumor suppressor gi:7145093	7×10^{-9}	*Homo* junctional adhesion molecule-1 gi: 13124448
AfAG259	GfAG092	1×10^{-18}	No matches	n.a.	No matches

Note: Q, query sequence; n.a., not applicable.

mately 30% of nonsynonymous SNPs in both species result in radical amino acid substitutions when classified by charge or polarity. However, when using the combined metric of polarity and volume, 80% and 47% of nonsynonymous SNPs result in radical changes in *Allonemobius* and *Gryllus*, respectively (Fig. 4).

Evolutionary analyses

Of the 27 unigenes showing significant similarity between *Allonemobius* and *Gryllus*, we were able to infer open reading frames for 17 of them. In an attempt to understand the evolution of cricket proteins, we analyzed nonsynonymous differences between *Allonemobius* and *Gryllus* for these genes. Using the methods of Zhang (2000), and classifying amino acids on the basis of polarity, we found a significant excess of radical amino acid substitutions in 5 of these genes. However, 2 of these 5 showed a significant excess of conservative substitutions when amino acids were classified by charge. Additionally, 4 distinct genes exhibited an excess of conservative substitutions when classified by the combined metric of polarity and volume (Table 3).

Structural protein model

Of the cricket homologs with an inferred open reading frame, we were able to fit a structural protein model to one.

© 2006 NRC Canada

	BLASTx (non-redundant database)			
P	Q: *Allonemobius*	*P*	Q: *Gryllus*	*P*
5×10^{-12}	*Anopheles* gi:31206651	7×10^{-38}	*Homo* dipeptidase 1 gi:4758190	2×10^{-51}
n.a.	*Aplysia* fasciclin-like protein gi:20799320	1×10^{-14}	*Anopheles* gi: 31209485	1×10^{-56}
2×10^{-5}	*Anopheles* gi:31216614	2×10^{-15}	*Astracus* carboxylpeptidase B gi:115881	5×10^{-22}
2×10^{-5}	No matches	n.a.	*Litopenaeus* chymotrypsin BII precursor gi:544112	5×10^{-7}
7×10^{-5}	No matches	n.a.	gi:48096253	6×10^{-11}
4×10^{-26}	*Biphyllus* ribosomal protein gi:50344449	8×10^{-67}	*Biphyllus* ribosomal protein p0 gi:50344450	3×10^{-94}
0	No matches	n.a.	*Homo* similar to rRNA intron-encoded homing endonuclease gi:41125794	2×10^{-6}
2×10^{-47}	*Heliothis* QM protein gi:14010642	6×10^{-77}	*Apis* QM protein gi:48107652	5×10^{-66}
0	No matches	n.a.	No matches	n.a.
5×10^{-5}	No matches	n.a.	No matches	n.a.
n.a.	No matches	n.a.	No matches	n.a.
0	*Anopheles* gi:31204457	0	*Anopheles* gi:31204457	1×10^{-76}
n.a.	No matches	n.a.	*Anopheles* gi:31241693	4×10^{-21}
0	No matches	n.a.	No matches	n.a.
n.a.	No matches	n.a.	No matches	n.a.
3×10^{-99}	No matches	n.a.	*Pisum* putative senescence protein gi:13359451	1×10^{-17}
2×10^{-7}	gi:18253043	9×10^{-18}	gil18253043	4×10^{-25}
3×10^{-7}	*Plutella* ribosomal protein P2 gi:49532906	1×10^{-16}	*Spodoptera* ribosomal protein P2 gi:18253045	4×10^{-18}
n.a.	No matches	n.a.	No matches	n.a.
0	No matches	n.a.	gi:50312719	8×10^{-12}
n.a.	*Apis* gi:48098822	6×10^{-44}	*Apis* similar to serine protease gi:48098816	5×10^{-48}
n.a.	*Candida* unnamed protein gi:50289503	4×10^{-4}	No matches	n.a.
0	No matches	n.a.	*Danio* Vangl2 protein gi: 41946856	2×10^{-6}
n.a.	gi:50289503	7×10^{-7}	No matches	n.a.
n.a.	No matches	n.a.	No matches	n.a.
2×10^{-19}	No matches	n.a.	No matches	n.a.
n.a.	No matches	n.a.	*Oryza* VsaA-like protein gi:34910776	6×10^{-4}

That sequence (AfAG021/GfAG231) fit the structure of a hydrolase. Figure 5 shows the 3-D model with differences between species mapped. Three differences between *Allonemobius* and *Gryllus* occur within the active site.

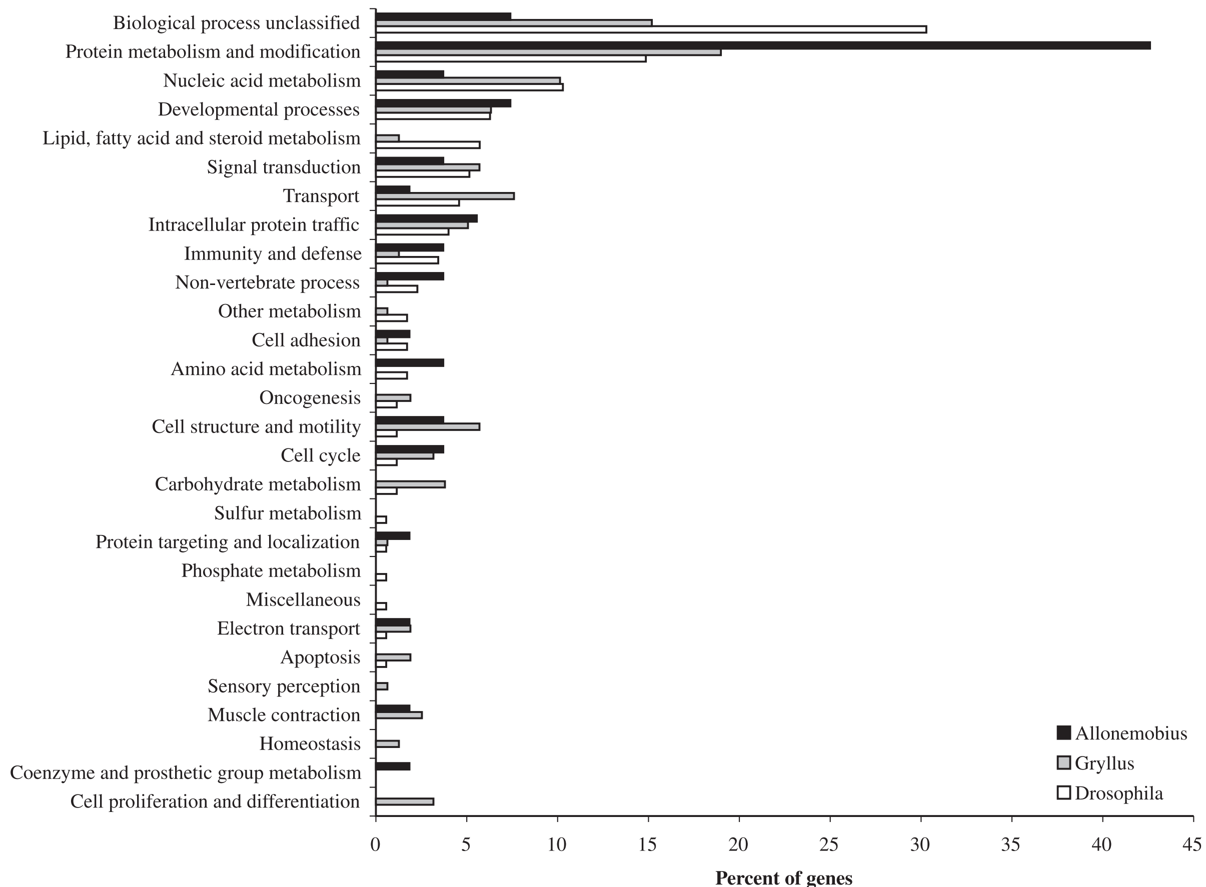
Discussions

The studies reported here were undertaken to extend our understanding of Acps beyond *Drosophila* and because extensive evidence indicates that post-insemination barriers to fertilization play important roles in the reproductive isolation of closely related cricket species (Howard and Gregory 1993; Gregory and Howard 1994; Howard et al. 1998). Thus, studies of Acps in crickets may lead to the identification and characterization of genes that mediate reproductive isolation. Here we are interested in describing the kinds of proteins found in cricket accessory glands and comparing these proteins to those found in *Drosophila*.

One of the first discoveries to emerge from sequencing the EST libraries constructed from accessory glands of *Allonemobius* and *Gryllus* was that a single sequence dominated these libraries. The abundance of the EST and its repetitive sequence suggests that the encoded protein is a structural component of the spermatophore, the sperm and seminal fluid containing packet passed to the female during copulation (Paesen et al. 1992; Feng and Happ 1996).

© 2006 NRC Canada

Fig. 2. Distribution of biological process ontology among *Allonemobius*, *Gryllus*, and *Drosophila*. Information concerning *Drosophila* accessory reproductive gland proteins is from Swanson et al. (2001).



After screening to enrich for male-expressed genes and to remove the presumed spermatophore protein-encoding sequence, more than 650 clones were sequenced for each library. These sequences were edited and compressed into 183 and 247 unigenes in *Allonemobius* and *Gryllus*, respectively. The majority of cricket accessory gland unigenes represent novel (or highly diverged) sequences that are not shared between the 2 species and are not found in *Drosophila* or in any other taxonomic group represented in GenBank. This finding is consistent with data from *Drosophila*, suggesting that Acps evolve rapidly (Swanson et al. 2001).

In a recent EST analysis of *Drosophila* accessory glands, 24% of the unigenes encoded proteins with predicted signal sequences (Swanson et al. 2001). Similar proportions of unigenes from the accessory glands of *Allonemobius* and *Gryllus* (16.4% and 20.2%, respectively) encode proteins with predicted signal sequences, indicating that they are destined for secretion (Table 1). Conversely, transmembrane domains, which bind proteins within the cell membrane, are less common in the unigenes of both cricket species.

The majority of unigenes from cricket accessory gland libraries have no homologs in the *Drosophila* database, and therefore cannot be assigned a biological process or a molecular function. As a result, interspecific gene ontology comparisons must be viewed as preliminary. However, the major biological processes associated with unigenes in *Allonemobius* that can be assigned a gene ontology are protein metabolism and modification, developmental processes, unclassified processes, and intracellular protein traffic. The same processes are associated with the unigenes of *Gryllus*, but the rank order of abundance is different. The majority of the biological processes associated with cricket accessory gland ESTs are found in *Drosophila* accessory gland ESTs. Despite variation in rank order abundance, the overall signal is one of similarity in biological processes associated with ESTs from insect accessory glands (Fig. 2).

Molecular functions of accessory gland ESTs are more variable among the 3 species, although several molecular functions are well represented in all 3 species (Fig. 3). It is particularly interesting that proteases and hydrolases occur in the accessory glands of both *Allonemobius* and *Gryllus*. Proteases are common among *Drosophila* Acps (Mueller et al. 2004) and may modify proteins of the female reproductive tract, protect sperm, act as coagulation factors (Leadley 2001), or be involved in proteolytic cascades that influence sperm competition and fertilization success (Park and Wolfner 1995; Wolfner 1997). In particular, proteolytic processes are known to regulate *Drosophila* Acp26Aa, a peptide hormone (Park and Wolfner 1995). Hydrolases, too, may be important in mediating the fertilization success of males. Indeed, 5 lipases in the *Drosophila* accessory glands have $\frac{dN}{dS}$ ratios greater than one, indicating that divergence has been driven by positive Darwinian selection (Swanson et al. 2001).

© 2006 NRC Canada

Fig. 3. Distribution of molecular function ontology among *Allonemobius*, *Gryllus*, and *Drosophila*. Information concerning *Drosophila* accessory reproductive gland proteins is from Swanson et al. (2001).

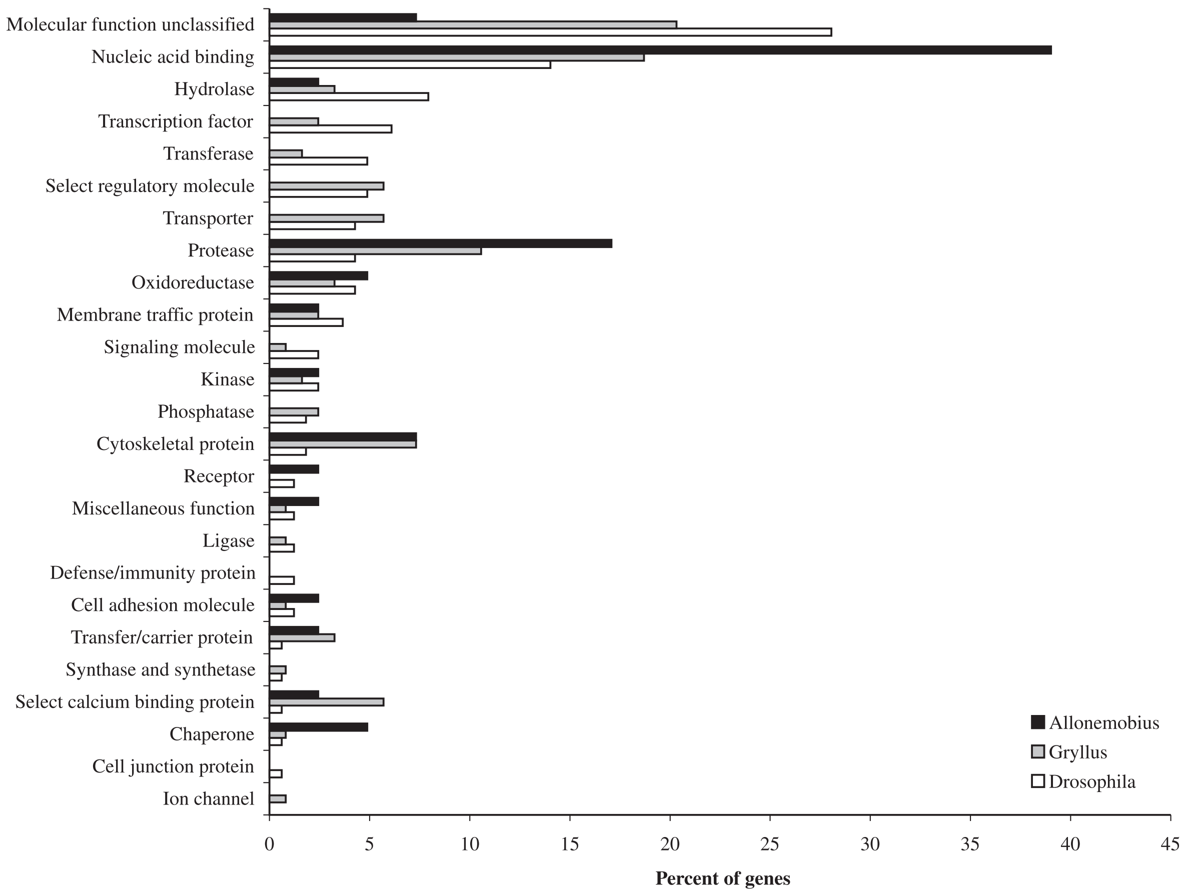
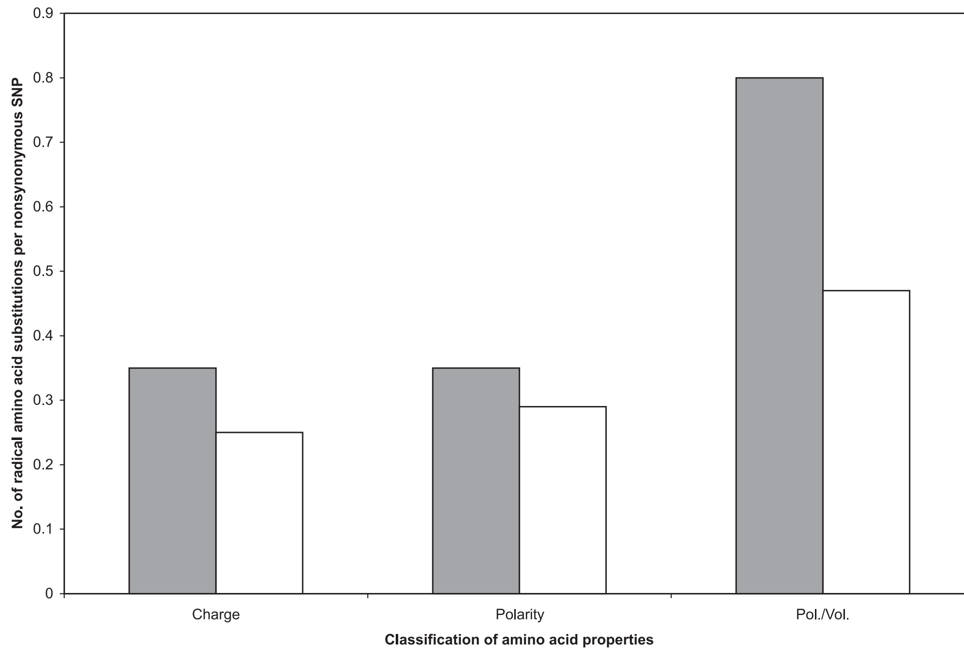


Fig. 4. Proportion of nonsynonymous SNPs causing radical amino acid substitutions based on different classifications of amino acid properties. Shaded bars represent *Allonemobius* substitutions, whereas white bars represent *Gryllus* substitutions.



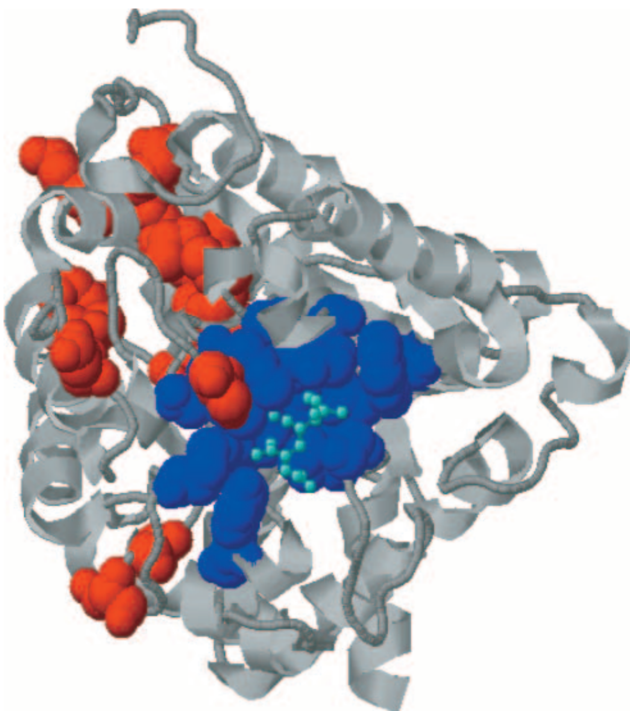
© 2006 NRC Canada

Table 3. Summary of radical versus conservative analyses.

Alignment		Radical and (or) conservative substitutions		
Allonemobius	*Gryllus*	Charge	Polarity	Polarity and volume
AfAG021	GfAG213	0.46	0.66	**0.40**
AfAG044	GfAG145	0.96	0.85	0.81
AfAG086	GfAG188	0.76	1.01	**0.58**
AfAG163	GfAG130	**0.78**	**1.34**	1.07
AfAG165	GfAG074	0.70	0.59	1.06
AfAG178	GfAG072	0.71	1.98	1.04
AfAG190	GfAG042	0.97	1.13	0.88
AfAG198	GfAG085	0.91	1.10	0.92
AfAG209	GfAG098	0.62	**1.94**	0.94
AfAG223	GfAG235	0.68	**1.19**	0.93
AfAG227	GfAG236	0.00	**1.89**	0.31
AfAG232	GfAG131	**0.78**	**1.33**	0.84
AfAG233	GfAG029	0.98	1.03	0.88
AfAG237	GfAG101	0.96	1.21	0.82
AfAG253	GfAG076	1.27	0.89	**0.69**
AfAG254	GfAG084	0.83	1.16	**0.68**
AfAG258	GfAG063	1.15	0.67	1.04

Note: Numbers in bold represent significant deviation of the radical to conservative substitution ratio from expectation.

Fig. 5. Three-dimensional protein structural model of a hydrolase found in both *Allonemobius* (AfAG021) and *Gryllus* (GfAG213) EST libraries. Blue spacefill represents the location of the active site, red spacefill shows differences between *Allonemobius* and *Gryllus*, and the bound substrate is shown in cyan ball and stick.



These hydrolases may function to modify cell membranes, provision nutrients to females, or even provide energy to sperm, affecting sperm motility or viability.

Because reproductive accessory glands are small, each library was constructed using mRNA pooled from multiple individuals. As a result, each EST represents one of several possible alleles. We took advantage of this opportunity to assess allelic variation within cricket unigenes. We found that the majority of polymorphism exists in the form of SNPs, and a large proportion of that variation occurs at nonsynonymous sites. This result is consistent with patterns of variation observed in *Drosophila* Acp loci (Begun et al. 2000). On the basis of all SNP and indel data, *Gryllus* is much more variable at EST loci than *Allonemobius* although both taxa exhibit similar levels of nonsynonymous variation. The increased variation in *Gryllus* is unlikely due to larger effective population sizes given the huge population sizes often observed in *Allonemobius* (Howard and Harrison 1984); however, increased variation may be an outcome of greater time since a selective sweep or population bottleneck, or of reduced selective pressures.

Some evidence exists to support the view that selection pressures differ between *Allonemobius* and *Gryllus*. First, sequences orthologous between *Allonemobius* and *Gryllus* acquire different numbers of “hits” in GenBank searches. Of the 27 orthologous sequences identified between *Allonemobius* and *Gryllus*, 11 have significant matches in GenBank when BLAST queried by the *Gryllus* sequence and no match when queried by its *Allonemobius* counterpart. Moreover, only 3 of the 27 have significant matches when queried by the *Allonemobius* sequence and no significant matches when queried by its *Gryllus* ortholog (Table 2). These GenBank “hits” occur in taxa that are of equivalent evolutionary distance from both *Allonemobius* and *Gryllus*, suggesting that *Allonemobius* sequences may evolve more rapidly than *Gryllus* sequences, thereby losing similarity to GenBank sequences while retaining similarity to the more closely related *Gryllus* sequences. Moreover, preliminary data suggest that unlike *Allonemobius*, premating barriers may be much more important than postmating barriers in *Gryllus* (Harrison 1986; Harrison and Rand 1989; Ross and Harrison 2002; L.S. Maroja, J.A. Andres, and R.G. Harrison, unpublished data); supporting the idea that

© 2006 NRC Canada

post-copulatory sexual selection may be less intense in *Gryllus*. Indeed, it appears that *Allonemobius* has proportionately more functional variation, as there are more radical amino acid substitutions per nonsynonymous polymorphic site in *Allonemobius* than in *Gryllus* (Fig. 4). Together, these results suggest that *Allonemobius* and *Gryllus* unigenes may experience unique selection pressures. While none of these results conclusively support the selection hypothesis, it is consistent with all the data and explains these disparate results.

Assessing whether amino acid substitutions are radical or conservative provides insight into the evolutionary forces that drive the divergence of proteins in cases where other measures, such as $\frac{dN}{dS}$ ratios, are not informative because of the high level of genetic divergence between the species being compared (Zhang 2000; Smith 2003). Depending on the physio-chemical properties used to define radical and conservative amino acids, we find evidence of an excess of radical or conservative amino acid replacements between *Allonemobius* and *Gryllus* in several Acps. Using polarity, we find a significant excess of radical substitutions in 5 of the 17 homologs. However, 2 of these 5 homologs demonstrate a significant excess of conservative substitutions when amino acids are classified by charge instead of polarity. Four other genes show a significant excess of conservative substitutions when polarity and volume are considered. While an excess of conservative substitutions has been touted as indicating negative (purifying) selection, the power of radical substitutions to indicate positive (diversifying) selection is less clear (Smith 2003).

To begin assessing the functional impact of protein divergence between *Allonemobius* and *Gryllus* we attempted to model the 3-dimensional protein structure of proteins encoded by homologous unigenes. We were able to model a predicted protein, encoded by AfAG021/GfAG231, which fit the structure of a hydrolase. By mapping amino acid differences between *Allonemobius* and *Gryllus* onto the 3-dimensional model, we found that 3 differences exist within the active site of the protein. These differences are clear candidates for selection studies, as they are likely to alter protein function.

Conclusions

Accessory gland ESTs of *Allonemobius* and *Gryllus* have much in common with those of *Drosophila.* As in *Drosophila*, many accessory gland ESTs in these 2 cricket species possess signal sequences indicating that they are secreted, whereas signals of transmembrane domains are few. Also as in *Drosophila*, most of the accessory gland EST sequences from crickets are novel, show no similarity to sequences in GenBank, and are not shared between the 2 cricket species studied here. The overall impression is one of rapid divergence of accessory gland ESTs in crickets. At the same time, gene ontology assignments show that the processes and functions associated with accessory gland ESTs are broadly similar across insect groups.

This overall signal of low sequence similarity yet conservation in biological processes and molecular functions echoes the results of a recent study by Mueller et al. (2004). By using comparative structural modeling, Mueller et al. (2004) were able to infer functional similarities between *Drosophila* Acps and known protein classes. The results of their structural models suggest that, despite rapid evolution at the sequence level, *Drosophila* Acps fall into the same general protein classes seen in mammalian seminal fluid. Thus, it appears that although Acps have great latitude to change through time, there are limits to this divergence, or at least there are limits to the types of proteins that are passed from males to females.

There are at least 2 ways in which sequence divergence coupled with functional conservation may be achieved. First, as sequences diverge or are lost from the accessory glands, there may be continual co-option of new proteins of certain functional classes into the accessory gland expression pathway. Alternatively, evolutionary processes responsible for intraspecific divergence of Acps may simultaneously explain sequence divergence and functional conservation. Specifically, reproductive proteins are involved in intimate interactions with other molecules from the same ejaculate, with molecules from another male's ejaculate present in the female reproductive tract, and with molecules from the female herself. These intimate interactions may dictate the functional conservation of Acps even as they diverge in sequence due to male-male competition, to female choice, or to sexual conflict.

Acknowledgements

We thank L.M. Birge and anonymous reviewers for their valuable comments on a previous draft of this manuscript. This research was supported by NSF grant IRCEB 0111613 to WJS, RGH, and DJH and by an EPA STAR fellowship to WEB. Sequence data from this article have been deposited with the EMBL–GenBank Data Libraries under accession Nos. EB409379–EB409998 and EG018565–EG019055.

References

Alexander, R.D., and Otte, D. 1967. The evolution of genitalia and mating behavior in crickets (Gryllidae) and other Orthoptera. Misc. Publ. Mus. Zool. Univ. Mich. **133**: 1–62.

Begun, D.J., Whitley, P., Todd, B.L., Waldrip-Dail, H.M., Clark, A.G. 2000. Molecular population genetics of male accessory gland proteins in *Drosophila*. Genetics, **156**: 1879–1888. PMID:11102381.

Bendtsen, J.D., Nielsen, H., von Heijne, G., and Brunak, S. 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. **340**: 783–795. PMID:15223320.

Borodovsky, M., and McIninch, J. 1993. GeneMark: parallel gene recognition for both DNA strands. Comput. Chem. **17**: 123–133. doi:10.1016/0097-8485(93)85004-V.

Clark, A.G., Aguade, M., Prout, T., Harshman, L.G., and Langley, C.H. 1995. Variation in sperm displacement and its association with accessory gland protein loci in *Drosophila melanogaster*. Genetics, **139**: 189–201. PMID:7705622.

Feng, X., and Happ, G.M. 1996. Isolation and sequencing of the gene encoding Sp23, a structural protein of spermatophore of the mealworm beetle, *Tenebrio molitor*. Gene, **179**: 257–262. doi:10.1016/S0378-1119(96)00372-1. PMID:8972909.

Gillott, C. 2003. Male accessory gland secretions: modulators of female reproductive physiology and behavior. Annu. Rev. Entomol. **48**: 163–184. doi:10.1146/annurev.ento.48.091801.112657. PMID:12208817.

© 2006 NRC Canada

Gregory, P.G., and Howard, D.J. 1994. A postinsemination barrier to fertilization isolates two closely related ground crickets. Evolution, **48**: 705–710. doi:10.2307/2410480.

Gregory, P.G., and Howard, D.J. 1996. Multiple mating in natural populations of ground crickets. Entomol. Exp. Appl. **78**: 353–356.

Grimnes, K.A., Bricker, C.S., and Happ, G.M. 1986. Ordered flow of secretion from accessory glands to specific layers of the spermatophore of mealworm beetles: demonstration with a monoclonal antibody. J. Exp. Zool. **240**: 275–286. doi:10.1002/jez.1402400214. PMID:3794623.

Harrison, R.G. 1986. Pattern and process in a narrow hybrid zone. Heredity, **56**: 337–349.

Harrison, R.G., and Rand, D.M. 1989. Mosaic hybrid zones and the nature of species boundries. *In* Speciation and its consequences. *Edited by* D. Otte and J.A. Endler. Sinauer, Sunderland, Mass. pp 111–133.

Herndon, L.A., and Wolfner, M.F. 1995. A *Drosophila* seminal fluid protein, Acp26Aa, stimulates egg laying in females for 1 day after mating. Proc. Natl. Acad. Sci. U.S.A. **92**: 10114–10118. doi:10.1073/pnas.92.22.10114. PMID:7479736.

Howard, D.J., and Gregory, P.G. 1993. Post-insemination signaling systems and reinforcement. Philos. Trans. R. Soc. Lond. B Biol. Sci. **340**: 231–236.

Howard, D.J., and Harrison, R.G. 1984. Habitat segregation in ground crickets: the role of interspecific competition and habitat selection. Ecology, **65**: 69–76. doi:10.2307/1939459.

Howard, D.J., Reece, M., Gregory, P.G., Chu, J., and Cain, M.J. 1998. The evolution of barriers to reproduction between closely related organisms. *In* Endless forms: species and speciation. *Edited by* D.J. Howard and S.H. Berlocher. Oxford University Press, New York, N.Y. pp. 279–290.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L.L. 2001. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J. Mol. Biol. **305**: 567–580. doi:10.1006/jmbi.2000.4315. PMID:11152613.

Leadley, R.J., Jr. 2001. Coagulation factor Xa inhibition: biological background and rationale. Curr. Top. Med. Chem. **1**: 151–159. PMID:11899250.

MacDonald, R.J., Swift, G.H., Przbyla, A.E., and Chirgwin, J.M. 1987. Isolation of RNA using guanidium salts. Methods Enzymol. **152**: 219–227. PMID:2443794.

Mays, D.L. 1971. Mating behavior of nemobiine crickets—*Hygronemobius, Nemobius*, and *Pteronemobius.* Fla. Entomol. **54**: 113–126.

McGuffin, L.J., Bryson, K., and Jones, D.T. 2000. The PSIPRED protein structure prediction server. Bioinformatics, **16**: 404–405. doi:10.1093/bioinformatics/16.4.404. PMID:10869041.

Mi, H., Lazareva-Ulitsky, B., Loo, R., Kejariwal, A., Vandergriff, J., Rabin, S., et al. 2005. The PANTHER database of protein families, subfamilies, functions and pathways. Nucleic Acids Res. **33**(database issue): , D284–D288. PMID:15608197.

Mueller, J.L., Ripoll, D.R., Aquadro, C.F., and Wolfner, M.F. 2004. Comparative structural modeling and inference of conserved protein classes in *Drosophila* seminal fluid. Proc. Natl. Acad. Sci. U.S.A. **101**: 13542–13547. doi:10.1073/pnas.0405579101. PMID:15345744.

Paesen, G.C., Schwartz, M.B., Peferoen, M., Weyda, F., and Happ, G.M. 1992. Amino acid sequence of Sp23, a structural protein of the spermatophore of the mealworm beetle, *Tenebrio molitor.* J. Biol. Chem. **267**: 18852–18857. PMID:1527013.

Park, M., and Wolfner, M.F. 1995. Male and female cooperate in the prohormone-like processing of a *Drosophila melanogaster* seminal fluid protein. Dev. Biol. **171**: 694–702. doi:10.1006/dbio.1995.1315. PMID:7556947.

Ross, C.L., and Harrison, R.G. 2002. A fine-scale spatial analysis of the mosaic hybrid zone between *Gryllus firmus* and *Gryllus pennsylvanicus.* Evolution, **56**: 2296–2312. PMID:12487359.

Sadowski, J.A., Grace, J.L., and Moore, A.J. 2002. Complex courtship behavior in the stripped ground cricket, *Allonemobius socius* (Orthoptera: Gryllidae): does social environment affect male and female behavior. J. Insect Behav. **15**: 69–84. doi:10.1023/A:1014432112801.

Sambrook, J., Russel, D.W., and Sambrook, J. 2001. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Smith, N.G.K. 2003. Are radical and conservative substitution rates useful statistics in molecular evolution? J. Mol. Evol. **57**: 467–478. doi:10.1007/s00239-003-2500-z. PMID:14708579.

Swanson, W.J., Clark, A.G., Waldrip-Dail, H.M., Wolfner, M.F., and Aquadro, C.F. 2001. Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila.* Proc. Natl. Acad. Sci. U.S.A. **98**: 7375–7379. doi:10.1073/pnas.131568198. PMID:11404480.

Tregenza, T., and Wedell, N. 2002. Polyandrous females avoid costs of inbreeding. Nature (London), **415**: 71–73. doi:10.1038/415071a. PMID:11780118.

Wagner, W.E., Jr., Kelley, R.J., Tucker, K.R., and Harper, C.J. 2001. Females receive a life span benefit from male ejaculates in a field crickets. Evolution, **55**: 994–1001. PMID:11430659.

Wolfner, M.F. 1997. Tokens of love: functions and regulation of *Drosphila* male accessory gland products. Insect Biochem. Mol. Biol. **27**: 179–192. PMID:9090115.

Wolfner, M.F. 2002. The gifts that keep on giving: physiological functions and evolutionary dynamics of male seminal proteins in *Drosphila.* Heredity, **88**: 85–93. doi:10.1038/sj.hdy.6800017. PMID:11932766.

Zhang, J. 2000. Rates of conservative and radical nonsynonymous nucleotide substitutions in mammalian nuclear genes. J. Mol. Evol. **50**: 56–68. PMID:10654260.

© 2006 NRC Canada