J. Microbiol. Biotechnol. (2016), 26(11), 1863–1870
http://dx.doi.org/10.4014/jmb.1608.08029

jmb

# Transcriptome Profiling and In Silico Analysis of the Antimicrobial Peptides of the Grasshopper *Oxya chinensis sinuosa*

**In-Woo Kim[1†], Kesavan Markkandan[2†], Joon Ha Lee[1], Sathiyamoorthy Subramaniyam[2], Seungil Yoo[2], Junhyung Park[2*], and Jae Sam Hwang[1*]**

[1]*Department of Agricultural Biology, National Institute of Agricultural Sciences, Rural Development Administration, Wanju 55365, Republic of Korea*

[2]*Theragen ETEX Bio Institute, Theragen Etex Inc., Suwon 16229, Republic of Korea*

*Received: August 11, 2016*
*Revised: August 19, 2016*
*Accepted: August 30, 2016*

First published online
September 2, 2016

**Corresponding authors*
J.S.H.
Phone: +82-63-238-2974;
Fax: +82-63-238-3833;
E-mail: hwangjs@korea.kr
J.P.
Phone: +82-31-888-9318;
Fax: +82-31-888-9335;
E-mail: junhyung.park@theragenetex.com

†These authors contributed equally to this work.

Supplementary data for this paper are available on-line only at http://jmb.or.kr.

pISSN 1017-7825, eISSN 1738-8872



Antimicrobial peptides/proteins (AMPs) are present in all types of organisms, from microbes and plants to vertebrates and invertebrates such as insects. The grasshopper *Oxya chinensis sinuosa* is an insect species that is widely consumed around the world for its broad medicinal value. However, the lack of available genetic information for this species is an obstacle to understanding the full potential of its AMPs. Analysis of the *O. chinensis sinuosa* transcriptome and expression profile is essential for extending the available genetic information resources. In this study, we determined the whole-body transcriptome of *O. chinensis sinuosa* and analyzed the potential AMPs induced by bacterial immunization. A high-throughput RNA-Seq approach generated 94,348 contigs and 66,555 unigenes. Of these unigenes, 36,032 (54.14%) matched known proteins in the NCBI database in a BLAST search. Functional analysis demonstrated that 38,219 unigenes were clustered into 5,499 gene ontology terms. In addition, 26 cDNAs encoding novel AMPs were identified by an in silico approach using public databases. Our transcriptome dataset and AMP profile greatly improve our understanding of *O. chinensis sinuosa* genetics and provide a huge number of gene sequences for further study, including genes of known importance and genes of unknown function.

**Keywords:** Grasshopper, *Oxya chinensis*, transcriptome, antimicrobial peptides, insect-derived AMP

## Introduction

*Oxya chinensis sinuosa* is a grasshopper species belonging to the phylum Arthropoda (Order: Orthoptera; Family: Acrididae; subfamily: Oxyinae) that is edible and widely consumed. There are approximately 1,900 edible insects, and among these, insects belonging to the Orthoptera are the fourth most commonly consumed at its mature stage [3, 19]. Orthopteran insects are also commonly used for entomotherapy [5], and are used to treat various human ailments, including enuresis in women, scorpion stings, anemia, violent headaches, foot inflammation, fertility, hypertension, asthma, stroke, and skin diseases [7]. These conditions are ameliorated by the combination of biochemicals, including proteins, minerals, and fatty acids, that are present in insect hemocytes and fat body mass.

Antimicrobial peptides (AMPs) are small peptides/proteins (~100 aa) that are secreted/triggered by the host innate immune system in response to external microbial infection. AMPs are important components of the host defense system in all invertebrates. In addition, AMPs are considered as an alternative to conventional antibiotics [18]. Recently, AMPs have been predicted in insect transcriptomes, and these transcriptome profiles revealed that defensins, cecropins, and attacins are widely distributed in insects [15]. In the genomic era, most of the bases present

in cells are sequenced by next-generation sequencing (NGS) technologies and annotated using bioinformatics methods [10]. To uncover the hidden benefits of insects in ecology and human health, the insect research community has collectively undertaken two molecular meta-data projects, the 1K Insect Transcriptome Evolution (1KITE) and 5,000 Insect Genome Project (i5k) [17]. The largest sequenced insect genome (~6 Gb) belongs to an Orthopteran (*Locusta migratoria*) [23]. Currently, only a few *Oxya* molecular studies have been published, on heavy metal stresses and nutrient supplements. Specifically, *Oxya hyla hyla* and *Oxya chinensis* have been studied as potential heavy metal bioindicators in industrial agricultural fields [1, 3, 27], and *O. fuscovittata* and *O. hyla hyla* have been studied as nutrient supplements for fish [9] and poultry birds [19].

Since *O. chinensis sinuosa* has long been used as a food source for the South Korean population and it possesses antioxidant and antimicrobial properties [9], we sought to examine its transcriptome. Therefore, in the present study, we set out to obtain high-throughput data of the *O. chinensis sinuosa* transcriptome using Illumina-based NGS. In addition, we obtained the first transcriptome data for a member of the Oxyinae subfamily, and predicted the AMPs as preliminary data to support future molecular studies.

## Materials and Methods

### Insects and Treatment

Adult *O. chinensis sinuosa* were obtained from Jeonnam Agricultural Research & Extension Services, South Korea. For immunization, each grasshopper was injected with log phase *Escherichia coli* ($4 \times 10^4$ colony forming units) suspended in 10 µl of autoclaved 10 mM sodium phosphate buffer (pH 7.4). Non-immunized and immunized grasshoppers were reared at 25°C ± 1°C for 18 h before total RNA isolation.

### *E. coli* Strain and Growth Conditions

*E. coli* KACC 13821 (ATCC 11775) was purchased from the Korean Agricultural Culture Collection (KACC). Bacteria were cultivated overnight in tryptic soy broth (TSB; Difco, USA) in a shaking incubator (200 rpm, 37°C) until the stationary phase. Then, the bacteria were cultivated in fresh TSB medium under the same conditions until log phase (for 3 h). Bacteria were stored in 15% glycerol at −70°C until use.

### Library Preparation and Sequencing

Total RNA was extracted from each sample with the RNeasy Lipid Tissue Kit (Qiagen, Germany) according to the manufacturer's instructions after treatment with RNase-free DNase I (Qiagen) to eliminate genomic DNA. The concentration and integrity of the RNA were assessed with a Thermo Scientific NanoDrop 8000 Spectrophotometer and Agilent 2100 Bioanalyzer, respectively (Agilent Technologies, USA). RNA with an $OD_{260/280}$ ≥1.8 and an RNA integrity number ≥7.0 was used in subsequent experiments. Equal amounts of high-quality RNA from tissues were then pooled for cDNA synthesis and sequencing. The cDNA library was prepared with ~2.5 µg of total RNA according to the Illumina TrueSeq RNA Sample Preparation Kit (Illumina) protocol. The library was then amplified, and the final library yielded ~400 ng of cDNA, with an average fragment size of ~300 bp. The resulting cDNA libraries (for all four samples) were then paired-end sequenced (2 × 150 bp) with NextSeq (Illumina).

### Preprocessing, De Novo Assembly, and Annotations

Paired-end sequence files from four samples (Fastq: R1, R2) were obtained and subjected to processing using Trimmomatic-0.32, with the following parameter settings: leading, 5; trailing, 5; sliding window, 4:15; and minlen, 30. Processed sequences were checked for bacterial contamination using an in-house bacterial database that was constructed from NCBI GenBank. Preprocessed clean reads were mapped to the bacterial database using Bowtie2 with default parameters, and the mapped reads and their respective pairs were removed. From this point on, these sequences are called preprocessed. Total preprocessed sequences from NextSeq were pooled and assembled with Trinity assembler ver. 2.0.6 [11] using default values. To remove redundant sequences, CD-HIT-EST [13] was used, with a 95% sequence similarity cutoff. Finally, transcripts greater than 500 bp were selected for inclusion in the reference transcriptome. The reference transcriptomes were subjected to functional annotation using BLASTX mapping (e-value cutoff $1e^{-5}$) against the UniProt KB (Metazoa) database, and the Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway maps were determined using Blast2GO [6]. GO annotations were classified using the WEGO Web server [25].

### Identification of Differentially Expressed genes

Differentially expressed genes (DEGs) were measured by counting tags in immunized samples and comparing the number with that in non-immunized *O. chinensis sinuosa* samples and were normalized using the RNA Sequence Expected Maximization method [16]. Reads from all samples were mapped to the reference transcriptome, and differential expression was assessed using Trinity utility scripts (align_and_estimate_abundance.pl and abundance_estimates_to_matrix.pl) as instructed (http://trinityrnaseq.github.io/). From the edgeR statistics files, regulated transcripts across libraries were filtered with default parameters (*i.e.*, $1 \leq \log_2$ (FC), FDR < 0.01) using Python scripts. To determine the differential expression pattern from the GenBank datasets, the same procedures were followed.

### Antimicrobial Peptide Prediction and Classification

The deduced amino acid sequences were subjected to AMP prediction using a modified bioinformatics strategy. Peptide

characteristics such as (i) molecular propensity (physicochemical properties) (ii) aggregation propensity (in vitro and in vivo), and (iii) AMP prediction were assessed by using a predefined bioinformatics strategy with the given parameters [24]. In addition, the allergenic propensity of the peptides was predicted using Allerdictor [8, 22]. Finally, the AMPs were mapped with the CAMP database [22] and classified as novel or known. To classify the AMPs as novel, sequences were matched to the CAMP database with two programs, PatMatch (no mismatch) for sequences ≤20 amino acids [19] and BLASTP (1E-05) for sequences ≥20 amino acids. The BLAST results were filtered using a similarity score ≥90. Similar sequences, using the given cutoff, were considered as known AMPs, and the others were considered novel AMPs.

### Data Deposition

The raw reads from *O. chinensis sinuosa* were submitted to the NCBI Sequence Read Archive under Accession No. SRP080832.

## Results and Discussion

### Transcriptome Sequencing and Assembly

The de novo transcriptome of *O. chinensis sinuosa* was processed as depicted in Fig. 1. To obtain a more comprehensive view of the transcriptome, pooled cDNA samples from the non-immunized control and *E. coli*-immunized *O. chinensis sinuosa* were isolated. Illumina sequencing produced an average of 40,809,315 and 40,232,081 clean reads, representing 5.52 Gb and 5.41 Gb of the non-immunized control and immunized samples, respectively (Table 1). The quality of the transcriptome sequence was high, with a $Q_{20}$ percentage (percentage of sequences with a sequencing error rate lower than 1%) of

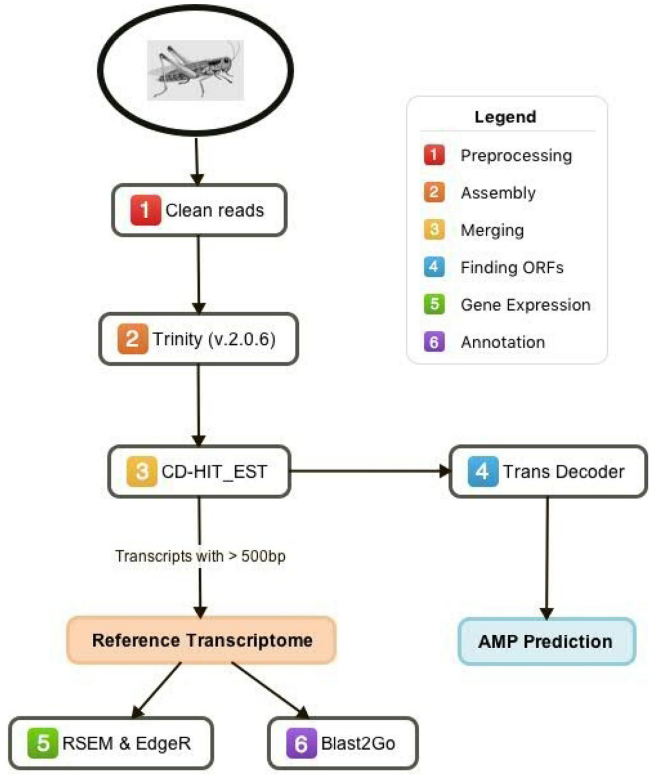


**Fig. 1.** Workflow of the transcriptome assembly and analysis of *Oxya chinensis sinuosa* high-throughput sequencing data.

85.74% and 45.16% G+C. These short reads were assembled into 94,348 contigs with a mean length of 1,439 bp. We obtained 66,555 unigenes, with a mean size of 1,286 bp (range, 500–26,614 bp), implying that the pipeline used to assemble the *O. chinensis sinuosa* transcriptome libraries

**Table 1.** Sequencing, assembly, and annotation statistics.

| A. Sequencing | | | | |
|---|---|---|---|---|
| | Control (non-immunized) | | Immunized | |
| Total sequenced bases | 13,720,373,521 | 100% | 13,550,796,905 | 100% |
| Total preprocessed bases | 11,053,358,720 | 81% | 10,834,378,027 | 80% |
| **B. Assembly** | | | | |
| | Contigs | | Bases | |
| Trinity | 94,383 | 100% | 135,864,252 | 100% |
| CD-HIT-EST | 64,047 | 67.8% | 82,380,545 | 60.6% |
| **C. Annotation** | | | | |
| | No. of Contigs | | % | |
| No. of hits | 30,523 | | 45.6% | |
| BLAST | 36,032 | | 54.1% | |
| Gene Ontology | 18,991 | | 29.6% | |
| KEGG | 3,115 | | 4.8% | |

was satisfactory (Fig. 1). Of these unigenes, 28,049 (42.14%) were longer than 1,000 bp. BLAST searching matched 36,032 unigenes (54.14%) to known genes, suggesting that the assembly quality was high (Table 1). The remaining unigenes could not be matched to any known gene, as there is currently no genome information available for *O. chinensis sinuosa*. Recently, the transcriptome profiles of some insect species in the Acrididae family, including *Schistocerca gregaria*, *Locusta migratoria*, and *Epacromius coerulipes*, were investigated using cDNA libraries [2, 4, 14].

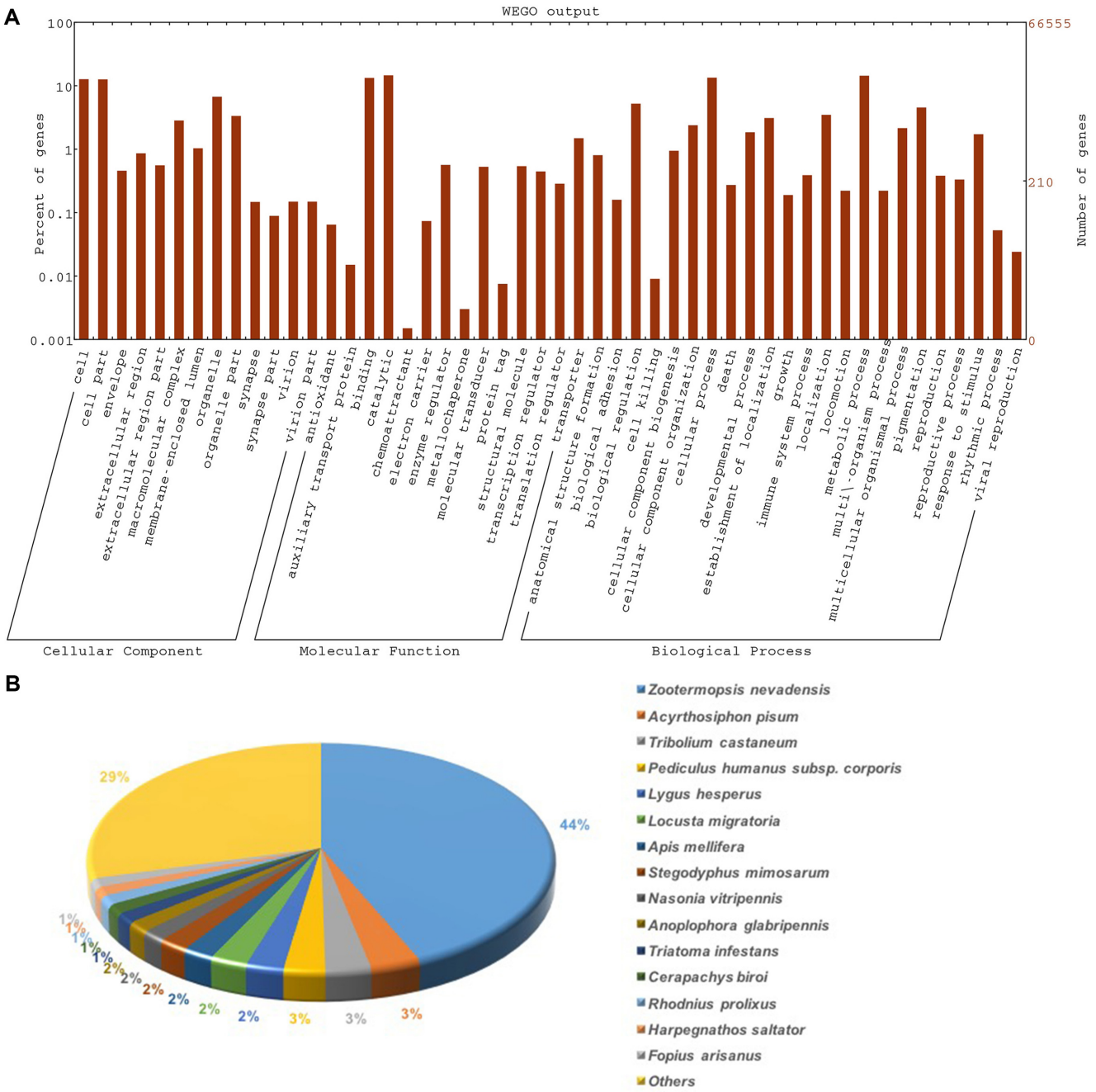


**Fig. 2.** Annotation of the *Oxya chinensis sinuosa* transcriptome.

(**A**) Histogram of the GO classifications. The *O. chinensis sinuosa* transcriptome was annotated in three ontology categories: "Biological Processes," "Cellular Component," and "Molecular Function." (**B**) Species distribution of the BLASTX matches to transcriptome unigenes against the nr protein database (cutoff value $E < 10^{-5}$) and the proportion of matches in each species.

In addition, the antennal transcriptome of the odorant-binding proteins in the grasshopper *Oedaleus asiaticus* has also been studied [26]. However, there has been no published research on bacteria-immunized grasshoppers. The current transcriptome sequencing of *O. chinensis sinuosa* provided significant amounts of information (66,555 unigenes) compared with that available for other grasshoppers, revealing more detailed genetic and gene expression data that were produced using whole-body transcriptome sequencing.

**Annotation of Predicted Proteins**

To determine the putative annotations of the reference transcripts, distinct sequences were first searched by BLASTX against the GenBank non-redundant protein database, with an E-value cutoff of $10^{-5}$ using Blast2GO, as described in the Materials and Methods section. A total of 36,032 distinct sequences (54.14% of the unigenes) were matched to known genes that encode functional proteins. The percentage of unigenes matching known proteins was much higher than those reported in the whole-body transcriptomes of *S. gregaria* and *L. migratoria* [2, 4]. Of the total unigenes, almost half (30,523) shared no significant similarity to known genes; therefore, these may be novel or fast-evolving sequences. Studying these transcripts will provide information about insects in the family Acrididae. In addition, our results will also be useful for future studies, such as the cloning and characterization of *O. chinensis sinuosa* genes. In the functional annotation, 5,499 GO terms were assigned, which were subsequently categorized into three level 2 category observations; biological processes (13,736), cellular component (8,834), and molecular function (15,649) (Fig. 2A). Within the biological processes category, genes encoding metabolic processes (16.8%) and cellular processes (16.7%) were the most enriched. Proteins related to cells (14.4%) and cell part (14.3%) were enriched in the cellular component category. With regard to the category of molecular function, catalytic (27.0%) and binding (26.6%) were the most highly represented categories. It is not very surprising that numerous sequences were classified into every GO category. These are more general GO terms that comprise the basic processes required to maintain a living organism. The top BLASTX hit showed that *O. chinensis sinuosa* is highly homologous to *Zootermopsis nevadensis* (44%), followed by *Acyrthosiphon pisum* (3%). *O. chinensis sinuosa* shared only 2% homology with *Locusta migratoria* (Fig. 2B). Our data are the first ever transcriptome profile of an organism belonging to the genus *Oxya*, and the above results suggest more representative collections of *O. chinensis sinuosa* genes in this study.

**Analysis of Differentially Expressed Genes**

Pairwise comparisons of the non-immunized control and immunized *O. chinensis sinuosa* for differential expression analysis revealed a total of 2,887 DEGs (Table S1). The MA plot showed significant DEGs (blue) against all nonsignificant DEGs (red) (Fig. 3). Among the identified DEGs, 1,800 were expressed at significantly higher levels in the control, whereas 1,087 genes were expressed at significantly higher levels in the immunized sample. In addition, 2,117 genes showed more than a 10-fold difference between the non-immunized control and immunized samples (Fig. 3). Of the DEGs, 1,053 could not be annotated using any database, and 419 of these were more highly expressed in immunized

| Fold change cut-off | Oxya control vs. Immunized | | |
|---|---|---|---|
| | Up | Down | Total |
| FC >= 2 | 1,800 | 1,087 | 2,887 |
| FC >= 4 | 1,756 | 936 | 2,692 |
| FC >= 10 | 1.505 | 612 | 2,117 |

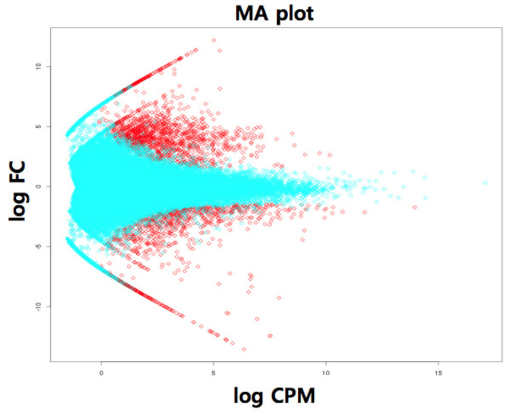


**Fig. 3.** MA plot of differentially expressed genes in the transcriptome of non-immunized control and *E. coli*-immunized *O. chinensis sinuosa*.

Data are the individual gene responses plotted as $\log_2$ fold-change versus base mean fold-change >2 ($p < 0.05$), with negative changes representing downregulated genes and positive changes representing upregulated genes.

**Table 2.** In silico functional characterization of identified AMPs.

| Filter | Propensity | Method | Cutoff | No. of total peptide Seq |
|---|---|---|---|---|
| Step 1 | Molecular | Pepstats | Total peptides (≤100) | 58,696 |
| | | | Charge >0(+) | 55,073 |
| | | | 8 ≤ pI ≤ 12 | 41,207 |
| | Aggregation | TANGO (in vivo) | AGG ≤ 500 | 42,866 |
| | | | 0 ≤ HELIX ≤ 25 | 41,612 |
| | | | 25 ≤ BETA ≤ 100 | 18,134 |
| | Homologs | BLAST | Known (similarity >90%) | 1,627 |
| | Novel | | Novel (no BLAST score) | 57,069 |
| | Allergen | Allerdictor | Non-allergen | 58,261 |
| | Filter 1 total (common) | | | 9,015 |
| Step 2 | Aggregation | Aggrescan (in vitro) | -40 ≤ Na4vSS ≤ 60 | 8,065 |
| | CAMP | SVM | AMP | 1,526 |
| | | RF | AMP | 1,908 |
| | | ANN | AMP | 2,934 |
| | | DA | AMP | 2,024 |
| | | (3 out of 4) | ≥ 3 | 1,496 |
| | Final peptides (Filter 2 common) | | | 1,346 |

samples. Of the 2,887 DEGs, 512 were annotated as homologous genes in *Zootermopsis nevadensis*, a dampwood termite that is closely related to *O. chinensis sinuosa*. The number of genes with higher expression levels in immunized *O. chinensis sinuosa* was lower than that in the control, which clearly showed that the bacterial infection was significant.

**In Silico Identification of AMPs**

Insects produce a greater number of AMPs than any other taxonomic group, and the number of individual AMPs produced by each insect species varies substantially [18, 21]. A vast number of studies have been performed that have identified more than 100 insect-specific AMPs by high-throughput, forward genetics approaches and various screening procedures [12]. Our comprehensive transcriptome dataset from non-immunized control and bacteria-immunized *O. chinensis sinuosa* was screened extensively for cDNAs encoding AMPs using public databases (Table 2). We identified 26 novel AMPs (non-allergen) belonging to diverse families and functional classes (Table S2). Among these, nine are involved in various biological functions, whereas the remaining 17 have unknown/uncharacterized functions. Furthermore, our DEG results revealed that most of the AMPs were differentially expressed, and this was visualized by preparing a heat map to compare the normalized mapped read (RPKM) values of each AMP between the non-immunized control and immunized samples (Fig. 4).

High-throughput analyses, such as RNA-Seq, have led to the discovery of not only many immune-related genes but also the corresponding networks and pathways involved in pathogen recognition, signal transduction, and effector functions, including AMPs [20]. Previous reports have identified several potential AMPs in insects of various orders [18]. Recently, the transcriptome of the ladybird beetle, *Harmonia axyridis*, was analyzed for AMPs induced against *E. coli*, *Micrococcus luteus*, and *Saccharomyces cerevisiae* [21]. Several novel AMPs in *Periplaneta americana* were identified in an *E. coli*-immunized de novo transcriptome and were validated in vitro. Indeed, most of these AMPs are involved in defense and protein binding [15].

In conclusion, transcriptome sequencing and annotation can be used to provide remarkable information for analyzing the molecular basis of the economically important traits of an organism. Here, we described the first comprehensive investigation of the *O. chinensis sinuosa* transcriptome. In this study, we characterized the transcriptome of *E. coli*-immunized *O. chinensis sinuosa* and identified a significant number of AMPs through an in silico approach. The transcriptome data assembled in this study will be a valuable resource for future studies, including gene expression and annotation studies of the *O. chinensis sinuosa* genome.

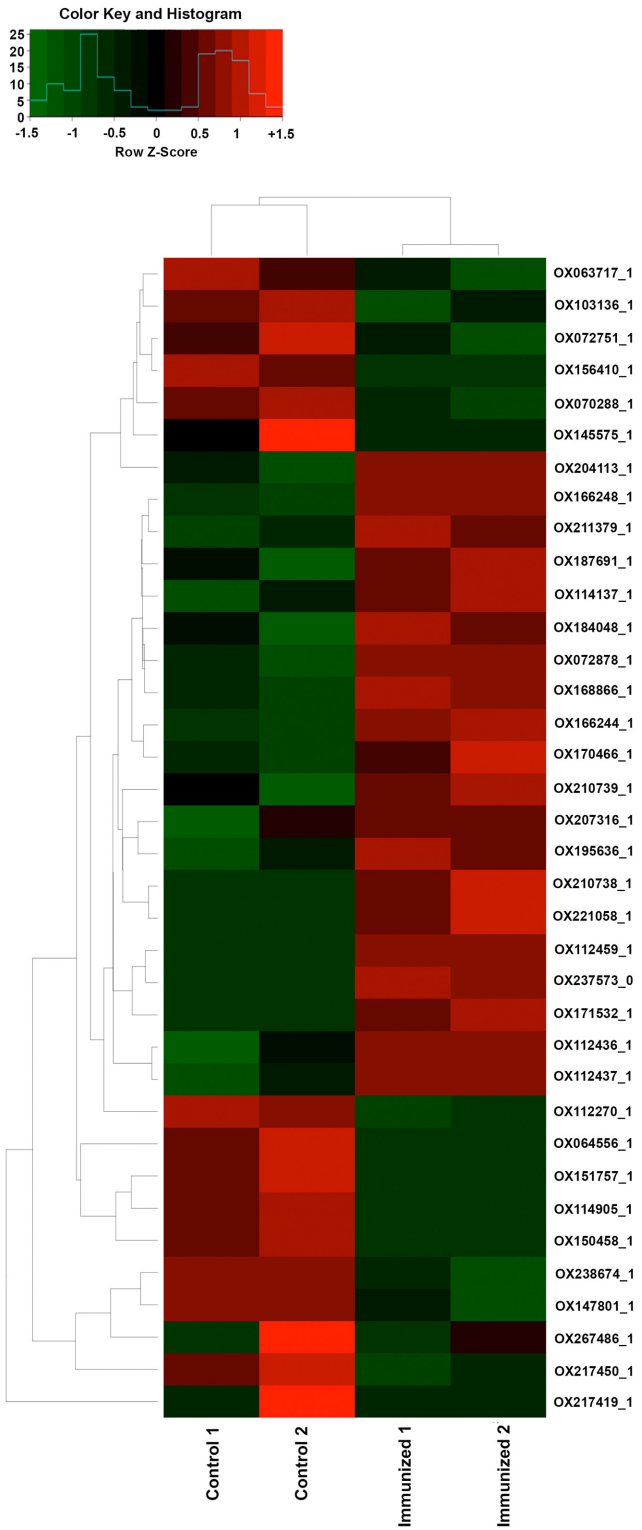


**Fig. 4.** Heat map showing relative expression levels of differentially expressed *O. chinensis sinuosa* antimicrobial peptides. The data are $\log_2$-transformed RPKM values (green, lower-expressed genes; red, highly expressed genes).

Promising insect-derived AMPs are currently being employed in various medicinal applications. Therefore, our future work will focus on producing *Oxya*-derived AMPs in the quantities required for topical therapeutic applications.

## Acknowledgments

This research was supported by a grant (PJ01099301) from the Agenda Program, Rural Development Administration.

## References

1. Azam I, Afsheen S, Zia A, Javed M, Saeed R, Sarwar MK, Munir B. 2015. Evaluating insects as bioindicators of heavy metal contamination and accumulation near industrial area of Gujrat, Pakistan. *Biomed. Res. Int.* **2015:** 942751.
2. Badisco L, Huybrechts J, Simonet G, Verlinden H, Marchal E, Huybrechts R, *et al.* 2011. Transcriptome analysis of the desert locust central nervous system: production and annotation of a *Schistocerca gregaria* EST database. *PLoS One* **6:** e17274.
3. Badman J, Harrison JF, McGarry MP. 2007. Grasshoppers in research and education: methods for maintenance and production. *Lab Anim. (NY)* **36:** 27.
4. Chen S, Yang P, Jiang F, Wei Y, Ma Z, Kang L. 2010. De novo analysis of transcriptome dynamics in the migratory locust during the development of phase traits. *PLoS One* **5:** e15633.
5. Cherniack EP. 2010. Bugs as drugs, Part 1: Insects: the "new" alternative medicine for the 21st century. *Altern. Med. Rev.* **15:** 124-135.
6. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21:** 3674-3676.
7. Costa-Neto EM. 2005. Entomotherapy, or the medicinal use of insects. *J. Ethnobiol.* **25:** 93-114.
8. Dang HX, Lawrence CB. 2014. Allerdictor: fast allergen prediction using text classification techniques. *Bioinformatics* **30:** 1120-1128.
9. Ganguly A, Chakravorty R, Sarkar A, Mandal DK, Haldar P, Ramos-Elorduy J, Moreno JMP. 2014. A preliminary study on *Oxya fuscovittata* (Marschall) as an alternative nutrient supplement in the diets of *Poecillia sphenops* (Valenciennes). *PLoS One* **9:** e111848.
10. Goodwin S, McPherson JD, McCombie WR. 2016. Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* **17:** 333-351.
11. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, *et al.* 2013. De novo transcript sequence reconstruction from RNA-Seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8:** 1494-1512.

12. Hadley EB, Hancock RE. 2010. Strategies for the discovery and advancement of novel cationic antimicrobial peptides. *Curr. Top. Med. Chem.* **10:** 1872-1881.
13. Huang Y, Niu B, Gao Y, Fu L, Li W. 2010. CD-HIT Suite: a Web server for clustering and comparing biological sequences. *Bioinformatics* **26:** 680-682.
14. Jin Y, Cong B, Wang L, Gao Y, Zhang H, Dong H, Lin Z. 2016. Differential gene expression analysis of the *Epacromius coerulipes* (Orthoptera: Acrididae) transcriptome. *J. Insect Sci.* **16:** 42.
15. Kim I-W, Lee JH, Subramaniyam S, Yun E-Y, Kim I, Park J, Hwang JS. 2016. De novo transcriptome analysis and detection of antimicrobial peptides of the American cockroach *Periplaneta americana* (Linnaeus). *PLoS One* **11:** e0155304.
16. Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12:** 1.
17. Misof B, Liu S, Meusemann K, Peters RS, Donath A, Mayer C, *et al.* 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science* **346:** 763-767.
18. Mylonakis E, Podsiadlowski L, Muhammed M, Vilcinskas A. 2016. Diversity, evolution and medical applications of insect antimicrobial peptides. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **371:** 20150290.
19. Van Huis A, Van Itterbeeck J, Klunder H, Mertens E, Halloran A, Muir G, Vantomme P. 2013. *Edible Insects: Future Prospects for Food and Feed Security*. FAO Forestry, Rome.
20. Vilcinskas A. 2013. Evolutionary plasticity of insect immunity. *J. Insect Physiol.* **59:** 123-129.
21. Vilcinskas A, Mukherjee K, Vogel H. 2013. Expansion of the antimicrobial peptide repertoire in the invasive ladybird *Harmonia axyridis*. *Proc. Biol. Sci.* **280:** 20122113.
22. Waghu FH, Barai RS, Gurung P, Idicula-Thomas S. 2015. CAMPR3: a database on sequences, structures and signatures of antimicrobial peptides. *Nucleic Acids Res.* **44:** D1094-D1097.
23. Wang X, Fang X, Yang P, Jiang X, Jiang F, Zhao D, *et al.* 2014. The locust genome provides insight into swarm formation and long-distance flight. *Nat. Commun.* **5:** 2957.
24. Yan T, Yoo D, Berardini TZ, Mueller LA, Weems DC, Weng S, *et al.* 2005. PatMatch: a program for finding patterns in peptide and nucleotide sequences. *Nucleic Acids Res.* **33:** W262-W266.
25. Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, *et al.* 2006. WEGO: a Web tool for plotting GO annotations. *Nucleic Acids Res.* **34:** W293-W297.
26. Zhang S, Pang B, Zhang L. 2015. Novel odorant-binding proteins and their expression patterns in grasshopper, *Oedaleus asiaticus*. *Biochem. Biophys. Res. Commun.* **460:** 274-280.
27. Zhang Y, Sun G, Yang M, Wu H, Zhang J, Song S, *et al.* 2011. Chronic accumulation of cadmium and its effects on antioxidant enzymes and malondialdehyde in *Oxya chinensis* (Orthoptera: Acridoidea). *Ecotoxicol. Environ. Saf.* **74:** 1355-1362.