

Proteome Analysis of Urticating Setae From *Thaumetopoea pityocampa* (Lepidoptera: Notodontidae)

L. Berardi,¹ M. Pivato,¹ G. Arrigoni,^{2,3} E. Mitali,¹ A. R. Trentin,¹ M. Olivieri,⁴ C. Kerdelhué,⁵ F. Dorkeld,⁵ S. Nidelet,^{5,6} E. Dubois,⁶ A. Battisti,^{1,7} and A. Masi¹

¹Department of Agronomy Food Natural Resources Animals and Environment (DAFNAE), University of Padova, Viale dell'Università 16, 35020 Legnaro, Padova, Italy (laura.berardi2@gmail.com; micaela.pivato@gmail.com; ettore.mitali@gmail.com; annarita.trentin@unipd.it; andrea.battisti@unipd.it; antonio.masi@unipd.it), ²Proteomics Center, University of Padova and Azienda Ospedaliera di Padova, via G. Orus 2/B, 35129 Padova, Padova, Italy (giorgio.arrigoni@unipd.it), ³Department of Biomedical Sciences, University of Padova, via U. Bassi 58/B, 35131 Padova, Padova, Italy, ⁴Occupational Medicine, University Hospital of Verona, Piazzale L.A. Scuro 10, 37134 Verona, Verona, Italy (mario.olivieri@ospedaleuniverona.it), ⁵INRA, UMR CBGP (INRA/CIRAD/IRD/Montpellier Supagro), 755 avenue du Campus Agropolis, CS30016, F-34988 Montpellier-sur-Lez Cedex, Montpellier-sur-Lez, France (carole.kerdelhue@inra.fr; franck.dorkeld@inra.fr; sabine.nidelet@inra.fr), ⁶MGX-Montpellier GenomiX, Institut de Génétique Fonctionnelle, 141 rue de la cardonille, 34094 Montpellier Cedex 05, Montpellier, France (emeric.dubois@gmx.cnrs.fr), and ⁷Corresponding author, e-mail: andrea.battisti@unipd.it

Subject Editor: Brenda Beerntsen

Received 18 April 2017; Editorial decision 23 June 2017

Abstract

Thaumetopoea pityocampa (Denis & Schiffermüller) (Lepidoptera: Notodontidae) is harmful to conifer trees because of defoliation and to public health because of the release of urticating setae from the caterpillars. Contact with setae by humans and domestic animals induces dermatitis, usually localized to the exposed areas. Recent studies demonstrated the presence of a complex urticating mechanism where proteins present in the setae may play a role as activators of immune responses. Yet, limited information is available at present about the proteins occurring in the setae of *T. pityocampa*. Using a refined method for protein extraction from the setae, and a combination of liquid chromatography tandem-mass spectrometry (LC-MS/MS), de novo assembly of transcriptomic data, and sequence similarity searches, an extensive data set of 353 proteins was obtained. These were further categorized by molecular function, biological process, and cellular location. All the 353 proteins identified were found to match through BLAST search with at least one Lepidoptera sequence available in databases. We found the previously known allergens Tha p 1 and Tha p 2 described from *T. pityocampa*, as well as enzymes involved in chitin biosynthesis, one of the principal components of the setae, and serine proteases that were responsible for inflammatory and allergic reactions in other urticating Lepidoptera. This new proteomic database may allow for a better understanding of the complexity of allergenic reactions due to *T. pityocampa* and to other Lepidoptera sharing similar defense systems.

Key words: caterpillar, hair, allergen, transcriptome

Thaumetopoea pityocampa (Denis & Schiffermüller, 1775) (Lepidoptera: Notodontidae), a common pest of conifer trees, is also a threat to public health in southern Europe, the Middle East, and northern Africa (Roques 2015). Caterpillars feed on needles of different species of pine and cedar and use special hairs (setae) to protect themselves against predators; when disturbed, they release the setae that can be blown by the wind over several kilometers, and persist in the environment for a long time (Battisti et al. 2017). The nature of the setae is thus very different from other defensive hairs described from caterpillars of some Lepidoptera, which are part of the integument and require direct contact with the caterpillar to cause the reaction. However, the urticating setae of *T. pityocampa*

are similar to those produced by other Lepidoptera (e.g., the brown tail moths *Euproctis* spp. and the *Hylesia* spp. moths) and some American spiders (Theraphosidae) (Battisti et al. 2011; Fig. 1).

The seta penetration in the skin and the mucosa is guided by hook-like structures, and it results in strong and painful reactions in humans (Fig. 1). Domestic and farm animals can be also affected, either through direct contact with caterpillars or their tents, or by ingesting fodder contaminated with setae (Battisti et al. 2017). In Europe, the incidence in humans in outbreak areas was as high as 18%, and 60% of veterinary practitioners in France had experience of reaction to setae (Moneo et al. 2015). Among the forestry workers of Verona (Italy), chainsaw operators were

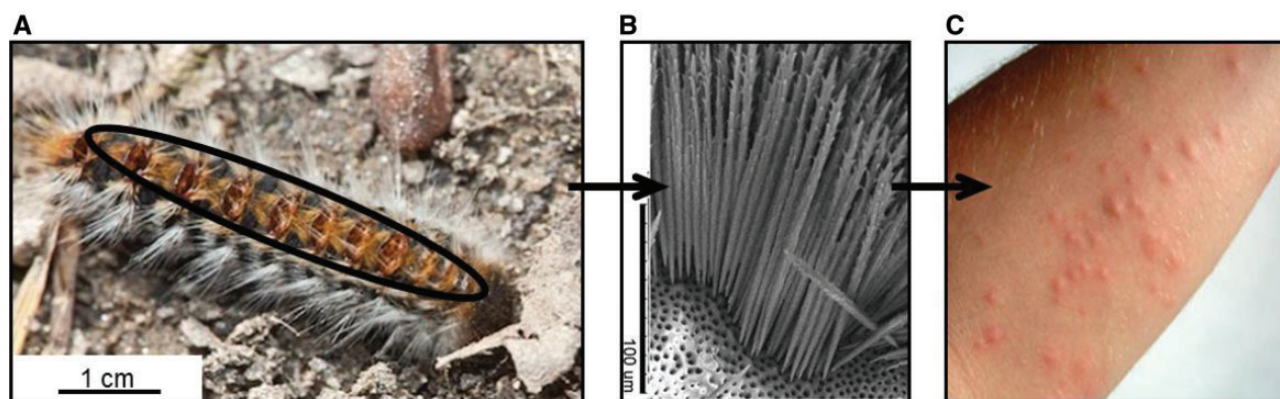


Fig. 1. *Thaumetopoea pityocampa*: (A) caterpillar L5 instar, the oval includes the integument dorsal areas of the abdomen where the urticating setae occur; (B) close-up of urticating setae, with a part of them removed from their sockets; (C) typical reaction on human skin.

identified as being most at risk of skin and respiratory disorders, suggesting the need for a thorough medical surveillance of this occupational allergy (Olivieri et al. 2012). In Australia, where the setae produced by *Ochrogaster lunifer* Herrich-Schäffer are associated with equine amnionitis and fetal loss on horse farms (Perkins et al. 2016), the cost of a similar moth species to equine health has been estimated at AU\$27–43 million per annum (Carrick et al. 2014).

Proteins associated with the urticating setae of Lepidoptera have been described by Bleumink et al. (1982) for the caterpillars of *Euproctis* spp. (Lepidoptera: Erebididae) and by Lundberg et al. (2007) for the female moths of *Hylesia metabus* (Cramer) (Lepidoptera: Saturniidae). In both cases, an enzymatic activity (serine protease) seems to occur and appears to be responsible for the typical reaction associated with the contact of human skin with the setae. The presence of several serine proteases has been confirmed in recent proteomic studies by Cabrera et al. (2016, 2017) in the setae isolated from the female moths of *H. metabus*, which are putatively thought to protect the eggs and the young caterpillars from vertebrate predation. This family of proteases represents 65% of the total proteins extracted from the setae of *H. metabus*.

Proteins associated with setae of *T. pityocampa* were first described by Lamy et al. (1985, 1986). Setae were immersed in liquid nitrogen and then crushed in a saline solution. In total, 16 proteins were detected, and one of them, a 28 kDa protein composed of two subunits of 13 and 15 kDa, was recognized as an antigen and called thaumetopoein. Later, Moneo et al. (2003) obtained a whole caterpillar extract after maceration in a saline solution, agitation for 24 h, and precipitation of the supernatant in ethanol, from which they were able to identify a 15 kDa protein named Tha p 1 because of its putative allergenic potential. Rodríguez-Mahillo et al. (2012) extracted proteins from setae immersed in PBS, followed by sonication on ice. This method allowed the identification of about 70 proteins, seven of which were recognized by sera of exposed persons. One of them (14 kDa) was found to be a major allergen different from Tha p 1 and was then called Tha p 2. Cultured human lymphocytes of persons previously exposed to a closely related species of *Thaumetopoea pinivora* (Treitschke), proliferated when in contact with setae or their extracts, indicating the occurrence of a cell-mediated immune response (Holm et al. 2014). It has also been speculated that chitin, considered to be associated with proteins in the integument (Andersen et al. 1994), may have an adjuvant role in immune reactions (Battisti et al. 2011). Using a protocol adapted from Zhang and Zhu (2006), it was possible to show that chitin makes up to 37% of the urticating setae of *T. pityocampa* (A.B.,

unpublished data). This value suggested that a considerable part of the remaining percentage might be composed of proteins.

Genomic and transcriptomic data are valuable resources to finely analyze the sequences of putatively important proteins, and to search for homologies in model species or in other well-studied organisms. Concerning *T. pityocampa*, a partial transcriptome based on 454 sequencing technology was recently released (Gschloessl et al. 2014). Yet, it was acknowledged that the corresponding data were incomplete and that many transcripts were probably missing, in particular the ones with low levels of expression, or for which expression was restricted to specific organs and tissues and thus possibly undersampled. More, the corresponding resource was obtained from a peculiar population of *T. pityocampa* showing a shifted phenology and occurring only in Portugal. Obtaining an extensive transcriptomic resource from different life stages and sequencing using a high-throughput technology is thus now required.

The aims of this study were to de novo assemble and release a high-quality specific transcriptome and to make a survey of the protein composition of the urticating setae of the caterpillars of *T. pityocampa*. A proteomic characterization of the setae would help to identify the putative allergen(s) and elucidate the mechanisms of the allergic reaction.

Materials and Methods

Transcriptome Analysis

We produced a de novo transcriptome assembly for *T. pityocampa*, obtained from RNA of two *T. pityocampa* populations from northern Italy (Cimolais, 12° 27' E, 46° 19' N; Tregnago, 11° 09' E, 45° 30' N) reared under laboratory conditions (Berardi et al. 2015b). Note that one of these populations was the same as for the proteomic analysis of the setae (see below). Caterpillars, pupae, and adults were sampled during the rearing and preserved in RNAlater (Thermo Fisher Scientific, Pittsburgh, CA) at -80°C . A sample of eight caterpillars from different instars, four pupae, and four adults was used for the RNA extraction using a Trizol extraction procedure. RNA quality was checked through migration on an agarose gel, and RNA purity was evaluated by NanoDrop analysis. Whenever the 260/280 OD ratio was <1.7 , samples were purified using the Qiagen RNeasy mini kit. RNA concentrations were estimated using the Qubit procedure (Quant-it RNA assay kit, Thermo Fisher Scientific, www.thermofisher.com, accessed 10 June 2017). Fifteen libraries were constructed using the TruSeq stranded mRNA sample prep kit (Illumina, ref. RS-122-2101, San Diego, CA), according to the manufacturer's instructions. Briefly, poly-A RNAs were purified

using oligo-d(T) magnetic beads. The selected RNAs were fragmented and reverse transcribed using random hexamers, SuperScript II (Life Technologies, ref. 18064-014, Thermo Fisher Scientific) and Actinomycin D. Double-stranded cDNAs were adenylated at their 3' ends and ligation was then performed using Illumina's indexed adapters. Ligated cDNAs were finally amplified by PCR. Libraries were validated using a DNA1000 chip (Agilent, ref. 5067-1504, Santa Clara, CA) on an Agilent Bioanalyzer and quantified by qPCR. The 15 libraries were sequenced using 3 lanes of an Illumina HiSeq2000, using the paired end protocol (2 by 100 nt) at the Montpellier GenomiX (MGX) sequencing core facility. Image analysis and base calling were performed using the HiSeq Control Software and Real-Time Analysis component provided by Illumina. Demultiplexing was performed using Illumina's sequencing analysis software (CASAVA 1.8.2). The quality of the data was assessed using FastQC and FastQ Screen from the Babraham Institute, and the Illumina software SAV (Sequence Analysis Viewer).

Reads produced by high-throughput sequencing must be preprocessed to retain only high-quality reads used in reconstruction. The raw sequences were trimmed using Trimmomatic (Bolger et al. 2014) using the following parameters: ILLUMINACLIP: adaptors_file.fa: 2: 40: 15; HEADCROP: 12; SLIDINGWINDOW: 4: 15 and MINLEN: 30 combined with PRINSEQ-lite (Schmieder and Edwards 2011) to eliminate polyA tails. The cleaned reads, obtained after the quality filtering step described above, were used for de novo transcript assembly using Trinity (Grabherr et al. 2011), with the normalization option on and default kmer value. The completeness of the transcriptome assembly was assessed using the Core Eukaryotic Genes Mapping Approach (CEGMA, Parra et al. 2007). After alignment with hidden Markov models, we thereby obtained the percentage of core genes that were found complete or partial in the dataset (i.e., transcript reconstruction). To further evaluate the quality of the assembly, we used the following metrics: number of transcripts assembled, total length of the assembly (in bp), mean transcript length, N50 value, and number of transcripts longer than 1 kb. We mapped the filtered reads back to assembled transcripts, and we also used the number of such reads (RBMT) as a quality assessment estimator. We used TransDecoder, provided by the Trinity package, to identify potential coding regions from the assembled transcripts. TransDecoder identifies candidate protein-coding regions based on nucleotide composition, open reading frame (ORF) length, and Pfam domain content.

Setae Collection

Twenty tents containing colonies of *T. pityocampa* were collected in March 2013 from *Pinus nigra* trees in Tregnago (the same site from which individuals were collected for the transcriptome analysis described above). Each caterpillar (either in the fourth or fifth instar) was removed from the tent in a vented hood and put in an Eppendorf tube at -20°C . The setae were then removed from thawing caterpillars with forceps using a stereomicroscope at $40\times$ magnification, inside a hood, with special attention to avoid wounds that could contaminate the setae with body fluids. If such a contamination occurred, the material was excluded from subsequent analysis. Because one caterpillar carries about 1 million setae (Petrucco-Toffolo et al. 2014), by collecting all the setae in one caterpillar from each colony, it was estimated that the sample consisted of about 20 million urticating setae. Setae were collected into Eppendorf tubes with 100 μl of acetone, which allows a quick detachment of the setae from the forceps. Setae were frozen in liquid nitrogen and stored in Eppendorf tubes at -20°C for a maximum of 1 wk before protein extraction (Fig. 2).

Protein Extraction and Digestion

Setae in Eppendorf tubes with acetone were frozen in liquid nitrogen and sonicated on ice three times, 3 min each at 18 W, then manually crushed and centrifuged for 2 min at 28,672 g-force. Following acetone removal, the pellet was put in 1.5 ml extraction buffer (50 mM Hepes pH 8, 1% Triton X-100, 1 M NaCl, and 1 mM PMSF/Benzamidine). The extract was then stirred for 4 h at room temperature to allow protein solubilization, and centrifuged for 20 min at 21,952 g-force at 4°C . One part of the clear supernatant was transferred in another tube and four parts of ice-cold acetone were added (final 80% acetone concentration); following overnight protein precipitation at -20°C , samples were centrifuged again at 16,128 g-force for 10 min. The supernatant was discarded, while the pellet was washed with 80% cold acetone and 20% water and suspended in Laemmli (1970) solution. Protein content was measured according to the Bradford method (Kruger 1994). Electrophoresis was carried out on a 13% acrylamide gel, and proteins were visualized by Coomassie staining (Fig. 2). The lanes were cut into bands, which were then placed in separate Eppendorf tubes, washed several times with 50 mM ammonium bicarbonate pH 8.0 and dried under vacuum after a short wash with 100% acetonitrile. Cysteines were reduced with 10 mM freshly prepared dithiothreitol in 50 mM ammonium bicarbonate for 1 h at 56°C , and alkylated with 55 mM iodoacetamide in 50 mM ammonium bicarbonate for 45 min at room temperature in the dark. Gel pieces were then washed with alternate steps of ammonium bicarbonate and acetonitrile, and finally dried under vacuum. Proteins were digested in situ with sequencing grade modified trypsin (Promega, Madison, WI) at 37°C overnight (12.5 ng/ μl trypsin in 50 mM ammonium bicarbonate). Peptides were extracted with three steps of 50% acetonitrile/0.1% formic acid. Extracted peptides were dried under vacuum and stored at -20°C until liquid chromatography tandem-mass spectrometry (LC-MS/MS) analysis was performed.

LC-MS/MS Analysis

The extract was suspended in $\text{H}_2\text{O}/0.1\%$ formic acid and analyzed by LC-MS/MS (Tolin et al. 2012). The MS analyses were conducted with a LTQ-Orbitrap XL mass spectrometer (Thermo Fisher Scientific) coupled online with a nano-HPLC Ultimate 3000 (Dionex—Thermo Fisher Scientific). Samples were loaded onto a 10-cm chromatographic column packed into a pico-frit (75 μm id, 15 μm tip, New Objectives) with C18 material (ReproSil, 300 $\text{\AA}$, 3 μm). Peptides were separated at a flow rate of 250 nl/min using a linear gradient of acetonitrile/0.1% formic acid from 3% to 50% in 30 min. Spray voltage was set to 1.3 kV with an ion source capillary temperature of 200°C . The instrument was operated in data-dependent mode with a top-ten acquisition method (a full scan at 60,000 resolutions on the Orbitrap followed by MS/MS scans of the 10 most intense ions acquired in the linear trap). Raw LC-MS/MS files were analyzed using Proteome Discoverer 1.4 (Thermo Fisher Scientific), connected to a Mascot Search Engine server (Matrix Science, London, United Kingdom). Raw LC-MS/MS spectra were searched against the transcriptome database obtained for *T. pityocampa* and concatenated with a database of the most common protein contaminants found in proteomics experiments. Enzyme specificity was set to trypsin with one missed cleavage, and peptide and fragment tolerance was set to 10 ppm and 0.6 Da, respectively. Carbamidomethylation of cysteine residues and methionine oxidation were set as fixed and variable modification, respectively. Based on the search against the corresponding randomized database, the algorithm Percolator was used to assess false discovery rates (FDR):

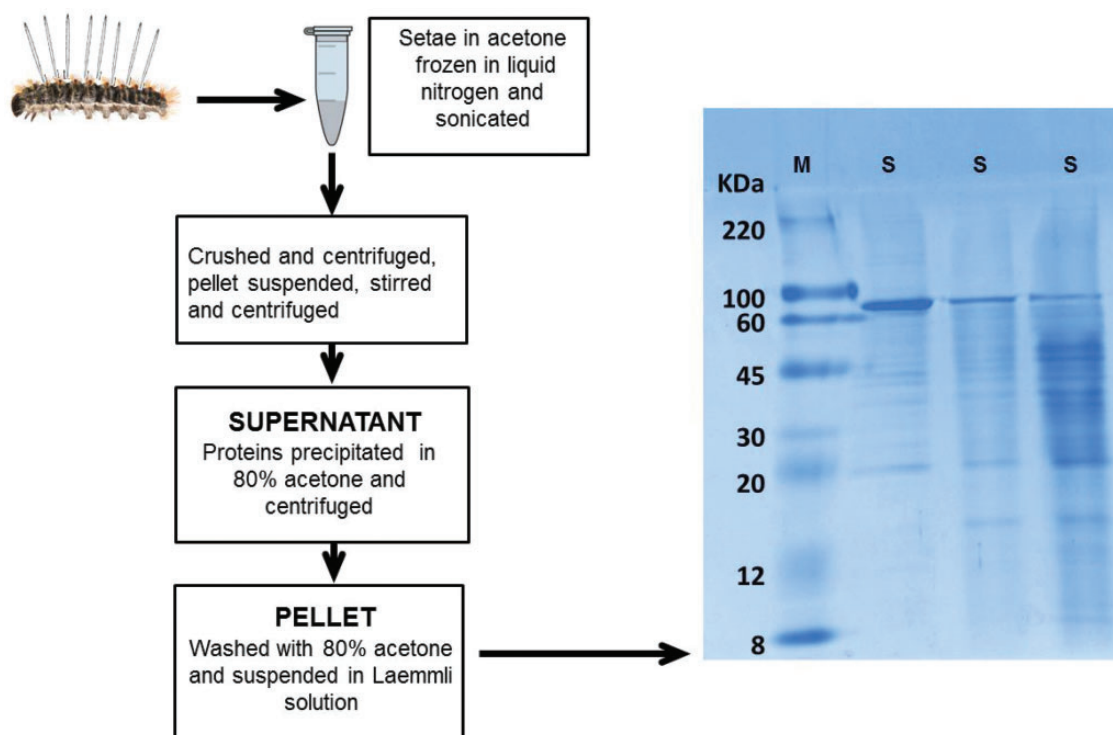


Fig. 2. Experimental flow chart for protein extraction from *T. pityocampa* urticating setae and SDS-PAGE separation. All bands of three samples (S) in stained gels were cut in bands which were then subjected to in-gel trypsin digestion followed by LC-MS/MS analysis. M—marker.

only proteins identified with a q value < 0.01 (99% confidence) and at least two unique peptides were considered as positive hits. Data were prefiltered to exclude MS/MS spectra containing < 5 peaks or with a total ion count < 50 .

Bioinformatics Analyses

Blast2GO 3.1 (Conesa et al. 2005) was used to analyze gene ontology (GO) at UniProtKB, www.uniprot.org, accessed 10 June 2017) terms after BLAST, mapping, and annotation of the amino acid sequences. The BLAST analyses against the NCBI BioSystems database permitted the identification of possible homologous proteins. Protein sequences found in the LC-MS/MS analysis were converted in FASTA format and analyzed (blastp) with the threshold E-value cutoff at $1.0E-3$ and number of BLAST hits, 4. Mapping and annotation were performed with the default values, and after that, protein sequences were assigned into three standard classifications: molecular function, biological process, and cellular component, and summarized according to GO (Gene Ontology) criteria. The metabolic pathways were analyzed with KEGG (Kyoto Encyclopedia of Genes and Genomes; Kanehisa et al. 2012) implemented in the Blast2GO program.

Results

For each library (i.e., each development stage), from 12 to 37 million clean read pairs were obtained after filtering. The total number of quality filtered reads was 532,900,683. The de novo transcriptome assembly obtained using Trinity consisted of 233,098 transcripts that were clustered into 188,610 potentially different genes (i.e., unigenes). In particular, the high number of core genes identified with the CEGMA procedure (94.35% when considering only

complete genes, and 98.39% when considering partial and complete genes), and the high proportion for RBMT (75%) suggested that the reconstruction was rather robust. Transcripts reconstructed with Trinity were processed using TransDecoder, which retrieved 41,933 coding sequences and 24,992 complete peptides.

The protein electrophoresis revealed the presence of many bands, suggesting a relative complex protein composition even if the proteome was only obtained from a fraction of larval tissues, i.e., setae (Fig. 2). The LC-MS/MS analysis and database search against the reference transcriptome identified a total of 353 proteins. All 353 proteins identified were found to match through BLAST search with at least one Lepidoptera sequence available in the NCBI database. Among the 353 identified proteins, 120 showed a high degree of identity to *Bombyx mori* (L.) proteins, 34 to *Papilio xuthus* L., 32 to *Danaus plexippus* (L.), and 11 to *Papilio polytes* L. proteins. The remaining proteins were associated with other Lepidoptera (Supp. Table 1 [online only]). The BLAST analyses identified the proteins Tha p 1 and Tha p 2 of *T. pityocampa* (Supp. Table 1 [online only]).

Based on GO functional annotations and categories, 224 of the 353 identified proteins were annotated with at least one common GO. The cellular location graph indicated that the majority of urticating setae proteins belonged to the categories of cell, cell part, organelle, macromolecular complex, and membrane (Fig. 3A). A small percentage corresponded to membrane proteins and extracellular regions. Within the classification of biological process, a large number of proteins were assigned to metabolic and cellular or single-organism process (Fig. 3B). The Blast2GO analyses detected two principal functions: catalytic and binding role (Fig. 3C). Looking deeper into the KEGG pathway analyses, enzymes were mapped onto different metabolic pathways, showing that particular proteins were likely to be involved in gluconeogenesis and glutathione metabolism, like glucose-6-phosphate isomerase (5.3.1.9) or

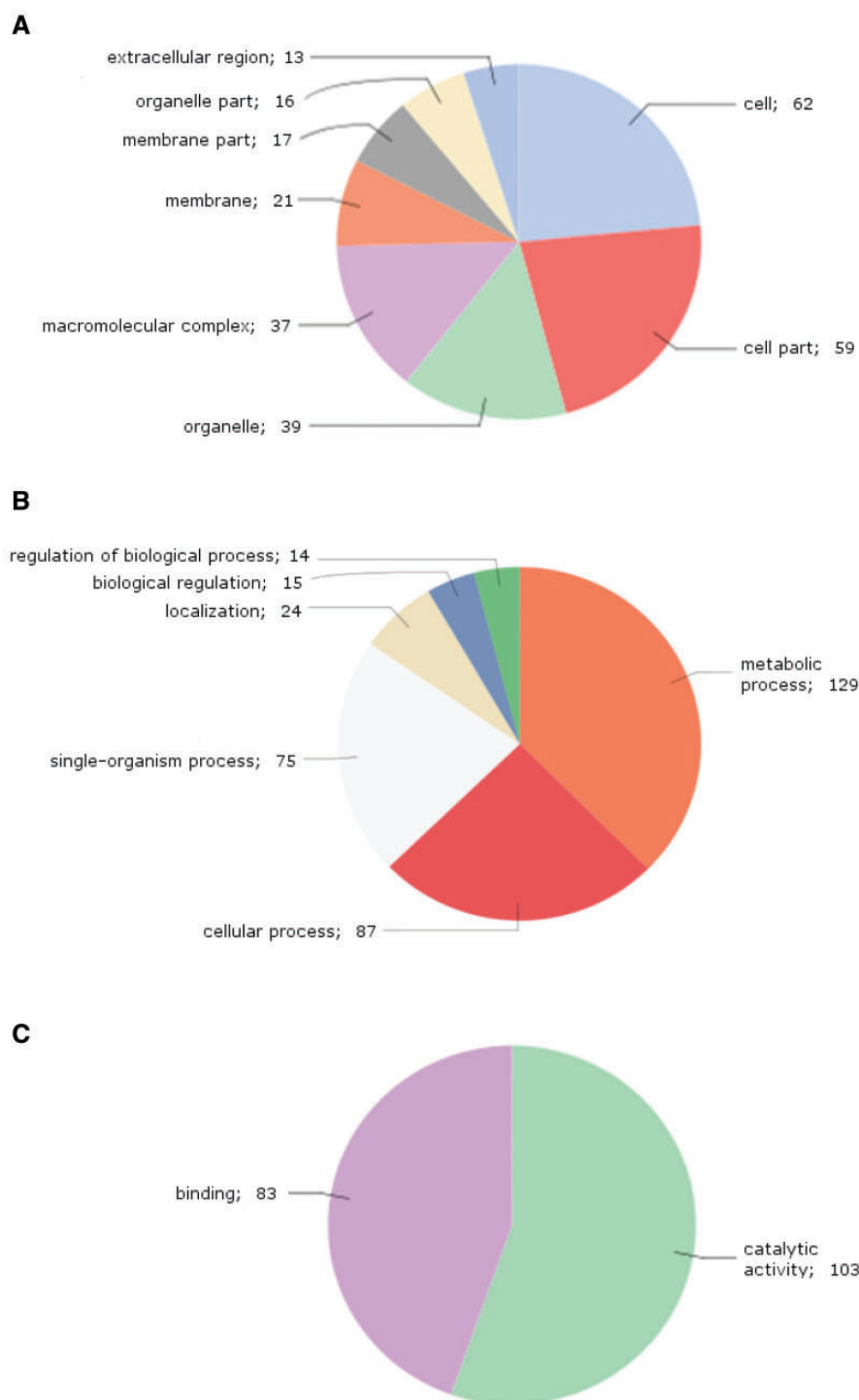


Fig. 3. Proteins as cellular component (A), biological process (B), and molecular function (C). Numbers indicate the proteins within each category (see text for details).

glyceraldehyde-3-phosphate dehydrogenase (1.2.1.12). Members of the chitin production pathway were significantly overrepresented, as well as several serine proteases.

Discussion

We here provide a comprehensive transcriptome assembly for Italian populations of *T. pityocampa*. As the expressed transcripts

were obtained from all developmental stages (from larva to adult, through most caterpillar instars and pupae), and deep sequenced using Illumina technology, we obtained a very valuable resource to analyze the resulting proteomic data. All quality assessments suggested that the transcriptomic resource produced here was of good quality and reached a high level of completeness, as suggested by the high proportion of core genes identified through the CEGMA analysis.

Because of the production of this transcriptomic database of *T. pityocampa*, new information about the quality and number of proteins associated with the setae was obtained. The large number of identifications is striking, if compared with the previous work from Rodríguez-Mahillo et al. (2012), where authors demonstrated that the setae contain a complex mixture of about 70 proteins and showed that one of them, Tha p 2, is allergenic in humans. This protein can be now considered as a major component of the seta proteome in *T. pityocampa*, as judged by the high score and the high number of peptide spectrum matches (PSMs) obtained from the LC-MS/MS analyses (see Supp. Table 1 [online only]). It is a glycine-, serine-, and cysteine-rich protein highly conserved among processionary moths (Berardi et al. 2015a). Another protein, called Tha p 1, was previously described as an allergen (Moneo et al. 2003), although it was obtained from an extract of the whole body, and it was not possible to know whether or not it was a setae protein. A later matching of a partial sequence of this protein indicated a possible similarity with chemosensory proteins of other insect species (Larsson and Backlund 2009; Moneo et al. 2015), although our finding associates it with setae, which do not have sensory functions (Battisti et al. 2011). The isolation of this putative allergen and its sequencing would be required to clarify the nature of Tha p 1.

The presence of a larger number of intracellular rather than extracellular proteins may have a dual explanation. First, cytoplasmic proteins of the epidermal cells producing the urticating setae are likely to be stored in the setae as remnants of this biosynthetic process. It is noteworthy that setae formation represents a major metabolic effort, leading to a large number of setae that have to be formed in a short time during each molt (up to 1,000,000 in a mature caterpillar of the *T. pityocampa*, Petruccio-Toffolo et al. 2014, and up to 2,000,000 in the bag-shelter moth *O. lunifer*, Perkins et al. 2016). Alternatively, it could be that hemolymph proteins can pass through membranes and enter the epidermis and, in turn, into the setae.

Because serine proteases have been found to have an important role in mediating the reactions of exposed humans and guinea pigs to the urticating setae of the caterpillars of *Euproctis* spp. (Bleumink et al. 1982) and of the female moths of *H. metabus* (Lundberg et al. 2007; Cabrera et al. 2016, 2017), it would be interesting to assess their role in the *T. pityocampa* system as well. Serine proteases do occur in the proteome of this species and show some similarity with proteins occurring in *Bombyx*, *Helicoverpa*, and *Mamestra* (Supp. Table 1 [online only]). Interestingly, these proteins are also present in the defensive droplets of Zygaenid caterpillars (Pentzold et al. 2016).

This full proteomic analysis of setae, highlighted by a comprehensive and specific transcriptomic resource, can be considered as a reference protein data set, and it will be of great importance to compare it with future studies on other species. It is likely that among the full list of proteins shown here, there are a few more that are directly responsible for immune reactions, as evidenced by Moneo et al. (2015). Enzymes involved in the chitin synthesis pathway were found in the setae. The chitin role in allergenic reaction is not clear yet, although it has been shown to stimulate in vitro human T-lymphocyte proliferation (Holm et al. 2014) and act as an adjuvant in immune reactions (Battisti et al. 2011). This new proteomic resource may enable a better understanding of the complexity of allergic reactions due to the *T. pityocampa* and to other Lepidoptera sharing similar defense systems. In the future, there is the need to correlate allergenic reactions to the identified proteins and to extend the extraction of the urticating setae to other species of Lepidoptera. Genomic and transcriptomic resources could be produced for these other species, as next-generation sequencing technologies now allow to relatively easily obtain high-quality resources.

Acknowledgments

We warmly acknowledge all the people that have provided help for this work, and especially Simone Vincenzi of Padova University for the analysis of the chitin component of the setae, Bernhard Gschloessl from INRA Montpellier for helping us work with the transcriptome of *T. pityocampa*, and Laure Sauné for RNA extraction and quality checking. We thank Myron Zalucki for useful comments on an earlier version of the manuscript, and two anonymous reviewers and the editor for comments on the submitted version. This work was achieved thanks to a grant of the University of Padova and the French National Research Agency (ANR-10-JCJC-1705-01 GENOPHENO). We wish to thank the Cassa di Risparmio di Padova e Rovigo (Cariparo) Holding for funding the acquisition of the LTQ-Orbitrap XL mass spectrometer.

Reference Cited

- Andersen, S. O., P. Hojrup, and P. Roepstorff. 1994. Insect cuticular proteins. *Ins. Biochem. Mol. Biol.* 25: 153–176.
- Battisti, A., G. Holm, B. Fagrell, and S. Larsson. 2011. Urticating hairs in arthropods: their nature and medical significance. *Annu. Rev. Entomol.* 56: 203–220.
- Battisti, A., S. Larsson, and A. Roques. 2017. Processionary moths and associated urtication risk: global-change driven effects. *Annu. Rev. Entomol.* 62: 323–342.
- Berardi, L., A. Battisti, and E. Negrisola. 2015a. The allergenic protein Tha p 2 of processionary moths of the genus *Thaumetopoea* (Thaumetopoeinae, Notodontidae, Lepidoptera): characterization and evolution. *Gene* 574: 317–324.
- Berardi, L., M. Branco, M. R. Paiva, H. Santos, and A. Battisti. 2015b. Development time plasticity of the pine processionary moth (*Thaumetopoea pityocampa*) populations under laboratory conditions. *Entomologia* 3: 273.
- Bleumink, E., M.C.J.M. de Jong, F. Kawamoto, G. T. Meyer, A. J. Kloosterhuis, and I. J. Slijper-Pal. 1982. Protease activities in the spicule venom of *Euproctis* caterpillars. *Toxicon* 20: 607–613.
- Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.
- Cabrera, G., V. Salazar, R. Montesino, Y. Támbara, W. B. Struwe, E. Leon, D. J. Harvey, A. Lesur, M. Rincón, B. Domon, et al. 2016. Structural characterization and biological implications of sulfated N-glycans in a serine protease from the neotropical moth *Hylesia metabus*, (Cramer [1775]) (Lepidoptera: Saturniidae). *Glycobiology* 26: 230–250.
- Cabrera, G., U. Lundberg, A. Rodríguez-Ulloa, M. Herrera, W. Machado, M. Portela, S. Palomares, L. A. Espinosa, Y. Ramos, R. Durán, et al. 2017. Protein content of the *Hylesia metabus* egg nest setae (Cramer [1775]) (Lepidoptera: Saturniidae) and its association with the parental investment for the reproductive success and lepidopterism. *J. Proteomics* 150: 183–200.
- Carrick, J. B., N. R. Perkins, and M. P. Zalucki. 2014. Causes of abortion in Australia (2005–2012)—proportion of cases due to equine amnionitis and fetal loss (EAFL). *J. Equine Vet. Sci.* 34: 212–214.
- Conesa, A., S. Götz, J. M. García-Gómez, J. Terol, M. Talón, M. Robles. 2005. Blast2go: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21: 3674–3676.
- Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29: 644–652.
- Gschloessl, B., H. Vogel, C. Burban, D. Heckel, R. Streiff, and C. Kerdelhué. 2014. Comparative analysis of two phenologically divergent populations of the pine processionary moth (*Thaumetopoea pityocampa*) by de novo transcriptome sequencing. *Insect Biochem. Mol. Biol.* 46: 31–42.
- Holm, G., M. Andersson, M. Ekberg, B. Fagrell, J. Sjöberg, M. Bottai, and M. Björkholm. 2014. Setae from larvae of the northern processionary moth (*Thaumetopoea pinivora*, TP) stimulate proliferation of human blood lymphocytes in vitro. *PLoS One* 9: e113977.
- Kanehisa, M., S. Goto, Y. Sato, M. Furumichi, and M. Tanabe. 2012. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* 40: D109–D114.
- Kruger, N. J. 1994. The Bradford method for protein quantization. *Methods Mol. Biol.* 32: 9–15.

- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680–685.
- Lamy, M., M. H. Pastureaud, and G. Ducombs. 1985. Thaumetopoein, an urticating protein of the processionary setae of the caterpillar (*Thaumetopoea pityocampa* Schiff) (Lepidoptera, Thaumetopoeidae). *C. R. Acad. Sci. III* 301: 173–176.
- Lamy, M., M. H. Pastureaud, F. Novak, G. Ducombs, P. Vincendeau, J. Maleville, and L. Texier. 1986. Thaumetopoein: an urticating protein from the setae and integument of the pine processionary caterpillar (*Thaumetopoea pityocampa* Schiff., Lepidoptera, Thaumetopoeidae). *Toxicon* 24: 347–356.
- Larsson, S., and A. Backlund. 2009. Regarding the putative identity of a moth (*Thaumetopoea* spp.) allergen. *Allergy* 64: 493.
- Lundberg, U., V. Salazar, M. Tovar, and J. Rodríguez. 2007. Isolation and partial characterization of proteins with vasodegenerative and proinflammatory properties from the egg nests of *Hylesia metabus* (Lepidoptera: Saturniidae). *J. Med. Entomol.* 44: 440–449.
- Monco, I., J. M. Vega, M. L. Caballero, J. Vega, and E. Alday. 2003. Isolation and characterization of Tha p 1, a major allergen from the pine processionary caterpillar *Thaumetopoea pityocampa*. *Allergy* 58: 34–37.
- Monco, I., A. Battisti, B. Dufour, J. C. Garcia-Ortiz, M. Gonzalez-Muñoz, F. Moutou, P. Paolucci, E. Petrucco Toffolo, J. Rivière, A.-I. Rodríguez-Mahillo, et al. 2015. Medical and veterinary impact of the urticating processionary larvae, pp. 359–410. In A. Roques (ed.), *Processionary moths and climate change: an update*. Springer-Quae, Dordrecht, NL.
- Olivieri, M., G. L. Lazzarini, I. Goio, S. Schinella, L. Romeo, and L. Perbellini. 2012. Occupational risk by the pine processionary moth *Thaumetopoea pityocampa* in the forestry workers of Verona. *G. Ital. Med. Lav. Ergon.* 34: 420–422.
- Parra, G., K. Bradnam, and I. Korf. 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* 23: 1061–1067.
- Pentzold, S., M. Zagrobelny, B. Khakimov, S. B. Engelsens, H. Clausen, B. L. Petersen, J. Borch, B. L. Møller, and S. Bak. 2016. Lepidopteran defence droplets—a composite physical and chemical weapon against potential predators. *Sci. Rep.* 6: 22407.
- Perkins, L. E., M. P. Zalucki, N. R. Perkins, A. J. Cawdell-Smith, and B. W. Cribb. 2016. Urticating setae of *Ochrogaster lunifer*, an Australian processionary caterpillar of veterinary importance. *Med. Vet. Entomol.* 30: 241–245.
- Petrucco-Toffolo, E., D. Zovi, C. Perin, P. Paolucci, A. Roques, A. Battisti, and H. Horvath. 2014. Size and dispersion of urticating setae in three species of processionary moths. *Integr. Zool.* 9: 320–327.
- Rodríguez-Mahillo, A. I., M. González-Muñoz, J. M. Vega, J. A. López, A. Yart, C. Kerdellhué, E. Camafeita, J. C. García Ortiz, H. Vogel, E. P. Toffolo, et al. 2012. Setae from the pine processionary moth (*Thaumetopoea pityocampa*) contain several relevant allergens. *Contact Dermatitis* 67: 367–374.
- Roques, A. 2015. Processionary moths and climate change: an update. Springer-Quae, Dordrecht NL.
- Schmieder, R., and R. Edwards. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27: 863–864.
- Tolin, S., G. Pasini, B. Simonato, F. Mainente, and G. Arrigoni. 2012. Analysis of commercial wines by LC-MS/MS reveals the presence of residual milk and egg white allergens. *Food Control* 28: 321–326.
- Zhang, J., and K. Y. Zhu. 2006. Characterization of a chitin synthase cDNA and its increased mRNA level associated with decreased chitin synthesis in *Anopheles quadrimaculatus* exposed to diflubenzuron. *Insect Biochem. Mol. Biol.* 36: 712–725.

SUPPLEMENTAL MATERIAL

Table S1. Total identified proteins with relative homology with other species and description. ID%: identification on the blast analyses, i.e. is the percentage of the similarity with sequence of the organism. Pep: number of unique peptides identified by mass-spectrometry. PSM: number of peptide spectrum matches, i.e. total number of identified peptide sequences for the protein, including those redundantly identified.

Name	Accession no.	Organism	Sequence description	Id. %	Pep.	PSM
cds.c12407	NP001037082	<i>Bombyx mori</i>	protein translation factor SUI1 homolog	100	2	2
cds.c40997	XP004922895	<i>Bombyx mori</i>	myophilin-like	92	4	5
cds.c4501	XP011554790	<i>Plutella xylostella</i>	uncharacterized protein	65	5	9
cds.c51703	HJ78721	<i>Danaus plexippus</i>	electron-transfer-flavoprotein beta polypeptide	93	10	16
cds.c62029	XP004928837	<i>Bombyx mori</i>	protein windbeutel-like	72	3	3
cds.c62449	BAM18526	<i>Papilio xuthus</i>	ubiquitin carboxy-terminal hydrolase	81	3	5
cds.c63060	BK21653	<i>Picea sitchensis</i>	unknown	57	5	12
cds.c6527	NP001108408	<i>Bombyx mori</i>	binding protein precursor	90	7	13
cds.c6609	XP008213876	<i>Nasonia vitripennis</i>	triosephosphate isomerase isoform X1	92.71	2	5
cds.c68702	BAM18677	<i>Papilio xuthus</i>	proteasome beta7 subunit	86.86	2	2
cds.c71063	P006287	<i>Agrotis segetum</i>	ORF57	42.79	10	29
cds.c71654	XP004931018	<i>Bombyx mori</i>	cysteine-rich venom protein ENH1-like	64.79	9	26
cds.c7180	XP011555192	<i>Plutella xylostella</i>	carbonyl reductase family member 4-like	56.72	10	30
cds.c71879	GG56522	<i>Antheraea yamamai</i>	protein kinase c inhibitor	90.62	4	6
cds.c7214	86356.1 APL3HE	<i>Heliothis virescens</i>	apolipophorin-3	82.86	17	167
cds.c72249	BAM17810	<i>Papilio xuthus</i>	obstructor-A	91.98	9	39
cds.c72418	XP004930219	<i>Bombyx mori</i>	gamma-interferon-inducible lysosomal thiol reductase-like	52	2	2
cds.c72884	EX07268	<i>Helicoverpa armigera</i>	CSP7	37	3	3
cds.c72894	HJ71814	<i>Danaus plexippus</i>	signal peptidase complex subunit 2	85.33	3	3
cds.c73072	ID66689	<i>Agrotis segetum</i>	enoyl-CoA hydratase	90.20	17	32
cds.c73203	HJ72513	<i>Danaus plexippus</i>	hypothetical protein KGM_09534	72.50	3	3
cds.c73519	HJ69853	<i>Danaus plexippus</i>	cuticular protein analogous to peritrophins 3-A2	81.36	4	5
cds.c73548	8242	<i>Heliothis virescens</i>	juvenile hormone binding protein	54.92	2	3
cds.c73724	EZ98895	<i>Tribolium castaneum</i>	hypothetical protein TcasGA2_TC004510	69.20	3	11
cds.c74587	XP004929459	<i>Bombyx mori</i>	ubiquitin-conjugating	100	2	2

Name	Accession no.	Organism	Sequence description	Id. %	Pep.	PSM
			enzyme E2			
cds.c74796	IL29316	<i>Cnaphalocrocis medinalis</i>	glutathione S-transferase omega 1	84.65	6	11
cds.c74814	BAM17916	<i>Papilio xuthus</i>	adenosylhomocysteinase at 13	94.19	11	45
cds.c75004	XP004931809	<i>Bombyx mori</i>	3-oxoacyl-acyl-carrier-	63	2	4
cds.c75009	XP004928934	<i>Bombyx mori</i>	fructose-bisphosphate	81.08	2	10
cds.c75213	XP004932946	<i>Bombyx mori</i>	10 kDa heat shock protein	91.26	6	6
cds.c75258	NP001166827	<i>Bombyx mori</i>	H ⁺ transporting ATP synthase O subunit isoform 2	89.29	2	3
cds.c75269	XP003999137	<i>Felis catus</i>	histone H3.3	100	4	15
cds.c75311	NP001040393	<i>Bombyx mori</i>	ubiquitin conjugating enzyme E2	100	3	3
cds.c75351	BAM19454	<i>Papilio xuthus</i>	unknown secreted protein, partial	50.65	3	4
cds.c75394	BAM18072	<i>Papilio xuthus</i>	WAP four-disulfide core domain protein 2 precursor	41.82	5	7
cds.c75678	FC76152	<i>Quadrastichus erythrinae</i>	heat shock protein 90	95.71	9	48
cds.c75827	XP004923023	<i>Bombyx mori</i>	proteasome subunit beta type-6-like	92	4	4
cds.c75854	NP001040460	<i>Bombyx mori</i>	proteasome beta 3 subunit	94.15	2	2
cds.c75977	BAM18637	<i>Papilio xuthus</i>	S-formylglutathione hydrolase	90.24	11	38
cds.c76065	EP40969	<i>Spodoptera exigua</i>	catalase	87.18	30	169
cds.c76149	BAM19834	<i>Papilio xuthus</i>	unknown unsecreted protein, partial	79.18	4	12
cds.c76321	NP001038957	<i>Bombyx mori</i>	ubiquinol-cytochrome C reductase complex 14kD subunit	83.33	2	2
cds.c76517	XP004926040	<i>Bombyx mori</i>	uncharacterized protein LOC101739782	43.81	4	20
cds.c76711	XP004930569	<i>Bombyx mori</i>	protein CREG1-like isoform X1	70.76	7	7
cds.c76822	XP004923378	<i>Bombyx mori</i>	insecticyanin-A-like	67.61	6	12
cds.c76883	NP001040525	<i>Bombyx mori</i>	6-phosphogluconate dehydrogenase	84.89	8	14
cds.c76938	XP004922754	<i>Bombyx mori</i>	phosphatidylethanolamine-binding protein homolog F40A3.3	85.95	7	28
cds.c77007	YP001110948	<i>Heliothis virescens</i> ascovirus 3e	hypothetical protein HVAV3e_gp095	32.41	2	2
cds.c77308	XP011569353	<i>Plutella xylostella</i>	transcription factor BTF3 homolog 4-like	81.25	4	8
cds.c77389	XP004929168	<i>Bombyx mori</i>	adenylate kinase isoenzyme	86	5	10
cds.c77408	GR84224	<i>Melitaea cinxia</i>	heat shock protein 70-2	92.44	3	24

Name	Accession no.	Organism	Sequence description	Id. %	Pep.	PSM
cds.c77540	BH09735	<i>Trichoplusia ni</i>	heat shock protein	96	22	132
cds.c77652	XP004921729	<i>Bombyx mori</i>	circadian clock-proteinlike	41.25	10	11
cds.c77657	XP011568761	<i>Plutella xylostella</i>	40S ribosomal protein S5	96.80	2	2
cds.c77733	NP001040431	<i>Bombyx mori</i>	cyclic AMP-regulated protein	94.48	10	17
cds.c77809	O65575	<i>Manduca sexta</i>	molting	26.64	4	6
cds.c77945	DT80677	<i>Euphydryas aurinia</i>	ribosomal protein SA	88.03	4	6
cds.c78007	HJ75809	<i>Danaus plexippus</i>	hypothetical protein	72.14	9	22
cds.c78188	B55603	<i>Spodoptera litura</i>	moderately methionine rich storage protein	74.77	20	70
cds.c78326	R98920	<i>Ostrinia nubilalis</i>	trypsin-like proteinase T2a precursor	58.25	4	4
cds.c78521	DK55517	<i>Spodoptera litura</i>	heat shock protein 90 cognate	86.36	3	11
cds.c78536	DE75591	<i>Antheraea pernyi</i>	abnormal wing disc protein	94.12	10	42
cds.c78554	CN67046	<i>Helicoverpa armigera</i>	GTP-binding nuclear protein ran	100	2	2
cds.c78584	BAM17873	<i>Papilio xuthus</i>	elongation factor 1 delta	75.28	2	2
cds.c78593	XP004933631	<i>Bombyx mori</i>	acetyl-CoA hydrolase-like	85.18	2	2
cds.c78608	BAM18142	<i>Papilio xuthus</i>	unknown unsecreted protein	68.07	2	4
cds.c78631	XP004923642	<i>Bombyx mori</i>	B-cell receptor-associated protein 31-like	86.61	3	3
cds.c78697	NP001037010	<i>Bombyx mori</i>	small GTP-binding protein	96.06	6	11
cds.c78703	XP011559092	<i>Plutella xylostella</i>	acidic leucine-rich nuclear phosphoprotein 32 family member A	74.90	2	4
cds.c78716	CT52824	<i>Chilo suppressalis</i>	heat shock protein 60	95.69	4	7
cds.c78816	DT80645	<i>Euphydryas aurinia</i>	ribosomal protein L23A	82.43	2	2
cds.c78878	NP001166675	<i>Danaus plexippus</i>	cuticular protein RR-2 motif 87 precursor	63.19	2	3
cds.c78914	HB18378	<i>Spodoptera exigua</i>	glutathione S-transferase epsilon	62.44	7	7
cds.c78955	II21974	<i>Sesamia inferens</i>	aldo-ketose reductase 5, partial	75.74	5	10
cds.c78975	ID66701	<i>Agrotis segetum</i>	enoyl-CoA delta isomerase 1	77.82	2	3
cds.c78984	K92181.1 AF40020	<i>Spodoptera frugiperda</i>	ribosomal protein S12	99.28	2	2
cds.c78991	XP004932543	<i>Bombyx mori</i>	hydroxyacylglutathione hydrolase, mitochondrial-like	78.91	2	3
cds.c79067	NP001164652	<i>Bombyx mori</i>	Nuc. excision repair prot.	81	3	4
cds.c79071	NP001037618	<i>Bombyx mori</i>	ras-related GTP-binding protein Rab11	96.71	4	4
cds.c79093	NP001093091	<i>Bombyx mori</i>	H ⁺ transporting ATP	88.82	4	6

Name	Accession no.	Organism	Sequence description	Id. %	Pep.	PSM
cds.c79103	XP011549057	<i>Plutella xylostella</i>	protein takeout-like	49.80	6	7
cds.c79110	NP001036859	<i>Bombyx mori</i>	ecdysteroid-inducible angiotensin-converting enzyme-related gene product precursor	82.97	5	6
cds.c79113	HJ63750	<i>Danaus plexippus</i>	replication protein A1	28.30	2	2
cds.c79123	IH07591	<i>Spodoptera litura</i>	glutathione S-transferase sigma 4	75	18	96
cds.c79133	K92188.1 AF40021	<i>Spodoptera frugiperda</i>	ribosomal protein S19	98.04	2	3
cds.c79152	XP004931115	<i>Bombyx mori</i>	retinal dehydrogenase 1	75.72	5	31
cds.c79156	GH70102	<i>Spodoptera exigua</i>	odorant binding protein 6	49.32	8	18
cds.c79163	HJ78046	<i>Danaus plexippus</i>	hypothetical protein KGM_13066	75.11	10	14
cds.c79170	BAM17938	<i>Papilio xuthus</i>	phosphoglyceromutase	92.94	13	25
cds.c79184	NP001036999	<i>Bombyx mori</i>	glutathione peroxidase	87.94	8	15
cds.c79193	P29504	<i>Manduca sexta</i>	ubiquitin-40S ribosomal protein S27a	100	5	38
cds.c79247	XP004932961	<i>Bombyx mori</i>	selenium-binding protein 1-like isoform X2	79.66	2	5
cds.c79291	BAM19138	<i>Papilio polytes</i>	peptidyl-prolyl cis-trans isomerase	83.90	11	40
cds.c79294	XP004928974	<i>Bombyx mori</i>	malate dehydrogenase	84.21	17	78
cds.c79334	XP011561164	<i>Plutella xylostella</i>	ninjurin-2-like	78.95	2	2
cds.c79425	R23825	<i>Antheraea pernyi</i>	dopa-decarboxylase	90.36	5	6
cds.c79439	XP004926522	<i>Bombyx mori</i>	dipeptidyl peptidase 3-like isoform X1	82.90	5	6
cds.c79473	XP011557739	<i>Plutella xylostella</i>	rho GDP-dissociation inhibitor 1-like	84.88	3	4
cds.c79527	HJ67827	<i>Danaus plexippus</i>	putative citrate synthase	91.81	3	8
cds.c79555	XP011551231	<i>Plutella xylostella</i>	uncharacterized oxidoreductase	76.23	2	2
cds.c79569	DX87348	<i>Heliconius melpomene</i>	yellow-c	74.94	15	25
cds.c7957	17.3 FABP2MA	<i>Manduca sexta</i>	fatty acid-binding protein 2	48.87	5	6
cds.c79581	K64362.1 AF33628	<i>Galleria mellonella</i>	glutathione-S-transferase- like protein	87.91	14	71
cds.c79587	NP001040337	<i>Bombyx mori</i>	aspartate aminotransferas	87.4	14	30
cds.c79638	BAM18236	<i>Papilio xuthus</i>	lactoylglutathione lyase	83.14	3	3
cds.c79650	DK55520	<i>Spodoptera litura</i>	small heat shock protein	88.33	8	22
cds.c79655	BG43001	<i>Bombyx mori</i>	splicing factor proline- and glutamine-rich	72.87	2	2
cds.c79713	XP004921989	<i>Bombyx mori</i>	glutaminy-peptide	80.99	2	2
cds.c79769	CY95350	<i>Manduca sexta</i>	ribosomal protein L27	98.51	2	2
cds.c79776	EL33703	<i>Spodoptera littoralis</i>	carboxylesterase CXE30	71.84	13	39
cds.c79789	HJ73760	<i>Danaus plexippus</i>	hypothetical protein	53.24	9	12

Name	Accession no.	Organism	Sequence description	Id. %	Pep.	PSM
			KGM_09817			
cds.c79873	NP001040544	<i>Bombyx mori</i>	transaldolase	88.55	15	34
cds.c79889	BU98620	<i>Helicoverpa armigera</i>	unknown	27.18	7	14
cds.c79940	C02870	<i>Helicoverpa assulta</i>	fatty acid binding protein	53.03	6	8
cds.c79945	NP001103774	<i>Bombyx mori</i>	receptor expression enhancing protein isoform 1	84	3	3
cds.c79950	BK29496	<i>Helicoverpa armigera</i>	nuclear transport factor 2	96.95	2	3
cds.c79975	F21123	<i>Spodoptera litura</i>	transferrin	83.80	14	20
cds.c80007	XP002135702	<i>Drosophila pseudoobscura</i>	GA27615	66.09	6	10
cds.c80023	NP001040432	<i>Bombyx mori</i>	L-xylulose reductase	77.88	2	2
cds.c80048	NP001037668	<i>Bombyx mori</i>	cytosolic juvenile hormone binding protein 36 kDa subunit	86.62	3	4
cds.c80057	BAM18976	<i>Papilio polytes</i>	GST-containing FLYWCH zinc-finger protein	78.97	5	6
cds.c80124	XP011551636	<i>Plutella xylostella</i>	repetitive proline-rich cell wall protein 2-like	80.78	2	2
cds.c80127	FP55240	<i>Cerura vinula</i>	arylphorin type 2	72.02	66	1649
cds.c80128	XP011551175	<i>Plutella xylostella</i>	gelsolin, partial	74.37	12	27
cds.c80149	BK90823	<i>Spodoptera exigua</i>	cathepsin B-like cysteine proteinase	80.18	3	7
cds.c80152	XP004924081	<i>Bombyx mori</i>	phytanoyl-CoA dioxygenase, peroxisomal-like isoform X1	69.43	2	3
cds.c80159	BAM76811	<i>Mythimna separata</i>	prophenoloxidase 1	78.36	18	100
cds.c80305	BAM17865	<i>Papilio xuthus</i>	fructose-bisphosphate aldolase	94.23	11	58
cds.c80319	XP011553552	<i>Plutella xylostella</i>	probable 3-hydroxyacyl-CoA dehydrogenase B0272.3	87.08	2	2
cds.c80327	XP004933454	<i>Bombyx mori</i>	ras-like GTP-binding protein Rho1-like isoform X2	100	3	7
cds.c80332	XP001998948	<i>Drosophila mojavensis</i>	GI24242	30	7	70
cds.c80335	XP004925887	<i>Bombyx mori</i>	L-xylulose reductase-like isoform X1	93.85	8	9
cds.c80338	HJ75579	<i>Danaus plexippus</i>	superoxide dismutase	88.96	6	44
cds.c80343	GH14258	<i>Spodoptera frugiperda</i>	glyceraldehyde-3-phosphate dehydrogenase	95.78	16	88
cds.c80355	62R9.1 RS10SP	<i>Spodoptera frugiperda</i>	RecName: Full=40S ribosomal protein S10	96.20	6	12
cds.c80372	XP004930886	<i>Bombyx mori</i>	venom serine carboxypeptidase-like	50.44	2	6

Name	Accession no.	Organism	Sequence description	Id. %	Pep.	PSM
cds.c80388	XP004924995	<i>Bombyx mori</i>	protein canopy homolog 1-like isoform X1	79.77	4	5
cds.c80396	XP011548877	<i>Plutella xylostella</i>	elongation factor 1-gamma	78.04	3	5
cds.c80452	BAM19092	<i>Papilio polytes</i>	unknown secreted protein	79.81	4	36
cds.c80476	NP001153668	<i>Bombyx mori</i>	histone H2B-like protein	100	7	25
cds.c80502	BAM73876	<i>Bombyx mori</i>	hypothetical protein, partial	97.30	2	2
cds.c80566	BAM18260	<i>Papilio xuthus</i>	proteasome 29kD subunit	96.14	3	3
cds.c80591	BAM18895	<i>Papilio polytes</i>	chitin binding peritrophin-A	90.84	6	10
cds.c80610	HJ75590	<i>Danaus plexippus</i>	H ⁺ transporting ATP synthase subunit d	78.74	6	6
cds.c80648	XP004933462	<i>Bombyx mori</i>	hemocyte protein	75.95	4	7
cds.c80651	NP001040113	<i>Bombyx mori</i>	cystathionine gamma-lyase	79.85	5	11
cds.c80692	HJ69215	<i>Danaus plexippus</i>	tetraspanin E118	66.29	2	2
cds.c80697	HJ66953	<i>Danaus plexippus</i>	proteasome beta-subunit	95.30	2	2
cds.c80717	BAM17904	<i>Papilio xuthus</i>	membrane steroid binding protein	73.99	7	9
cds.c80726	2I1	<i>Spodoptera Frugiperda</i>	chain A, Moesin	93.96	4	13
cds.c80763	XP011567055	<i>Plutella xylostella</i>	mitochondrial enolase superfamily member 1-like	80.48	2	2
cds.c80778	NP001040233	<i>Bombyx mori</i>	ATP synthase	98	13	29
cds.c80782	NP001037057	<i>Bombyx mori</i>	BCP inhibitor precursor	47.47	9	15
cds.c80785	NP001040134	<i>Bombyx mori</i>	isocitrate dehydrogenase	92.16	14	53
cds.c80797	NP001040372	<i>Bombyx mori</i>	transgelin	92.55	8	17
cds.c80800	D85119	<i>Spodoptera exigua</i>	juvenile hormone epoxide hydrolase	69.83	2	4
cds.c80802	K52495.1 AF36148	<i>Bombyx mori</i>	chymotrypsin inhibitor CI-8A	59.95	16	126
cds.c80835	XP011566745	<i>Plutella xylostella</i>	lambda-crystallin	79.58	12	40
cds.c80869	BAM17778	<i>Papilio xuthus</i>	calreticulin	91.46	8	41
cds.c80871	XP004932565	<i>Bombyx mori</i>	aconitate hydratase, mitochondrial-like	92.25	5	7
cds.c80913	XP004924438	<i>Bombyx mori</i>	glyoxylate reductase	78.01	9	12
cds.c80921	XP004930780	<i>Bombyx mori</i>	uncharacterized protein LOC101741978	83.11	7	21
cds.c80933	XP011550540	<i>Plutella xylostella</i>	ecdysteroid-regulated 16 kDa protein	57.35	5	14
cds.c80958	BS59653	<i>Spodoptera exigua</i>	prophenoloxidase	82.27	20	127
cds.c80965	ID66694	<i>Agrotis segetum</i>	hydroxyacyl-CoA dehydrogenase	75.18	4	5
cds.c81036	HJ63999	<i>Danaus plexippus</i>	lopap	83.25	6	9
cds.c81079	II22000	<i>Sesamia inferens</i>	alcohol dehydrogenase AD2	91.47	9	12
cds.c81091	EI30868	<i>Mamestra brassicae</i>	chitin deacetylase 1	95.92	4	4

Name	Accession no.	Organism	Sequence description	Id. %	Pep.	PSM
cds.c81091	XP011556129	<i>Plutella xylostella</i>	uncharacterized protein LOC105387140 isoform X1	92.11	4	3
cds.c81098	CR78449	<i>Heliothis virescens</i>	growth-blocking peptide	31.21	3	6
cds.c81161	BAM17678	<i>Papilio xuthus</i>	ribosomal protein LP0	92.41	7	8
cds.c81169	BAM20528	<i>Papilio xuthus</i>	6-pyruvoyltetrahydropterin synthase purple	88.02	2	2
cds.c81171	BAM19217	<i>Papilio polytes</i>	takeout/JHBP like protein	63.93	4	4
cds.c81174	CT88125	<i>Drosophila melanogaster</i>	AT15141p	100	5	5
cds.c81177	NP001037090	<i>Bombyx mori</i>	serine protease inhibitor 4 precursor	63.72	4	4
cds.c81257	DD22718	<i>Helicoverpa armigera</i>	arginine kinase	97.18	22	211
cds.c81276	HJ63776	<i>Danaus plexippus</i>	hypothetical protein KGM_14041	79.82	7	18
cds.c81283	FC87786	<i>Bombyx mori</i>	yellow-d	68.5	14	28
cds.c81313	XP004928620	<i>Bombyx mori</i>	phosphoglycerate kinase-like	93.01	8	14
cds.c81330	XP004932371	<i>Bombyx mori</i>	uncharacterized protein LOC101739845 isoform X1	81.44	5	16
cds.c81334	V91442	<i>Lonomia obliqua</i>	hypothetical protein 29	44.50	8	10
cds.c81355	BAM19642	<i>Papilio xuthus</i>	dj-1 protein	68.44	5	9
cds.c81384	CY95330	<i>Manduca sexta</i>	ribosomal protein L31	100	2	4
cds.c81498	XP011556783	<i>Plutella xylostella</i>	fasciclin-3 isoform X7	62.85	5	5
cds.c81502	00.1 VATAMA	<i>Manduca sexta</i>	V-type proton ATPase catalytic subunit A	96.92	9	18
cds.c81530	BAH03386	<i>Ostrinia scapularis</i>	acetyltransferase 1	91.16	5	6
cds.c81637	31190	<i>Manduca sexta</i>	lysozyme	77.86	2	2
cds.c81694	XP004933858	<i>Bombyx mori</i>	THAP domain-containing protein 4-like	80.62	2	3
cds.c81742	NP001040416	<i>Bombyx mori</i>	leukotriene A4 hydrolase	75.88	2	5
cds.c81765	HJ77299	<i>Danaus plexippus</i>	small GTP binding protein RAB5	92.02	2	4
cds.c81765	NP001040381	<i>Bombyx mori</i>	fructose-1,6-bisphosphatase	89.58		
cds.c81786	BAM17986	<i>Papilio xuthus</i>	ATP synthase-gamma chain	86.82	4	5
cds.c81798	GW15345	<i>Spodoptera exigua</i>	glycerol-3-phosphate dehydrogenase, partial	96.88	6	7
cds.c81813	XP001865168	<i>Culex quinquefasciatus</i>	histone 1	100	4	8
cds.c81813	XP002430908	<i>Pediculus humanus corporis</i>	histone H4, putative	100	4	8
cds.c81813	NP001153666	<i>Bombyx mori</i>	histone H2A-like protein 2	99.19	4	5
cds.c81813	2NQ	<i>Drosophila melanogaster</i>	chain C, Drosophila Nucleosome Structure	99.19	7	8

Name	Accession no.	Organism	Sequence description	Id. %	Pep.	PSM
cds.c81816	IW61834	<i>Spodoptera litura</i>	thioredoxin 1	96.23	5	23
cds.c81832	BAM18144	<i>Papilio xuthus</i>	pyruvate kinase	86.60	4	10
cds.c81845	CDJ26240	<i>Oreta rosea</i>	histone H1	82.76	8	44
cds.c81853	AI06088	<i>Manduca sexta</i>	putative annexin IX-C	95.33	16	65
cds.c81891	II01042	<i>Dendrolimus kikuchii</i>	chemosensory protein	83.33	3	3
cds.c81913	BAM20137	<i>Papilio xuthus</i>	unknown unsecreted protein	85.33	2	3
cds.c81955	CC91897	<i>Heliothis virescens</i>	hemolin	74.12	12	24
cds.c81959	NP001040451	<i>Bombyx mori</i>	vacuolar ATP synthase subunit E	93.81	7	10
cds.c82003	BAM17994	<i>Papilio xuthus</i>	endoU protein	84.57	8	10
cds.c82013	CB54949	<i>Helicoverpa armigera</i>	fatty acid-binding protein 2	60.90	11	40
cds.c82013	CB54949	<i>Helicoverpa armigera</i>	fatty acid-binding protein 2	52.59	6	22
cds.c82015	63957	<i>Spodoptera frugiperda</i>	CALNUC	72.78	2	5
cds.c82018	XP004922166	<i>Bombyx mori</i>	gelsolin, cytoplasmic-like	68.58	3	6
cds.c82031	HJ66584	<i>Danaus plexippus</i>	prohibitin protein WPH	95.62	3	3
cds.c82060	XP004927505	<i>Bombyx mori</i>	prefoldin subunit 2-like	73.10	2	2
cds.c82065	K39637	<i>Manduca sexta</i>	iron regulatory protein 1	89.57	6	8
cds.c82104	XP004931117	<i>Bombyx mori</i>	bifunctional purine biosynthesis protein	83.62	2	3
cds.c82108	NP001037378	<i>Bombyx mori</i>	sterol carrier protein	88.25	4	5
cds.c82139	FK28500	<i>Helicoverpa armigera</i>	malate dehydrogenase	93.66	8	14
cds.c82147	XP004929181	<i>Bombyx mori</i>	protein homolog isoform X1	66.23	2	5
cds.c82157	BAM18017	<i>Papilio xuthus</i>	nascent polypeptide associated complex protein alpha subunit	95.73	2	4
cds.c82249	BU98620	<i>Helicoverpa armigera</i>	unknown	65.93	16	105
cds.c82264	HJ79131	<i>Danaus plexippus</i>	ubiquitin-conjugating enzyme E2L	98.70	3	4
cds.c82301	XP011560620	<i>Plutella xylostella</i>	aldehyde dehydrogenase, mitochondrial	81.41	5	13
cds.c82301	HJ74714	<i>Danaus plexippus</i>	chemosensory protein	50.98	5	16
cds.c82320	HJ71114	<i>Danaus plexippus</i>	glucosidase	82.80	5	5
cds.c82345	NP001037570	<i>Bombyx mori</i>	ribosomal protein S15A	100	2	2
cds.c82362	NP001104835	<i>Bombyx mori</i>	aminopeptidase N3 precursor	70.79	5	12
cds.c82375	HJ71918	<i>Danaus plexippus</i>	hypothetical protein KGM_06034	98.61	3	3
cds.c82402	DO79970	<i>Antheraea pernyi</i>	phosphoserine aminotransferase	79.12	15	37
cds.c82421	XP004931712	<i>Bombyx mori</i>	voltage-dependent anion-	95.74	10	17

Name	Accession no.	Organism	Sequence description	Id. %	Pep.	PSM
			selective channel-like isoform X4			
cds.c82467	XP004931507	<i>Bombyx mori</i>	TPPP family protein CG4893-like isoform X1 protein CG4893-like isoform X3	87.80	6	11
cds.c82481	BAM20654	<i>Papilio polytes</i>	unknown secreted protein	47.20	8	29
cds.c82489	CI32835	<i>Helicoverpa armigera</i>	serine protease-like protein 1	64.63	6	12
cds.c82492	CR56864	<i>Mamestra brassicae</i>	serine protease inhibitor-1A	62.50	4	226
cds.c82492	N71634	<i>Mamestra configurata</i>	serine protease inhibitor serpin 1c	58.97	2	239
cds.c82502	XP004926634	<i>Bombyx mori</i>	carbonyl reductase NADPH 1-like	75.19	7	7
cds.c82515	XP004922450	<i>Bombyx mori</i>	uncharacterized protein LOC101742045	82.07	2	3
cds.c82520	BAM17839	<i>Papilio xuthus</i>	dihydropteridine reductase	84.62	11	52
cds.c82574	NP001040386	<i>Bombyx mori</i>	peroxiredoxin	85.11	10	25
cds.c82598	IN34682	<i>Agrotis segetum</i>	fatty alcohol acetyltransferase	86.49	6	8
cds.c82631	FP36379	<i>Spodoptera frugiperda</i>	profilin, partial	96.83	5	15
cds.c82633	NP001036964	<i>Bombyx mori</i>	beta-tubulin	100	17	111
cds.c82664	BAM17725	<i>Papilio xuthus</i>	ribosomal protein S18	99.34	2	4
cds.c82703	BAM18892	<i>Papilio polytes</i>	translationally controlled tumor protein	93.02	9	27
cds.c82705	XP011563498	<i>Plutella xylostella</i>	cofilin/actin-depolymerizing factor homolog	99.32	12	40
cds.c82716	NP001091755	<i>Bombyx mori</i>	ADP-ribosylation factor	100	7	9
cds.c82749	O21505.1 AF41306	<i>Manduca sexta</i>	serpin 3a	65.58	3	4
cds.c82758	BAM17793	<i>Papilio xuthus</i>	protein disulfide	84.04	24	176
cds.c82759	XP011562283	<i>Plutella xylostella</i>	pyrroline-5-carboxylate reductase-like	85.61	2	5
cds.c82809	XP004931482	<i>Bombyx mori</i>	ribonuclease UK114-like isoform X1	72.25	2	4
cds.c82839	XP011552685	<i>Plutella xylostella</i>	muscle-specific protein 20	91.12	6	10
cds.c82857	XP004929915	<i>Bombyx mori</i>	mitochondrial-like isoform X2	91.94	14	53
cds.c82884	DQ89807	<i>Bombyx mori</i>	aldo-keto reductase	79.87	3	31
cds.c82900	BAE43845	<i>Papilio xuthus</i>	ebony	74.45	5	9
cds.c82912	HJ74479	<i>Danaus plexippus</i>	transitional endoplasmic reticulum ATPase TER94	98.43	2	5
cds.c82916	CM78945	<i>Spodoptera exigua</i>	heat shock protein 70	95.81	3	4
cds.c82921	HJ70611	<i>Danaus plexippus</i>	putative alcohol dehydrogenase	51.10	4	5
cds.c82930	HJ64409	<i>Danaus plexippus</i>	hypothetical protein KGM_18350	99.52	18	112

Name	Accession no.	Organism	Sequence description	Id. %	Pep.	PSM
cds.c82943	XP004924122	<i>Bombyx mori</i>	4-hydroxyphenylpyruvate dioxygenase-like	87.10	2	2
cds.c82961	GQ42771	<i>Helicoverpa armigera</i>	small heat shock protein G7	58.26	2	2
cds.c82974	BAM20677	<i>Papilio polytes</i>	similar to CG1648	81.74	3	4
cds.c83002	XP011564545	<i>Plutella xylostella</i>	aldehyde dehydrogenase X, mitochondrial-like	86.27	22	147
cds.c83005	GZ62936	<i>Helicoverpa armigera</i>	transferin	67.23	3	7
cds.c83012	XP011557565	<i>Plutella xylostella</i>	14-3-3 protein epsilon	97.71	13	62
cds.c83046	BU62829	<i>Spodoptera exigua</i>	serpin-2	73.83	8	10
cds.c83057	K58066	<i>Heliothis virescens</i>	110 kDa aminopeptidase	71.85	5	14
cds.c83120	XP004923039	<i>Bombyx mori</i>	aspartate aminotransferase, cytoplasmic-like isoform X1	83.74	18	58
cds.c83123	HJ75804	<i>Danaus plexippus</i>	putative selenoprotein M	68.52	5	7
cds.c83127	BX54737	<i>Spodoptera exigua</i>	elongation factor 1 beta'	89.24	4	5
cds.c83138	01.1 VATBMA	<i>Manduca sexta</i>	V-type proton ATPase subunit B	98.77	11	21
cds.c83157	BAM19032	<i>Papilio polytes</i>	cuticular protein PpolCPH30	73.54	3	7
cds.c83187	JG43843	<i>Bicyclus anynana</i>	translation elongation factor EF-1 alpha subunit	98.49	12	86
cds.c83193	9321	<i>Manduca sexta</i>	methionine-rich storage protein 2	72.41	57	893
cds.c83204	BAM19068	<i>Papilio polytes</i>	probable methylmalonate-semialdehyde dehydrogenase	86.59	2	2
cds.c83247	DL09420	<i>Antheraea pernyi</i>	myosin light chain 2	96.69	5	7
cds.c83293	XP011559595	<i>Plutella xylostella</i>	luciferin 4-monooxygenase-like	50.83	3	8
cds.c83302	NP001041705	<i>Bombyx mori</i>	H ⁺ transporting ATP synthase beta subunit isoform 2	96.09	15	47
cds.c83395	BAM17721	<i>Papilio xuthus</i>	stress-sensitive B	95.96	2	2
cds.c83417	Q75437	<i>Helicoverpa armigera</i>	cathepsin L-like protease	81.23	3	5
cds.c83422	NP001037068	<i>Bombyx mori</i>	chemosensory protein 7 precursor	61.16	4	8
cds.c83430	XP011558612	<i>Plutella xylostella</i>	calcyphosin-like protein isoform X7	76.26	3	5
cds.c83508	BW96360	<i>Helicoverpa armigera</i>	thioredoxin peroxidase	94.87	10	18
cds.c83508	BM55645	<i>Maconellicoccus hirsutus</i>	putative cytosolic thioredoxin peroxidase	76.44	7	9
cds.c83513	HB18379	<i>Spodoptera exigua</i>	glutathione S-transferase omega	64.61	8	33
cds.c83517	NP001093267	<i>Bombyx mori</i>	phosphatidylethanolamine	81.62	3	6

Name	Accession no.	Organism	Sequence description	Id. %	Pep.	PSM
			binding protein isoform 2			
cds.c83555	GQ45613	<i>Agrotis ipsilon</i>	aldo-keto reductase	65.58	21	126
cds.c83555	XP004928315	<i>Bombyx mori</i>	alpha-actinin, sarcomeric-like isoform X1	97.89		
cds.c83620	XP004931781	<i>Bombyx mori</i>	cytochrome c-like isoform X1	96.30	5	9
cds.c83624	XP011563810	<i>Plutella xylostella</i>	multifunctional protein ADE2	90.35	8	22
cds.c83632	CR07788	<i>Heliothis virescens</i>	14-3-3 zeta	98.99	9	67
cds.c83641	AU84716	<i>Helicoverpa armigera</i>	triosephosphate isomerase	89.52	10	35
cds.c83648	BAM18569	<i>Papilio xuthus</i>	unknown secreted protein	54.23	4	8
cds.c83689	XP011567327	<i>Plutella xylostella</i>	electron transfer flavoprotein subunit alpha, mitochondrial	89.70	12	22
cds.c83690	NP001040287	<i>Bombyx mori</i>	vacuolar ATP synthase subunit G	90.60	4	8
cds.c83708	NP001037299	<i>Bombyx mori</i>	Mn superoxide dismutase	85.58	7	14
cds.c83709	NP001040382	<i>Bombyx mori</i>	FK506-binding protein	94.44	2	3
cds.c83709	XP005798317	<i>Xiphophorus maculatus</i>	peptidyl-prolyl cis-trans isomerase FKBP1B-like isoform X1	78.70	2	3
cds.c83717	F51482	<i>Bombyx mori</i>	sorbitol dehydrogenase	83.24	7	18
cds.c83726	AAT76805	<i>Helicoverpa armigera</i>	hexamerine	61.83	27	204
cds.c83729	NP001040462	<i>Bombyx mori</i>	serine proteinase-like protein precursor	62.68	7	16
cds.c83737	XP011548147	<i>Plutella xylostella</i>	40S ribosomal protein S14	99.34	2	2
cds.c83740	XP004932381	<i>Bombyx mori</i>	uncharacterized protein	48.73	5	5
cds.c83742	GQ53952	<i>Spodoptera litura</i>	enolase	94.46	25	130
cds.c83763	HJ66241	<i>Danaus plexippus</i>	ecdysteroid 22-kinase	44.30	5	5
cds.c83766	CN78408	<i>Spodoptera exigua</i>	glycogen phosphorylase	94.17	10	12
cds.c83779	XP011552532	<i>Plutella xylostella</i>	eukaryotic initiation factor	97	3	9
cds.c83798	XP004932673	<i>Bombyx mori</i>	protein takeout-like Bombyx mori	55.97	3	4
cds.c83809	XP004930717	<i>Bombyx mori</i>	regucalcin-like isoform X2	77.30	10	38
cds.c83836	HJ69803	<i>Danaus plexippus</i>	saposin-related protein	59.27	7	23
cds.c83853	XP011565728	<i>Plutella xylostella</i>	nucleoplasmin-like protein isoform X1	91.49	3	11
cds.c83872	HJ71719	<i>Danaus plexippus</i>	putative RNA polymerase II transcriptional coactivator	82.24	5	8
cds.c83876	NP001163865	<i>Bombyx mori</i>	translation elongation factor 2 isoform 2	97.16	11	30
cds.c83908	NP001040146	<i>Bombyx mori</i>	proteasome zeta subunit	97.53	3	4
cds.c83908	NP001037197	<i>Bombyx mori</i>	glutathione S-transferase epsilon 1	70.59	6	6

Name	Accession no.	Organism	Sequence description	Id. %	Pep.	PSM
cds.c83913	NP001037063	<i>Bombyx mori</i>	chemosensory protein 3 precursor	67.72	4	11
cds.c83982	NP001243939	<i>Bombyx mori</i>	immune-related protein 2 precursor	75.30	4	5
cds.c84008	XP004928724	<i>Bombyx mori</i>	delta-1-pyrroline-5-carboxylate dehydrogenase	82.57	7	7
cds.c84090	XP004926024	<i>Bombyx mori</i>	counting factor associated protein D-like isoform X1	78.12	2	2
cds.c84124	XP011558699	<i>Plutella xylostella</i>	tumor protein D54 isoform X3	80.57	4	4
cds.c84193	DK55523	<i>Spodoptera litura</i>	small heat shock protein	67.37	6	8
cds.c84197	BAM17842	<i>Papilio xuthus</i>	similar to CG9796	71.77	4	4
cds.c84199	HJ73228	<i>Danaus plexippus</i>	fumarylacetoacetase	82.94	8	11
cds.c84214	NP001037358	<i>Bombyx mori</i>	time interval measuring enzyme-esterase A4 precursor	58.62	12	97
cds.c84246	XP004930719	<i>Bombyx mori</i>	regucalcin-like isoform X1	65.03	5	7
cds.c84293	BAM20615	<i>Papilio polytes</i>	unknown secreted protein	60	2	2
cds.c84379	YP008003662	<i>Mythimna separata entomopoxvirus 'L'</i>	unknown similar to AMEV109	59.79	17	59
cds.c84390	XP004930607	<i>Bombyx mori</i>	myosin light chain alkali-like isoform X2	92	3	7
cds.c84431	BAM18395	<i>Papilio xuthus</i>	thioredoxin	79.87	3	4
cds.c84438	GC23338	<i>Chilo suppressalis</i>	heat shock protein 21.4	96.79	5	7
cds.c84526	GQ45606	<i>Agrotis ipsilon</i>	alcohol dehydrogenase,	67	4	11
cds.c84539	CY95351	<i>Manduca sexta</i>	ribosomal protein S15	100	2	4
cds.c84563	BAK82317	<i>Bombyx mori</i>	apolipoprotein precursor protein	59	213	1828
cds.c84582	XP004925737	<i>Bombyx mori</i>	flavin reductase (NADPH)-like isoform X1	74	7	14
cds.c84600	XP004931115	<i>Bombyx mori</i>	retinal dehydrogenase 1-like	76	17	99
cds.c84614	HJ65739	<i>Danaus plexippus</i>	allantoinase	65	13	37
cds.c84684	XP004923742	<i>Bombyx mori</i>	delta(3,5)-Delta(2,4)-dienoyl-CoA isomerase, mitochondrial-like	76	2	2
cds.c84710	GQ45614	<i>Agrotis ipsilon</i>	aldo-keto reductase	83	19	118
cds.c84770	BAM17991	<i>Papilio xuthus</i>	peptidyl-prolyl cis-trans isomerase f	92	12	54
cds.c84780	NP001036972	<i>Bombyx mori</i>	ras protein	91	2	2
cds.c84781	NP001103762	<i>Bombyx mori</i>	muscle LIM protein isoform 1	93	2	3
cds.c84809	AHW40461	<i>Zygaena filipendulae</i>	actin	100	22	357
cds.c84839	GK24580	<i>Chilo suppressalis</i>	odorant-binding protein 4	64	2	2
cds.c84853	XP004928342	<i>Bombyx mori</i>	mitochondrial genome	74	4	4

Name	Accession no.	Organism	Sequence description	Id. %	Pep.	PSM
			maintenance exonuclease 1-like isoform X1			
cds.c84869	44717.1 L4712	<i>Manduca sexta</i>	ferritin	79	2	2
cds.c84869	G41120	<i>Galleria mellonella</i>	26kDa ferritin subunit	81	4	4
cds.c84925	XP004921981	<i>Bombyx mori</i>	mesencephalic astrocyte-derived neurotrophic factor homolog	86	7	15
cds.c84940	XP011556713	<i>Plutella xylostella</i>	cytochrome b5-like	76	2	2
cds.c85008	XP004930044	<i>Bombyx mori</i>	ornithine aminotransferase, mitochondrial-like	89	2	2
cds.c85028	IL29312	<i>Cnaphalocrocis medinalis</i>	glutathione S-transferase epsilon 3	68	9	29
cds.c85038	BAG30740	<i>Papilio xuthus</i>	muscle myosin heavy chain	94	12	19
cds.c85486	CB54950	<i>Helicoverpa armigera</i>	fatty acid-binding protein 3	93	13	33
cds.c87973	AD23378	<i>Manduca sexta</i>	juvenile hormone diol kinase	68	5	5
cds.c9472	HJ69256	<i>Danaus plexippus</i>	putative rab gdp-dissociation inhibitor	94	5	6
	ADK47876	<i>Thaumetopoea pityocampa</i>	Tha p 1 protein	100	5	13
	P86360	<i>Thaumetopoea pityocampa</i>	Tha p 2 protein	100	6	402