Etebari, Kayvan; Palfreyman, Robin W.; Schlipalius, David; Nielsen, Lars K.; Glatz, Richard Vernon; Asgari, Sassan

Deep sequencing-based transcriptome analysis of Plutella xylostella larvae parasitized by Diadegma semiclausum, BMC Genomics, 2011; 12:446

© 2011 Etebari et al; licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The electronic version of this article is the complete one and can be found online at: http://www.biomedcentral.com/1471-2164/12/446

PERMISSIONS

http://www.biomedcentral.com/about/license

Anyone is free:

• to copy, distribute, and display the work;
• to make derivative works;
• to make commercial use of the work;

Under the following conditions: Attribution

• the original author must be given credit;
• for any reuse or distribution, it must be made clear to others what the license terms of this work are;
• any of these conditions can be waived if the authors gives permission.

16 November 2012

http://hdl.handle.net/2440/70557
Deep sequencing-based transcriptome analysis of *Plutella xylostella* larvae parasitized by *Diadegma semiclausum*

Kayvan Etebari1, Robin W Palfreyman2, David Schlipalius3, Lars K Nielsen2, Richard V Glatz4 and Sassan Asgari1*

**Abstract**

**Background:** Parasitoid insects manipulate their hosts' physiology by injecting various factors into their host upon parasitization. Transcriptomic approaches provide a powerful approach to study insect host-parasitoid interactions at the molecular level. In order to investigate the effects of parasitization by an ichneumonid wasp (*Diadegma semiclausum*) on the host (*Plutella xylostella*), the larval transcriptome profile was analyzed using a short-read deep sequencing method (Illumina). Symbiotic polydnaviruses (PDVs) associated with ichneumonid parasitoids, known as ichnoviruses, play significant roles in host immune suppression and developmental regulation. In the current study, *D. semiclausum* ichnovirus (*DsIV*) genes expressed in *P. xylostella* were identified and their sequences compared with other reported PDVs. Five of these genes encode proteins of unknown identity, that have not previously been reported.

**Results:** De novo assembly of cDNA sequence data generated 172,660 contigs between 100 and 10000 bp in length; with 35% of > 200 bp in length. Parasitization had significant impacts on expression levels of 928 identified insect host transcripts. Gene ontology data illustrated that the majority of the differentially expressed genes are involved in binding, catalytic activity, and metabolic and cellular processes. In addition, the results show that transcription levels of antimicrobial peptides, such as gloverin, cecropin E and lysozyme, were up-regulated after parasitism. Expression of ichnovirus genes were detected in parasitized larvae with 19 unique sequences identified from five PDV gene families including vankyrin, viral innexin, repeat elements, a cysteine-rich motif, and polar residue rich protein. Vankyrin 1 and repeat element 1 genes showed the highest transcription levels among the *DsIV* genes.

**Conclusion:** This study provides detailed information on differential expression of *P. xylostella* larval genes following parasitization, *DsIV* genes expressed in the host and also improves our current understanding of this host-parasitoid interaction.
highly characterized of the known mutualistic viruses [10], replicating only within the calyx cells of the reproductive tract of female wasps [11]. No virus replication occurs in parasitized larvae; however expression of encapsidated PDV genes induces different physiological modifications such as interruption of the larval endocrine system and suppression of the host immune system [12-15].

PDVs are classified into two genera based on their host wasp families; *Ichnovirus* containing ichtnoviruses (IVs) and *Bracovirus* containing bracoviruses (BVs) [11,16]. These genera are morphologically distinct and their gene functions vary. Analysis of virion structural components from IVs indicates that the set of structural genes is conserved among wasps associated with IVs and might originate from an ancestral virus [17]. Recently, Bigot et al. suggested that IVs originated from ascoviruses by lateral transfer of ascoviral genes into wasp genomes [18]. Interestingly, the DNA that is encapsidated within PDVs appears more similar to that of eukaryotes than that of other viral genomes [19]. It has been demonstrated that less than 2% of encapsidated PDV genes have homologs in other viruses in contrast to the proviral or structural genes [20]. Most of the PDV DNA appear to be non-coding, except for some groups of genes that are involved in host immune suppression pathways [21]. IV genomes generally encompass more than 20 circular DNA segments ranging in size from 2-28 Kb with estimated total genome sizes ranging from 75 Kb to greater than 250 Kb [20].

Research so far has concentrated mainly on individual genes or small defined groups of host or PDV genes, to explore their function or differential expression following parasitization. Deep sequencing data can provide extensive information about host-parasitoid interactions at the transcriptome level. The large amount of sequencing data that can readily be produced by next-generation sequencing platforms, such as the Illumina GAII, reduces the need for prior sequence knowledge for gene expression profiling and are now making direct sequencing approaches the method of choice for whole transcriptome analysis in many species [22-26]. The large numbers of short reads produced by next-generation sequencers provide opportunities for development of new applications where sequencing only a portion of a molecule is sufficient. However, the analysis of transcriptome data produced by these technologies for organisms with limited genomic information still presents challenges because the sequenced fragments must be aligned against existing good quality reference genomes [27].

In this study, we used deep sequencing to explore the impact of *D. semiclausum* parasitization on its host, Diamondback moth *Plutella xylostella* L. (Lepidoptera, Plutellidae), a notorious pest of cruciferous plants. *P. xylostella* has developed resistance to many groups of chemical insecticides and also *Bacillus thuringiensis* endotoxin [28,29]. This has made *P. xylostella* one of the world’s most destructive insect pests and the estimated global cost of controlling this insect is around US $1 billion annually [28]. This emphasizes the necessity for the continued development of innovative alternative control measures and resistance management strategies. Parasitoids or parasitoid-produced regulatory molecules can be used to improve conventional pest control strategies in sustainable agriculture. We have used a deep sequencing approach to give a comprehensive view of immune- and metabolic-related genes that are differentially expressed in parasitized versus non-parasitized *P. xylostella* larvae, revealing a significant number of *D. semiclausum* ichtnovirus genes (DsIV) which have not been reported previously. This type of study may facilitate new controls for pest larvae by identifying molecules that are crucial for larval immune defence, development, pesticide resistance and other important metabolic regulatory functions.

Results and Discussion

**P. xylostella** transcriptome profile

A transcriptome is the complete set of expressed RNA transcripts in one or more cells. Transcriptome profiling of organisms under stress or parasitization challenge helps us to obtain a better understanding of subsequent related cellular activities in organisms including growth, development, and immune defence. Recently, the newly developed deep sequencing approaches have significantly changed how the functional complexity of the transcriptome can be investigated [22,25,30].

To analyse the transcriptome of *P. xylostella* host larvae following parasitization by *D. semiclausum*, RNA samples isolated from the whole host larvae at various time points after parasitization were pooled together prior to sequencing. This may lead to bias in the results (e.g. a gene may first be up-regulated and down-regulated subsequently); however, we aimed at obtaining an overview of what occurs during parasitism and generating a transcriptome of *P. xylostella* which has not been previously available and to isolate as many DsIV genes as possible. Illumina GAII RNAseq deep sequencing analysis produced approximately 26.6 and 27.1 million single-end reads from RNA extracted from the whole body of non-parasitized (control) and parasitized larvae of *P. xylostella*, respectively. *De novo* assembly using the CLC Genomic Workbench produced 172,660 contigs with a minimum contig size of 100 bp; of these, 66% were between 100-199 bp and only 59,255 (34%) of the contigs were above 200 bp (Table 1). Only contigs above 200 bp were selected for further analysis and 6%
Table 1 Summary of contig statistics resulting from Illumina deep sequencing of parasitized and non-parasitized *P. xylostella* larvae

| Parameters | Number |
|------------|--------|
| Total *de novo* assembled contigs with CLC software | 172,660 |
| Contigs used for BLAST (cut-off: above 200 bp length) | 59,255 |
| Best BLAST matches (cut-off: E-value > E-6) | 26,119 |
| No Hits (contigs without any BLAST match) | 33,136 |
| Not mapped with any Gene Ontology (GO) database | 6087 |
| Annotated sequences (cut-off: GO weight 5E value > E-6) | 20,704 |

of these had nucleotide lengths above 1000 bp (Figure 1). We found 49,922 contigs that shared their greatest homology with *Tribolium castaneum* (Coleoptera) genes with a minimum E-value of 1e-06. *Bombyx mori* (silkworm) was another species with which a range of *P. xylostella* genes showed high levels of homology. One reason for the higher number of hits against the beetle genome is that three times more *T. castaneum* genes than *B. mori* genes are reported in databases.

The short-read data generated by deep sequencing may be problematic when annotating alternative splice variants [24]. However, many gene expression profiling studies that use high-throughput sequencing can also provide valuable annotation information, such as existence of novel genes, exons or splice events which can be used to annotate putative gene sequences for other species. For instance, 39% of the reads from the brain transcriptome of the wasp, *Polistes metricus*, were matched to the honeybee (*Apis mellifera*) genome sequence and EST resources, for annotation [31]. Using a series of filtering and critical cut-off values for BLAST E-value and gene ontology (GO) weighting, 20,704 sequences were annotated by B2GO software http://www.blast2go.org through UniProt KB/TrEMBL and other available databases. GO-annotated consensus sequences were assigned to biological process clusters such as cellular component and molecular function, and distributed among various sub-categories such as metabolism, growth, development, apoptosis, immune defense, molecular processing, signal transduction, transcription regulator activity, catalytic activity etc. (Figure 2).

Comparison of the transcriptome pattern of *P. xylostella* for eight different GO terms (molecular function - Level 2) with those of silkworm http://www.silkdb.org showed high similarity in the distribution of genes across GO categories indicating that the transcriptome analysed was not biased towards particular categories (Figure 3).

Overall, from the assembled contigs of over 200 bp in length, 44% showed similarity with genes or proteins in the NCBI database. The rest may represent unknown genes, non-coding RNA or misassembled contigs that are expected due to the presence of large repetitive or duplicated regions. David *et al.* (2010) suggested that a significant proportion of transcript signatures detected outside predicted genes represent regulatory non-coding RNAs, because these large numbers of non-coding RNA can be antisense, intergenic or overlapping with protein-coding genes [22].

**Effects of parasitism on the transcription of host immune-related genes**

Our sequencing data analyses indicate that parasitism has a significant impact on the transcriptome profile of *P. xylostella* larvae. The ‘Oases’ package was used to assemble data for differential expression analysis. Initially, reads from the parasitized and non-parasitized larvae were cleaned and combined, before *de novo* transcriptome assembly was carried out using Oases 0.1.18. [32]. The individual sets of reads were then mapped back to the previously assembled contigs and counted as a proxy for gene expression. After filtering our dataset using criteria such as number of reads, contig length, E-value (for nearest homolog identity) and greater than 2-fold change, 928 contigs were shortlisted. Figure 4 shows that most of the differentially expressed transcripts for the selected GO terms (Level 2), molecular function and biological process, were up-regulated. Among the contigs, only those related to immunity and development, which were differentially transcribed after *D. semiclausum* parasitization, are displayed in Tables 2 and 3, respectively, since these are the genes most relevant to parasitism. In instances where differential expression of *P. xylostella* genes following *D. semiclausum* parasitization may not be consistent with proteomic observations in other host-parasitoid systems previously reported (e.g. [2]), it is
Figure 2 GO annotation of consensus sequences (Level 2). 20704 sequences were annotated by B2GO software. This program categorized 14075 contigs in biological process, 18127 contigs in molecular function and 10048 contigs in cellular components. The data from InterPro terms, EC codes and KEGG were merged with GO terms for a wide functional range cover in B2GO annotation.

Figure 3 Comparison of P. xylostella (DBM; diamondback moth) transcriptome pattern with that of B. mori (silkworm), based on 8 different molecular function Gene Ontology (GO) terms (Level 2) Silkworm data was obtained from http://www.silkdb.org.
possible that post-transcriptional inhibitory effects of parasitoids’ maternal factors (e.g. PDVs or venom) contribute to these differences. It has been reported that PDV genes are able to interrupt translation of host genes with their host translation inhibitory factors (HTIF) which was initially characterized from *Campolestis sonorensis* IV (CsIV) [33]. Accordingly, in *Heliothis virescens* it was shown that lysozyme activity declined after parasitization by *C. sonorensis* or injection of CsIV; however, the transcript levels of the gene increased after parasitization. This suggested that CsIV may regulate host cell gene expression at the translation level. We also found that lysozyme transcript levels increased following parasitization (Table 2). Another study also showed that lysozyme concentration and activity in *P. xylostella* larvae parasitized by *Cotesia plutellae* was decreased [2]; however, transcript levels were not measured. Recently, Barandoc and Kim [34] showed that the translation of storage protein in *P. xylostella* was inhibited by two BV genes (CpBV15a and CpBV15b). In another study, it was shown that expression of antimicrobial peptides (dipetecin, cecropin A and drosomycin) were either unchanged or minimally induced in parasitized *Drosophila melanogaster* larvae by *Leptopilina boulardi* [35]. The authors concluded that antimicrobial genes are regulated differently independent of those mediating cellular encapsulation.

In addition to post-transcriptional effects of PDV infection, there may also be variations in different host-parasitoid interactions with regard to host transcripts affected by parasitization/PDV infection. In a very recent study, it was shown that in *Spodoptera frugiperda* larvae infected with *Hyposoter didymator* ichnovirus (HdIV) or *Microplois demolitor* bracovirus (MdBV), the differentially expressed transcripts in hemocytes and fat body largely differed depending on the PDV injected suggesting that the tissues responded differently to the different viruses [36]. However, there were a number of host genes that responded similarly when infected with HdIV or MdBV. Based on this study, HdIV affected transcript levels in both hemocytes and fat body, whereas MdBV mostly affected gene expression in the fat body.

In our study, we found that antimicrobial peptides such as gloverin, moricin, lysozyme II and cecropin were up-regulated after parasitoid attack, compared to transcription levels in control unparasitized larvae. Among these, lysozyme II and gloverin had the highest transcription levels in parasitized *P. xylostella* larvae relative to other antimicrobial peptides (Figure 5; Table 2). Transcription levels for cecropin 1 and pxCecropin E increased by more than 3-fold after *D. semiclasmus* attack (Table 2). It has been reported that immune-related genes were also up-regulated in *P. xylostella* larvae in response to microbial challenge (e.g. pxCecropin 5.94 fold, hemolin 3.18 fold and cecropin E 4.99 fold) [37]. Barandoc *et al.* measured pxCecropin expression levels by quantitative RT-PCR (qRT-PCR) in parasitized *P. xylostella* larvae and showed that pxCecropin was suppressed by *C. plutellae* BV [12]. Some antimicrobial peptides, such as moricin and gloverin, were previously found not to be induced by bacterial challenge of lepidopteran larvae, while lysozyme and cecropins are well-known inducible antimicrobial peptides [12,38]. Recently, it has been reported that gloverin expression was changed after bacterial infection in *P. xylostella* [39]. Our data shows that immune responses after parasitoid attack appear to be different from microbial challenge responses, because gloverin and moricin-like peptide were up-regulated about 7 and 11 fold, respectively, after *D. semiclasmus* attack.

Other related genes which were up-regulated in parasitized larvae included a serine protease and serine protease inhibitor (pxSerpin 3) (Table 2). Serine proteases are major immune regulatory proteins, which are found in a wide range of species from insects to mammals. In contrast, a proteomics analysis showed that pxSerpin 2 was suppressed in *P. xylostella* larval plasma during parasitism by *C. plutellae* [7]. Beck *et al.* demonstrated that the ovarian calyx fluid of the ichneumonoid endoparasitoid *Venturia canescens* has the potential to suppress the host immune system due to a putative serpin activity [40]. Here, we found that serine protease inhibitors (JL943746, JL943748, JL943775, JL943864) were over-expressed two- to seven-fold, after parasitism in *P. xylostella*. Thus, the induction of serine protease inhibitors upon immune challenge in parasitized larvae could be part of an endoparasitoid immune suppressive strategy. Many insect protease inhibitors are known to inactive enzymes isolated from entomopathogenic fungi, and their involvement in insect-pathogen interactions has been widely postulated [41]. Aguilar *et al.* reported that five serine proteases involved in a single metabolic cascade were up-regulated in *Anopheles gambiae* upon microbial challenge, but they suggested a role for the proteases in protecting the mosquito from detrimental effects of an uncontrolled spread of immune reaction [42].

Serine proteases also play a significant role in the activation of the prophenoloxidase (proPO) cascade. In insects, proPO is activated upon injury or invasion, which results in localized melanization of the wound area and/or melanotic capsules capturing invading microorganisms and parasites [43]. In the current study, proPO activating protease (PAP) transcription was up-regulated in parasitized larvae of *P. xylostella*. In a genome-wide microarray study of *D. melanogaster*, several genes encoding enzymes of the melanization cascade were found to be up-regulated by *L. boulardi*. 
parasitization [44], consistent with this study. Asgari et al. (2003) reported that venom protein (Vn50) from *Cotesia rubecula* is homologous to serine protease homologs [43]. It is likely that the injection of this protein (or putative homologs thereof) by parasitoid wasps into the host body, may interfere with the proteolytic cascade that leads to the activation of proPO. cDNA microarray analysis of *S. frugiperda* hemocytes and fat body 24 hours after injection of *HdIV* revealed differential expression of several host genes [45]. Among these, eight immune-related genes showed differential expression in hemocytes with proPO-1 and proPO-2 showing up-regulation, while PAP transcript levels declined. Other immune-related genes that were differentially expressed in the hemocytes were galectin, which showed up-regulation, whereas scavenger R, immulec-tin-2, lysozyme and calreticulin showed down-regulation [45]. A recent study confirmed up-regulation of proPOs...
| Locus ID | TSA Accession No. | Nt. Length | Fold change | Protein Description | Species | E value | Nt. ID (%) | Accession no. |
|----------|------------------|------------|-------------|---------------------|---------|---------|-----------|--------------|
| Locus 2992 | JL943792 | 517 | 11 | Monic-like peptide C2 | *Galleria mellonella* | 5.00E-12 | 75 | ABQ42576.1 |
| Locus 11864 | JL943746 | 452 | 7.06 | Serine proteinase inhibitor | *Procambarus clarkii* | 5.00E-14 | 41 | AAGQ2777.1 |
| Locus 13987 | JL943753 | 281 | 6.97 | Gloverin | *Plutella xylostella* | 3.00E-49 | 100 | AC69342.1 |
| Locus 42625 | JL943821 | 487 | 6.71 | Beta-1,3-glucan recognition protein 3 | *Helicoverpa armigera* | 5.00E-30 | 45 | AC31282.1 |
| Locus 652 | JL943855 | 803 | 6.25 | Proline-rich protein | *Galleria mellonella* | 6.00E-18 | 38 | ACQ99193.1 |
| Locus 10530 | JL943744 | 1440 | 6.09 | Transferrin | *Plutella xylostella* | 0 | 99 | BAF36818.1 |
| Locus 19881 | JL943772 | 503 | 4.96 | Odorant binding protein | *Heliothis virescens* | 1.00E-49 | 66 | ACX53743.1 |
| Locus 8118 | JL943865 | 992 | 4.51 | Serine proteinase | *Samia cynthia ricini* | 6.00E-89 | 55 | BAF36818.1 |
| Locus 4070 | JL943818 | 449 | 4.37 | Cecropin E | *Plutella xylostella* | 8.00E-17 | 100 | BAF36818.1 |
| Locus 6056 | JL943850 | 650 | 4.3 | Cecropin 1 (antibacterial peptide) | *Plutella xylostella* | 1.00E-16 | 98 | ACX53743.1 |
| Locus 10530 | JL943744 | 1440 | 6.09 | Transferrin | *Plutella xylostella* | 0 | 99 | BAF36818.1 |
| Locus 11896 | JL943747 | 1453 | 3.28 | Prophenoloxidase-activating proteinase 3 | *Plutella xylostella* | 0 | 93 | BAF36824.1 |
| Locus 4541 | JL943828 | 843 | 3 | Trypsin-like serine proteinase 1 | *Plutella xylostella* | 2.00E-37 | 42 | ADK66277.1 |
| Locus 15717 | JL943753 | 333 | 2.73 | Peptidoglycan recognition protein 56 | *Bombyx mori* | 3.00E-58 | 61 | NP_001036858.1 |
| Locus 5206 | JL943842 | 712 | 3.67 | Hemolin | *Plutella xylostella* | 7.00E-131 | 97 | ACN69054.1 |
| Locus 2270 | JL946205 | 1040 | 6.29 | Lysozyme | *Galleria mellonella* | 6.00E-17 | 38 | ACQ99193.1 |
| Locus 8756 | JL943866 | 1480 | 6.09 | Transferrin | *Plutella xylostella* | 1.00E-49 | 66 | ACX53743.1 |
| Locus 8756 | JL943866 | 1480 | 6.09 | Transferrin | *Plutella xylostella* | 1.00E-49 | 66 | ACX53743.1 |
| Locus 14548 | JL943754 | 580 | 2.87 | Beta-1,3-glucan recognition protein 2 | *Plutella xylostella* | 2.00E-40 | 49 | NP_001036858.1 |
| Locus 15771 | JL943757 | 333 | 2.73 | Peptidoglycan recognition protein 56 | *Bombyx mori* | 3.00E-19 | 57 | AAL31707.1 |
| Locus 13921 | JL943752 | 602 | 2.68 | Serine protease 33 | *Mamestra configurata* | 2.00E-36 | 50 | ACR15983.2 |
| Locus 12724 | JL943748 | 1288 | 2.51 | Serine protease inhibitor 7 | *Bombyx mori* | 5.00E-94 | 46 | NP_001139701.1 |
| Locus 43519 | JL943823 | 256 | 2.5 | Ceramidase | *Aedes aegypti* | 2.00E-21 | 62 | XP_001658093.1 |
| Locus 48995 | JL943836 | 207 | 2.5 | Toll receptor | *Tribolium castaneum* | 2.00E-19 | 59 | XP_971999.1 |
| Locus 1006 | JL943743 | 1154 | 2.41 | Lipocalin | *Bombyx mori* | 1.00E-12 | 75 | ABQ42576.1 |
| Locus 6114 | JL943851 | 1293 | 2.38 | Lysoosomal acid lipase | *Tribolium castaneum* | 1.00E-65 | 39 | XP_972957.2 |
| Locus 28994 | JL943788 | 911 | 2.33 | Pattern recognition serine proteinase | *Manduca sexta* | 3.00E-87 | 51 | AAR29602.1 |
| Locus 20345 | JL943774 | 322 | 2.26 | Trypsin T6 | *Heliothis virescens* | 9 E-16 | 54 | ABR88249.1 |
| Locus 17873 | JL943765 | 412 | 2.25 | NADP+ oxidoreductase | *Danio rerio* | 1.00E-56 | 76 | NP_980858.1 |
| Locus 7994 | JL943864 | 1305 | 2.24 | Serine protease inhibitor (pSerpin 3) | *Plutella xylostella* | 2.00E-99 | 50 | BAF36821.1 |
| Locus 2093 | JL943776 | 2890 | 2.21 | Apolipoporphins | *Manduca sexta* | 0 | 67 | Q54901.0 |
| Locus 20404 | JL943775 | 251 | 2.19 | Serine protease inhibitor (Serpin 13) | *Bombyx mori* | 4.00E-14 | 49 | NP_001139705.1 |
| Locus 41604 | JL943819 | 282 | 2.16 | Broad-Complex isoform Z2 | *Bombyx mori* | 1.00E-30 | 98 | BAD24051.1 |
| Locus 35475 | JL943806 | 363 | 2.15 | Heat Shock Protein (HSP70) | *Acrythosiphon pisum* | 3.00E-12 | 44 | XP_001950064.1 |
| Locus 29819 | JL943791 | 461 | 2.14 | Beta-1,3-glucan recognition protein 2a | *Helicoverpa armigera* | 1.00E-53 | 64 | ACI32826.1 |
| Locus 4518 | JL943827 | 303 | 2.13 | Serine protease 1 | *Lonomia obliqua* | 4.00E-22 | 51 | AAV91432.2 |
in *S. frugiperda* larvae injected with *HdIV*; however, in the same host injected with *MbBV*, proPO transcript levels declined [36] which suggested differential responses of the host to different PDVs.

Transcripts of most proteins, which are involved in the Toll pathway such as Relish, Dorsal, Pelle, Cactus and Toll receptor, were found in our deep sequencing analysis, but only transcription levels of proteins that showed similarity to the Toll receptor were up-regulated [23]. Gene ontology analysis, but only transcription levels of proteins that showed similarity to the Toll receptor were up-regulated [23].

Transcription levels of host development-related genes on **Table 2**

| Locus   | JL943817 | 299 | 2.12 | 1-phosphatidylinositol-4,5-bisphosphate | *Bombyx mori* | 6.00E-45 | 86 | NP_001165393.1 |
|---------|----------|-----|------|---------------------------------------|---------------|----------|----|----------------|
| Locus   | JL943870 | 627 | 2.11 | Peripheral-type benzodiazepine receptor | *Bombyx mori* | 6.00E-34 | 57 | NP_001040343.1 |
| Locus   | JL943771 | 1136| 2.1  | Peroxidasin                             | *Tribolium castaneum* | 1.00E-134 | 62 | XP_968570.1     |
| Locus   | JL943781 | 437 | 2.04 | Hemolymph proteinase 8                  | *Manduca sexta* | 4.0E-49 | 71 | AAA91006.1      |
| Locus   | JL943852 | 1032| -2   | Heat Shock Protein (HSP90)              | *Plutella xylostella* | 8.00E-140 | 100| BAE48742.1      |
| Locus   | JL943766 | 351 | -2.06| Vasorin                                 | *Culex quinquefasciatus* | 1.00E-54 | 85 | XP_001870087.1 |
| Locus   | JL943755 | 615 | -2.07| Arylalkylamine N-acetyltransferase       | *Antheraea pernyi* | 1.00E-29 | 38 | ABD17803.1      |
| Locus   | JL943768 | 253 | -2.12| Flap endonuclease                       | *Caruckia barmesi* | 6.00E-25 | 83 | ACY74444.1      |
| Locus   | JL943789 | 213 | -2.12| Putative lysozyme                       | *Bombyx mori* | 4.0E-22 | 67 | ADA67927.1      |
| Locus   | JL943861 | 3056| -2.14| Fascin                                  | *Tribolium castaneum* | 0 | 79 | XP_972494.1     |
| Locus   | JL943751 | 701 | -2.4 | Hemocyte protease-1                     | *Bombyx mori* | 6 E-75 | 60 | BAG70409.1      |
| Locus   | JL943873 | 364 | -3.25| Trypsin                                | *Helicoverpa armigera* | 5 E-16 | 43 | ACB54039.1      |
| Locus   | JL943783 | 211 | -3.5 | Black (DOPA-deC-like)                   | *Papilio xuthus* | 1.00E-29 | 87 | BAI87832.1      |
| Locus   | JL943843 | 527 | -3.65| Catalase                                | *Takifugu obscurus* | 1.00E-17 | 35 | ABV24056.1      |
| Locus   | JL943800 | 289 | -3.78| Immune-related Hdd1                     | *Hypanthria cunea* | 3.00E-11 | 39 | AAD29279.1      |
| Locus   | JL943750 | 452 | -5.81| Putative defense protein Hdd11          | *Hypanthria cunea* | 8.00E-55 | 67 | O96382.1        |

Fold changes were generated by comparing average read depths for each contig.

The results of our transcriptome analysis also indicated that genes known to be involved in insecticide resistance/detoxification are up-regulated following parasitization (Table 2). In agreement, it has previously been reported that cytochrome P450 (CYP) and glutathione-S-transferase (GST) activities increased in parasitized *P. xylostella* larvae [29]. Takeda et al. (2006) suggested that parasitoid larvae contributed to CYP activity enhancement since the parasitoid hatched two days after oviposition and the CYP activity was significantly increased three days after parasitization [29].

Our analyses also detected a considerable number of other immune-related genes whose transcription levels altered after parasitization by *D. semiclausum*. However, only a small group affected by parasitoid attack were found to be altered enough to be statistically biologically relevant (i.e. showing greater than two-fold change).

**Transcription levels of host development-related genes**

Generally, larval endoparasitoids lay their eggs into the host hemocoel, and their progenies develop by consuming host hemolymph and tissues. As a consequence, the parasitoid’s larval growth also fully depends on the host’s development [1,51,52]. In *P. xylostella* larvae [15,50], Fath-Goodin *et al.* (2009) reported that CsIV vankyrin genes also encode proteins with sequence homology to the inhibitory domains of NF-κB transcription factor inhibitors [13].
parasitized by *D. semiclausum*, development is arrested at the prepupal stage [53]. Developmental arrest before pupation is one of the most common effects of PDVs and/or other maternal factors injected by many endoparasitoids into their hosts [14,54-56]. In these interactions, the parasitoid larvae and PDVs are responsible for increasing the juvenile hormone (JH) titre in host larvae and preventing ecdysteroid levels from rising sufficiently to allow host pupation [57-61].

Based on our transcriptome data, parasitism by *D. semiclausum* leads to down-regulation of genes associated with ecdysteroid activities; for example, the transcription level of ecdysteroid regulated protein was down-regulated more than 26 times in parasitized larvae (Table 3). Considering that ecdysteroids are required to trigger expression of ecdysteroid regulated protein [62], and that parasitism in general leads to reductions in ecdysteroid titres [54,63-66], down-regulation of ecdysteroid regulated protein is expected.

Juvenile hormone binding protein (JHBP) may protect JH from non-specific degradation and adsorption by preventing exposure of JH to epoxide hydration by JH epoxide hydrolase (JHEH), which generates the hormonally inactive JH diol [67]. In agreement, JHBP transcript levels were up-regulated more than 2 times and interestingly, transcription levels of JHEH were down-regulated more than 2 times (Table 3). In all the reported host-parasitoid systems, it seems that JH is maintained at high levels during parasitoid larval development [58,66,68]. Therefore, an increase in JHBP, and corresponding decrease in JHEH transcription levels, after parasitism, seems logical to maintain higher JH levels during parasitism.

As indicated above, previous studies have shown that PDVs may inhibit translation of specific storage or growth-associated proteins despite up-regulation (or steady-state) of transcript levels of the encoding genes, following parasitism or injection of PDVs [69-71]. In this study, we found that arylphorin and methionine-rich storage proteins were over-transcribed in parasitized larvae (Table 3); however, their translation may be affected similarly to mechanisms in the reports discussed above, which needs to be experimentally shown.

**Quantitative RT-PCR validation of transcriptome analysis**

To validate our deep sequencing data, nine differentially regulated *P. xylostella* genes were selected from immune- and development-related genes (Tables 2 and 3) for qRT-PCR analysis, using the same RNA samples as for deep sequencing. These were cuticle protein, ecdysteroid regulated protein, JH epoxide hydrolase, insulin receptor, methionine-rich storage protein 1, gloverin, hemolin, pxSerpin 3 and Toll receptor. The qRT-PCR results confirmed the data obtained from deep sequencing analysis showing similar trends in up- or down-regulation of host genes (Figure 6). For example, based on deep sequencing analysis, gloverin, Toll receptor, and pxSerpin 3 were up-regulated 6.9, 2.5 and 2.2 fold, respectively (Table 2), and showed 5.5, 2.6 and 2.7 fold changes, respectively in qRT-PCR analyses (Figure 6).

Since the samples analysed above were pools of RNA from various time points, and it is known that expression of host genes may vary at different periods after parasitization, we further isolated RNA from 2nd instar *P. xylostella* larvae at 16, 24 and 48 hrs after parasitization, and analyzed three associated genes by qRT-PCR. These were hemolin, gloverin and the ecdysteroid regulated protein. For each time point, a mixture of 10 larvae was used. As expected, there were fluctuations in the expression levels of these genes following parasitization (Figure 7). For example, gloverin expression was highly induced at 16 hrs after parasitization, subsequently declined at 24 hrs, and then further reduced to the same level as unparasitized larvae at 48 hrs post-parasitization (Figure 7). Hemolin was only up-regulated at 48 hrs after parasitization. Expression of ecdysteroid regulated proteint was initially down-regulated at 16 hrs after parasitization, but increased at 24 hrs before subsequently declining by 48 hrs post-parasitization (Figure 7).

**Diadegma semiclausum ichnovirus genes**

PDV genes are divided into three groups based on whether they are expressed in the carrier wasp (class I), the infected host larvae (class II) or both (class III) [72]. Among these genes, class II genes have received the greatest attention and have been studied more than other groups [73]. The genomes of some IVs, such as those found in the wasps *C. sonorensis*, *C. chloridiae*, *Hyposoter fugitivus*, *H. didymator*, and *Tranosema rosstrale*, have been sequenced and resultant data are available on public databases [21,72,74]. Six conserved gene families: repeat element, cysteine motif, viral innexin, viral ankyrin, N-family and the polar-residue-rich proteins (a newly defined putative family), have been reported in most IV genomes [72].

Here, expression of a range of *DsIV* genes were detected in parasitized larvae. In our analysis, 19 unique sequences were identified from five PDV gene families including vankyrin, viral innexin, repeat elements, cysteine-rich motif, and polar residue rich protein families (Table 4). In addition, five other putative virus protein sequences with unknown function were identified in parasitized larvae, and showed more than 50% similarity with some parts of IV reference genomes, but no putative specific protein domain homologies were detected in their sequences (Table 4). The online open reading frame finder tool at the NCBI website http://www.ncbi.nlm.nih.gov/projects/gorf was used for prediction of full-length sequences in *DsIV* genes and only
### Table 3 Developmental- and non-immune metabolism-related transcripts of *P. xylostella*, which were differentially expressed after *D. semiclausum* parasitization

| Sequence ID | TSA Accession No. | Nt. Length | Fold change | Protein Description | Species | E value | Nt. ID (%) | Accession no. |
|-------------|-------------------|------------|-------------|---------------------|---------|---------|-----------|--------------|
| Locus 29717 | JL943790          | 504        | 27.86       | Endonuclease-reverse transcriptase | *Bombyx mori* | 1E-46   | 67 AD61832.1 |
| Locus 3567  | JL943808          | 277        | 20.01       | Glucose dehydrogenase precursor | *Pediculus humanus corporis* | 1E-15   | 52 XP_002429706.1 |
| Locus 784   | JL943863          | 1943       | 12.5        | Methionine-rich storage protein 1 | *Plutella xylostella* | 1E-11   | 63 CBH09254.1 |
| Locus 50007 | JL943839          | 234        | 8.31        | Putative RecQ Helicase | *Heliconus melpomene* | 1E-11   | 63 CBH09254.1 |
| Locus 3712  | JL943813          | 1744       | 7.8         | Methionine-rich storage protein 2 | *Plutella xylostella* | 1E-11   | 63 CBH09254.1 |
| Locus 1098  | JL943745          | 724        | 4.4         | Arylphorin-like hexamerin-2 | *Heliconus melpomene* | 1E-11   | 63 CBH09254.1 |
| Locus 9922  | JL943871          | 1790       | 4.31        | 44 kDa zymogen (serine protease) | *Tenebrio molitor* | 9E-61   | 35 BAG14262.1 |
| Locus 2239  | JL943782          | 2186       | 3.46        | Methionine-rich storage protein | *Spodoptera exigua* | 1E-11   | 63 CBH09254.1 |
| Locus 352   | JL943805          | 1023       | 3.43        | Arylphorin-like hexamerin-1 | *Plutella xylostella* | 2E-172  | 89 BAF32561.1 |
| Locus 50007 | JL943839          | 234        | 3.33        | Syntaxin | *Culex quinquefasciatus* | 3E-21   | 60 XP_001865470.1 |
| Locus 1575  | JL943758          | 1738       | 3.12        | Phosphoribosylaminomimidazole carboxylase | *Bombyx mori* | 1E-34   | 59 NP_001037011.1 |
| Locus 33387 | JL943799          | 375        | 3.11        | Insulin receptor | *Bombyx mori* | 1E-34   | 59 NP_001037011.1 |
| Locus 47369 | JL943832          | 347        | 3          | Leucine-rich transmembrane protein | *Pediculus humanus corporis* | 1E-24   | 60 NP_001037011.1 |
| Locus 7126  | JL943859          | 1075       | 2.98        | Sugar transporter 4 | *Bombyx mori* | 2E-104  | 64 NP_001165395.1 |
| Locus 35539 | JL943807          | 592        | 2.55        | Torso-like protein | *Tribolium castaneum* | 2E-31   | 38 NP_001107843.1 |
| Locus 36082 | JL943809          | 735        | 2.44        | Reverse transcriptase | *Aedes aegypti* | 3E-36   | 39 AA21523.1 |
| Locus 199   | JL943773          | 1588       | 2.4         | S-adenosyl-L-homocysteine hydrolase | *Plutella xylostella* | 2E-101  | 62 AAV91003.1 |
| Locus 5556  | JL943845          | 1186       | 2.33        | Yellow-fa | *Bombyx mori* | 2E-141  | 69 NP_001037424.1 |
| Locus 4588  | JL943829          | 205        | 2.24        | Endoprotease FURIN | *Spodoptera frugiperda* | 4E-28   | 87 CA93116.1 |
| Locus 4894  | JL943835          | 861        | 2.24        | Lipase | *Helicoverpa armigera* | 3E-50   | 54 AGB54943.1 |
| Locus 4689  | JL943830          | 2429       | 2.21        | Cathepsin L precursor | *Tribolium castaneum* | 2E-31   | 38 NP_001107843.1 |
| Locus 3008  | JL943793          | 950        | 2.21        | Hemolymph proteinase 5 | *Manduca sexta* | 2E-101  | 62 AAV91003.1 |
| Locus 20931 | JL943777          | 661        | 2.17        | Gamma-glutamyl transferase | *Bombyx mori* | 1E-62   | 56 NP_001165385.1 |
| Locus 21405 | JL943780          | 705        | 2.14        | Juvenile hormone binding protein | *Manduca sexta* | 2E-27   | 35 AA25736.2 |
| Locus 3797  | JL943815          | 1653       | 2.13        | Imaginal disk growth factor | *Plutella xylostella* | 0       | 100 BAF36822.1 |
| Locus 5088  | JL943841          | 1476       | 2.12        | Cathepsin B-like cysteine proteinase | *Spodoptera exigua* | 1E-146  | 81 ABK90823.1 |
| Locus 1887  | JL943767          | 1872       | 2.1         | Cathepsin D isoform 1 | *Tribolium castaneum* | 7E-165  | 73 XP_966517.1 |
| Locus 47019 | JL943831          | 207        | 2.09        | DNA-binding protein Ewg putative | *Pediculus humanus corporis* | 3E-22   | 93 XP_002430412.1 |
| Locus 16109 | JL943872          | 515        | 2.08        | Nesrin-1 | *Pediculus humanus corporis* | 3E-26   | 36 XP_002427810.1 |
| Locus 41758 | JL943820          | 222        | 2.06        | Neuroglian | *Mythimna separata* | 4E-29   | 78 BAA9425.1 |
| Locus 6830  | JL943856          | 427        | 2.04        | Cytochrome P450 | *Plutella xylostella* | 2E-67   | 96 ABW34440.1 |
four of these genes are reported here as full-length (Table 4).

Presence of multiple sequences for each PDV gene family is common in other IVs and we identified four sequences with vankyrin domain and two viral innexin transcripts, which showed high similarity with *H. fugitivus* IV innexins (Table 4). Repeat element domains were also identified in six sequences with translated lengths of between 209-244 amino acids. Transcription levels (or gene expression values) were found to be different among different members within each gene family. Vankyrin 1 and repeat element 1 had the highest transcription levels in their respective families, and also relative to other *Ds* IV genes (Figure 8).

Hierarchical cluster analysis of vankyrin gene sequences of all reported PDV vankyrins (protein sequences at NCBI) using a neighbour-joining algorithm, classified BVs and IVs into two major groups (Figure 9). All four *Ds* IV vankyrins that were identified in this study have high similarities with other IVs and were distributed into four separate clusters indicating that these members of the ankyrin family possibly originated from different segments of the *Ds* IV genome or at least there has been some ancient gene duplication and/or differential selection even if encoded on the same circle. In addition, the separate clustering of vankyrin genes suggests that they are not closely related, and therefore did not likely undergo recent duplication to form similar paralogs.

**Conclusion**

Overall, this study provides the first comprehensive analysis of the impact of a parasitoid wasp on its host
at the transcriptomic level, using RNA deep sequencing technique. The results showed differential expression of a large number of *P. xylostella* genes, including immune-related genes, upon parasitization by *D. semicaudum*. In addition, although presence of DsIV particles has been reported in parasitized larvae, our results provide evidence for expression of 19 DsIV genes expressed in the host, which have not been previously reported. Analysis of these sequences indicated the presence of conserved genes that belong to major IV class II genes. The transcriptome profiling data sets obtained in this study provide a basis for future research in this under-explored host-parasitoid interaction. In addition, the identified immune-, development- and detoxification-related genes may be targets for *P. xylostella* control and allow manipulation of host-parasite interactions.

**Methods**

**Insects and parasitization**

*P. xylostella* and the parasitoid wasp (*D. semicaudum*) were raised on cabbage plants and host larvae, respectively, at 25°C. Twenty five 3rd and 4th instar *P. xylostella* larvae each were exposed to wasps until parasitization was observed. Individual larvae that had been attacked by the parasitoid were collected and fed on fresh cabbage leaves. Larval samples were taken at
four different time intervals after parasitization (6, 12, 24 and 48 hrs post-parasitization) and the samples were kept at -80°C until RNA isolation. The same numbers of mixed larval instars (3rd and 4th) of non-parasitized larvae were collected as the control treatment. It is worth mentioning that *P. xylostella* larvae parasitized at 3rd instar continue to develop to 4th instar.

**Sample preparation, deep sequencing and de novo transcriptome assembly**

Total RNA was extracted from all larval samples using Tri-Reagent™ (Molecular Research Center Inc.). RNA extracted from larvae at various time points post-parasitization were pooled and therefore temporal expression data was lost. This was also performed for non-parasitized samples. The pooled RNA sample concentrations were measured using a spectrophotometer and integrity was ensured through analysis on a 1% (w/v) agarose gel. The samples with total concentration of 3.9 and 4.1 μg/μl for parasitized and non-parasitized larvae, respectively, were used for cDNA library production.

The cDNA library was prepared by using 5 μg of starting material for the Illumina mRNA Sequencing Sample preparation procedure (kit RS-930-1001). This involved purification and fragmentation of mRNA, first strand cDNA synthesis, second strand cDNA synthesis, end repair, addition of "A" bases to 3' ends, ligation of adapters, purification of ligated products, and PCR amplification to enrich cDNA templates. The library was validated, quantified and subjected to deep sequencing using a Genome Analyzer IIx Next generation sequencer on a 66 cycle single-end sequencing run, following the supplier’s instructions (Geneworks, Adelaide). The GAII analyzer data were output as sequence tags of 65 bases. Sequence.txt files (in FASTQ format) were generated using Illumina Pipeline version 1.5.1. The CLC Genome Workbench (version 4.0.2) [75] algorithm for de novo sequence assembly was used to assemble contigs from a pooling of all the short-read data, using default parameters (similarity = 0.8, length fraction = 0.5, insertion/deletion cost = 3, mismatch cost = 3).

**RNA sequence analysis**

The contigs arising from the de novo assembly were then used as a reference set of transcripts for RNAseq analysis. Short-read sequence data from parasitized and non-parasitized larvae were separately mapped against the reference set of assembled transcripts using the CLC Genome Workbench RNAseq function (min. length fraction = 0.9, maximum mismatches = 2). The relative transcript levels were output as RPKM (Reads Per Kilo-base of exon model per Million mapped reads) values, which take into account the relative size of the transcripts and only uses the mapped-read datasets (i.e. excludes the non-mapped reads), to determine relative transcript abundance. In this way, the output for each dataset can be directly compared as the number of mapped reads per dataset and transcript size has already been taken into account.

Reads from parasitized and non-parasitized larvae were cleaned and combined, before de novo transcriptome assembly was carried out using Oases 0.1.18 [32]. The individual sets of reads were then mapped back to the transcripts using BWA 0.5.8a [76]. The average read depth (proportional to expression level) for each transcript was then calculated using SAMtools 0.1.8 [77]. The transcripts that had a greater than two-fold average read depth difference between the parasitized and non-parasitized sets were counted as being statistically biologically relevant and were selected for annotation. We used both CLC and Oases to compare assembly of contigs. In general, Oases produced similar contigs to CLC, although contig lengths produced by Oases were in some instances longer.

**BLAST homology search and annotation**

BLASTX algorithm [78] with an E-value cut off of 10^-6 was applied to the National Centre for Biotechnology
Table 4  *D. semiclausum* IV transcripts which were detected in parasitized *P. xylostella* larvae

| Protein            | Length (nt) | Accession Number | Similarity (Protein/virus/Accession No.) | length (aa) | E value     | Nt. ID % | Conserved Domains |
|--------------------|-------------|------------------|------------------------------------------|-------------|-------------|----------|------------------|
| Vankyrin 1         | 671*        | JI257593         | vankyrin-b17 (HfIV) AAS90270.1           | 170         | 3.56E-52    | 61       | Yes              |
| Vankyrin 2         | 519*        | JI257594         | vankyrin-d8.3 (HfIV) BAF45734.1          | 159         | 6 E-64      | 80       | Yes              |
| Vankyrin 3         | 257         | JI257595         | hypothetical protein_2 (HdIV) AAR99845.1 | 126         | 3.58E-23    | 61       | Yes              |
| Vankyrin 4         | 361         | JI257596         | vankyrin-b1 (HfIV) AAX24120.1            | 167         | 2 E-39      | 76       | Yes              |
| Viral Innexin 1    | 1370        | JI257597         | viral innexin-b5.1 (HfIV) BAF45654.1     | 354         | 5 E-95      | 60       | Yes              |
| Viral Innexin 2    | 576         | JI257598         | viral innexin-c16 (HfIV) AASS8041.1      | 371         | 2 E-81      | 75       | Yes              |
| Repeat element 1   | 833         | JI257599         | repeat element protein_7 (HdIV) AAR89179.1 | 224         | 4.32E-73    | 67       | Yes              |
| Repeat element 2   | 697         | JI257600         | repeat element protein-d10.1 (HfIV)      | 244         | 1.53E-72    | 64       | Yes              |
| Repeat element 3   | 947         | JI257601         | repeat element protein-c18.1 (HfIV)     | 244         | 9.52E-94    | 71       | Yes              |
| Repeat element 4   | 859*        | JI257602         | repeat element protein-e2.3 (HfIV)      | 209         | 3.40E-76    | 67       | Yes              |
| Repeat element 5   | 656         | JI257603         | repeat element protein (HdIV) AAO16959.1 | 225         | 1.21E-73    | 59       | Yes              |
| Repeat element 6   | 500         | JI257604         | repeat element protein-d11.2 (HfIV)     | 244         | 1E-68       | 72       | Yes              |
| Cysteine rich motif| 630         | JI257605         | cysteine motif gene-d9.1 (HfIV) BAF45736.1 | 311         | 2.99E-55    | 64       | Yes              |
| Polar residue rich | 671         | JI257606         | polar residue rich protein-b13.2 (HfIV) | 159         | 1.24E-12    | 42       | No               |
| Unknown Protein    | 256         | JI257607         | c7.2.1 (TrIV) BAF45559.1                 | 119         | 2.90E-17    | 68       | Yes              |
| Unknown Protein    | 2143*       | JI257608         | c12.1 (HfIV) AAS68099.1                  | 432         | 2E-107      | 50       | No               |
| Unknown Protein    | 418         | JI257609         | P12 (HdIV) AAS83461.1                    | 106         | 6E-18       | 58       | No               |
| Unknown Protein    | 584         | JI257610         | c10.1 (HfIV) AAS90272.1                  | 385         | 9.15E-50    | 51       | No               |
| Unknown Protein    | 208         | JI257611         | b5.3 (HfIV) BAF45655.1                   | 115         | 9.57E-13    | 69       | No               |

* Full length
Hf: Hyposoter fugitives
Hd: Hyposoter didymator
Tr: Tranosema rostrale

---

**Figure 8** Relative gene expression values based on average read depth for all detected *D. semiclausum* ichnovirus genes. RPKM normalized values were used to generate the data.
Information (NCBI) non-redundant protein sequence database, to determine the homology of sequences with known genes. In the absence of *P. xylostella* and *D. semiclausum* genome sequences, we discarded annotations that showed similarity to hymenopteran genes and tried to use annotations that showed the highest similarity to lepidopteran genes. Gene ontology and annotation were performed on all assembled contigs greater than 200 bp length by BLAST2GO software http://www.blast2go.org[79]. For gene ontology mapping, Blast2GO
(which performs four different mapping strategies) was used, and the program defaults were applied for all annotation steps [79]. BLAST2GO allows the selection of a significance level for the False Discovery Rate (FDR), which was used as a cut-off at the 0.05% probability level. The data from InterPro terms [80], enzyme classification codes (EC), and metabolic pathways (KEGG, Kyoto Encyclopedia of Genes and Genomes) were merged with GO terms for a wide functional range cover in annotation.

For some of the identified *D. semiclausum* ichnovirus (DsIV) genes, ORFs were predicted and identified by using ORF finder at NCBI http://www.ncbi.nlm.nih.gov/gorf/gorf.html. Predicted ORFs with highest BLASTp E-values in internal comparisons involving other IV genes, were accepted for further analyses.

**Quantitative RT-PCR (qRT-PCR) validation of deep sequencing data**

Quantitative RT-PCR technique was used on the same RNA samples which were used for transcriptome profiling to verify deep sequencing results using three replicates, each obtained from a pool of 10 larvae. In addition, to observe gene expression levels at different time points after parasitization for a selected group of genes, another experiment was performed by parasitizing 2nd instar *P. xylostella* larvae. For each time point after parasitization, a pool of 10 larvae was used. The RNA samples were extracted from larvae at 16, 24 and 48 hrs after parasitization.

First strand cDNA was synthesized from 1 μg of RNA using M-MuLV reverse transcriptase (New England BioLabs). The qPCR reaction consisted of 2 μl of diluted cDNA (10 ng), 5 μl of Platinum SYBR Green Super-Mix-UDG with ROX (Invitrogen), and 1 μM of each primer (Table 5) in 10 μl total volume. Reactions were performed in triplicates in a Rotor-Gene thermal cycler (Qiagen) under the following conditions: 50°C for 2 min; 95°C for 2 min; and 40 cycles of 95°C for 10 s, 60°C for 15 s, and 72°C for 20 s, followed by melting curve generation (68°C to 95°C). Melting curves for each sample were analyzed to check the specificity of amplification. Gene copy numbers were calculated using the Rotor-Gene software, and an endogenous actin reference gene was used for normalization.

**Table 5 Primers used for qRT-PCR analyses to validate deep sequencing data**

| Gene            | Forward primer | Reverse Primer |
|-----------------|----------------|----------------|
| Storage protein 1 | CAAGACACGCCTACGACGC | GTGGCGCATGCGAGTACC |
| Insulin receptor  | GTACTCCGTGATCTCGG  | CACAGTGCAAGGGAACCC |
| JH epoxide       | AGGATCTACCGGAGGGGC | TGGTACACCCCTCCTGAC |
| Ecdysteroid regulated protein | ACCCCGAAAGAAGCGCAAG  | CTCTGTAATGCTGGTCTAC |
| Cuticle protein  | CAGATAGGAGATCTCTGCTG  | GCTGCTCGTCTATCTAC |
| Toll receptor    | CTCCCGGCAACCGGCTTCA  | CGGACAGAAATTCGAGG |
| Hemolin          | AGCCTCAGAAGACTACC  | GTGGTGATTGAGAACATT |
| Glovemin         | AGCTAGCCGCGGCATCGCG | GCAGCTGACCGCGCGCTTAC |
| PspSerpin 3     | GAAATGCTTACTACAGGC  | TGTAGAAGGGATCTCGTAC |
| Actin           | ATGGAGAAAGATCTGGCAC  | GGAGCTCCGGTGAGCAC |
12. Barandoc KP, Kim J, Kim Y. Cotesia plutellae bracovirus suppresses expression of an antimicrobial peptide, Cecropin, in the Diamondback Moth, Plutella xylostella, challenged by bacteria. J Microbiol 2010, 48(1):117-123.

13. Fath-Goobin A, Kroeber JA, Webb BA. The Ccompotesi sonorensis ichnovirus vankpin protein P-vank-1 inhibits apoptosis in insect S9F cells. Insect Mol Biol 2009, 18(4):497-506.

14. Kwon B, Song S, Choi JY, Je YH, Kim Y. Transient expression of specific Cotesia plutellae bracoviral segments induces prolonged larval development of the diamondback moth, Plutella xylostella. J Insect Physiol 2010, 56(6):650-658.

15. Tian SP, Zhang JH, Wang CZ. Cloning and characterization of two Cotesia plutellae chlorideich ichnovirus vankpin genes expressed in parasitized host Helicoverpa armigera. J Insect Physiol 2007, 53(7):699-707.

16. Beckage NE. Parasitoids and polydnaviruses - An unusual mode of symbiosis in which a DNA virus causes host insect immunosuppression and allows the parasitoid to develop. Bioscience 1998, 48(4):305-311.

17. Volkoff AN, Jouan V, Urbach S, Samain S, Bengon M, Wimckner P, Demetre E, Cousserans F, Provost B, Coulbaey F, Legeai F, Belliveau C, Cussen M, Gyapay G, Diezen JM. Analysis of virion structural components reveals vestiges of the ancestral Ichnovirus genome. PLoS Pathog 2010, 6(5):e1000293.

18. Ego T, Samain S, Auge-Goullou C, Federici BA. Molecular evidence for the evolution of ichnoviruses from ascoviruses by symbiogenesis. BMC Evol Biol 2008, 8:253.

19. Webb B, Fisher T, Nusawatvanit T. The Natural Genetic Engineering of Polydnaviruses. In Natural Genetic Engineering and Natural Genome Editing Edited by: Wittny Z 2009, 1178:146-156.

20. Webb BA. Polydnaviruses biology, genome structure, and evolution. In The insect viruses. Edited by: Miller LK, Ball LA. New York: Plenum Press; 1998:105-139.

21. Webb BA, Strand MR, Dickey SE, Beck MH, Hilgarth RS, Barney WE, Kadas K, Kroeber JA, Lindstrom KG, Ratanadechakul W, Shelby KS, Thoektiakku H, Tumbull MW, Witherell RA. Polydnaviruses reflect their dual roles as mutualists and pathogens. Virology 2006, 347(1):160-174.

22. David JP, Coissac E, Melodelima C, Poupardin R, Rio MA, Chandor-Proust A, Reynaud S. Transcriptome response to pollutants and insecticides in the dengue vector Aedes aegypti using next-generation sequencing technology. BMC Genomics 2010, 11:216.

23. Heigendorz Z, Zakrzewska A, Agoston VC, Oradas A, Racz P, Mink M, Spahn KP, Meijer AH. Deep sequencing of the zebrafish transcriptome response to mycobacterium infection, Mal Inf Mol Biol 2009, 4(1):2918-2939.

24. Morozova O, Marra MA. Applications of next-generation sequencing technologies in functional genomics. Genomics 2008, 92(5):255-264.

25. Xiang L, He D, Dong W, Zhang Y, Shao J. Using expressed sequence tags and deep sequencing-based expression profiling analysis revealed changes of multiple signaling pathways involved in immunity in the large yellow croaker during Aeromonas hydrophila infection. BMC Genomics 2010, 11:506.

26. Chen M, Zhao JZ, Collins HL, Earle ED, Cao J, Shelton AM. A critical assessment of the effects of Bt transgenic plants on parasitoids. PLoS One 2008, 3(5):e2284.

27. Takeda T, Nakamoto Y, Tanaka T. Parasitization by Cotesia plutellae enhances detoxifying enzyme activity in Plutella xylostella. Pestic Biochem Physiol 2006, 86(1):15-22.

28. Mu YN, Ding F, Gou P, Ao JQ, Hu SN, Chen XH. Transcriptome and expression profiling analysis revealed changes of multiple signaling pathways involved in immunity in the large yellow croaker during Aeromonas hydrophila infection. BMC Genomics 2010, 11:506.

29. Bryant MR. Steady progress and recent breakthroughs in the automated genome annotation. Nat Rev Genet 2006, 9(1):62-73.

30. Oases: a transcriptome assembler for very short reads. [http://www.ebi.ac.uk/~zebino/oases].
Etebari et al. BMC Genomics 2011, 12:446
http://www.biomedcentral.com/1471-2164/12/446

Page 18 of 18

Heliolthis virens larvae by calyx fluid from the parasitoid Campopleis sonorensis. J Insect Physiol 1987, 33(5):333-338.
55. Soller M, Lanzrein B: Polydnavirus and venom of the egg-larval parasitoid Chelonus inanitus (Braconidae) induce developmental arrest in the prepupa of its host Spodoptera littoralis (Noctuidae). J Insect Physiol 1996, 42(5):471-481.
56. Webb BA, Dahlman DL: Developmental pathology of Heliolthis virens larvae parasitized by Microplitis croceipes parasite mediated host developmental arrest. Arch Insect Biochem Physiol 1985, 21(2):131-143.
57. Lee S, Kim Y: Juvenile hormone esterase of diamondback moth, Plutella xylostella, and parasitism of Cotesia plutella. J Asia-Pacific Entomol 2004, 7(3):283-287.
58. Li S, Falabella P, Kuriachan I, Vinson SB, Borst DW, Malva C, Pennacchio F: Juvenile hormone synthesis, metabolism, and resulting haemolymph titre in Heliolthis virens larvae parasitized by Toxoneuron nigricipes. J Insect Physiol 2003, 49(11):1021-1030.
59. Pfister-Wilhelm R, Lanzrein B: Stage dependent influences of polydnaviruses and the parasitoid larva on host ecdysteroids. J Insect Physiol 2009, 55(8):707-715.
60. Schafellner C, Markl RC, Schopf A: Inhibition of juvenile hormone esterase activity in Lymantria dispar (Lepidoptera, Lymantriidae) larvae parasitized by Glyptapanteles liparidis (Hymenoptera, Braconidae). J Insect Physiol 2007, 53(8):858-868.
61. Zhu JY, Ye GY, Dong SZ, Fang Q, Hu C: Venom of Peromatus pumarum (Hymenoptera: Peromataidae) induced endocrine changes in the hemolymph of its host, Pieris rapae (Lepidoptera:Pieridae). Arch Insect Biochem Physiol 2009, 71(1):45-53.
62. Mieszaro S, Morton DB: Isolation and partial characterization of a gene from trachea of Manduca sexta that is required and is negatively regulated by ecdysteroids. Cave Biol 1994, 16(2):619-630.
63. Dahlman DL, Coar DL, Koller CN, Neary TJ: Alterations in the ecdysteroid and juvenile hormone titers in Heliothis virescens (Lepidoptera: Noctuidae) by infection with a symbiotic polydnavirus: Comparison with bracoviruses and ichnoviruses. J Gen Virol 1997, 78(9):263-270.
64. Touhara K, Prestwich GD: Juvenile hormone epoxide hydrolase, an enzyme associated with ecdysteroid metabolism in the hornworm eggs. J Biol Chem 1993, 268(26):19604-19609.
65. Pruijssers AJ, Falabella P, Eum JH, Pennacchio F, Brown MR, Strand MR: Infection by a symbiotic polydnavirus induces wasting and inhibits metamorphosis of the moth Pseudoplasia includens. J Exp Biol 2009, 212(18):3998-3006.
66. Strand MR, Dover BA, Johnson JA: Alterations in the ecdysteroid and juvenile hormone esterase profiles of Trichoplusia ni parasitized by the polyembryonic wasp Copidosoma floridanum. Arch Insect Biochem Physiol 1990, 13(1-2):41-51.
67. Touhara K, Prestwich GD: Juvenile hormone epoxide hydrolase, photoaffinity labeling, purification and characterization from tobacco hornworm eggs. J Biol Chem 1993, 268(26):19604-19609.
68. Dover BA, Menon A, Brown MR, Strand MR: Suppression of juvenile hormone esterase in Heliothis virescens by Microplitis demolitor calyx fluid. J Insect Physiol 1995, 41(9):809-817.
69. Nalini M, Kim Y: A putative protein translation inhibitory factor encoded by Cotesia plutellae bracovirus suppresses host hemocyte-spreading behavior. J Insect Physiol 2007, 53(12):1293-1292.
70. Shelby KS, Webb BA: Polydnavirus infection inhibits synthesis of an insect plasma protein arylphorin. J Gen Virol 1994, 75:2285-2292.
71. Shelby KS, Webb BA: Polydnavirus infection inhibits translation of specific growth-associated host proteins. Insect Biochem Mol Biol 1997, 27(3):263-270.
72. Tanaka K, Lapointe R, Barney WE, Makkay AM, Stoltz D, Cusson M, Webb BA: Shared and species-specific features among ichnovirus genomes. Virology 2007, 363(1):26-35.
73. Strand MR: Polydnaviruses. In Insect Virology. Edited by: Asgari S, Janson K. Norfolk: Caister Academic Press; 2010:171-197.
74. Lapointe R, Tanaka K, Barney WE, Whitfield JB, Banks JC, Belliveau C, Stoltz D, Webb BA, Cusson M: Genomic and morphological features of a banchine polydnavirus: Comparison with bracoviruses and ichnoviruses. J Virol 2007, 81(12):6491-6501.
75. CLC Bio: CLC genomics workbench. [http://www.clcbio.com].
76. Li H, Durbin R: Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009, 25(14):1754-1760.
77. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data P: The sequence alignment/map format and SAMTools. Bioinformatics 2009, 25(15):177-178.
78. Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nuc Acids Res 1997, 25(17):3389-3402.
79. Conesa A, Gotz S. Blast2GO: A comprehensive suite for functional analysis in plant genomics. Int J Plant Genomics 2008, 2008:619832.
80. Hunter S, Apweiler R, Attwood TK, Baroja A, Bateman A, Birns D, Bork P, Das U, Daugherty L, Duquenne L, Finn RD, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Laurgaard A, Letunic I, Lonsdale D, Lopez R, Madera M, Madden J, McAnulla C, McDowall J, Misty J, Mitchell A, Mulder N, Natale D, Orengo C, Quinn AF, Selengut JD, Signst CJA, Thimma M, Thomas PD, Valentin F, Wilson D, Wu CH, Yeats C: InterPro: the integrative protein signature database. Nuc Acids Res 2009, 37:D211-D215.

Cite this article as: Etebari et al.: Deep sequencing-based transcriptome analysis of Plutella xylostella larvae parasitized by Diadegma semiclausum. BMC Genomics 2011 12:446.

doi:10.1186/1471-2164-12-446

Submit your next manuscript to BioMed Central and take full advantage of:

• Convenient online submission
• Thorough peer review
• No space constraints or color figure charges
• Immediate publication on acceptance
• Inclusion in PubMed, CAS, Scopus and Google Scholar
• Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit