

# Nutrigenomics in *Arma chinensis*: Transcriptome Analysis of *Arma chinensis* Fed on Artificial Diet and Chinese Oak Silk Moth *Antheraea pernyi* Pupae

Deyu Zou<sup>1,2</sup>, Thomas A. Coudron<sup>3</sup>, Chenxi Liu<sup>1,2</sup>, Lisheng Zhang<sup>1,2</sup>, Mengqing Wang<sup>1,2</sup>, Hongyin Chen<sup>1,2\*</sup>

**1** Key Laboratory of Integrated Pest Management in Crops, Ministry of Agriculture, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China, **2** Sino-American Biological Control Laboratory, USDA-Agricultural Research Service, Beijing, China, **3** Biological Control of Insects Research Laboratory, USDA-Agricultural Research Service, Columbia, Missouri, United States of America

## Abstract

**Background:** The insect predator, *Arma chinensis*, is capable of effectively controlling many pests, such as Colorado potato beetle, cotton bollworm, and mirid bugs. Our previous study demonstrated several life history parameters were diminished for *A. chinensis* reared on an artificial diet compared to a natural food source like the Chinese oak silk moth pupae. The molecular mechanisms underlying the nutritive impact of the artificial diet on *A. chinensis* health are unclear. So we utilized transcriptome information to better understand the impact of the artificial diet on *A. chinensis* at the molecular level.

**Methodology/Principal Findings:** Illumina HiSeq2000 was used to sequence 4.79 and 4.70 Gb of the transcriptome from pupae-fed and artificial diet-fed *A. chinensis* libraries, respectively, and a *de novo* transcriptome assembly was performed (Trinity short read assembler). This resulted in 112,029 and 98,724 contigs, clustered into 54,083 and 54,169 unigenes for pupae-fed and diet-fed *A. chinensis*, respectively. Unigenes from each sample's assembly underwent sequence splicing and redundancy removal to acquire non-redundant unigenes. We obtained 55,189 unigenes of *A. chinensis*, including 12,046 distinct clusters and 43,143 distinct singletons. Unigene sequences were aligned by BLASTx to nr, Swiss-Prot, KEGG and COG (E-value <math>10^{-5}</math>), and further aligned by BLASTn to nt (E-value <math>10^{-5}</math>), retrieving proteins of highest sequence similarity with the given unigenes along with their protein functional annotations. Totally, 22,964, 7,898, 18,069, 15,416, 8,066 and 5,341 unigenes were annotated in nr, nt, Swiss-Prot, KEGG, COG and GO, respectively. We compared gene expression variations and found thousands of genes were differentially expressed between pupae-fed and diet-fed *A. chinensis*.

**Conclusions/Significance:** Our study provides abundant genomic data and offers comprehensive sequence information for studying *A. chinensis*. Additionally, the physiological roles of the differentially expressed genes enable us to predict effects of some dietary ingredients and subsequently propose formulation improvements to artificial diets.

**Citation:** Zou D, Coudron TA, Liu C, Zhang L, Wang M, et al. (2013) Nutrigenomics in *Arma chinensis*: Transcriptome Analysis of *Arma chinensis* Fed on Artificial Diet and Chinese Oak Silk Moth *Antheraea pernyi* Pupae. PLoS ONE 8(4): e60881. doi:10.1371/journal.pone.0060881

**Editor:** Zhanjiang Liu, Auburn University, United States of America

**Received:** December 19, 2012; **Accepted:** March 4, 2013; **Published:** April 11, 2013

**Copyright:** © 2013 Zou et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This research was supported by the 973 Program (2013CB127602), Special Fund for Agro-scientific Research in the Public Interest (no. 201103002), Project of International Cooperation of Agriculture of China (no. 2011-G4) and Sino-America Biocontrol International Cooperation Program (no. 58-4001-4-053). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: hongyin@gmail.com

## Introduction

*Arma chinensis* is a predaceous insect species that preys upon a large variety of species, and can effectively suppress agricultural and forest pests in the orders Lepidoptera, Coleoptera, Hymenoptera and Hemiptera [1–8]. Colorado potato beetle, *Leptinotarsa decemlineata*, one of the most important coleopteran exotic pests in Asia, has developed resistance to most insecticides used for its control [9], and has the potential to develop resistance to *Bacillus thuringiensis* (Bt) toxins [10,11]. The application of transgenic Bt crops has suppressed lepidopteran insects such as the cotton bollworm *Helicoverpa armigera*. However, Lu *et al.* concluded that mirid bugs (Heteroptera: Miridae) in northern China have progressively increased in population size and acquired pest status in cotton and multiple other crops, in association with a regional increase in Bt cotton adoption [12]. Thus, the release of *A. chinensis*

in association with transgenic crops may be a sustainable biocontrol strategy to decrease dependence on insecticides.

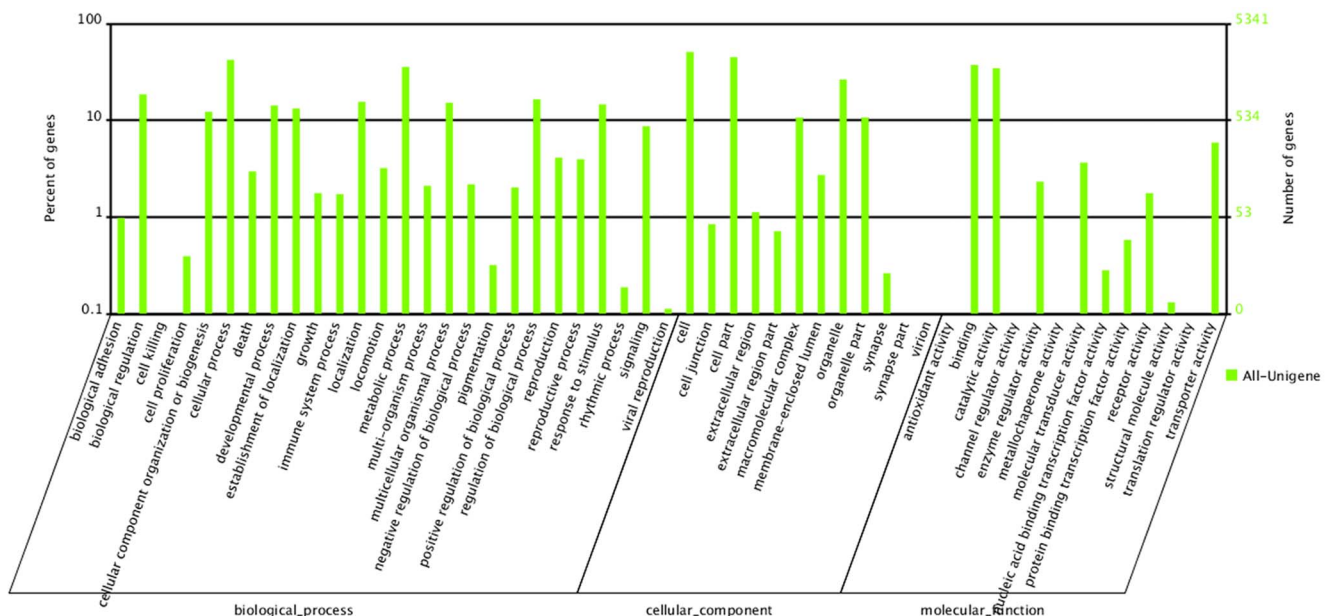
Mass rearing of biocontrol insects is important given the environmental, health and resistance issues associated with the use of chemical insecticides [13]. To achieve an effective level of control, however, requires a large number of beneficial insects be available at low cost for their augmentative and inoculative release. Conventional rearing methods require raising natural prey and (or) host insects on host plants. The extensive input necessary for this method makes the use of beneficial insects economically unfeasible. However, the development of artificial diets could considerably reduce the mass propagation costs of beneficial insects [14–18]. An insect-free artificial diet comprised of pig liver and tuna was developed for *A. chinensis*. Fecundity and egg viability were lower for diet-fed *A. chinensis* compared to *A. chinensis* reared on pupae of the natural prey, the Chinese oak silk moth *Antheraea*

*pernyi*. Developmental time from 2nd instar to adult and the preovipositional period were significantly longer for diet-fed *A. chinensis*. Nymphal weight, body length, adult longevity, survival from 2nd instar to adult, and fertility increased, while sex ratio (♀/♂) decreased, with the rearing of consecutive generations on the diet. In particular, the longevity of adults reared on the artificial diet was significantly longer than of those reared on pupae. Additionally, the diet-fed *A. chinensis* released less defensive odor than did the pupae-fed insects.

The current method for optimizing diet is to measure a few pre-selected biochemical and (or) physiological parameters to test the effect of changes in diet formulation on insect performance [18–21]. Typically, diet components are changed one at a time and insect performance is tested after each change. This endeavor is time-consuming, taking years to decades to optimize a diet, with many attempts ending in failure. To accelerate diet development, a more direct method that can provide informative feedback to target deficiencies in diet formulation is required. Nutrigenomics examines how nutrition affects gene expression patterns and offers not only a means to measure an insect's response to changes in the food stream but also provides information on diet limitations [22]. Using suppressive subtractive hybridization, Yocum *et al.* discovered two artificial diet up-regulated and two prey up-regulated transcript fragments in the predatory pentatomid *Perillus bioculatus*, and a BLASTx search found similarities for two diet up-regulated clones, i.e., the tyrosine-3-monooxygenase gene and the chitin binding protein gene, Gasp [22]. Coudron *et al.* reared groups of oriental fruit fly, *Bactrocera dorsalis* separately on media either devoid of, or supplemented with, wheat germ oil and identified one gene encoding receptor for activated C kinase 1 that increased in expression by 6.8-fold in eggs from adults reared on media supplemented with wheat germ oil. The receptor for activated C kinase 1 is an essential component of at least three intracellular signal transduction pathways, making it a good candidate molecular marker of lipid deficiency in fruit flies and possibly many other insect species [23]. Alaux *et al.* compared the

transcriptome of bees fed with pollen and sugar and bees restricted to a sugar diet, and found that pollen activated nutrient-sensing and metabolic pathways. In addition, those nutrients had a positive influence on genes affecting longevity and the production of some antimicrobial peptides [24]. Coudron *et al.* found that trace element levels in *Podisus maculiventris* were substantially affected by food source and could vary significantly from levels in the food source [25]. These studies demonstrated the feasibility of using nutrigenomics to assist in analyzing insect responses to nutritional changes and dietary quality with the intent of improving insect diets.

Currently, next generation high-throughput sequencing techniques (Solexa/Illumina, Roche 454) provide a unique opportunity for genomic exploration in insect species where little or no molecular knowledge is available [26]. This technology greatly increases the quantity of data that can be generated in a short time at a reduced cost [27]. For example, Illumina sequencing has been applied in *Nilaparvata lugens* [28], *Apis mellifera* [24], *Bemisia tabaci* [29,30], *Sogatella furcifera* [31], *B. dorsalis* [32] and *Pogonus chalceus* [33] research. The use of 454-sequencing technology has enabled the application of functional genomics to a broad range of insect species including *Cimex lectularius* [34], *Melitta cinxia* [35], *Zygaena filipendulae* [36], *Chrysomela tremulae* [37], *Aphis glycines* [38], *Manduca sexta* [39,40], *Laodelphax striatellus* [41], *Stomoxys calcitrans* [42], *Dermaecentor variabilis* [43], *Erynnis properties* and *Papilio zelicaon* [44] and *Agilus planipennis* [45]. To date, however, transcriptome analysis has not yet been utilized to assist in artificial diet formulation for a beneficial predatory insect. In this study, we applied Illumina sequencing technology to obtain the transcriptome of *A. chinensis* and compared the differentially expressed genes between *A. chinensis* fed with Chinese oak silk moth pupae and those fed with an artificial diet. Our results will allow for a better understanding of the food-driven molecular processes in *A. chinensis* and the identification of potentially useful candidate genes in gene-assisted diet formulation.



**Figure 1. Gene Ontology (GO) categories of the unigenes.** Unigenes were annotated in three categories: biological processes (26 sub-categories, 12,428 sequences), cellular components (12 sub-categories, 7,977 sequences), and molecular functions (13 sub-categories, 4,625 sequences).

doi:10.1371/journal.pone.0060881.g001

## Results and Discussion

### Sequencing and Sequence Assembly

After removal of adaptor sequences, ambiguous reads and low-quality reads ( $Q_{20} < 20$ ), a total of 53,224,704 (SRA accession number SRR617645) and 52,244,538 (SRA accession number SRR618073) high-quality clean reads comprised of 4,790,223,360 nucleotides (4.79 Gb) and 4,702,008,420 nucleotides (4.70 Gb) from the Chinese oak silk moth pupae-fed (CY\_1) and artificial diet-fed (AD\_1) *A. chinensis* libraries were generated, respectively. All high-quality reads were assembled *de novo* into 112,029 (CY\_1) and 98,724 (AD\_1) contigs using the Trinity program [46], with an N50 of 318 bp (CY\_1) and 494 bp (AD\_1) (i.e. 50% of assembled bases were incorporated into contigs of 318 bp and 494 bp (or longer) for pupae-fed and diet-fed insects (Table S1).

A total of 112,029 contigs of CY\_1 were assembled, with a total length of 28,024,077 nt and mean length of 250 bp. Similarly, a total of 98,724 contigs of AD\_1 were assembled, with a total length of 30,459,937 nt and mean length of 309 bp (Table S1). Size distribution indicated that the lengths of the 4,578 and 6,773 contigs were more than 1000 bp for CY\_1 and AD\_1, respectively (Figure S1C and S1A). Using paired-end reads and gap-filling, these contigs were further assembled and clustered into unigenes. Finally, we obtained 54,083 and 54,169 unigenes of CY\_1 and AD\_1, including 8,636 and 10,315 distinct clusters and 45,447 and 43,854 distinct singletons with a mean length of 427 and 541 bp, respectively (Table S1). Among these unigenes, the lengths

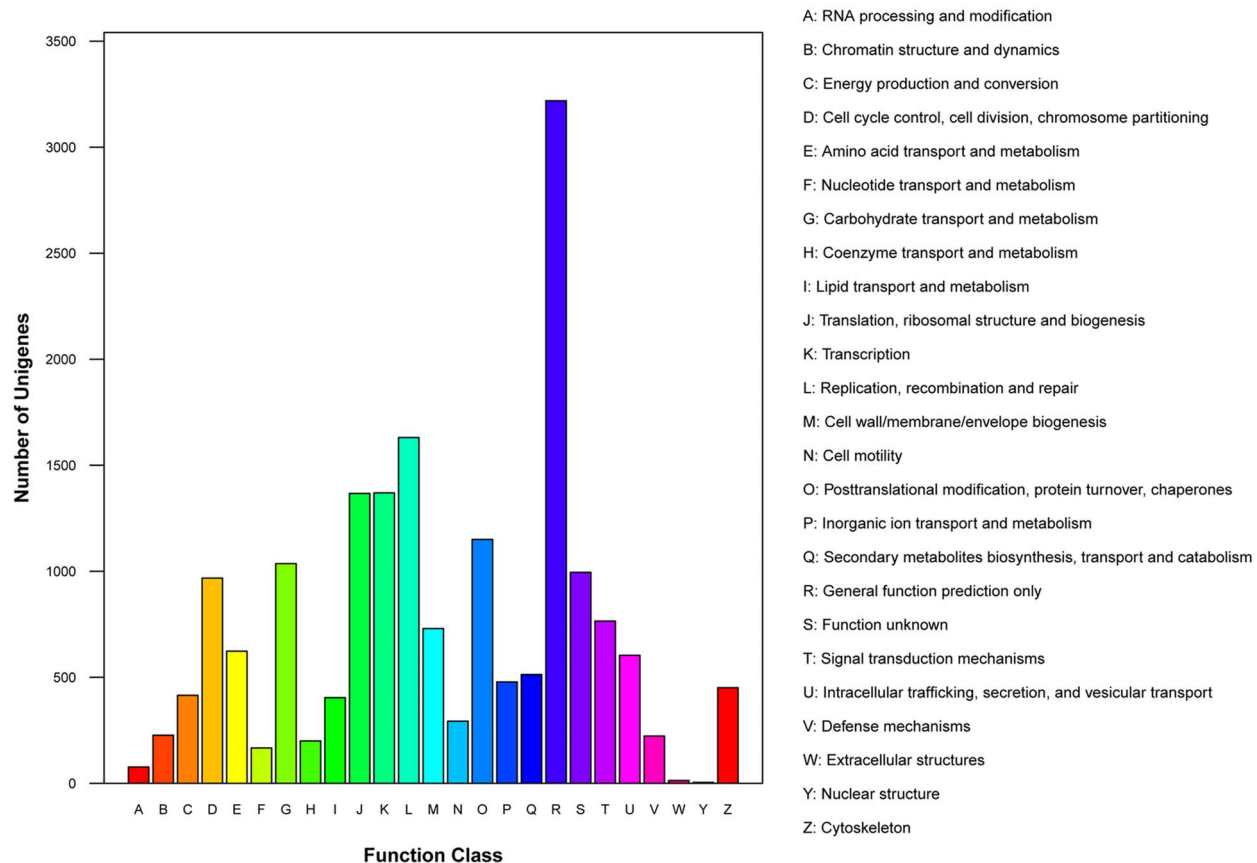
of the 5,334 CY\_1 unigenes and 8,681 AD\_1 unigenes were more than 1000 bp (Figure S1D and S1B). Unigenes from each sample assembly underwent sequence splicing and redundancy removal to acquire longer non-redundant All-unigenes, given these two samples were sequenced from the same species.

### Summary of Annotation Results

Functional annotation of unigenes includes information on proteins, COG, and Gene Ontology (GO). Unigene sequences were first aligned by BLASTx to protein databases nr, Swiss-Prot, KEGG and COG (E-value  $< 10^{-5}$ ), and then aligned by BLASTn to nucleotide databases nt (E-value  $< 10^{-5}$ ), retrieving proteins of highest sequence similarity with the given unigenes and their protein functional annotations.

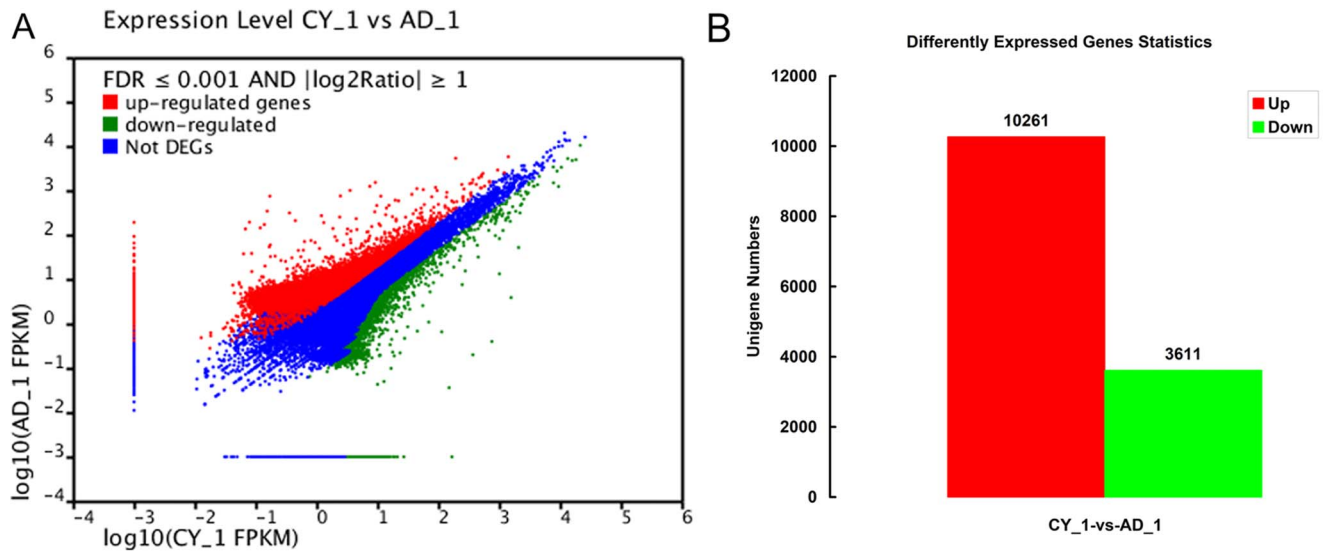
We obtained 55,189 unigenes of *A. chinensis* from both food treatments, including 12,046 distinct clusters and 43,143 distinct singletons with a mean length of 583 bp (Table S1). Size distribution indicated that the lengths of the 9,262 unigenes were more than 1000 bp (Figure S2A). Of those, 22,964 (41.61%), 7,898 (14.31%), 18,069 (32.74%), 15,416 (27.93%), 8,066 (14.62%) and 5,341 (9.68%) unigenes were annotated in nr, nt, Swiss-Prot, KEGG, COG and GO, respectively. A total of 24,187 unigenes were annotated in one or more of the databases (43.83% of all unigenes), suggesting they have relatively well conserved functions.

### COG Function Classification of All-Unigene.fa Sequence



**Figure 2. Clusters of Orthologous Groups (COG) functional classification.** From 22,964 nr hits, 8,066 sequences were classified into 25 COG categories (E-value  $< 1.0^{-5}$ ).

doi:10.1371/journal.pone.0060881.g002

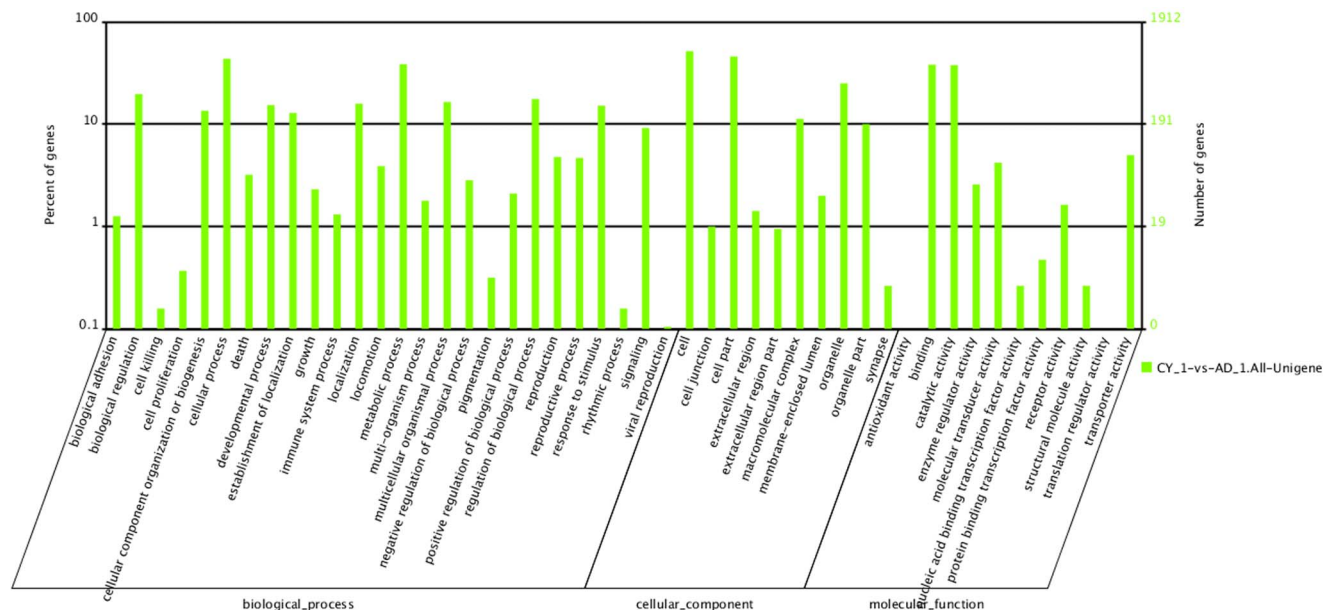


**Figure 3. Expression level and statistics of the differentially expressed genes (DEG) in artificial diet-fed (AD<sub>1</sub>) vs. prey-fed (CY<sub>1</sub>) treatments.** (A) Expression level of DEG. (B) Statistics of DGE. In total, the expression levels of 54,977 genes were affected by the artificial diet. Of which, 13,872 DEGs had significantly differential expression levels ( $FDR \leq 0.001$  and  $|\log_2 \text{Ratio}| \geq 1$ ). Among them, 10,261 and 3,611 genes were up-regulated and down-regulated, respectively in AD<sub>1</sub> vs. CY<sub>1</sub>. doi:10.1371/journal.pone.0060881.g003

#### Annotation of Predicted Proteins

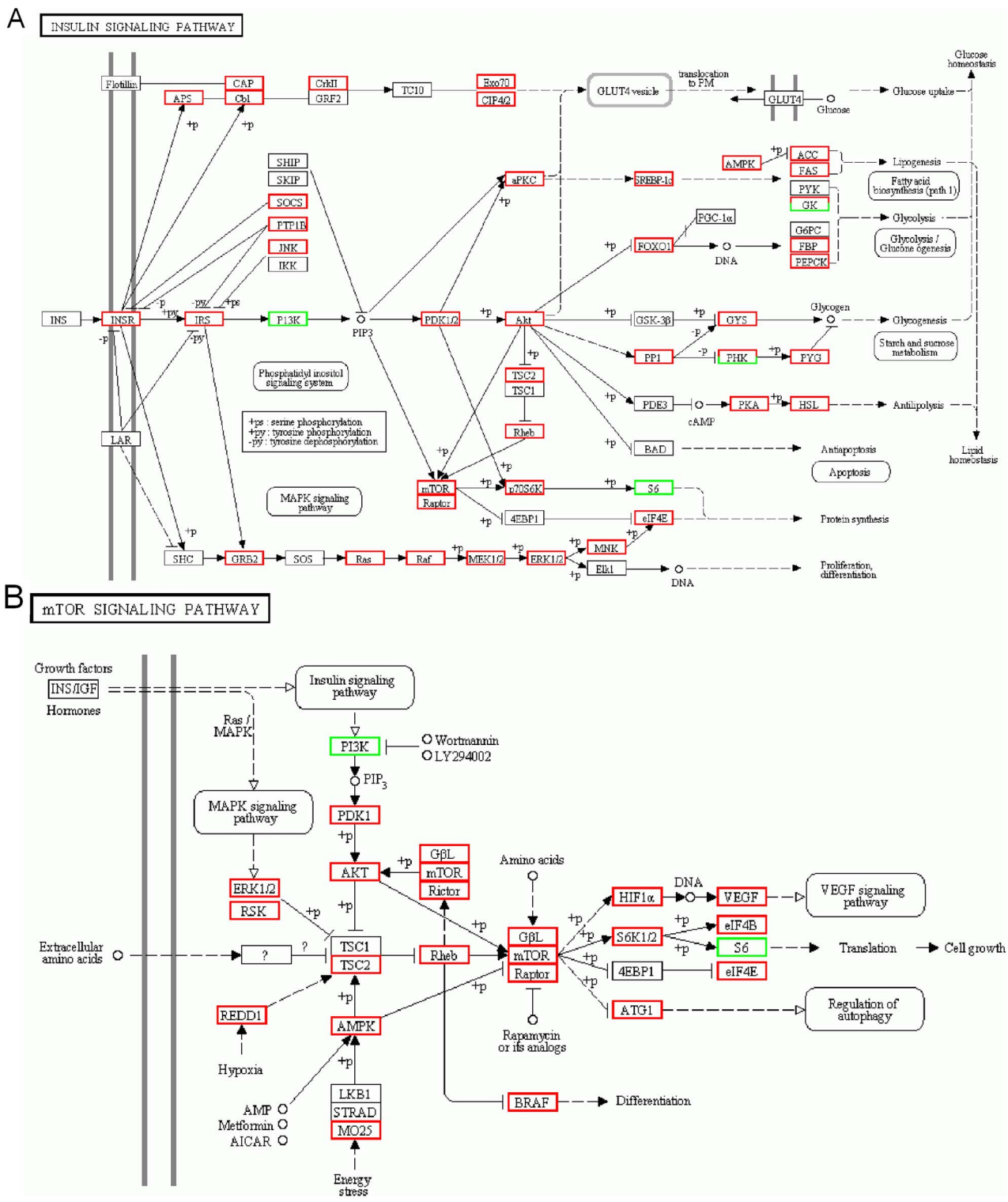
For functional annotation, distinct gene sequences were searched using BLASTx against nr NCBI nucleotide database with a cut-off E-value of  $10^{-5}$ . A total of 22,964 genes returned an above cut-off BLAST result, representing 41.61% of all distinct sequences. The E-value distribution of the top hits in the nr database showed that 39.2% of the mapped sequences had strong homology (smaller than  $1.0E^{-45}$ ), whereas 60.7% of the homolog sequences ranged between  $1.0E^{-5}$  to  $1.0E^{-45}$  (Figure S3A). Likewise, similarity distribution showed that 29.3% of the

sequences had a similarity higher than 60%, while 70.7% had a similarity ranging from 16% to 60% (Figure S3B). The highest percentage of *A. chinensis* sequences were matched to *Tribolium castaneum* (14.8%), followed by *Acyrtosiphon pisum* (11.5%), *Nasonia vitripennis* (6.2%), *Camponotus floridanus* (5.4%), *Acromyrmex echinator* (5.1%), *Harpegnathos saltator* (4.9%) and *Bombus impatiens* (4.4%) (Figure S3C).



**Figure 4. Gene Ontology (GO) categories of the differentially expressed genes (DEG) in artificial diet-fed (AD<sub>1</sub>) vs. prey-fed (CY<sub>1</sub>) treatments.** In total, 1,912 DEGs of *A. chinensis* were annotated in three categories: biological processes (26 sub-categories, 4,713 sequences), cellular components (10 sub-categories, 2,853 sequences) and molecular functions (11 sub-categories, 1,725 sequences). doi:10.1371/journal.pone.0060881.g004





**Figure 5. Insulin and mTOR signaling pathways affected by artificial diet feeding.** (A) Insulin signaling pathway. (B) mTOR signaling pathway. Red and green frames indicate genes and enriched functions that were up- and down-regulated in artificial diet-fed (AD<sub>1</sub>) and prey-fed (CY<sub>1</sub>) treatments. doi:10.1371/journal.pone.0060881.g005

**GO Assignments**

In total 5,341 transcripts of *A. chinensis* were assigned to GO terms based on BLAST matches with previously known sequences (Figure 1, Table S2). These transcripts were

associated with biological processes (26 sub-categories, 12,428 sequences), cellular components (12 sub-categories, 7,977 sequences) and molecular functions (13 sub-categories, 4,625 sequences). Among the molecular function assignments, a high

**Table 1.** Predicted single nucleotide polymorphisms (SNP) in Chinese oak silk moth pupae-fed (CY\_1) and artificial diet-fed (AD\_1) *Arma chinensis* sequences.

SNP type	CY_1	AD_1
<b>Number of transition</b>	37,661	50,116
A-G	18,635	24,945
C-T	19,026	25,171
<b>Number of transversion</b>	22,759	27,776
A-C	5,216	6,290
A-T	8,654	10,507
C-G	3,551	4,586
G-T	5,338	6,393
<b>Total</b>	60,420	77,892

doi:10.1371/journal.pone.0060881.t001

percentage of genes were associated with binding functions (1995, 43.14%), predominantly heat shock proteins (Hsp) and catalytic activity (1844, 39.87%). In a recent study of *C. lectularius*, the transcript levels for Hsp 70 and Hsp 90 were elevated when bugs were subjected to various stress factors (heat, cold and dehydration) suggesting that these proteins may play an important role during environmental stress and could potentially play a role in control strategies [47–50]. Cellular component sequences in the present study showed a significant percentage of genes assigned to cellular (2720, 34.10%) and cell component (2397, 30.05%) functions, whereas biological process

sequences were associated predominantly with cellular processes (2254, 18.14%) such as proteolysis, carbohydrate metabolic processes and oxidation reduction utilization, and metabolic processes (1906, 15.34%). Similar observations for metabolic processes have been reported in transcriptomic studies of other insects [28,29,32,34,40,45].

### COG Classification

To further evaluate the completeness of our transcriptome library and the effectiveness of our annotation process, we searched the annotated sequences for genes involved in COG classifications. From 22,964 nr hits, 8,066 sequences had a COG classification (Figure 2). Among the 25 COG categories, the cluster for ‘General function prediction only’ represented the largest group (3,219, 39.91%) followed by ‘Replication, recombination and repair’ (1,631, 20.22%), ‘Transcription’ (1,370, 16.98%), ‘Translation, ribosomal structure and biogenesis’ (1,367, 16.95%) and ‘Posttranslational modification, protein turnover, chaperones’ (1,150, 14.26%). Nuclear structure (4, 0.05%), extracellular structures (13, 0.16%) and RNA processing and modification (77, 0.95%) represented the smallest groups (Figure 2).

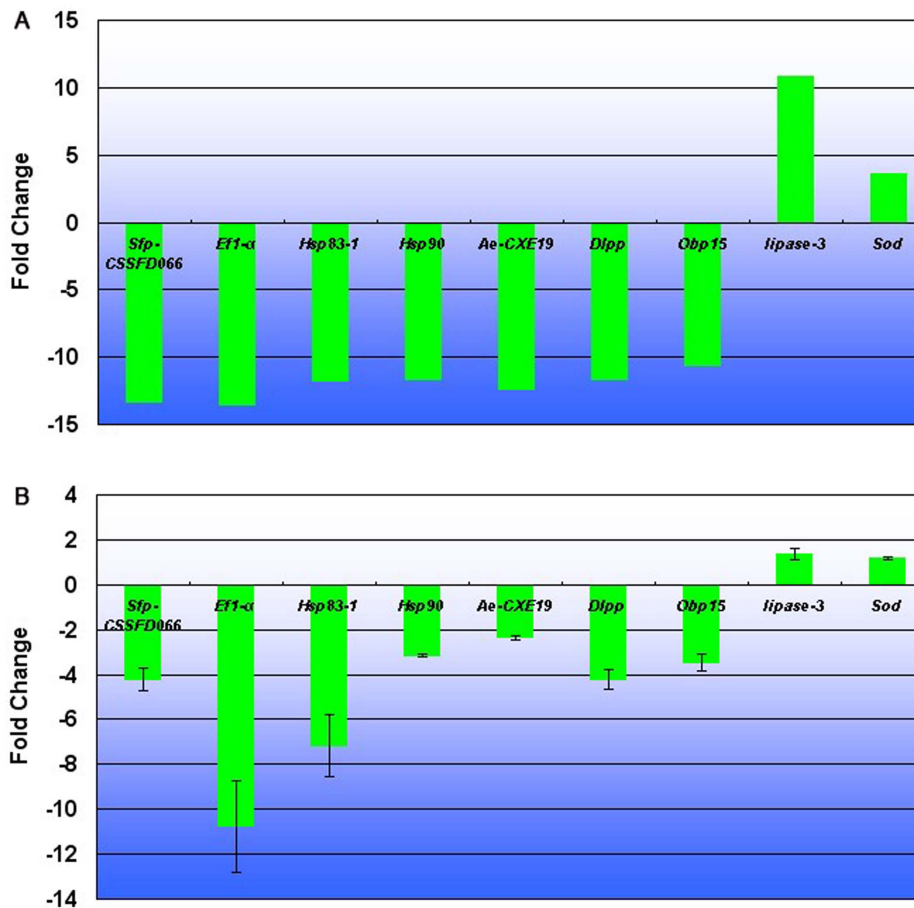
### KEGG Analysis

To identify active biological pathways in *A. chinensis*, we mapped the 22,964 annotated sequences to the reference canonical pathways in KEGG [51]. In total, we assigned 15,416 sequences to 240 KEGG pathways. The pathways most represented by the unique sequences were metabolic pathways (2,333, 15.13%), regulation of actin cytoskeleton (568, 3.68%), focal adhesion (543, 3.52%), pathways in cancer (540, 3.5%) and purine metabolism (473, 3.07%). The smallest groups were

**Table 2.** Summary of microsatellite loci predicted in *Arma chinensis* sequences.

Number of repeats	Mononucleotide repeats	Dinucleotide repeats	Trinucleotide repeats	Tetranucleotide repeats	Pentanucleotide repeats	Hexanucleotide repeats
4	0	0	0	0	28	13
5	0	0	403	13	4	1
6	0	338	142	5	0	0
7	0	146	34	1	2	0
8	0	83	9	0	0	0
9	0	31	1	0	0	0
10	0	39	1	0	0	0
11	0	53	0	0	0	0
12	897	6	0	0	0	0
13	554	0	0	0	0	0
14	353	0	1	0	0	0
15	230	0	0	0	0	0
16	149	0	0	0	0	0
17	101	0	0	0	0	0
18	102	0	0	0	0	0
19	109	0	0	0	0	0
20	114	0	0	0	0	0
21	85	1	0	0	0	0
22	41	0	0	0	0	0
23	24	0	1	0	0	0
Subtotal	2,759	697	592	19	34	14

doi:10.1371/journal.pone.0060881.t002



**Figure 6. Verification of differentially expressed genes (DEG) by qRT-PCR.** (A) DEG data in transcriptome analysis. The fold changes of the genes were calculated as the log<sub>2</sub> value of each AD<sub>1</sub>/CY<sub>1</sub> comparison and are shown on the y-axis. (B) The qRT-PCR analysis of gene expression data. Expression ratios of selected genes in AD-1 compared to CY<sub>1</sub>. doi:10.1371/journal.pone.0060881.g006

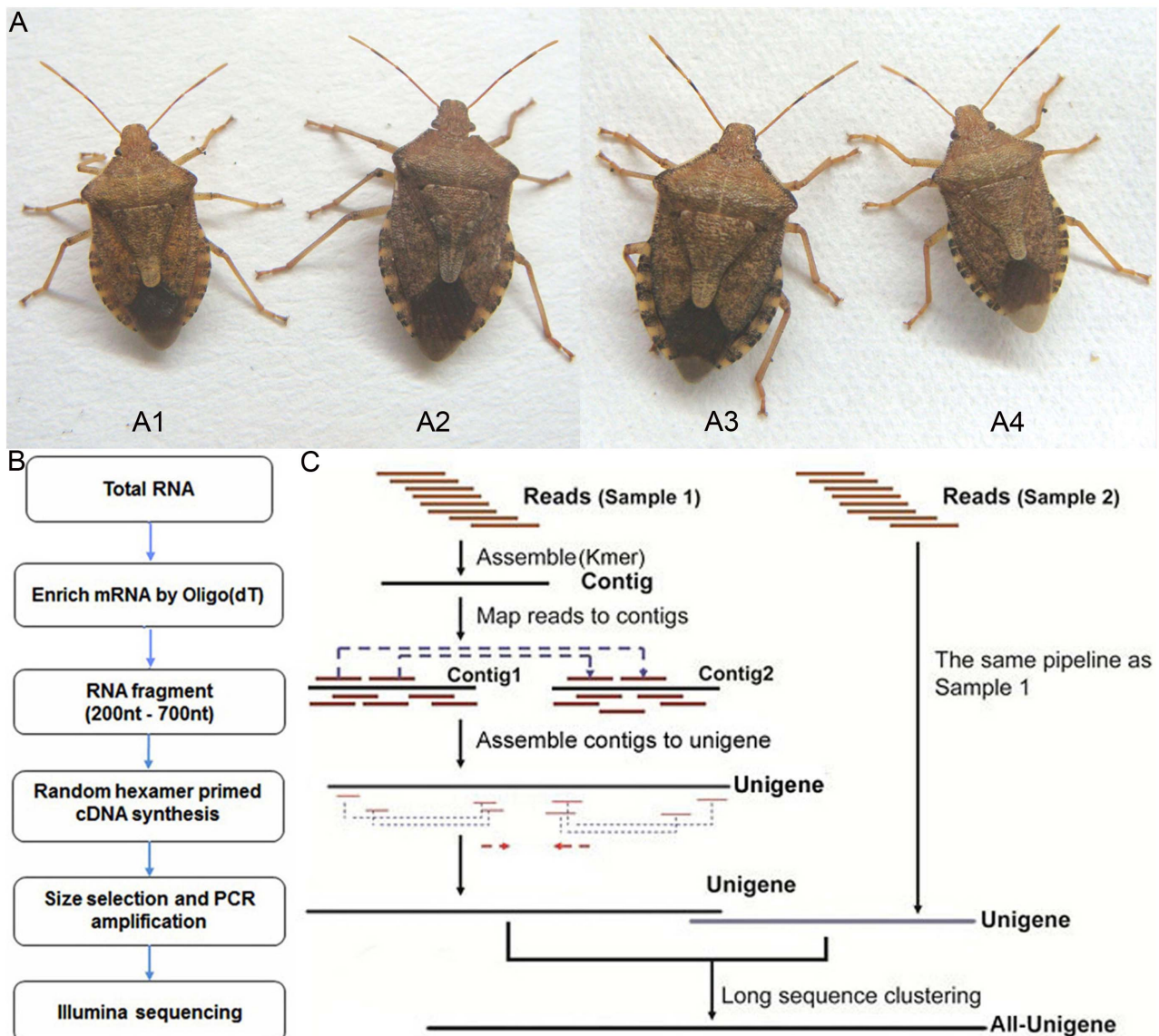
lysine biosynthesis (3, 0.02%), asthma (3, 0.02%), polyketide sugar unit biosynthesis (1, 0.01%) and intestinal immune network for Immunoglobulin A (IgA) production (1, 0.01%) (Table S3). These annotations provide a valuable resource for investigating specific processes, functions, and pathways during nutrigenomics research of *A. chinensis*.

### Protein Coding Region Prediction

Unigenes were first aligned by BLASTx (E-value  $<10^{-5}$ ) to protein databases in the priority order of nr, Swiss-Prot, KEGG, and COG. Unigenes that aligned to a higher priority database were not aligned to a lower priority database. Proteins with the highest BLAST ranks were selected to decide the coding region sequences of unigenes, which were then translated into amino acid sequences with the standard codon table. Both nucleotide (5′–3′) and amino acid sequences of the unigene coding region were acquired. We obtained a total of 23,124 significant BLAST hits (41.90% of all unigenes). The size distribution for the protein coding sequences (CDS) and predicted proteins are shown in Figure S2B and S2D. Unigenes not aligned to any database were scanned by ESTScan [52], producing nucleotide sequence (5′–3′) directions and amino acid sequences of the predicted coding region. A total of 4,292 unigenes were analyzed using ESTScan (size distributions of the ESTs and proteins are shown in Figure S2C and S2E).

### Differentially Expressed Genes (DEG) Statistics

The expression level of 54,977 genes were affected by the artificial diet, in which, 13,872 DEGs had significant differentially expressed levels ( $FDR \leq 0.001$  and  $|\log_2 \text{Ratio}| \geq 1$ ) between the two food treatments. Among those, 10,261 were up-regulated and 3,611 were down-regulated in artificial diet-fed vs. prey-fed *A. chinensis* (Figure 3). Of the 30 most differentially up-regulated genes, 12 had defined functions, i.e., six energy metabolic related genes (*ATP synthase F0 subunit 6*, *NADH dehydrogenase subunit 1*, 2, 3, 4, and 6), two mitochondrial genes (*cytochrome c oxidase subunit I* and *cytochrome oxidase subunit II*), a structural cuticular protein (*apidermin 1 precursor*), a transmembrane receptor (*tollo*), *kinesin heavy chain* and *alpha 1 S haptoglobin*. Of the 30 most differentially down-regulated genes, 13 had defined functions, i.e., one energy metabolic related gene (*ATP synthase F0 subunit 6*), two mitochondrial genes (*cytochrome b* and *cytochrome oxidase subunit II*), three ribosome genes (*ribosomal protein l3*, *40S ribosomal protein S27*, and *60S ribosomal protein L6*), one translation initiation factor 2 (*PvLEA1 protein*), one globular protein (*beta tulin*), one RNA transport gene (*elongation factor 1-alpha*, Unigene41798\_All), one coactivator (*MBF2*), two serine protease inhibitors (*pacifastin light chain precursor* and *serine protease inhibitor 2θ*), and one seminal fluid protein (*seminal fluid protein CSSFP066*). A total of 35 genes among the 60 differentially expressed genes had unknown functions or no annotations (Table S4). These top 30 most differentially up- and down-regulated genes demonstrated



**Figure 7. *Arma chinensis* fed with different diets and schematics of the transcriptome sequencing analysis.** (A) Female adults fed with artificial diet (A1) and Chinese oak silk moth pupae (A2); male adults fed with artificial diet (A4) and Chinese oak silk moth pupae (A3). (B) Experiment pipeline of transcriptome. (C) Assembly process of the data. doi:10.1371/journal.pone.0060881.g007

that the nutritional differences represented by Chinese oak silk moth pupae and artificial diet caused changes in a broad range of genes, yet both supported complete development of *A. chinensis*.

### GO Assignments of DEG

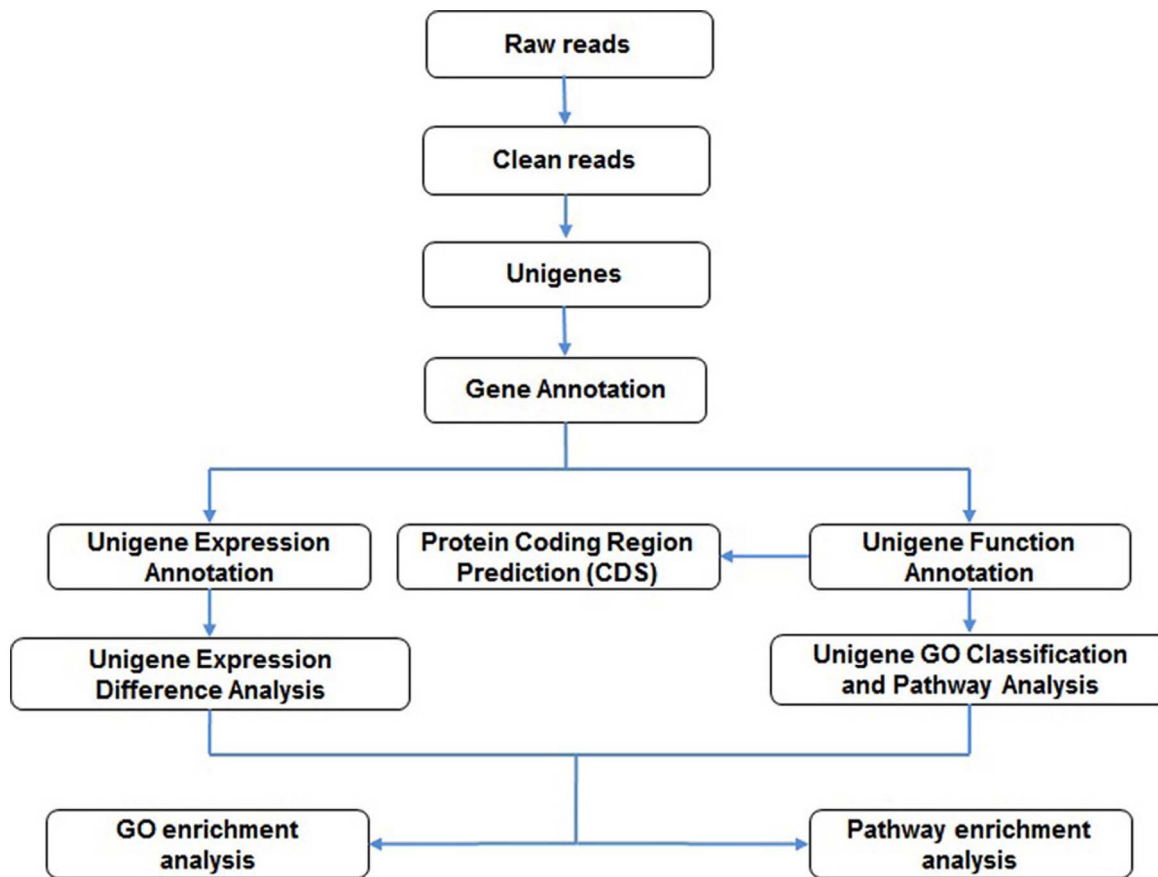
In total, 1,912 DEGs of *A. chinensis* were assigned to GO terms based on BLAST matches with sequences of known function (Figure 4, Table S5). These transcripts were assigned to biological processes (26 sub-categories, 4,713 sequences), cellular components (10 sub-categories, 2,853 sequences) and molecular functions (11 sub-categories, 1,725 sequences). Among the biological process terms, a high percentage of genes were assigned to cellular (832, 17.65%) and metabolic (735, 15.60%) processes. The cellular component terms showed a significant percentage of genes assigned to cells (987, 34.60%) and cell parts (875, 30.67%), whereas molecular function assignments were predominantly

associated with binding (728, 42.20%) and catalytic activity (721, 41.80%).

### KEGG of DEG

To identify the biological pathways associated with genes differentially expressed between the food treatments, we mapped the 13,872 DEGs ( $FDR \leq 0.001$  and  $|\log_2 \text{Ratio}| \geq 1$ ) and annotated sequences to the reference canonical pathways in KEGG [51]. In total, we assigned 5,879 sequences to 239 KEGG pathways. The pathways most represented by the DEGs were metabolic pathways (891, 15.16%), pathways in cancer (215, 3.66%), regulation of actin cytoskeleton (201, 3.42%), focal adhesion (178, 3.03%), influenza A (166, 2.82%), RNA transport (157, 2.67%), endocytosis (156, 2.65%), protein processing in endoplasmic reticulum (150, 2.55%), ubiquitin mediated proteolysis (150, 2.55%) and neuroactive ligand-receptor interaction (144, 2.45%). The smallest groups were intestinal immune network for





**Figure 8. Pipeline of bioinformatics analysis.**

doi:10.1371/journal.pone.0060881.g008

IgA production (1, 0.02%), D-glutamine and D-glutamate metabolism (1, 0.02%), lysine biosynthesis (1, 0.02%), asthma (1, 0.02%), lipoic acid metabolism (2, 0.03%), D-arginine and D-ornithine metabolism (3, 0.05%), primary immunodeficiency (3, 0.05%), phenylalanine, tyrosine and tryptophan biosynthesis (3, 0.05%), thiamine metabolism (3, 0.05%) and biotin metabolism (4, 0.07%). These annotations provide a valuable resource for investigating specific processes, functions and pathways during nutrigenomics research of *A. chinensis*.

#### DEG Related to Different Biological Characteristics

Several biological characteristics differed between pupae-fed and diet-fed *A. chinensis*, such as reduced fecundity, lower egg viability, prolonged nymphal development time, longer lifespan and higher cannibalism. We found several DEGs ( $FDR \leq 0.001$  and  $|\log_2\text{Ratio}| \geq 2$ ) related to these biological characteristics (Table S6). For reduced fecundity, three differentially expressed genes between the food treatments, specifically two *heat shock proteins 83-1* and *90*, associated with progesterone-mediated oocyte maturation, were markedly down-regulated in diet-fed insects.

In insects, seminal fluid proteins (Sfps) produced in the male accessory glands significantly increase male fitness by promoting sperm storage, temporarily increasing female egg-laying rate, and decreasing female sexual receptivity [53], thus increasing progeny production and delaying sperm displacement/competition [54]. Post-copulatory sexual selection can select for sperm allocation strategies in males [55] but males can also strategically allocate non-sperm components of the ejaculate [56,57] such as Sfps.

Thus, Sfps can influence the extent of post-copulatory sexual selection [58–60]. Using *Drosophila melanogaster*, Wigby *et al.* demonstrated that Sfps were strategically allocated to females in response to the potential level of sperm competition [55]. Males who were able to produce and transfer larger quantities of specific Sfps had a significant competitive advantage. Large male accessory glands also significantly increased competitive reproductive success. Quantitative variation in specific Sfps may play an important role in post-copulatory sexual selection and investment in Sfp production is essential for male fitness in a competitive environment [55]. Thus, the down-regulated *seminal fluid protein CSSFP066* in the artificial diet-fed *A. chinensis* could also contribute to reduced fecundity, lower egg viability, and a sex ratio shift in favor of males.

Among the DEGs for insect hormone biosynthesis, the gene *cytochrome P450 302a1* (ko: ecdysteroid 22-hydroxylase) (molting hormone) was up-regulated, as were juvenile hormone genes, *esterase FE4-like* (ko: juvenile-hormone esterase), *esterase FE4-like isoform 1* (ko: juvenile-hormone esterase), *beta-esterase 2 precursor* (ko: juvenile-hormone esterase), *venom carboxylesterase-6-like* (ko: juvenile-hormone esterase) and *pheromone-degrading enzyme 2* (ko: juvenile-hormone esterase). These up-regulated juvenile hormone genes may have contributed to the observed prolonged nymphal development time. Since the nutrition of the artificial diet increased longevity in *A. chinensis*, we looked at *Drosophila* data for insight on the expression pattern of genes found to affect longevity, as discovered by functional analysis from DAVID 6.7 bioinformatic resources [61] and GoToolbox [62]. Collectively, 20

genes known to effect longevity [24] were found, of which three were up-regulated in diet-fed *A. chinensis*, *superoxide dismutase [Cu-Zn]-like precursor* (Swiss-Prot: superoxide dismutase [Cu-Zn]) and *Cu, Zn-superoxide dismutase* encoding a cytoplasmic Cu-Zn superoxide dismutase, which increase lifespan [63,64]. *Atalpha (sodium pump alpha subunit)* with catalytic function was also up-regulated in diet-fed insects. Additionally, four DEGs, *antennal esterase CXE19*, *sensory appendage protein 1*, *defensin-like protein precursor* and *odorant binding protein 15*, related to higher cannibalism, were all down-regulated in the diet-fed insects, suggesting sensory ability to odors was possibly linked to the observed reduction in defense odors.

### KEGG of DEG Related to Artificial Diets

Honeybees require ten essential amino acids in their diet for their adult development: arginine, histidine, lysine, tryptophan, phenylalanine, methionine, threonine, leucine, isoleucine and valine [65]. Less is known about essential amino acids in *A. chinensis*. However, there was a trend in the diet-fed *A. chinensis* of up-regulation of sequences associated with the metabolism of all ten of the amino acids essential to the honey bee. Of those, DEGs enriched at a significant level ( $Q < 0.05$ ) only in one pathway (alanine, aspartate and glutamate metabolism) (Table S7).

Most DEGs enriched in the seven pathways related to fat metabolism were up-regulated in diet-fed insects, including adipocytokine signaling pathway, pyruvate metabolism, fatty acid biosynthesis, glycerolipid metabolism, fat digestion and absorption, fatty acid metabolism and fatty acid elongation (Table S7). This indicated that the artificial diet may contain excess lipids, suggesting that dietary reductions in tuna, chicken egg and pig liver were required. Most DEGs enriched in the four pathways related to starch and sugar metabolism were up-regulated in diet-fed insects, including carbohydrate digestion and absorption, and fructose and mannose metabolism (Table S7). Because sucrose was the main source of sugar, sucrose reductions were likely required in the diet. Most DEGs enriched in the ten metabolic pathways related to vitamins were up-regulated in diet-fed insects, including ascorbate and aldarate metabolism, vitamin digestion and absorption, folate biosynthesis, pantothenate and coenzyme A (CoA) biosynthesis, nicotinate and nicotinamide metabolism, biotin metabolism, retinol metabolism, thiamine metabolism, vitamin B6 metabolism and riboflavin metabolism (Table S7). Because vitamins, especially vitamin B, were added to the diet separately, the concentration of each vitamin in the diet could be reduced independently.

Zinke *et al.* categorized differentially expressed nutrient-controlled genes in *Drosophila* larvae into groups reflecting distinct physiological pathways mediating sugar metabolism such as lipase 3, glucose transporter, insulin receptor, as well as fatty acid synthase, fat metabolism such as acetyl CoA carboxylase, acyl CoA thioesterhydrolase, ATP-citrate lyase, *Zwischenferment*, glucose-6-phosphate dehydrogenase, and triacylglycerol lipases [66]. For diet-fed *A. chinensis*, we also found some differentially expressed nutrient-controlled genes, including sugar-lipase-3, glucose transporter, insulin receptor, fatty acid synthase and acetyl CoA carboxylase, which were all up-regulated in the diet-fed insects (Table S8). These differentially expressed nutrient-controlled genes again demonstrated that fat and sugar should be reduced in the artificial diet.

Insulin signaling plays a very important role in the regulation of glucose and lipid metabolism [67], and the insulin/TOR pathway is a conserved signaling cascade that functions as a nutrient sensing pathway by linking food-intake to animal growth and metabolism, including reproduction and lifespan [68,69]. In honey bees, this pathway plays a major role in the regulation of aging of individuals

[70]. Most genes in the insulin and mTOR signaling pathway were up-regulated in diet-fed insects (Figure 5), again indicating excess amounts of lipids and sugar in the diet. However, the PI 3-kinases (or PI3Ks) and the ribosomal protein S6 were down-regulated in diet-fed insects. Many PI3Ks have been linked to a diverse group of cellular functions, including cell growth, proliferation, differentiation, motility, survival and intracellular trafficking. Further, PI3Ks are a key component of the insulin signaling pathway [71], which regulates glucose uptake through a series of phosphorylation events.

The target-of-rapamycin (TOR) has been shown to respond to the presence of amino acids and induce up-regulation of ribosome biogenesis, translation [72–76] and energy metabolism [77,78] required for tissue growth. Representing an extensively studied effector of the TOR complex 1 (TORC1), S6K possesses an important yet incompletely defined role in cellular and organismal physiology. TORC1 functions as an environmental sensor by integrating signals derived from diverse environmental cues to promote anabolic and inhibit catabolic cellular functions. Mammalian TORC1 (mTORC1) phosphorylates and activates S6K1 and S6K2, whose first identified substrate was ribosomal protein S6. The mTORC1-S6K1 axis controls fundamental cellular processes, including transcription, translation, protein and lipid synthesis, cell growth/size and metabolism, glucose homeostasis, insulin sensitivity, adipocyte metabolism, body mass and energy balance, tissue and organ size [79]. For diet-fed *A. chinensis*, S6 in the mTOR signaling pathway was down-regulated (Figure 5), which may have caused delayed translation and cell growth, and probably contributed to the longer developmental time for nymphs.

In *D. melanogaster*, neurosecretory insulin-like peptide-producing cells (IPCs), analogous to mammalian pancreatic beta cells, are involved in glucose homeostasis. Extending those findings, Haselton *et al.* developed an oral glucose tolerance test in the adult fly and demonstrated that IPCs were responsible for executing an acute glucose clearance response. To further develop *D. melanogaster* as a relevant system for studying age-associated metabolic disorders, Haselton *et al.* determined the impact of adult-specific partial ablation of IPCs (IPC knockdown) on insulin-like peptide (ILP) action, metabolic outcomes and longevity. The findings showed a significant increase in stored glycogen and triglyceride levels as well as elevated levels of circulating lipids, increased resistance to starvation, impaired female fecundity, increased life span and decreased mortality; all of which demonstrated it was possible to modulate ILP action in adult flies to achieve life span extension without insulin resistance [80]. Based on this, we speculate that the lower fecundity and longer lifespan in diet-fed *A. chinensis* adults were probably linked to the high sucrose level in the artificial diet.

### Putative Molecular Markers

We predicted a total of 60,420 putative single nucleotide polymorphisms (SNPs) from CY\_1 *A. chinensis* libraries wherein 22,759 were transversions and 37,661 were transitions; a total of 77,892 SNPs from AD\_1 *A. chinensis* libraries wherein 27,776 were transversions and 50,116 were transitions (Table 1, Table S9, Table S10, Table S11, Table S12). In total, 4,115 simple sequence repeats (SSRs or microsatellites) including 697 (16.94%) dinucleotide, 592 (14.39%) trinucleotide repeats, and 34 (0.83%) pentanucleotide repeats were identified from All-unigenes of *A. chinensis*. Additionally, 3,814 All-unigenes contained SSRs, in which 269 (7.05%) had more than 1 SSR (Table 2, Table S13). Molecular markers identified in the current study could lay a platform for better understanding the adaptation/ecology of *A.*

*chinensis*. However all the predicted molecular markers need to be validated to rule out false positives and sequencing errors.

### Quantitative Real-time PCR (qRT-PCR) Validation

To validate the differentially expressed genes (DEG) determined by the transcriptome results, we compared expression profiles of the CY\_1 and AD\_1 using qRT-PCR. We selected nine genes randomly, all of which demonstrated a concordant direction of change for both DEG and qRT-PCR. But both methods indicated only two genes, *elongation factor 1-alpha (Efl- $\alpha$ )*, *heat shock protein 83-1 (Hsp83-1)*, were down-regulated in AD\_1 by ca. 10 times the amount found in CY\_1. Other seven genes, *seminal fluid protein CSSFP066 (Sfp-CSSFD066)*, *heat shock protein 90 (Hsp90)*, *antennal esterase CXE19 (Ae-CXE19)*, *defensin-like protein precursor (Dlpp)*, *odorant binding protein 15 (Obp15)*, *lipase-3*, and *Cu, Zn-superoxide dismutase (Sod)*, were up- or down-regulated in AD\_1 with lower values for the qRT-PCR method compared to transcriptome sequencing (Figure 6, Table S14). This difference might be caused by a lower sensitivity of qRT-PCR than transcriptome sequencing, and read coverage may be uneven across the transcript length, owing to sequencing biases. Nevertheless, qRT-PCR analysis confirmed the direction of change detected by transcriptome analysis, indicating that our results are reliable.

### Conclusions

In the present study, we sequenced and characterized the transcriptome in artificial diet-fed and Chinese oak silk moth pupae-fed *A. chinensis*. Among 13,872 DEGs with significantly differential expression levels ( $FDR \leq 0.001$  and  $|\log_2 \text{Ratio}| \geq 1$ ), more genes (10,261) were up-regulated in diet-fed insects. Additionally, many metabolic pathways related to nutrition were up-regulated in diet-fed *A. chinensis*, in some cases indicating excess quantities of specific nutrients in the diet. With this study, we showed that a nutrigenomic approach holds promise for deciphering the impact of dietary changes in insects and for improving diet formulations. We showed that changes in gene expression caused by dietary changes were correlated to physiological differences observed in diet-fed and pupae-fed *A. chinensis*. Several differentially expressed genes related to these different physiological properties were found, such as heat shock protein 90 (reduced fecundity), seminal fluid protein (lower egg viability), juvenile hormone esterase (prolonged nymphal development time), Cu, Zn-superoxide dismutase (longer lifespan in adults), antennal esterase CXE19, and odorant binding protein 15 (higher cannibalism in adults). It is worth noting that transcriptome analyses enabled determination of effects on male performance (e.g., expression of seminal fluid proteins), which is not easily attained via life history analyses. Also of importance, we found some metabolic pathways related to nutrition and differentially expressed nutrient-controlled genes, from which a more informative feedback for diet formulation was obtained and the artificial diet could be more efficiently optimized. In addition, a number of SNPs and microsatellite markers were predicted, which upon validation could facilitate the identification of polymorphisms within *A. chinensis* populations.

## Materials and Methods

### Insects

The *A. chinensis* colonies were originally obtained from the Institute of Forest Protection, Jilin Provincial Academy of Forestry Sciences, Changchun, China. The insects used in this experiment (Figure 7A) were reared at  $27 \pm 1^\circ\text{C}$  and RH of  $75 \pm 5\%$ , and a 16:8 (L:D) h photoperiod. The secondary prey, Chinese oak silk

moth *A. pernyi* pupae, was purchased from a supermarket in Beijing and the artificial diet was comprised of pig liver, chicken egg and tuna but devoid of insect components as previously described. First to 5th instar nymphs were distinguished by appearance and body size. Adults used in this study were about 15 to 20 days old and were fertile, as verified by hatch of their eggs. A total of 200 eggs, and approximately 150 nymphs and adults were collected for RNA extraction. Total RNA was extracted from freshly sacrificed insects.

### cDNA Library Preparation and Illumina Sequencing for Transcriptome Analysis

Total RNA was extracted using RNAiso Plus (TaKaRa, Dalian, China) according to the manufacturer's protocols. Gene expression information was obtained from RNA samples from eggs, 1st instar nymphs, 2nd instar nymphs, 3rd instar nymphs, 4th instar nymphs, 5th instar nymphs, and male and female adults obtained from the diet-fed and prey-fed treatments of the F<sub>6</sub> and F<sub>7</sub> generations, respectively.

The RNA sequencing was performed by the Beijing Genome Institute (Shenzhen, China), using a 2100 Bioanalyzer (Agilent Technologies) and following the Illumina manufacturer's instructions. The poly(A)<sup>+</sup> RNA was purified from 10  $\mu\text{g}$  of pooled total RNA using oligo(dT) magnetic beads and fragmented into short sequences in the presence of divalent cations at  $94^\circ\text{C}$  for 5 min. The cleaved poly(A)<sup>+</sup> RNA was transcribed, and then second-strand cDNA synthesis was performed. After the end-repair and ligation of adaptors, the products were amplified by PCR and purified using the QIA quick PCR purification kit to create a cDNA library, which was sequenced on the Illumina sequencing platform (HiSeq™ 2000). The sequencing machine generated raw images, which were controlled by the HCS system. The raw images were transformed into a BCL file by RTA, and BclConvert software was then used to generate raw reads from BCL (Figure 7B).

After removal of low quality reads, processed reads with an identity value of 95% and a coverage length of 90 bp were assembled using Trinity [46] *de novo* software and clustered using TGI clustering tools [81]. Trinity connected the contigs, forming unigenes sequences that could not be extended on either end. Because these two treatments were conducted with the same species, unigenes from each treatment assembly were processed further with sequence clustering software to acquire longer non-redundant All-unigenes (Figure 7C).

For further analysis, we used BLASTx alignment (E-value  $< 10^{-5}$ ) to search the unigene sequences against various protein databases such as Nr, Swiss-Prot, KEGG and COG. The BLAST results were used to extract CDS from the unigene sequences, and translate them into peptide sequences. The CDS of unigenes with no BLAST hits were predicted by ESTScan [52] and then translated into peptide sequences. For Nr annotation, we used the BLAST2GO program [82] and WEGO software [83] to explore the macro-distribution of gene functions for this species. A schematic of the pipeline used for the bioinformatics analysis is shown in Figure 8.

The FPKM method (fragments per kb per million fragments) [84] was used to calculate unigene expression. False discovery rate (FDR) was used to determine the threshold P-value in multiple tests. A FDR, 0.001 and an absolute value of the  $\log_2$  ratio  $> 1$  were used as the threshold to determine significant differences in gene expression. The differentially expressed genes were used for GO and KO enrichment analyses. Enriched P values were calculated according to the hypergeometric test:

$$P = 1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i} \binom{N-M}{n-i}}{\binom{N}{n}}$$

where  $N$  represents the number of genes with GO/KO annotation,  $n$  represents the number of differentially expressed genes in  $N$ ,  $M$  represents the number of genes in each GO/KO term, and  $m$  represents the number of differentially expressed genes in each GO/KO term. For GO enrichment analysis, all  $P$ -values were performed with Bonferroni's correction. We selected a corrected  $P$ -value  $<0.05$  as the threshold to determine significant enrichment of the gene sets. For KO enrichment analysis, after multiple testing correction, we selected pathways with  $Q$  value  $\leq 0.05$ , which were significantly enriched in DEGs. We considered the  $Q$ -values to be the FDR analogues of the  $P$ -values. The  $Q$ -value of an individual hypothesis test was the minimum FDR at which the test may be called significant.

SSRs were identified with the Microsatellite identification tool (MISA) (<http://pgrc.ipk-gatersleben.de/misa/>). The script can identify both perfect and compound microsatellites, which are interrupted by a certain number of bases. SNPs were predicted using SOAPsnp software [85] with an arbitrary criterion of at least 2 reads supporting the consensus or variant.

### Quantitative Real-time PCR (qRT-PCR) Validation

Real-time PCR was performed on randomly selected genes expressed differentially (*heat shock protein 83-1*, *heat shock protein 90*, *seminal fluid protein CSSFP066*, *Cu*, *Zn-superoxide dismutase*, *antennal esterase CXE19*, *odorant binding protein 15*, *defensin-like protein precursor*, *Gl23315 (lipase 3)*, and *elongation factor 1-alpha*) with two biological replicates and three technical replications. Total RNA was extracted as described for the DEG library preparation and sequencing. The concentration of each RNA sample was adjusted to 1  $\mu\text{g}/\mu\text{l}$  with nuclease-free water, and ca. 6  $\mu\text{g}$  of total RNA was used as the template to synthesize first-strand cDNA in a 20  $\mu\text{l}$  reaction system using a Superscript III Reverse Transcriptase kit (Invitrogen) following the manufacturer's protocols. The sequences of the specific primer sets are listed in Table S15. The *bata actin* gene of *A. chinensis* was used as an internal gene. Quantitative real-time-PCR was performed using the SYBR(R) Green I Nucleic A kit (Invitrogen) according to the manufacturer's protocols. The cycling parameters were 95°C for 2 min followed by 40 cycles at 95°C for 10 s and 60°C for 30 s ending with a melting curve analysis (60°C to 95°C in increments of 0.5°C every 5 s) to check for nonspecific product amplification. Relative gene expression was analyzed by the  $2^{-\Delta\Delta\text{C}_T}$  method (Applied Biosystems 7500 Fast Real-Time PCR System).

### Supporting Information

**Figure S1 Overview of *Arma chinensis* transcriptome assembly (1).** (A) and (C) Size distribution of the contigs obtained from high-quality clean reads of AD\_1 (*A. chinensis* fed on artificial diet) and CY\_1 (*A. chinensis* fed on Chinese oak silk moth pupae), respectively. (B) and (D) Size distribution of the unigenes produced from further assembly of contigs from AD\_1 and CY\_1, respectively. (TIF)

**Figure S2 Overview of *Arma chinensis* transcriptome assembly (2).** (A) Size distribution of the All-unigenes produced from further assembly of AD\_1 and CY\_1 unigenes. (B) Size

distribution of the CDS produced by searching All-unigene sequences against various protein databases (Nr, Swiss-Prot, KEGG and COG, in order) using BLASTx (E-value  $<10^{-5}$ ). (C) and (E) Size distributions of the ESTs and proteins obtained from the ESTScan results. (D) Size distribution of the proteins predicted from the CDS sequences.

(TIF)

**Figure S3 Characteristics of homology search of Illumina sequences against the nr database.** (A) E-value distribution of BLAST hits for each unique sequence with a cut-off E-value of  $1.0\text{E}^{-5}$ . (B) Similarity distribution of the top BLAST hits for each sequence. (C) Species distribution of unigenes top BLASTx results against the nr protein database with a cutoff E-value of at least  $1.0\text{E}^{-5}$ .

(TIF)

**Table S1** Summary for the Chinese oak silk moth pupae-fed (CY\_1) and artificial diet-fed (AD\_1) *Arma chinensis* transcriptome. (DOC)

**Table S2** Gene Ontology of *Arma chinensis* sequences. (XLS)

**Table S3** KEGG summary of *Arma chinensis* sequences. (XLS)

**Table S4** Top thirty differentially expressed genes. (DOC)

**Table S5** GO of DEG in CY\_1-vs-AD\_1. (XLS)

**Table S6** Differentially expressed genes related to different biological characteristics. (DOC)

**Table S7** KEGG of DEG related to artificial diets. (DOC)

**Table S8** Differentially expressed nutrient-controlled genes (DENCg). (DOC)

**Table S9** Putative SNPs in CY\_1 *Arma chinensis* (1). (XLS)

**Table S10** Putative SNPs in CY\_1 *Arma chinensis* (2). (XLS)

**Table S11** Putative SNPs in AD\_1 *Arma chinensis* (1). (XLS)

**Table S12** Putative SNPs in AD\_1 *Arma chinensis* (2). (XLS)

**Table S13** Putative microsatellite loci in *Arma chinensis*. (XLS)

**Table S14** Verification of differentially expressed genes by qRT-PCR. (XLS)

**Table S15** Sequences of qRT-PCR primers. (XLS)

### Acknowledgments

We thank Xiaodan Lv and Jiang Li, affiliated to Beijing Genomics Institute (Shenzhen), Junmin Li, affiliated to Zhejiang Academy of Agricultural Sciences, and Xiaowei Wang, affiliated to Zhejiang University for assistance with some bioinformatics. We also extend our thanks to all reviewers for their useful comments. The mention of trade names or commercial products in this publication is solely for the purpose of



providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. The U.S. Department of Agriculture is an equal opportunity provider and employer.

## References

- Chai XM, He ZH, Jiang P, Wu ZD, Pan CR, et al. (2000) Studies on natural enemies of *Dendrolimus punctatus* in Zhejiang Province. *Journal of Zhejiang Forestry Science and Technology* 20: 1–56.
- Chen J, Zhang JP, Zhang JH, Tian YH, Xu ZC, et al. (2007) Study on functional response of *Arma chinensis* to the adults of *Monolepta hieroglyphica*. *Natural Enemies of Insects* 29: 149–154.
- Gao CQ, Wang ZM, Yu EY (1993) Studies on artificial rearing of *Arma chinensis* Fallou. *Journal of Jilin Forestry Science and Technology* 2: 16–18.
- Gao Z (2010) Studies on Biological Characteristic and Control Condition of *Arma chinensis* Fallou. Harbin: Heilongjiang University Press. p. 11.
- Liang ZP, Zhang XX, Song AD, Peng HY (2006) Biology of *Clostera anachoreta* and its control methods. *Chinese Bulletin of Entomology* 43: 147–152.
- Yan JH, Li HJ, Peng HY, Zhou XZ, Gao BD (2006) Bionomics of *Batrachomorphus dentatus* and its control. *Chinese Bulletin of Entomology* 43: 562–566.
- Yan JH, Tang WY, Zhang H, Wang HY (2006) Bionomics of the leafhopper *Macropsis matsudanis*. *Chinese Bulletin of Entomology* 43: 245–248.
- Zou DY, Wang MQ, Zhang LS, Zhang Y, Zhang XJ, et al. (2012) Taxonomic and bionomic notes on *Arma chinensis* (Fallou) (Hemiptera: Pentatomidae: Asopinae). *Zootaxa* 3382: 41–52.
- Grafius EJ (1995) Is local selection followed by dispersal a mechanism for rapid development of multiple insecticide resistance in the Colorado potato beetle? *American Entomologist* 41: 104–109.
- Ferro DN (1993) Potential for resistance to *Bacillus thuringiensis*: Colorado potato beetle (Coleoptera: Chrysomelidae), a model system. *American Entomologist* 39: 38–44.
- Whalon ME, Miller DL, Hollingworth RM, Grafius EJ, Miller JR (1993) Selection of a Colorado potato beetle (Coleoptera: Chrysomelidae) strain resistant to *Bacillus thuringiensis*. *Journal of Economic Entomology* 86: 226–233.
- Lu YH, Wu KM, Jiang YY, Xia B, Li P, et al. (2010) Mirid bug outbreaks in multiple crops correlated with wide-scale adoption of Bt cotton in China. *Science* 328: 1151–1154.
- Greany PD, Carpenter JE (1998) Culture Medium for Parasitic and Predaceous Insects. United States Patent. Patent number: 5799607.
- Glenister CS (1998) Predatory heteropterans in augmentative biological control: an industry perspective. In: Coll M, Ruberson JR, editors. *Predatory Heteroptera: Their Ecology and Use in Biological Control*. Lanham, MD: Proceedings, Thomas Say Publications in Entomology, Entomological Society of America. 199–208.
- Glenister CS, Hoffmann MP (1998) Mass-reared natural enemies: scientific, technological, and informational needs and considerations. In: Ridgway R, Hoffmann MP, Inscoc MN, Glenister CS, editors. *Mass-Reared Natural Enemies: Application, Regulation, and Needs*. Lanham, MD: Proceedings, Thomas Say Publications in Entomology, Entomological Society of America. 242–247.
- Ruberson JR, Coll M (1998) Research needs for the predaceous Heteroptera. In: Coll M, Ruberson JR, editors. *Predatory Heteroptera: Their Ecology and Use in Biological Control*. Lanham, MD: Proceedings, Thomas Say Publications in Entomology, Entomological Society of America. 225–233.
- Thompson SN (1999) Nutrition and culture of entomophagous insects. *Annual Review of Entomology* 44: 561–592.
- Wittmeyer JL, Coudron TA (2001) Life table parameters, reproductive rate, intrinsic rate of increase and estimated cost of rearing *Podisus maculiventris* (Heteroptera: Pentatomidae) on an artificial diet. *Journal of Economic Entomology* 94: 1344–1352.
- Adams TS (2000) Effects of diet and mating status on ovarian development in a predaceous stinkbug *Perillus bioculatus* (Hemiptera: Pentatomidae). *Annals of the Entomological Society of America* 93: 529–535.
- Coudron TA, Wittmeyer J, Kim Y (2002) Life history and cost analysis for continuous rearing of *Podisus maculiventris* (Say) (Heteroptera: Pentatomidae) on a zoophytophagous artificial diet. *Journal of Economic Entomology* 95: 1159–1168.
- Coudron TA, Kim Y (2004) Life history and cost analysis for continuous rearing of *Perillus bioculatus* (Heteroptera: Pentatomidae) on a zoophytophagous artificial diet. *Journal of Economic Entomology* 97: 807–812.
- Yocum GD, Coudron TA, Brandt SL (2006) Differential gene expression in *Perillus bioculatus* nymphs fed a suboptimal artificial diet. *Journal of Insect Physiology* 52: 586–592.
- Coudron TA, Chang CL, Goodman CL, Stanley D (2011) Dietary wheat germ oil influences gene expression in larvae and eggs of the oriental fruit fly. *Archives of Insect Biochemistry and Physiology* 76: 67–82.
- Alaux C, Dantec C, Parrinello H, Conte YL (2011) Nutrigenomics in honey bees: digital gene expression analysis of pollen's nutritive effects on health and varroa-parasitized bees. *BMC Genomics* 12: 496–508.
- Coudron TA, Mitchell LC, Sun R, Robertson JD, Pham NV, et al. (2012) Dietary composition affects levels of trace elements in the predator *Podisus maculiventris* (Say) (Heteroptera: Pentatomidae). *Biological Control* 61: 141–146.
- Gibbons JG, Janson EM, Hittinger CT, Johnston M, Abbot P, et al. (2009) Benchmarking next-generation transcriptome sequencing for functional and evolutionary genomics. *Molecular Biology and Evolution* 26: 2731–2744.
- Ansoorge WJ (2009) Next-generation DNA sequencing techniques. *New Biotechnology* 25: 195–203.
- Xue J, Bao YY, Li BL, Cheng YB, Peng ZY, et al. (2010) Transcriptome analysis of the brown planthopper *Nilaparvata lugens*. *PLoS One* 5: e14233.
- Wang XW, Luan JB, Li JM, Bao YY, Zhang CX, et al. (2010) *De novo* characterization of a whitefly transcriptome and analysis of its gene expression during development. *BMC Genomics* 11: 400.
- Wang XW, Luan JB, Li JM, Su YL, Xia J, et al. (2011) Transcriptome analysis and comparison reveal divergence between two invasive whitefly cryptic species. *BMC Genomics* 12: 458.
- Xu Y, Zhou WW, Zhou YJ, Wu JX, Zhou XP (2012) Transcriptome and comparative gene expression analysis of *Sogatella furcifera* (Horvath) in response to southern rice black-streaked dwarf virus. *PLoS One* 7: e36238.
- Shen GM, Dou W, Niu JZ, Jiang HB, Yang WJ, et al. (2011) Transcriptome analysis of the oriental fruit fly (*Bactrocera dorsalis*). *PLoS One* 6: e29127.
- Van Bellegheem SM, Roelofs D, Van Houdt J, Hendrickx F (2012) *De novo* transcriptome assembly and SNP discovery in the wing polymorphic salt marsh beetle *Pogonys chalcus* (Coleoptera, Carabidae). *PLoS One* 7: e42605.
- Bai XD, Mamidala P, Rajarapu SP, Jones SC, Mittapalli O (2011) Transcriptomics of the bed bug (*Cimex lectularius*). *PLoS One* 6: e16336.
- Vera JC, Wheat CW, Fescemyer HW, Frilander MJ, Crawford DL, et al. (2008) Rapid transcriptome characterization for a non-model organism using 454 pyrosequencing. *Molecular Ecology* 17: 1636–1647.
- Zagrobely M, Scheibye-Alsing K, Jensen NB, Moller BL, Gorodkin J, et al. (2009) 454 pyrosequencing based transcriptome analysis of *Zygaena filipendulae* with focus on genes involved in biosynthesis of cyanogenic glucosides. *BMC Genomics* 10: 574.
- Pauchet Y, Wilkinson P, Van Munster M, Augustin S, Pauron D, et al. (2009) Pyrosequencing of the midgut transcriptome of the poplar leaf beetle *Chrysomela tremulae* reveals new gene families in Coleoptera. *Insect Biochemistry and Molecular Biology* 39: 403–413.
- Bai XD, Zhang W, Ornates L, Jun TH, Mittapalli O, et al. (2010) Combining next-generation sequencing strategies for rapid molecular resource development from an invasive aphid species, *Aphis glycines*. *PLoS One* 5: e11370.
- Zou Z, Najaf F, Wang Y, Roe B, Jiang HB (2008) Pyrosequencing analysis of expressed sequence tags for *Manduca sexta* hemolymph proteins involved in immune responses. *Insect Biochemistry and Molecular Biology* 38: 677–682.
- Pauchet Y, Wilkinson P, Vogel H, Nelson DR, Reynolds SE, et al. (2010) Pyrosequencing the *Manduca sexta* larval midgut transcriptome: messages for digestion, detoxification and defence. *Insect Molecular Biology* 19: 61–75.
- Zhang FJ, Guo HY, Zheng HJ, Zhou T, Zhou YJ, et al. (2010) Massively parallel pyrosequencing-based transcriptome analysis of small brown planthopper (*Laodelphax striatellus*), a vector insect transmitting rice stripe virus (RSV). *BMC Genomics* 11: 303.
- Olafson PU, Lohmeyer KH (2010) Analysis of expressed sequence tags from a significant livestock pest, the stable fly (*Stomoxys calcitrans*), identifies transcripts with a putative role in chemosensation and sex determination. *Archives of Insect Biochemistry and Physiology* 74: 179–204.
- Jaworski DC, Zou Z, Bowen CJ, Wasala NB, Madden R, et al. (2010) Pyrosequencing and characterization of immune response genes from the American dog tick, *Dermacentor variabilis* (L.). *Insect Molecular Biology* 19: 617–630.
- O'Neil ST, Dzurisin JDK, Carmichael RD, Lobo NF, Emrich SJ, et al. (2010) Population-level transcriptome sequencing of non model organisms *Erynnis propertius* and *Papilio zelicaon*. *BMC Genomics* 11: 310.
- Mittapalli O, Bai XD, Mamidala P, Rajarapu SP, Bonello P, et al. (2010) Tissue-specific transcriptomics of the exotic invasive insect pest emerald ash borer. *PLoS One* 5: e13708.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, et al. (2011) Fulllength transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29: 644–652.
- Reinhardt K, Siva-Jothy MT (2007) The biology of bedbugs (Cimicidae). *Annual Review of Entomology* 52: 351–374.
- Hwang SW, Svoboda TJ, De Jong IJ, Kabasele KJ, Gogosis E (2005) Bed bug infestations in an urban environment. *Emerging Infectious Diseases* 11: 533–538.
- Doggett SL, Geary MJ, Russell RC (2004) The resurgence of bed bugs in Australia, with notes on their ecology and control. *Environmental Health* 4: 30–38.

## Author Contributions

Conceived and designed the experiments: DZ TC HC. Performed the experiments: DZ CL. Analyzed the data: DZ TC. Contributed reagents/materials/analysis tools: CL LZ MW. Wrote the paper: DZ TC HC.

50. Benoit JB, Lopez-Martinez G, Teets NM, Phillips SA, Denlinger DL (2009) Responses of the bed bug, *Cimex lectularius*, to temperature extremes and dehydration: levels of tolerance, rapid cold hardening and expression of heat shock proteins. *Medical and Veterinary Entomology* 23: 418–425.
51. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M (2004) The KEGG resource for deciphering the genome. *Nucleic Acids Research* 32: D277–280.
52. Iseli C, Jongeneel CV, Bucher P (1999) ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *Proceedings of the International Conference on Intelligent Systems for Molecular Biology*: 138–148.
53. Gillott C (2003) Male accessory gland secretions: modulators of female reproductive physiology and behavior. *Annual Review of Entomology* 48: 163–184.
54. Fricke C, Wigby S, Hobbs R, Chapman T (2009) The benefits of male ejaculate sex peptide transfer in *Drosophila melanogaster*. *Journal of Evolutionary Biology* 22: 275–286.
55. Wigby S, Sirot LK, Linklater JR, Buchner N, Calboli FCF, et al. (2009) Sperm fluid protein allocation and male reproductive success. *Current Biology* 19: 751–757.
56. Cameron E, Day T, Rowe L (2007) Sperm competition and the evolution of ejaculate composition. *American Naturalist* 169: E158–E172.
57. Hodgson DJ, Hosken DJ (2006) Sperm competition promotes the exploitation of rival ejaculates. *Journal of Theoretical Biology* 243: 230–234.
58. Chapman T (2001) Seminal fluid-mediated fitness traits in *Drosophila*. *Heredity* 87: 511–521.
59. Poiani A (2006) Complexity of seminal fluid: a review. *Behavioral Ecology and Sociobiology* 60: 289–310.
60. Ravi Ram K, Wolfner MF (2007) Seminal influences: *Drosophila* Acps and the molecular interplay between males and females during reproduction. *Integrative and Comparative Biology* 47: 427–445.
61. Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nature Protocols* 4: 44–57.
62. Martín D, Brun C, Remy E, Mouren P, Thieffry D, et al. (2004) GOToolBox: function analysis of gene datasets based on Gene Ontology. *Genome Biology* 5: R101.
63. Sun J, Tower J (1999) FLP recombinase-mediated induction of Cu/Zn superoxide dismutase transgene expression can extend the life span of adult *Drosophila melanogaster* flies. *Molecular and Cellular Biology* 19: 216–228.
64. Orr WC, Mockett RJ, Sohal RS (1999) Overexpression of glutathione reductase extends survival in transgenic *Drosophila melanogaster* under hyperoxia but not normoxia. *Faseb Journal* 13: 1733–1742.
65. De Groot AP (1953) Protein and amino acid requirements of the honey bee (*Apis mellifica* L.). *Physiologia Comparata et Oecologia* 3: 197–285.
66. Zinke I, Schütz CS, Katzenberger JD, Bauer M, Pankratz MJ (2002) Nutrient control of gene expression in *Drosophila*: microarray analysis of starvation and sugar-dependent response. *The European Molecular Biology Organization Journal* 21: 6162–6173.
67. Saltiel AR, Kahn CR (2001) Insulin signalling and the regulation of glucose and lipid metabolism. *Nature* 414: 799–806.
68. Colombani J, Raisin S, Pantalacci S, Radimerski T, Montagne J, et al. (2003) A nutrient sensor mechanism controls *Drosophila* growth. *Cell* 114: 739–749.
69. Oldham S, Hafen E (2003) Insulin/IGF and target of rapamycin signaling: a TOR de force in growth control. *Trends in Cell Biology* 13: 79–85.
70. Münch D, Amdam GV (2010) The curious case of aging plasticity in honey bees. *FEBS Letters* 584: 2496–2503.
71. Wikipedia website. Available: [http://en.wikipedia.org/wiki/Phosphoinositide\\_3-kinase](http://en.wikipedia.org/wiki/Phosphoinositide_3-kinase). Accessed 2013 Feb 26.
72. Grewal SS (2009) Insulin/TOR signaling in growth and homeostasis: a view from the fly world. *International Journal of Biochemistry & Cell Biology* 41: 1006–1010.
73. Li L, Edgar BA, Grewal SS (2010) Nutritional control of gene expression in *Drosophila* larvae via TOR, Myc and a novel cis-regulatory element. *BMC Cell Biology* 11: 7.
74. Hay N, Sonenberg N (2004) Upstream and downstream of mTOR. *Genes & Development* 18: 1926–1945.
75. Grewal SS, Evans JR, Edgar BA (2007) *Drosophila* TIF-IA is required for ribosome synthesis and cell growth and is regulated by the TOR pathway. *Journal of Cell Biology* 179: 1105–1113.
76. Guertin DA, Guntur KV, Bell GW, Thoreen CC, Sabatini DM (2006) Functional genomics identifies TOR-regulated genes that control growth and division. *Current Biology* 16: 958–970.
77. Tiefenbock SK, Baltzer C, Egli NA, Frei C (2009) The *Drosophila* PGC-1 homologue Spargel coordinates mitochondrial activity to insulin signalling. *EMBO Journal* 29: 171–183.
78. Baltzer C, Tiefenbock SK, Marti M, Frei C (2009) Nutrition controls mitochondrial biogenesis in the *Drosophila* adipose tissue through Delg and cyclin D/Cdk4. *PLoS One* 4: e6935.
79. Magnuson B, Ekim B, Fingar DC (2012) Regulation and function of ribosomal protein S6 kinase (S6K) within mTOR signalling networks. *Biochemical Journal* 441: 1–21.
80. Haselton A, Sharmin E, Schrader J, Sah M, Poon P, et al. (2010) Partial ablation of adult *Drosophila* insulin-producing neurons modulates glucose homeostasis and extends life span without insulin resistance. *Cell Cycle* 9: 3063–3071.
81. Pertea G, Huang X, Liang F, Antonescu V, Sultana R, et al. (2003) TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics* 19: 651–652.
82. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, et al. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21: 3674–3676.
83. Ye J, Fang L, Zheng HK, Zhang Y, Chen J, et al. (2006) WEGO: a web tool for plotting GO annotations. *Nucleic Acids Research* 34 (Web Server issue): W293–297.
84. Audic S, Claverie JM (1997) The significance of digital gene expression profiles. *Genome Research* 7: 986–995.
85. Li RQ, Li YR, Fang XD, Yang HM, Wang J, et al. (2009) SNP detection for massively parallel whole-genome resequencing. *Genome Research* 19: 1124–1132.