Computational identification of novel chitinase-like proteins in the Drosophila melanogaster genome

Qingsong Zhu1, Youping Deng2,*, Phaneendra Vanka2, Susan J. Brown2, Subbaratnam Muthukrishnan1 and Karl J. Kramer3

1Department of Biochemistry, Kansas State University, Manhattan, KS 66506, USA, 2Bioinformatics Program, Division of Biology, Kansas State University, Manhattan, KS 66506, USA and 3Grain Marketing and Production Research Center, ARS-USDA, Manhattan, KS 66502, USA

Received on November 9, 2003; accepted on November 11, 2003

ABSTRACT

Motivation: Multiple chitinases as well as lectins closely related to them have been characterized previously from many insect species and the corresponding genes/cDNAs have been cloned. However, the identification of the entire assortment of genes for chitinase family proteins and their differences in biochemical properties have not been carried out in any individual insect species. The completion of the entire DNA sequence of Drosophila melanogaster (fruit fly) genome and identification of open reading frames presents an opportunity to study the structures and functions of chitinase-like proteins, and also to identify new members of this family in Drosophila. We are, therefore, interested in studying the functional genomics of chitinase-like gene families in insects.

Methods: We searched the Drosophila protein sequences database using fully characterized insect chitinase sequences and BLASTP software, identified all the putative chitinase-like proteins encoded in Drosophila genome, and predicted their structures using domain analysis tools. A phylogenetic analysis of the chitinase-like proteins from Drosophila and several other insect species was carried out. The structures of these chitinases were modeled using homology modeling software.

Results: Our analysis revealed the presence of 18 chitinase-like proteins in the Drosophila protein database. Among these are seven novel chitinase-like proteins that contain four signature amino acid sequences of chitinases belonging to family 18 glycosylhydrolases, including both acidic and hydrophobic amino acid residues critical for enzyme activity. All the proteins contain at least one catalytic domain with one having four catalytic domains. Phylogenetic analysis of chitinase-like proteins from Drosophila and other insects revealed an evolutionary relationship among all these proteins, which indicated gene duplication and domain shuffling to generate the observed diversity in the encoded proteins. Homology modeling showed that all the Drosophila chitinase-like proteins contain one or more catalytic domains with a (α/β)8 barrel-like structure. Our results suggest that insects utilize multiple family 18 chitinolytic enzymes and also non-enzymatic chitinase-like proteins for degrading/ remodeling/binding to chitin in different insect anatomical extracellular structures, such as the cuticle, peritrophic membrane, trachea and mouth parts during insect development, and possibly for other roles including chitin synthesis.

Availability: Perl program and supplementary material are available at http://www.ksu.edu/bioinformatics/ supplementary.htm

Contact: ydeng@ksu.edu

INTRODUCTION

Chitin, a polysaccharide consisting of β(1,4)-linked N-acetyl-d-glucosamine residues, is an insoluble structural component found in the exoskeleton and gut linings of insects. Chitin can be hydrolyzed enzymatically by chitinases (EC 3.2.1.14), which have been identified in a wide range of organisms, such as insects, crustaceans, yeasts and other fungi as well as in organisms that do not contain chitin, such as bacteria, higher plants and vertebrates. In fungi, chitinases are involved in cell growth and division, whereas in insects and crustaceans they degrade chitin in the peritrophic membrane and exoskeletal cuticle or shell. There is growing evidence that chitinases may also have roles in defense of organisms against pathogenic fungi and parasites (Wang et al., 1996; Kramer et al., 1997; Ding et al., 1998; Neuhaus, 1999).

Chitinase genes have been isolated from various insects, including lepidopteran insects, such as Manduca sexta (Kramer et al., 1993), Bombyx mori (Kim et al., 1998, Mikitani et al., 2000; Abdel-Banat and Koga, 2001; Daimon et al., 2003), Spodoptera litura (Shinoda et al., 2001),

*To whom correspondence should be addressed.
Hyphantria cunea (Kim et al., 1998) and Choristoneura fumiferana (Zheng et al., 2002); dipteran insects, such as Aedes aegypti (de la Vega et al., 1998), Anopheles gambiae (Shen and Jacobs-Lorena, 1997), Glossina morsitans morsitans, Drosophila melanogaster (de la Vega et al., 1998), Chelonus sp. (Krishnan et al., 1994), Lutzomyia longipalpis (Ramalho-Ortigao and Traub-Cseko, 2003) and Chironomus tentans (Feix et al., 2000); coleopteran insects, such as Phaedon cochlereae (Shinoda et al., 2001) and Tenebrio molitor (Royer et al., 2002).

In most insect species, only a single chitinase gene has been identified (Kramer et al., 1993; Krishnan et al., 1994; Shen and Jacobs-Lorena, 1997; Kim et al., 1998; Shinoda et al., 2001; Royer et al., 2002; Zheng et al., 2002; Ramalho-Ortigao and Traub-Cseko, 2003). However, chitinases have been detected in molting fluid secreted by the epidermis and in midgut tissues of lepidopteran insects, and it has been assumed that a single gene is functional in both these types of tissues. Results from Southern blot analyses have tended to support this possibility. The hypothesis that chitinases are a multi-gene family has been proposed for dipterans, including Aedes, Anopheles and Drosophila, based on the fact that 2–4 different PCR products were resolved in A.aegypti, A.freeborni, A.gambiae, A.stephensi and D.melanogaster by using degenerate primers (de la Vega et al., 1998). The deduced protein sequences have high sequence similarity to chitinases from other species of insects. In B.mori, at least two chitinase genes were identified (Kim et al., 1998; Mikitani et al., 2000; Abdel-Banat and Koga, 2001; Daimon et al., 2003). Five cDNAs, which are thought to be derived from different chitinase or chitinase-like genes, were identified (Daimon et al., 2003) in a search for expressed sequence tags (ESTs). One of these genes, BmChi-h, displays a similar expression pattern to another well-studied B.mori chitinase gene (Daimon et al., 2003). The BmChi-h gene also shows sequence similarity to bacterial and baculoviral chitinases, suggesting that the ancestral species of B.mori acquired the chitinase gene from a bacterium or baculovirus. Nonetheless, we still do not know the total number of chitinase-like genes in any individual insect species.

With the completion of the genome sequences of several insect species, it is expected that most of the proteins with significant similarity to insect chitinases will be identified and annotated. The completed Drosophila genome provides us with a powerful resource with which to study proteins that are related to proteins with identified functions from other organisms. It is probable that all the Drosophila chitinase and chitinase-like genes can be identified using bioinformatics methods.

In this study, five well-characterized chitinases from different insect species were used as query sequences in a BLAST search of the D.melanogaster genome database. Potential Drosophila chitinase or chitinase-like sequences were identified. In this paper, we report 18 potential chitinase or chitinase-like sequences that were identified in the Drosophila genome database. We present our characterization and analysis of these sequences.

METHODS

BLAST searching of the D.melanogaster genome database

Well-known insect chitinase protein sequences were used as queries in a BLAST search of the Drosophila genome database (http://www.fruitfly.org/blast/). These sequences included M.sexta chitinase (GenBank accession no. A56596), which was the first insect chitinase identified and is functional in the molting stage (Choi et al., 1997); A.gambiae chitinase (GenBank accession no. AAB87764), a gut-specific chitinase h.fochleriae chitinase (GenBank accession no. CAA77014) (Shen and Jacobs-Lorena, 1997); A.aegypti chitinase (GenBank accession no. AAB81849); and B.mori chitinase (GenBank accession no. BAC67246) chitinase. The protein database was searched using the BLASTP program with default parameters.

Potential Drosophila chitinase identification from BLAST hits

A Perl program was used to search the conserved regions in the BLAST hit sequences. A protein was identified as a potential Drosophila chitinase or chitinase-like protein if it contained four of the conserved regions of amino acid sequence (Kramer and Muthukrishnan, 1997; de la Vega et al., 1998; Zhu, 1998).

Domain analysis of identified Drosophila chitinases

Pfam (http://pfam.wustl.edu/) (Sonhammer et al., 1998) and SMART (http://smart.embl-heidelberg.de/) (Schultz et al., 1998) domain analysis programs were used to predict the domain architecture of identified chitinases.

Multiple sequence alignment and phylogenetic analysis

Multiple sequence alignment was performed using the ClustalW program (http://www.ebi.ac.uk/clustalw/) available at the European Bioinformatics Institute web site (Thompson et al., 1994).

To investigate the evolutionary relationship between the putative Drosophila chitinases identified here and other insect chitinases, phylogenetic analysis was performed. A phylogenetic tree was constructed using the neighbor-joining algorithm as described by Saitou and Nei (1987). Protein sequences used for phylogenetic analysis were extracted from GenBank. Amino acid sequences used for analysis were: A.aegypti (GenBank: AAB81849), A.gambiae (GenBank: AAB87764), B.mori 1 (GenBank: AAB47538), B.mori 2 (GenBank: BAC67246), Chelonus sp. (GenBank: A53918), C.tentans (GenBank: CAA73685),
**RESULTS**

**BLAST search for identification of putative *Drosophila* chitinases**

Previous studies have shown that there are four conserved regions in the amino acid sequences of all known insect chitinases (Kramer and Muthukrishnan, 1997; de la Vega et al., 1998; Zhu, 1998). Conserved region I is KXXXXXGGW, where X is a non-specific amino acid. Conserved region II is FDGXDLDEWYP, which is believed to be located in the catalytic region of the enzyme and the residue E is a putative proton donor in the mechanism. Conserved regions III and IV are MXYDXXG and GXXXWXXDXD, respectively.

The sequences of five prototypic chitinases from different insect species were used to carry out a BLAST search for identification of putative *Drosophila* chitinase catalytic domain fragments.

**Homology modeling of *Drosophila* chitinase family**

The SWISS-MODEL program (http://www.expasy.org/swissmod/SWISS-MODEL.html) was used to generate the homology models of *Drosophila* chitinase catalytic domain fragments.

---

**Table 1. Summary of chitinase-like protein BLAST search**

<table>
<thead>
<tr>
<th>Query sequences</th>
<th>A.aegypti</th>
<th>A.gambiae</th>
<th>B.mori</th>
<th>Pocchleriae</th>
<th>M.sexta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hits</td>
<td>81</td>
<td>82</td>
<td>73</td>
<td>21</td>
<td>70</td>
</tr>
<tr>
<td>Hits of potential chitinase/chitinase-like</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>19</td>
<td>16</td>
</tr>
</tbody>
</table>

Different insect chitinase sequences were used as query sequences to BLAST the *Drosophila* genome database (http://www.fruitfly.org/blast/). The sequences used for BLAST were: *M. sexta* chitinase (GenBank accession no. A56596), *A. gambiae* chitinase (GenBank accession no. AAB87764), *Pocchleriae* morsitans chitinase (GenBank accession no. CAA77014), *A. aegypti* (GenBank accession no. AAB81849) and *B. mori* chitinase (GenBank accession no. BAC67246).

---

**Domain analysis**

Domain analysis of the potential *Drosophila* chitinase-like proteins showed that all of them contained at least one catalytic domain of family 18 chitinases. Some of them contained more than one catalytic domain. For example, the protein encoded by sequence Cht7 contains two catalytic domains and that encoded by sequence Cht10 contains four catalytic domains. The domain architecture of all 18 *Drosophila* chitinases and chitinase-like proteins is shown in Figure 1. The majority of these identified proteins appear to have signal peptides (and not signal anchors), suggesting an extracellular location for the corresponding proteins. Cht4, Cht5 and Cht12 have the typical insect chitinase domain architecture with a leader signal, one or more catalytic domains, and one or more chitin-binding domains. There is also a S–T-rich linker region between the catalytic domain and chitin-binding domain in Cht5 and a large S–T-rich region in the C-terminal of Cht6. Cht6 and Cht10 are two unusually large chitinases or chitinase-like proteins. Cht10 contains 2286 amino acid residues and Cht6 has 4498 residues. The smallest chitinase-like protein, Cht9 has only 368 residues.

Those proteins with DWEYP sequences are predicted to be active catalytically, including Cht1–Cht11, whereas those amino acid substitutions in those residues including Cht12–Cht14, and Imaginal Disc Growth Factor 1–4 (Idg1–Idg4) are predicted to be non-enzymatic and probably carbohydrate-binding proteins or lectins.

---

**Multiple sequence alignment and phylogenetic tree**

The catalytic domains and chitin-binding domains sequences were extracted from the amino acid sequences of the **C.fumiferana** (GenBank: AAM43792), **G.morsitans** (GenBank: AAL65401), **H.cunea** (GenBank: AAB47539), **L.longipalpis** (GenBank: AAN71763), **M.sexta** (GenBank: A56596), **Pocchleriae** (GenBank: CAA77014), **S.litura** (GenBank: BAB12678), **T.molitor** (GenBank: CAD31740) and the *Drosophila* sequences identified in this paper.

---

163
Table 2. Identified putative *Drosophila* chitinase-like proteins

<table>
<thead>
<tr>
<th>GenBank accession no.</th>
<th>Computed gene</th>
<th>Name</th>
<th>Length (aa)</th>
<th>Position of the conserved motif</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Region I</td>
</tr>
<tr>
<td>NP_523617.1</td>
<td>CG17682</td>
<td>Ch1</td>
<td>508</td>
<td>222</td>
</tr>
<tr>
<td>NP_477298.2</td>
<td>CG2054</td>
<td>Ch2</td>
<td>484</td>
<td>128</td>
</tr>
<tr>
<td>NP_523616.1</td>
<td>CG18140</td>
<td>Ch3</td>
<td>458</td>
<td>253</td>
</tr>
<tr>
<td>NP_524962.2</td>
<td>CG3986</td>
<td>Ch4</td>
<td>462</td>
<td>107</td>
</tr>
<tr>
<td>NP_650314.1</td>
<td>CG9307</td>
<td>Ch5</td>
<td>595</td>
<td>109</td>
</tr>
<tr>
<td>NP_572598.1</td>
<td>CG2989</td>
<td>Ch6</td>
<td>4498</td>
<td>133</td>
</tr>
<tr>
<td>NP_647768.2</td>
<td>CG1869</td>
<td>Ch7</td>
<td>983</td>
<td>206</td>
</tr>
<tr>
<td>NP_611542.1</td>
<td>CG9357</td>
<td>Ch8</td>
<td>460</td>
<td>98</td>
</tr>
<tr>
<td>NP_611543.3</td>
<td>CG10531</td>
<td>Ch9</td>
<td>368</td>
<td>102</td>
</tr>
<tr>
<td>EAA46011.1</td>
<td>CG18140-PA3</td>
<td>Ch10</td>
<td>2286</td>
<td>296</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1089</td>
</tr>
<tr>
<td>NP_572361.1</td>
<td>CG3044</td>
<td>Ch11</td>
<td>432</td>
<td>72</td>
</tr>
<tr>
<td>NP_726022.1</td>
<td>CG30293</td>
<td>Ch12</td>
<td>470</td>
<td>105</td>
</tr>
<tr>
<td>NP_477081.1</td>
<td>CG5210</td>
<td>Ch13</td>
<td>452</td>
<td>114</td>
</tr>
<tr>
<td>NP_611332.1</td>
<td>CG5154</td>
<td>Ch14</td>
<td>443</td>
<td>111</td>
</tr>
<tr>
<td>NP_477258.1</td>
<td>CG4472</td>
<td>Idgf1</td>
<td>439</td>
<td>103</td>
</tr>
<tr>
<td>NP_477257.2</td>
<td>CG4475</td>
<td>Idgf2</td>
<td>440</td>
<td>105</td>
</tr>
<tr>
<td>NP_723967.1</td>
<td>CG4559</td>
<td>Idgf3</td>
<td>441</td>
<td>108</td>
</tr>
<tr>
<td>NP_723734.1</td>
<td>CG1780</td>
<td>Idgf4</td>
<td>442</td>
<td>108</td>
</tr>
</tbody>
</table>

A Perl program was used to screen the BLAST hits. Four conserved regions were identified using the program. The positions show the last residues of the conserved regions. Conserved region I is KXXXXXGGW, region II is FDGXDLDWEYP, which is the catalytic region of the enzyme with residue E being a putative proton donor, region III is MXYDXXG and region IV is GXXXWXXDXD. X is a non-specific amino acid. aa, amino acid.

Fig. 1. Domain architecture of identified *Drosophila* chitinase-like proteins. The programs Pfam and SMART were used to analyze the identified sequences. A total of 18 protein sequences were analyzed. All the sequences contain at least one catalytic domain.

18 identified putative *Drosophila* chitinase-like proteins and were aligned with one another. Regions I–IV were very highly conserved among all the catalytic domains (Fig. 2). Figure 3 shows the alignment of the chitin-binding domains. There are six cysteine residues in the chitin-binding domains. Four of these, which are located in the middle of this domain, were completely conserved in all the sequences. The other two, near the ends, were highly (but not universally) conserved.

Genes encoding family 18 chitinases have been cloned from several insect species. To establish the phylogenetic relationships among these insect chitinase genes, a phylogenetic tree was constructed based on their protein sequences (Fig. 4). The insect chitinases grouped into two major clusters, one of which included most of the chitinases from Lepidoptera and the other consisting mostly of enzymes from Diptera. Twelve of the *Drosophila* chitinase-like proteins, Cht4, Cht6, Cht4, Cht8, Cht9 and Cht14 and Idgf1–Idgf4, group together the Dipteran chitinases. Three of the enzymes, Cht1, Cht3 and Cht10, are closely related to *T. molitor* chitinase.
Computational identification of novel chitinase-like proteins

Fig. 2. Multiple sequence alignment of the catalytic domains of Drosophila chitinase-like proteins. Catalytic domain sequences were extracted from the identified Drosophila chitinase-like proteins and the ClustalW program was used. The four conserved regions are boxed.

Homology modeling

All 18 of the Drosophila chitinase-like protein sequences were submitted to the automated comparative protein modeling server (http://www.expasy.org/swissmod/SWISS-MODEL.html). Those proteins with the highest identity in amino acid sequence were chosen as modeling templates. In Figure 5, Cht2 (NP_477298) is a published chitinase sequence (de la Vega et al., 1998), and Cht12 (NP_726022) is a novel chitinase-like protein which was determined in this study. These two have similar structure and both contain an

<table>
<thead>
<tr>
<th>Consensus</th>
<th>MPLFPIHNVVLYSDDNFDVWDLCGCFLQDSEQ</th>
<th>165</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cht1_1</td>
<td>(1) LVPQVHTR--NKRKYDCLNR-------TVQPVHNLNSDDNFDVWDLCGCFLQDSEQ</td>
<td>69</td>
</tr>
<tr>
<td>Cht3_2</td>
<td>(1) SNQQRPVPIKQQKLYSDDNFDVWDLCGCFLQDSEQ</td>
<td>69</td>
</tr>
<tr>
<td>Cht4_5</td>
<td>(1) LVPQVHTR--NKRKYDCLNR-------TVQPVHNLNSDDNFDVWDLCGCFLQDSEQ</td>
<td>69</td>
</tr>
<tr>
<td>Cht6_8</td>
<td>(1) LVPQVHTR--NKRKYDCLNR-------TVQPVHNLNSDDNFDVWDLCGCFLQDSEQ</td>
<td>69</td>
</tr>
<tr>
<td>Cht8_1</td>
<td>(1) LVPQVHTR--NKRKYDCLNR-------TVQPVHNLNSDDNFDVWDLCGCFLQDSEQ</td>
<td>69</td>
</tr>
<tr>
<td>Cht10_14</td>
<td>(1) LVPQVHTR--NKRKYDCLNR-------TVQPVHNLNSDDNFDVWDLCGCFLQDSEQ</td>
<td>69</td>
</tr>
</tbody>
</table>

Fig. 3. Multiple sequence alignment of chitin-binding domains of Drosophila chitinase-like proteins. Chitin-binding domain sequences were extracted from the identified Drosophila chitinases and analyzed using the ClustalW program. The six cysteine residues are conserved among the chitin-binding domains of Drosophila chitinases.
Phylogenetic analysis of *Drosophila* and other insect chitinase-like proteins. A phylogenetic tree was constructed by the neighbor-joining method based on the protein sequences of 25 insect chitinases (Saitou and Nei, 1987). Sequences used for the analysis are: *A. aegypti* (GenBank: AAB81849), *A. gambiae* (GenBank: AAB87764), *B. mori* 1 (GenBank: AAB47538), *B. mori* 2 (GenBank: BAC67246), *Chelonus* sp. Venom (GenBank: A53918), *C. tentans* (GenBank: CAA73685), *C. juniferana* (GenBank: AAM43792), *G. morsitans* (GenBank: AAL65401), *H. cunea* (GenBank: AAB47539), *L. longipalpis* (GenBank: AAN71763), *M. sexta* (GenBank: A56596), *P. cochleriae* (GenBank: CAA77014), *S. littura* (GenBank: BAB12678) and *T. molitor* (GenBank: CAD31740).

(β/α)₈ TIM barrel structure. Other known and novel chitinases also contain this structure (the structures are shown in the supplementary data).

**DISCUSSION**

Chitinases are glycosylhydrolases that catalyze the random hydrolysis of the β-(1,4)-glycosidic bonds in chitin. The insoluble polymeric chitin is digested and becomes soluble, yielding low molecular mass multimers of GlcNAc, such as chitotetraose, chitotriose and chitobiose (Kramer and Koga, 1986; Samuels and Reynolds, 1993). The multimers of GlcNAc are subsequently hydrolyzed to N-acetylgulcosamine by β-N-acetylgulcosaminidase (Fukamizo and Kramer, 1985; Filho et al., 2002). In insects, chitinases are present in molting fluids, venom glands and midguts. Chitinases participate in the periodic shedding of old exoskeletons and the turnover of peritrophic membranes.

Most of the previous studies of lepidopteran insects have identified only one or two genes encoding chitinases consistent with two major locations of chitin synthesis in these insects. The identification of 3–4 different chitinase gene sequences in dipterans (*Aedes, Anopheles* and *Drosophila*, see de la Vega et al., 1998) prompted us to carry out an extensive search of the *Drosophila* genome for chitinase-related protein genes. The computational identification of 18 genes for chitinase-like proteins in *Drosophila* reported here reveals for the first time that there is a rather high degree of complexity in the chitinase-like protein gene family in dipterans. Chitin is found in a number of different structures besides cuticle and PM. For example, chitin is associated with wing hinges, trachea and mouth hooks in *Drosophila* (Wilson and Cryan, 1997). Chitin synthesis occurs throughout insect development including embryonic, larval, pupal and adult stages. It will be interesting to determine whether there is developmental and tissue-specific regulation of expression of different chitinase-like protein genes in *Drosophila*. 
Computational identification of novel chitinase-like proteins

Fig. 5. Homology modeling of Drosophila chitinase catalytic domain. The SWISS-MODEL program was used to generate the models. (A) Cht2 (NP_477298). Human chitotriosidase (Fusetti et al., 2002) was used as template; (B) Cht12 (NP_726022). The templates were Mammalian lectin, Ym1 (Sun et al., 2001) and human chitotriosidase (Fusetti et al., 2002).

The search of the Drosophila EST database has revealed that each of the 18 chitinase-like genes is functional in this insect even though there are no formal reports on the expression of any one of these 18 genes. This computational search has led to the identification for the first time of all the members of the family of proteins closely related to insect chitinases. The putative Drosophila chitinases belong to glycosylhydrolase family 18 and contain conserved regions characteristic of this family. Cht12, Cht13 and Cht14 lack the glutamic acid residue that has been identified as the proton donor in the catalytic mechanism and these proteins are predicted to be devoid of enzyme activity. Most likely, these proteins are lectins that will bind to chitin or other carbohydrates containing N-acetylglucosamine. They may be involved in cell-to-cell communication or in insect immunity. The glutamate residue is replaced by glutamine in all the Idgf genes (Idgf1–Idgf4). It is believed that the Idgfs have evolved from an ancestral chitinase and have growth-promoting function (Kawamura et al., 1999).

Homology modeling showed that all these chitinases most probably contain the \((\alpha/\beta)_8\) TIM barrel structure, suggesting that this architecture may be crucial for the function of this enzyme. Indeed, several kinds of glycosylhydrolases have a similar \((\alpha/\beta)_8\) TIM barrel structure, which may indicate that this structure may be important for the hydrolysis of polymeric substrates. In contrast, the enzyme N-acetylglucosaminidase, which utilizes an exoglycosidic instead of an endoglycosidic mechanism, has a different domain organization and belongs to a different protein family (family 89 of the glycosylhydrolases; http://afmb.cnrs-mrs.fr/CAZY).

Two questions remain about the evolutionary relationships among Drosophila chitinase-like proteins. First, why are there so many chitinases or chitinase-like proteins in Drosophila? In the soil fungus, Trichoderma virens, three members of 42 kDa chitinases genes, Tv-ech1, Tv-ech2 and Tv-ech3, were identified. These three genes exhibit distinctive structural features and are regulated in response to different environments and/or developmental stages or expressed in different subcellular localizations (Kim et al., 2002). In insects, B.mori was the first species from which multiple chitinase genes were identified (Kim et al., 1998; Mikitani et al., 2000; Abdel-Banat and Koga, 2001; Daimon et al., 2003). Some of the putative chitinase genes in Drosophila may be pseudogenes. Harrison et al. (2003) found approximately 100 pseudogenes in the Drosophila genome. However, none of the 18 candidates identified in this study was included in those putative pseudogenes. Furthermore, there are EST’s corresponding to each of these 18 chitinase-like genes.

Second, why are some of the chitinases so large compared with other insect chitinases? Cht10 has a multidomain architecture, including four family 18 chitinolytic domains and four chitin-binding domains and Cht7 has duplicated catalytic domains. Cht6 has the typical three domain architecture and a long tail that is rich in Threonine and Serine residues. This protein is 4498 amino acids long with a predicted molecular mass of 493 kDa. A rather large chitinase with multiple catalytic domains and chitin-binding domains has been reported previously for the yellow mealworm beetle, T.molitor (Royer et al., 2002).

Another interesting prediction from the computational analysis of the chitinase-like proteins is that only six of them are expected to be extracellular, while the others, including the one with four catalytic domains, is expected to be intracellular. Although we have identified multiple Drosophila chitinase-like genes, in vivo experiments and in vitro studies with purified chitinases corresponding to each of these genes are necessary to establish their distinctive activities and biological roles. Specific primers could be designed and used to amplify these potential Drosophila chitinase-like protein cDNAs. Protein expression and chitinase activity assays...
as well as chitin-binding studies are needed to confirm that the sequences identified in this study are indeed chitinase or chitin-binding protein genes.

ACKNOWLEDGEMENTS
This project was supported by NIH grant P20 RR16475 from the BRIN Program of the National Center for Research Resources and by NSF grant IBN-0316963 to S.M. and K.J.K.

REFERENCES


