De Novo Transcriptome Analysis of Wing Development-Related Signaling Pathways in Locusta migratoria Manilensis and Ostrinia furnacalis (Guenée)

Suning Liu*, Wei Wei*, Yuan Chu, Long Zhang, Jie Shen*, Chunju An*
Department of Entomology, College of Agriculture and Biotechnology, China Agricultural University, Beijing, China

Abstract
Background: Orthopteran migratory locust, Locusta migratoria, and lepidopteran Asian corn borer, Ostrinia furnacalis, are two types of insects undergoing incomplete and complete metamorphosis, respectively. Identification of candidate genes regulating wing development in these two insects would provide insights into the further study about the molecular mechanisms controlling metamorphosis development. We have sequenced the transcriptome of O. furnacalis previously. Here we sequenced and characterized the transcriptome of L. migratoria wing discs with special emphasis on wing development-related signaling pathways.

Methodology/Principal Findings: Illumina Hiseq2000 was used to sequence 8.38 Gb of the transcriptome from dissected nymphal wing discs. De novo assembly generated 91,907 unigenes with mean length of 610 nt. All unigenes were searched against five databases including Nt, Nr, Swiss-Prot, COG, and KEGG for annotations using blastn or blastx algorithm with an cut-off E-value of 10^-5. A total of 23,359 (25.4%) unigenes have homologs within at least one database. Based on sequence similarity to homologs known to regulate Drosophila melanogaster wing development, we identified 50 and 46 potential wing development-related unigenes from L. migratoria and O. furnacalis transcriptome, respectively. The identified unigenes encode putative orthologs for nearly all components of the Hedgehog (Hh), Decapentaplegic (Dpp), Notch (N), and Wingless (Wg) signaling pathways, which are essential for growth and pattern formation during wing development. We investigated the expression profiles of the component genes involved in these signaling pathways in forewings and hind wings of L. migratoria and O. furnacalis. The results revealed the tested genes had different expression patterns in two insects.

Conclusions/Significance: This study provides the comprehensive sequence resource of the wing development-related signaling pathways of L. migratoria. The obtained data gives an insight into better understanding the molecular mechanisms involved in the wing development in L. migratoria and O. furnacalis, two insect species with different metamorphosis types.

Introduction
Insects are the only group of invertebrates that have evolved flight [1]. Their wings serve not only as organs of flight, but also may be adapted variously as protective covers [2], thermal collectors [3], gyroscopic stabilizers [4], sound producers [5], or visual cues for species recognition and sexual contact [6]. For those insects with wings during the adult stage, complete wings are not always visible throughout the life cycle. The insects undergoing incomplete metamorphosis have represented functional wings during the stage of nymph [7], while the insects going through complete metamorphosis only have wing discs inside the body in the larval stage [7]. Therefore, the comparative studies on how the insect wings are developed will be helpful to understand the insect metamorphosis. Additionally, given the high evolutionary conservation of the proteins involved in the wing development, the understanding of the molecular mechanisms involved in wing development also sheds light on organogenesis, tissue homeostasis, human disease and so on [8–11].

Insect wing development is controlled with amazing precision and complication. Current understanding of insect wing development mechanisms is mainly from the fruit fly Drosophila melanogaster [12–14]. The adult wing of fruit fly is derived from the wing imaginal disc, formed at the end of embryonic
development [15]. In *Drosophila*, the wing imaginal disc is subdivided into anterior (A) and posterior (P) compartments at very early stage and then further subdivided into dorsal (D) and ventral (V) compartments at second instar stage [16–18]. Organizers located in the A/P and D/V boundaries coordinate the patterning of the wing disc by secreting signal molecules including the long-range morphogens Decapentaplegic (Dpp; the vertebrate homolog of which is TGFβ) and Wingless (Wg; the vertebrate homolog of which is Wnt) [11,19,20], and short-range morphogen Hedgehog (Hh) [21]. The produced morphogens form gradients to regulate the expression of target genes and control all aspects of wing development at cell level via the specific signaling pathways [19,22–27]. These signaling pathways are highly evolutionarily conserved in many different animals. Some key genes for the wing disc development have also been identified in a few limited insect species, such as *Tribolium castaneum* [28]. However, knowledge about the identification and involvement of wing development-related genes in various insects, especially in non-model insects, is still unclear and incomplete.

Orthopteran migratory locust, *Locusta migratoria* manilensis, and lepidopteran Asian corn borer, *Ostrinia furnacalis* (Gueneé), are two different types of insects undergoing incomplete and complete metamorphosis, respectively [29,30]. Identification of candidate genes regulating wing development would provide insights into the further study about the molecular mechanisms controlling metamorphosis development. Traditionally, such gene identification in “non-model” insects relied on degenerate PCR, which is labor-intensive, expensive, prone to failure, and only produces incomplete fragments [31]. The introduction of novel high throughput sequencing technologies greatly facilitates the global analysis of the blueprint of development-related genes [32]. This technology has been used, for example, to characterize the wing development-related genes in the milkweed bug *Oncopeltus fasciatus* [33], the oriental fruit fly *Bactrocera dorsalis* [34], and the salt marsh beetle *Pogonus chalceus* [35] etc.

In this study, we combined the Illumina sequencing and de novo assembly to obtain and characterize the transcriptome of the wing discs of *L. migratoria* nymph. 91,907 unigenes were assembled and 23,359 ones were annotated to known databases. We also re-characterized the previous transcriptome of *O. furnacalis* larvae [36], with special emphasis on wing development-related genes. Overall, we identified 50 potential wing development-related unigenes from *L. migratoria* transcriptome, and 46 unigenes from *O. furnacalis* transcriptome, respectively. Additionally, we performed qRT-PCR analysis to investigate the gene expression profiles of several key wing development genes during the stage of rapid growth in *L. migratoria* and *O. furnacalis*. All these results provide valuable information for studying the molecular mechanisms involved in the insect wing development, and are useful resources for further exploring the mechanism how signaling pathways control wing development.

**Materials and Methods**

**Insect rearing**

Migratory locust *L. migratoria* manilensis was reared on fresh wheat seedlings at 28–30°C, 60% relative humidity with an 18L/6D photoperiod. Asian corn borer (*O. furnacalis* (Gueneé)) larvae were reared on an artificial diet at 28°C under a relative humidity of 70–90% and a photoperiod of 18L/6D [36].

**Dissection and total RNA extraction of wing discs**

The wing discs of *L. migratoria* fifth instar nymph were dissected along the wing root with small scissors under microscope. Thirty pairs of dissected wing discs were combined, and total RNA samples were prepared using TRizol Reagent (TIANGEN, Beijing, China) following the manufacturer’s instructions. Total RNA was dissolved in H2O and RNA quantity was determined on a Nanodrop ND-2000 spectrophotometer (NanoDrop products, Wilmington, DE, USA). RNA integrity was checked on Agilent 2100 BioAnalyzer (Agilent Technologies, Englewood, CO, USA).
Library construction and Illumina sequencing

Ten µg of total RNA was used to isolate mRNA using oligo(dT) magnetic beads. The cDNA library was constructed using NEBNext mRNA Library Prep Reagent Set (NEB, Ipswich, MA, USA) following the manufacturer’s protocols. Briefly, enriched poly(A) RNA of each sample was fragmented into 200–700 nt pieces with RNA Fragmentation Reagents. The cleaved RNA fragments were transcribed into the first-strand cDNA using random hexamer-primers, followed by second-strand cDNA synthesis. The resulting double-stranded cDNA (dsDNA) was purified with QiaQuick PCR extraction kit (Qiagen, Hilden, Germany) and dissolved in EB buffer. The purified dsDNA was treated with T4 DNA Polymerase and T4 Polynucleotide Kinase for end-repairing and dA-tailing. After that, they were ligated to sequencing adaptors with barcode using T4 DNA ligase. Finally, fragments with around 200 bp-length were purified with QiaQuick GelPurify Kit (Qiagen, Hilden, Germany), and used as templates for PCR amplification to create the cDNA library. The library was sequenced on Illumina HiSeq 2000 (Illumina, San Diego, CA, USA) in Baimaike company (Beijing, China).

Assembly and annotation of transcriptomes

Raw reads were filtered to remove low quality reads with Q20 less than 20 and the sequence reads containing adapters and poly-A/T tails. The resulting clean reads were assembled to produce unigenes using the short reads assembling program – Trinity [37]. The BLAST results were used to extract coding region sequences (CDS) from the unigene sequences, and translate them into peptide sequences. When a unigene happened to have no BLAST hits, ESTScan software [39] would be used to determine the sequence direction. In addition, we performed the Gene Ontology (GO) annotations for each unigene with Blast2GO program according to the GO association done by a BLASTX against the Nr database [40,41].

Identification and sequence analysis of wing development-related genes from *L. migratoria* and *O. fumacalis* transcriptome

The available wing development-related gene sequences from *Drosophila* were used as references to screen *L. migratoria* transcriptome database obtained above and *O. fumacalis* transcriptome obtained previously [36]. The potential candidates of *L. migratoria* and *O. fumacalis* wing development-related genes were confirmed by searching the BLASTX algorithm against the non-redundant (nr) NCBI nucleotide database using a cut-off E-value of $10^{-5}$.

For the sequence analysis of putative wing development-related genes identified above, the deduced protein domains were determined by using Pfam (http://www.sanger.ac.uk/Software/Pfam/) and SMART (http://smart.embl.de/). Analysis of deduced amino acid sequences, including prediction of signal peptide, molecular weight and isoelectric point, was carried out in the EXPASY (Expert Protein Analysis System) proteomics server (http://www.expasy.org). Sequence comparisons and phylogenetic analysis were performed by MEGA5 software [42]. Phylogenetic trees were constructed by the neighbor-joining
method, with statistical analysis by the bootstrap method, using 1000 repetitions.

Expression assay of several identified wing development-related genes

To investigate the expression profiles of several key wing development-related genes in forewings and hind wings, we dissected the wing discs as described above, and extracted total RNA independently from 3 biological replicates. DNase I-treated RNA (1 μg) was converted into first-strand cDNA using TIANScript RT Kit (TIANGEN, Beijing, China). The cDNA products were diluted 2 fold for use as template. Specific primers for each gene were designed and listed in Table S1. L. migratoria Actin and O. furnacalis ribosomal protein L8 (rpL8) was used as an internal standard to adjust the template amounts in a preliminary PCR experiment. The qRT-PCR was performed on an Applied Biosystems 7500 Real-time PCR system (Life Technologies, Grand Island, NY, USA) using the GoTaq qPCR Master (Promega, Madison, WI, USA), according to the manufacture’s instructions. The thermal cycling conditions for qRT-PCR and calculation methods were same as described previously [36].

Results and Discussion

Dissection and observation of wing discs in L. migratoria and O. furnacalis

All insects in the Pterygota undergo metamorphosis from immature to adult. Among them, insects with incomplete metamorphosis, such as migratory locust L. migratoria, have young nymph resembling the adult with visible forewings and hind wings. Meanwhile, insects with complete metamorphosis, such as corn borer O. furnacalis, go through four stage processes from egg, larva, pupa, and adult in which only adult has visible wings. Although there are some minor differences, the overall appearances of forewings and hind wings in L. migratoria nymph and adult are similar (Fig. 1). However, actual wings are only visible in O. furnacalis adult. In fifth-instar O. furnacalis larva, a pair of pea-like wing discs composed of large amount of cells was dissected from the position where the forewing and hind wing will be derived. There is no similarity on the appearance between the wing discs and adult wings (Fig. 1). The huge difference about the wing development in L. migratoria and O. furnacalis suggests that the molecular mechanisms controlling the wing development in insects with incomplete or complete metamorphosis might be largely different. As a first step to understand these molecular mechanisms, it is important and necessary to identify as many as possible genes functioning in the wing development. Previously, we have combined the Illumina sequencing and de novo assembly to obtain the high quality transcriptome from O. furnacalis larvae [36]. In this study, we also obtained the data for L. migratoria transcriptome, and we can comparatively characterize both transcriptomes and identify wing development-related genes. This work will provide useful information for studying the molecular basis involved in the wing development in L. migratoria and O. furnacalis, two insect species with different type of metamorphosis.

Sequencing and unigene assembly of L. migratoria transcriptome

In order to obtain detailed information about L. migratoria transcriptome, we prepared cDNA from the wing discs of fifth-instar nymph, and subjected it to Hiseq 2000 sequencing. After cleaning of dirty reads and quality checks, a total of 83,000,540 high-quality clean reads (SRA accession number SRX491784) with a cumulative length of 8,382,624,748 nucleotides (8.38 Gb) were generated from L. migratoria wing disc library. The GC percentage of the reads is 42.65%, which is comparable with genome sequence of other insects. Using Trinity software by the manner of paired-end joining and gap-filling, these reads were assembled into 91,907 unigenes longer than 200 nt (mean length of 610 nt and N50 of 1024 nt). It was significantly larger than that in previous
insect transcriptome projects [43,44]. The assembled sequences have been deposited in the NCBI Transcriptome Shotgun Assembly (TSA) Database under the accession GBDZ00000000.

Functional annotation and classification

In order to annotate the unigenes, the obtained sequences were first aligned by BLASTX to various protein databases of nr, Swiss-Prot, KEGG, COG, and GO (E-value $< 10^{-5}$), and then aligned by BLASTN to nucleotide database nt (E-value $< 10^{-5}$). Among 91,907 unigenes, 20,746 (22.6%), 8,737 (9.5%), 11,450 (12.5%), 5,312 (5.8%), 4,986 (5.4%), and 10,644 (11.6%) ones were annotated in nr, nt, Swiss-Prot, KEGG, COG, and GO, respectively (Table S2). A total of 23,359 unigenes (25.4%) were annotated to at least one database. The remaining 68,548 unigenes (74.6%) were not annotated to any referred databases.

The low annotated percentage might be due to the transcripts derived from the cDNA of untranslated regions, assembly errors, and nonconserved areas of proteins where homology is not detected [43]. Another possibility was that a large part of the genes in *L. migratoria* transcriptome database were with unknown functions, and the un-annotated unigenes were the potential sources of novel genes.

The functional annotations of unigenes were performed mainly based on the BLASTX results against the nr database. Among the 20,746 unigenes annotated to nr database, 7,883 (38%) showed strong homology (E-value smaller than 1e-50) (Fig. 2A). The identity comparison showed 5,809 (28%) unigenes have more than 60% identity with other insects (Fig. 2B). Among the biological process assignments, cellular processes (15.0%) and metabolic processes (14.4%) represented the most abundant subcategories. It indicated the importance of cell cycle, generation as well as metabolic activities in wing development stage. Under the category of cellular component, the top 3 sub-categories were cell part (21.0%), cell (20.2%) and organelle (16.0%). In the molecular function category, binding (43.0%) and catalytic activities (36.9%) were the most abundant (Fig. 3). The subcategories taking up the largest two proportions in each category were consistent with that in transcriptomic studies of other insects [45,46].

Identification and comparison of wing development-related genes in *L. migratoria* and *O. furnacalis* transcriptome

Wing development is a complicated biological process comprising multiple integrated signaling pathways. In particular, four signaling pathways – Notch, Hedgehog (Hh), Decapentaplegic (Dpp), Wingless (Wg) – play important roles in growth control and cell fate determination during the wing development.

[Figure 4. Clusters of orthologous groups (COG) classification of *L. migratoria* unigenes. A total of 6,508 produced functional annotations were among the 25 categories. The Y-axis shows the number of unigene in each COG term. doi:10.1371/journal.pone.0106770.g004]
Table 1. Summary of the wing development-related unigenes identified in *L. migratoria* and *O. furnacalis* transcriptome.

| Function-pathway | **Locusta migratoria** | **Ostrinia furnacalis** |
|------------------|------------------------|-------------------------|
| **Hedgehog pathway** |                        |                         |
| hedgehog (hh)    | L a37032 4051 3' 383 66 e-101 | Unigene16362 590 M 196 73 9e-064 |
| patched (ptc)    | R a57898 1327 5' 372 75 e-135 | Unigene153 3111 Y 775 62 0.0 |
| smoothened (smo) | R a59513 2125 3' 651 62 6e-073 | CL2764.Contig1 2937 Y 751 67 0.0 |
| cubitus interruptus (ci) | F a69651 650 M 216 79 4e-096 | Unigene15615 910 M 303 77 5e-098 |
| costal-2 (co-2)  | B a8431 3970 Y 843 52 e-102 |                         |
| Fused            | K a32509 2482 3' 650 74 7e-071 | Unigene11591 1346 M 448 64 1e-071 |
| suppressor of fused (su(fu)) | B a11681 2472 3' 488 55 3e-086 | Unigene16575 637 M 212 59 3e-042 |
| **Decapentaplegic pathway** |                      |                         |
| decapentaplegic (dpp) |                         |                         |
| glass bottom boat (gbb) | L a2659 1653 3' 332 64 5e-077 | CL9574.Contig1 1573 Y 448 61 8e-097 |
| saxophone (sax)   | R a14241 2261 Y 568 76 0.0 | Unigene21017 796 M 256 77 1e-098 |
| punt              | R a17942 5910 Y 503 69 e-144 | CL628.Contig1 2051 Y 547 49 1e-051 |
| thickveins (tkv)  | R a4697 2436 Y 512 72 e-174 | Unigene5852 2162 3' 335 82 e-137 |
| mothers against Dpp (mad) | F a1173 2860 Y 475 84 0.0 | Unigene16360 1775 Y 423 86 0.0 |
| medea (med)       | A a26380 1235 3' 356 67 1e-095 | Unigene6852 1575 Y 447 58 e-147 |
| a11651 3401 M 397 66 e-131 | Unigene17363 589 M 196 89 3e-080 |
| optomotor-blind (omb) | F a4984 803 M 267 96 e-121 | CL105.Contig5 1131 M 376 89 9e-142 |
| spalt major (salm) | F a6922 1073 5' 325 55 2e-070 | CL2350.Contig2 2278 M 759 54 2e-097 |
| spalt related (salr) | F a12971 2087 5' 651 47 2e-089 |                         |
| **Notch signaling Pathway** |                      |                         |
| Notch (n)         | R a1269 12296 Y 2484 72 0.0 | Unigene6905 2340 M 780 83 0.0 |
| delta (dl)        | L a123129 926 M 308 69 e-114 | Unigene3068 987 M 328 69 e-120 |
| serrate (ser)     | L a1910 7128 Y 1324 49 0.0 | CL7020.Contig1 1927 M 409 52 e-164 |
| Hairless (h)      | R a35245 1975 5' 538 61 5e-032 | Unigene3346 243 M 80 68 4e-015 |
| Functiona | Unigene ID | Nucleotide Length (nt) | Full lengthb | Protein length (aa)c | Identity to Best hit (%)d | P valuee | Location |
|-----------|------------|------------------------|-------------|---------------------|--------------------------|-----------|----------|
| suppressor of hairless (su(H)) | A | a4077 | 3351 | Y | 496 | 95 | 0.0 | Unigene16575 |
| mastermind (mam) | A | a14121 | 1140 | 5’ | 153 | 54 | 1e-025 | Unigene19131 |
| notchless | B | a6959 | 1804 | Y | 474 | 78 | 0.0 | CL2896.Contig1 |
| fringe | K | a17301 | 2466 | 3’ | 259 | 82 | e-111 | Unigene9365 |
| hairy | B | a4291 | 3634 | 3’ | 407 | 52 | 6e-070 | Unigene22517 |
| Wingless pathway |
| wingless (wg) | L | a40139 | 948 | 3’ | 227 | 85 | 7e-050 | Unigene1219 |
| frizzled (fz) | R | a1577 | 5642 | Y | 571 | 68 | 0.0 | CL4000.Contig1 |
| frizzled2 (fz2) | R | ? | | | | | |
| arrow (arr) | R | a39913 | 4189 | 5’ | 1381 | 65 | 0.0 | Unigene11432 |
| disheveled (dsh) | F | a6911 | 3503 | Y | 721 | 60 | e-154 | Unigene24227 |
| armadillo (arm) | F | a1041 | 3398 | Y | 820 | 87 | 0.0 | Unigene21327 |
| axin | A | a4962 | 3922 | 3’ | 676 | 64 | 2e-031 | Unigene17645 |
| adenomatous polyposis coli (APC) | A | a20732 | 4723 | 5’ | 1541 | 47 | e-160 | Unigene11483 |
| vestigial (vg) | F | a302963 | 243 | M | 81 | 62 | 4e-013 | Unigene4539 |
| scalloped (sd) | F | a1575 | 4661 | Y | 459 | 69 | 0.0 | CL1169.Contig4 |
| shaggy | K | a2857 | 10994 | Y | 1913 | 83 | 0.0 | Unigene16270 |
| Others |
| ultrabithorax (Ubx) | F | a12834 | 2418 | 3’ | 174 | 56 | 6e-053 | Unigene20941 |
| apterous (ap) | F | a188016 | 391 | M | 130 | 59 | 5e-044 | Unigene4961 |
| engrailed (en) | F | a1647 | 2530 | 3’ | 138 | 80 | 3e-052 | CL5152.Contig1 |
| homothorax (Hth) | F | a22527 | 721 | M | 240 | 75 | e-102 | CL5887.Contig2 |
| teashirt (Tsh) | F | a866117 | 347 | M | 115 | 55 | 8e-027 | ? |
| epidermal growth factor receptor (EGFR) | R | a1492 | 4431 | Y | 1394 | 53 | 0.0 | CL625.Contig2 |
| rhomboid (rho) | K | a36342 | 2527 | Y | 294 | 58 | 1e-072 | ? |
| dublin (Nub) | F | a3913 | 1608 | 3’ | 460 | 45 | 1e-080 | Unigene16843 |
| panner (Pnr) | F | a455482 | 478 | M | 159 | 71 | 7e-053 | Unigene160141 |
| noturn | K | a538614 | 267 | M | 88 | 62 | 9e-029 | Unigene26847 |
| fat (ft) | R | a9170 | 5093 | M | 1697 | 51 | 0.0 | Unigene18347 |
| four-jointed (fj) | K | a37684 | 381 | M | 127 | 66 | 9e-036 | Unigene6909 |
Table 1. Cont.

| Hh    | Unigene ID | Genes involved in the Hh signaling pathway. |
|-------|------------|--------------------------------------------|
| L. migratoria | a37032 | 
| O. furnacalis | Unigene16362 | 
| D. melanogaster | 

Hh signaling is mediated by a multi-component receptor complex in the cell membrane. This receptor complex consists of a 12-span transmembrane protein, Patched (Ptc) as the receptor and a 7-span transmembrane protein, Smoothened (Smo) as the obligatory signal transducer across the plasma membrane [51,58]. When extracellular Hh binds to and is inhibited by Ptc, Smo starts to accumulate, and inhibit the proteolytic cleavage of zinc-finger transcription factor Cubitus interruptus (Ci) which is normally bound by the kinesin-like protein Costal-2 (Cos2). The intact Ci protein then translocate into the nucleus, allowing the transcription of some genes such as dpp. In the absence of Hh, Ptc blocks Smo activity, and full-length Ci protein is degraded into Ci fragment (CiR) that functions as a transcriptional repressor to block the transcription of target genes (Fig. S3A) [50]. By searching L. migratoria and O. furnacalis transcriptome using Drosophila corresponding genes as reference, we have identified 46 unigenes with high potential to be related to the wing development (Table 1).
the potential 1:1 orthologs of PtC, Smo, Ci from both transcriptome databases (Table 1). It suggests that Hh signaling pathway is conserved in both L. migratoria and O. furnacalis.

**Genes involved in Dpp signaling pathway.** Dpp, directed by the Hh signaling, is a member of the bone morphogenetic protein (BMP) growth factor family [59]. Dpp forms a long-range dynamic and precise gradients to control cell survival, cell morphogenesis, cell proliferation, cell differentiation during the wing development [22,25,26,60]. When a cell receives a Dpp signal, its heteromeric receptor complex composed of type I receptor Thickveins (Tkv) and type II receptor Punt is activated [61,62]. Once activated, the receptors are able to phosphorylate an intracellular protein called mothers against Dpp (Mad) [63]. The phosphorylated Mad then associates with the Medea (Med), and the complex translocates to the nucleus where it binds to DNA and activates or suppresses the expression of the target genes in conjunction with other transcription factors [64]. Within this signaling pathway, the Drosophila Hox gene Ultrabithorax (Ubx) restricts both the transcription and the mobility of Dpp [65].

Figure 5. Alignments of potential Hhs in L. migratoria, O. furnacalis and Drosophila. The deduced amino acid sequences of L. migratoria a37032 and O. furnacalis Unigene16362 were compared with Drosophila Hh (Dm_Hh). Completely conserved amino acids are indicated by “*”, and conservative substitutions by “:” and “.” below the sequences. The conserved cysteine residue for an acyl rearrangement analogous to step 1 of the protein splicing pathway is boxed. The conserved motifs in Drosophila Hedgehog are underlined.

doi:10.1371/journal.pone.0106770.g005
BMP genes with TGF-beta homology (Fig. 8). The identified fragments, *L. migratoria* a136595 and *O. furnacalis* Unigene7326, reside in the *dpp* group with statistical support (reliability value = 98 and 99 for a136595 and Unigene7326, respectively). Therefore, they are postulated as potential *dpp* genes. We conducted qRT-PCR analysis to investigate the expression profile of *dpp*. *L. migratoria* dpp was expressed at high level in the first instar nymphs, then decreased significantly in the second, third, and fourth instar nymphs, but significantly increased in the hind wings of the fifth instar nymph. The mRNA level of *O. furnacalis* dpp remained unchanged in all tested stages (Fig. 7). The different expression patterns of dpp in two insects suggest that dpp might be related to the different metamorphosis.

In addition, it is interesting that only one *mad* gene was present in *Drosophila* but two transcripts encoding for potential *mad* were identified from both *L. migratoria* and *O. furnacalis* transcriptomes (Table 1). Two Mads in *L. migratoria* (a11173 and a12619) shared 60% identity while two *O. furnacalis* Mads (Unigene16360 and Unigene6852) had 61% identity in amino acid sequences. A similar situation happened to the identification of *med*. There is only one *med* gene in *Drosophila* but two transcript fragments were identified for both *L. migratoria* and *O. furnacalis* med, including a26380 and a11651 in *L. migratoria*, and unigene4133 and unigene17363 in *O. furnacalis*, respectively. The encoded amino acid sequences of a26380 and unigene4133 were highly similar to the N-terminus of *Drosophila* Med while a11651 and unigene17363 were highly similar to the C-terminus of *Drosophila* Med (Fig. S2). We doubted that the two unigenes from each transcriptome were just two partial fragments within *med* gene. However, no overlapped sequences were observed in a26380 and a11651, or in unigene4133 and unigene17363 (Fig. S2). One possible reason might be due to the sequencing errors because the predicted overlapping part was just located at the terminus of sequenced fragment. Further experiments are required to determine the full length of *med* in these two insect species.

Genes involved in the Notch signaling pathway. The Notch signaling pathway is highly conserved throughout the animal kingdom. It regulates cell-fate determination during development and maintains adult tissue homeostasis [47]. The key component of this signaling pathway is the Notch receptor. Notch is a 300 kDa single-pass transmembrane protein, composed of a large extracellular domain, a single transmembrane portion, and a small intracellular region [68–70]. After binding to its ligands, Delta and Serrate (known as Jagged in mammals), inactive Notch precursor undergoes two proteolytic cleavage events: the first cleavage is catalyzed by ADAM-family metalloproteases; the second cleavage is mediated by c-secretase which is an enzyme complex containing presenilin, nicastrin, PEN2 and APH1 [71].

Figure 6. Phylogenetic analysis of Hhs. The amino acid sequences of *L. migratoria* a37032 and *O. furnacalis* Unigene16362, together with 68 potential Hhs from other organisms were used to build the neighbor joining tree. *L. migratoria* a37032 and *O. furnacalis* Unigene16362 are boxed. Nematode Hh-related proteins and other eukaryote Hh proteins are indicated with brackets. The arrows at nodes denote bootstrap value greater than 700 from 1000 trials. doi:10.1371/journal.pone.0106770.g006
The second cleavage liberates the Notch intracellular domain (NICD), which then migrates into the nucleus and cooperates with the DNA-binding protein CSL (also named as CBF1, Suppressor of hairless (Su(H)), LAG-1, or RBP) and its co-activator Mastermind (Mam) to promote the transcription of downstream target genes, such as \(wg\) (Fig. S3C) [72]. Using the corresponding components in \(Drosophila\) as referred sequences, we have identified 9 transcripts for potential components in Notch signaling pathway from \(L.\ migratoria\) and \(O.\ furnacalis\) transcriptomes, respectively (Table 1). Four out of nine unigenes in \(L.\ migratoria\) are complete while only one unigene is complete in \(O.\ furnacalis\). The possible reason was that the sequencing quality of \(L.\ migratoria\) transcription was better than that of \(O.\ furnacalis\) transcriptome (8.38 Gb vs. 4.72 Gb). Given the importance of Notch in this signaling pathway, we selected it for further analysis. Similar to the case in other invertebrates, we only identified one transcript for Notch gene, a1269 from \(L.\ migratoria\) transcriptome, and Unigene6905 from \(O.\ furnacalis\) transcriptome (Table 1). The canonical Notch protein consists of three repeated sequence motifs: the extracellular domains contain 10–36 copies of an ~40-amino acid epidermal growth factor-like (EGFL) sequence motif and 3 copies of an ~40-amino acid \(lin/Notch/glp\) (LNG) sequence motif; the intracellular domains contain 6–7 copies of a \(cde10/5W16/ankyrin\) (CDC/ANK) sequence motif flanked by stretches of nonrepetitive sequences [73]. \(L.\ migratoria\) a1269 encodes a 2,484-amino acids full-length protein which has 63% similarity to \(Drosophila\) Notch. \(L.\ migratoria\) Notch contains 36 EGFL tandem repeats, 3 LNG repeats, and 7 tandem ANK repeats (Fig. 9A), suggesting it is a canonical Notch. \(O.\ furnacalis\) Unigene6905 encodes a 780-amino acids polypeptide with 70% similarity to \(Drosophila\) Notch. Only 20 tandem EGFL repeats are predicted in the current identified Unigene6905 fragment (Fig. 9A). It is unknown whether \(O.\ furnacalis\) Notch also contain classic LNG and ankyrin repeats because the current transcript is incomplete. We performed the phylogenetic analysis for \(L.\ migratoria\) and \(O.\ furnacalis\) Notch and 30 Notch sequences from other species to investigate the evolutionary relationship of the Notch protein family. As shown in Fig. 9B, the Notch from insects forms a separate group which includes \(L.\ migratoria\) and \(O.\ furnacalis\) Notch, and the Notch from vertebrates is clustered into another group. Notch from \(Ciona\), sea urchin, and amphioxus is grouped with vertebrate Notch, however, with a low bootstrap value of 82. \(Caenorhabditis\ elegans\) Notch is out of any group, showing a great differentiation from the other taxa. We analyzed the expression profiles of Notch using qRT-PCR methods. As shown in Fig. 7, the expression profile of Notch was similar to that of \(Hh\) in two insects. It kept unchanged from the first instar through the fifth instar stage in \(L.\ migratoria\), while it increased significantly in \(O.\ furnacalis\) fifth instar larva (Fig. 7).
**Figure 8. Phylogenetic analysis of Dpps.** Except for *L. migratoria* Unigene7326 (marked in red), the used amino acid sequences of other 33 Dpps are from mouse (Mm), zebrafish (Dr), lancet (Branchiostoma floridae; Bf), acorn worm (Pychodera flava; Pf), the sea urchins Strongylocentrotus purpuratus (Sp) and Lytechinus variegatus (Lv), fruit fly (Dm), flour beetle (Tribolium castaneum; Tc), sawfly (Athalia rosae; Ar), buckeye butterfly (Junonia coenia; Jc), the grasshoppers Schistocerca americana (Sa) and *S. gregaria* (Sg), cricket (Gryllus bimaculatus; Gb), Cupiennius salei (Cs), Ceratitis capitata (Cc), Blattella germanica (Bg), Polyrrhachis vicina (Pv), Episyrphus balteatus (Eb), Apis mellifera (Am), Pheidole morrisi (Pm), Biston betularia (Bb), Bombyx mori (Bm), Culex quinquefasciatus (Cq), Nasonia vitripennis (Nv). The branches specific for arthropod Dpps and TGF-beta homologs are shaded in blue and yellow, respectively. For explanation of the arrows see Fig. 6.

doi:10.1371/journal.pone.0106770.g008

**Genes involved in the wg signaling pathway.** Wg (the vertebrate homolog of which is Wnt) is transcriptionally activated by Notch signaling, and Wg signaling pathway plays critical roles in axis patterning, cell fate specification, cell proliferation, and cell migration etc [74]. During the development of *Drosophila* wing, canonical Wg signaling pathway specifies pattern formation along the dorsal/ventral (D/V) axis while Dpp signaling pathway play this role along the A/P axis [75]. The ligand Wg is the founding member of the Wnt family, and is a secreted lipid-modified signaling glycoprotein that has 350–400 amino acids in length [76]. Wg is expressed at the D/V boundary and forms a stable and long-range gradient by symmetrically diffusing at both sides of the boundary. When Wg binds to a receptor complex consisting of the seven-transmembrane protein Frizzled (Fz) and the single-pass transmembrane protein Arrow (Arr, homologous to murine and human low density lipoprotein (LDL) receptor-related protein 5 or 6 (LRP5/6)), the downstream cytoplasmic protein disheveled (Dsh in *Drosophila* and Dvl in vertebrates) is activated [77]. Dsh in turn inhibits glycogen synthase kinase (GSK)-3β in the β-catenin destruction complex, which mainly consists of Axin, GSK-3β, adenomatous polyposis coli (APC) and β-catenin (armadillo (arm) in *Drosophila*). Consequently, β-catenin accumulates in the cytoplasm, and stabilized β-catenin then translocates into the nucleus and acts together with the transcription factor Pangolin to regulate the transcription of Wg target genes (Fig. S3D) [77]. In this study, we identified 10 and 11 unigenes for the known components in the Wg signaling pathway from *L. migratoria* and *O. furnacalis* transcriptome, respectively (Table 1). No putative *Frizzled 2* (fz2) ortholog was identified in *L. migratoria* transcriptome. A possible reason is that fz2 gene is missing in *L. migratoria* because of the evolutionary event. The other reason with higher possibility is that the transcript level of fz2 is low and it is not captured in the RNA-seq. We attempted to perform further analysis for identified wg gene, a40139 in *L. migratoria* and Unigene1219 in *O. furnacalis*. However, unigene1219 only encodes a 97-amino acid peptide which is too short to have no common sequences with other Wg during the alignment. Therefore, we failed to conduct phylogenetic analysis to reveal the evolutionary relationship of Wg. Additionally, we identified another 12 and 9 unigenes from *L. migratoria* and *O. furnacalis* transcriptome, respectively (Table 1), which were potentially involved in the wing development, including homologs to *Drosophila apterous, engrailed, homothorax, teashirt, epidermal growth factor receptor, rhomboid, nubbin, panier, netum, fat, four-jointed, daily-like.*

**Conclusions**

In summary, we sequenced and characterized the transcriptome from the wing discs of *L. migratoria* nymph. The assembled sequence data comprising 91,907 unique transcripts provides a comprehensive sequence source for future *L. migratoria* study. We identified a large set of genes relevant to wing development with high significance, especially the genes involved in four signaling pathways – Notch, Hh, Dpp, and Wg signaling pathways, from *L. migratoria* transcriptome and another *O. furnacalis* transcriptome obtained previously. The explored wing development-related genes constitute an integrated picture of the development network, which provides the valuable clues for a better understanding of the wing development in *L. migratoria* and *O. furnacalis*. These development repertoire genes appear to be evolutionarily conserved to different extent. Functional analyses are necessary to verify our predictions. Nevertheless, the framework of information presented in this study should help to further
understand the complex molecular mechanisms involved in wing development in *L. migratoria* and *O. furnacalis*, two insect species with different type of metamorphosis.

**Supporting Information**

**Figure S1** Assembled unigene length distribution of *L. migratoria* transcriptome. The x-axis indicates unigene size and the y-axis indicates the number (left) or percentage (right) of unigenes of each size.

**Figure S2** Alignments of potential Meds in *L. migratoria*, *O. furnacalis* and *Drosophila*. Note: *L. migratoria* a26380 (green) and *O. furnacalis* Unigene4133 (purple) match well with the N-terminus of *Drosophila* Med (blue), while *L. migratoria* a11651 (black) and *O. furnacalis* Unigene17363 (red) are highly similar to the C-terminus of *Drosophila* Med (blue). The part without identity is shaded.

**Figure S3** Schematic drawing of the Hh (A), Dpp (B), Notch (C), and Wg (D) signaling pathways. (A) In the presence of Hh, Hh binds to its receptor Ptc and Ptc relieves Smo repression. Smo accumulates and is activated by phosphorylation. This promotes its association with the complex including Costal 2 (Cos2), Fused (Fu), Suppressor of Fu (Su(Fu)) and Cubitus interruptus (Ci). Uncleaved Ci is released from the complex and acts as transcriptional activator in the nucleus and induces expression of target genes. (B) Dpp binds to its receptor complex including the type I receptor Tkv and type II receptor Punt which induces the phosphorylation of Tkv and in turn phosphorylates the signal transducer Mad. Phosphorylated Mad forms complex with Med and then translocates to the nucleus and induces expression of
target genes. (C) Notch receptor is a transmembrane protein, with a large extracellular domain (NECD), a transmembrane domain and an intracellular domain (NICD). When its ligands DI or Ser bind to the Notch, two proteolytic cleavage events are activated: the first cleavage is catalyzed by ADAM-family TACE metalloproteinase; and an intracellular domain (NICD). When its ligands D1 or Ser binds to the Notch receptor, the second cleavage is mediated by γ-secretase which is an enzyme complex containing presenilins, nicastrin, PEN2 and APH1. The second cleavage liberates the NICD, which then migrates into the nucleus and cooperates with the DNA-binding protein CSL and its co-activator Mam to promote the transcription of downstream target genes. (D) Wg binds to the receptors Fz and Art. The Wg-Arr-Fz complex binds to and activates the target genes.

References

1. Dudley R (2002) The biomechanics of insect flight: form, function, evolution. Princeton University Press.
2. Wagner GP (2007) The developmental genetics of homology. Nat Rev Genet 8: 473–479.
3. Miaoalis IN, Heilman BD (1998) Butterfly thin films serve as solar collectors. Ann Entomol Soc Am 91: 122–127.
4. Sudo S, Tsuyuki K, Kanno K (2005) Wing characteristics and flapping behavior of flying insects. Exp Mech 45: 550–555.
5. Simmons RB, Conner WE (1996) Ultrasonic signals in the defense and courship of Euchuast eggl Detray and E bolteri Stretch (Lepidoptera: Arctiidae). J Insect Behav 9: 909–919.
6. Gibson G, Warren B, Russell J (2010) Humming in tune: sex and species recognition by mosquitoes on the wing. J Assoc Res Otolaryngol 11: 527–540.
7. Truman JW, Riffell JD (1999) The origin of insect metamorphosis. Nature 401: 447–452.
8. Alberts B, Johnson A, Lewis J, Walter P, riot M, et al. (2002) Organogenesis and the Patternning of Appendages. In: Alberts B, Johnson A, Lewis J, Walter P, riot M, et al. editors. Molecular Biology of the Cell 3rd Edition. Routledge. pp. 1347–1363.
9. Herrera SC, Martin R and Morata G (2013) Tissue homeostasis in the wing disc of Drosophila melanogaster: immediate response to massive damage during development. PLoS Genet 9: e1003446.
10. Fortini ME, Bonini NM (2000) Modeling human neurodegenerative diseases in Drosophila: on a wing and a prayer. Trends Genet 16: 161–167.
11. Neto-Silva RM, Wells BS, Johnston LA (2009) Mechanisms of growth and development in the wing imaginal disc disrupts epithelial architecture and promotes cell sorting. BMC Dev Biol 10: 23.
12. Richards S, Gibbs RA, Weinstock GM, Brown SJ, Derril D, et al. (2006) The genome of the model beetle and pest Tribolium castaneum. Nature 452: 949–955.
13. Waldron JA, Newbury SF (2012) The roles of miRNAs in wing imaginal disc development. PLoS Genet 8: e1002936.
14. Worley MI, Setiawan L, Hariharan IK (2012) Regeneration and transdifferentiation of the wing imaginal disc epithelium: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 28: 644–652.
15. Altshul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389–3402.
16. Zecca M, Basler K, Struhl G (1996) Direct and long-range action of a wingless morphogen. Cell 85: 971–980.
17. Zecca M, Basler K, Struhl G (1996) Direct and long-range action of a wingless morphogen. Cell 85: 971–980.
18. Zecca M, Basler K, Struhl G (1996) Direct and long-range action of a wingless morphogen. Cell 85: 971–980.
19. Zhang X, Luo D, Pflugfelder GO, Shen J (2013) Dpp signaling inhibits proliferation in the Drosophila wing by Omb-dependent regional control of bantam. Development 140: 2917–2922.
20. Shen J, Dahmann C (2005) Extrusion of cells with inappropriate Dpp signaling from Drosophila wing imaginal epithelium. Science 307: 1789–1790.
21. Sheng JA, Dahmann C, Pflugfelder GO (2010) Spatial discontinuity of optomotor-blinded nuclear EGFR-mediated signaling. Development 137: 1357–1368.
22. Zecca M, Basler K, Struhl G (1996) Direct and long-range action of a wingless morphogen. Cell 85: 971–980.

Table S1 Primers for RT-PCR analysis.

Table S2 The statistics of annotated unigenes.

Acknowledgments

We thank Dr. Yang Liu for some valuable suggestions on the transcriptome analysis.

Author Contributions

Conceived and designed the experiments: SL LZ JS CA. Performed the experiments: SL WW YC. Analyzed the data: SL JS CA. Contributed reagents/materials/analysis tools: SL WW YC. Contributed to the writing of the manuscript: SL LZ JS CA.
47. Rafel N, Milan M (2008) Notch signalling coordinates tissue growth and wing fate specification in Drosophila. Development 135: 3995–4001.
48. Brook WJ (2000) Hedgehog signaling and the axial patterning of Drosophila wings. Biochem Cell Biol 78: 565–591.
49. Neumann CJ, Cohen SM (1997) Long-range action of Wingless organizes the dorsal-ventral axis of the Drosophila wing. Development 124: 871–880.
50. Jiang J, Hui CC (2008) Hedgehog signaling in development and cancer. Dev Cell 15: 801–812.
51. Ingham PW, McMahon AP (2001) Hedgehog signaling in animal development: paradigms and principles. Genes Dev 15: 3059–3087.
52. Weedon MN, Lango H, Lindgren CM, Wallace C, Evans DM, et al. (2008) Genome-wide association analysis identifies 20 loci that influence adult height. Nat Genet 40: 575–583.
53. Tabata T, Kornberg TB (1994) Hedgehog is a signaling protein with a key role in patterning Drosophila imaginal discs. Cell 76: 89–102.
54. Perrimon N, Mahowald AP (1987) Multiple functions of segment polarity genes in Drosophila. Dev Biol 119: 387–400.
55. Varjosalo M, Taipale J (2008) Hedgehog: functions and mechanisms. Genes Dev 22: 2454–2472.
56. Perler FB (1998) Protein splicing of inteins and hedgehog autoproteolysis: structure, function, and evolution. Cell 92: 1–4.
57. Burghin TR (2008) Evolution of hedgehog and hedgehog-related genes, their origin from Hgo proteins in ancestral eukaryotes and discovery of a novel Hint motif. BMC Genomics 9: 127.
58. Carpenter D, Stone DM, Brush J, Ryan A, Armanini M, et al. (1998) Characterization of two patched receptors for the vertebrate hedgehog protein family. Proc Natl Acad Sci U S A 95: 13630–13634.
59. Spencer FA, Hoffmann FM, Gelbart WM (1992) Decapentaplegic: a gene complex affecting morphogenesis in Drosophila melanogaster. Cell 28: 451–461.
60. Akiyama T, Kamimura K, Firkus C, Takeo S, Shimmi O, et al. (2008) Dally regulates Dpp morphogen gradient formation by stabilizing Dpp on the cell surface. Dev Biol 313: 408–419.
61. Affolter M, Basler K (2007) The Decapentaplegic morphogen gradient: from pattern formation to growth regulation. Nat Rev Genet 8: 663–674.
62. Bangi E, Wharton K (2006) Dual function of the Drosophila Alk1/Alk2 ortholog Saxophone shapes the Bmp activity gradient in the wing imaginal disc. Development 133: 3295–3305.