Gene 637 (2017) 90–99

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Transcriptome profiling of the *Plutella xylostella* (Lepidoptera: Plutellidae) ovary reveals genes involved in oogenesis

Lu Peng[a,b,c,d,e,1], Lei Wang[a,b,c,d,e,1], Yi-Fan Yang[a,b,c,d,e,1], Ming-Min Zou[a,b,c,d,e], Wei-Yi He[a,b,c,d,e], Yue Wang[a,b,c,d,e], Qing Wang[a,b,c,d,e], Liette Vasseur[a,b,c,d,e,f], Min-Sheng You[a,b,c,d,e,*]

[a] State Key Laboratory of Ecological Pest Control for Fujian-Taiwan Crops and College of Life Science, Fujian Agriculture and Forestry University, Fuzhou 350002, China
[b] Institute of Applied Ecology, Fujian Agriculture and Forestry University, Fuzhou 350002, China
[c] Fujian-Taiwan Joint Innovation Centre for Ecological Control of Crop Pests, Fujian Agriculture and Forestry University, Fuzhou 350002, China
[d] Key Laboratory of Integrated Pest Management for Fujian-Taiwan Crops, Ministry of Agriculture, Fuzhou 350002, China
[e] Fujian Provincial Key Laboratory of Insect Ecology, Fujian Agriculture and Forestry University, Fuzhou 350002, China
[f] Department of Biological Sciences, Brock University, St. Catharines, Ontario L2S 3A1, Canada

ARTICLE INFO

Keywords:
Ovary development
Oogenesis
Transcriptome analysis
Reproductive regulation
Plutella xylostella

ABSTRACT

Background: As a specialized organ, the insect ovary performs valuable functions by ensuring fecundity and population survival. Oogenesis is the complex physiological process resulting in the production of mature eggs, which are involved in epigenetic programming, germ cell behavior, cell cycle regulation, etc. Identification of the genes involved in ovary development and oogenesis is critical to better understand the reproductive biology and screening for the potential molecular targets in *Plutella xylostella*, a worldwide destructive pest of economically major crops.

Results: Based on transcriptome sequencing, a total of 7.88 Gb clean nucleotides was obtained, with 19,934 genes and 1861 new transcripts being identified. Expression profiling indicated that 61.7% of the genes were expressed (FPKM ≥ 1) in the *P. xylostella* ovary. GO annotation showed that the pathways of multicellular organism reproduction and multicellular organism reproduction process, as well as gamete generation and chorion were significantly enriched. Processes that were most likely relevant to reproduction included the spliceosome, ubiquitin mediated proteolysis, endocytosis, PI3K-Akt signaling pathway, insulin signaling pathway, cAMP signaling pathway, and focal adhesion were identified in the top 20 'highly represented' KEGG pathways. Functional genes involved in oogenesis were further analyzed and validated by qRT-PCR to show their potential predominant roles in *P. xylostella* reproduction.

Conclusions: Our newly developed *P. xylostella* ovary transcriptome provides an overview of the gene expression profiling in this specialized tissue and the functional gene network closely related to the ovary development and oogenesis. This is the first genome-wide transcriptome dataset of *P. xylostella* ovary that includes a subset of functionally activated genes. This global approach will be the basis for further studies on molecular mechanisms of *P. xylostella* reproduction aimed at screening potential molecular targets for integrated pest management.

1. Background

Oogenesis is the process by which female gametes are developed and includes ovary and egg development, as well as reproductive regulation (Schuetz, 1985). This process involves a large number of genes and signal transduction pathways related to epigenetic programming

Abbreviation: bp, Base pair(s); BmCho, *Bombyx mori* chorion; cAMP, Cyclic adenosine 3′,5′-monophosphate; Cdks, Cyclin-dependent kinases; cDNA, DNA complementary to RNA; COG, Clusters of Orthologous Groups; FDR, False discovery rate; FPKM, Fragments Per Kilobase of transcript sequence per Million base pairs sequenced; Gb, Giga base; GO, Gene Ontology; HD, Head; HSPs, Heat shock proteins; KEGG, Kyoto Encyclopedia of Genes and Genomes; MG, Midgut; NE, Nuclear envelope; NPCs, Nuclear pore complexes; Nups, Nucleoporins; OV, Ovary; piRNA, Piwi-interacting RNA; *Pxds*, Peroxidases; Q20 percentages, Percentage of sequences with sequencing error rate lower than 1%; qRT-PCR, Quantitative Real-Time PCR; RABT, Reference Annotation Based Transcript; RNAi, RNA interference; RNA-seq, RNA Sequencing; TE, Testis; TX, Thorax; Vg, Vitellogenin; VgR, Vitellogenin receptor; VMPs, Vitelline membrane proteins; WEGO, Web Gene Ontology Annotation Plot; Δ, Deletion

* Corresponding author at: State Key Laboratory of Ecological Pest Control for Fujian-Taiwan Crops and College of Life Science, Fujian Agriculture and Forestry University, Fuzhou 350002, China.
E-mail address: msyou@iae.fjau.edu.cn (M.-S. You).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.gene.2017.09.020
Received 14 December 2016; Received in revised form 21 July 2017; Accepted 8 September 2017
Available online 13 September 2017

(Skora and Spradling, 2010), germ cell behavior (Dansereau and Lasko, 2008), cell cycle regulation (Bastock and St Johnston, 2008) and developmental patterning mechanisms (Lynch et al., 2010; Wilson et al., 2011). In insects, female reproductive genes, such as chorion, vitelline membrane proteins, Vg and VgR (Sdralia et al., 2012; Chen et al., 2015; Upadhyay et al., 2016), and signal transduction pathways including insulin and hormone pathways (Fox et al., 2011; Jindra et al., 2013) have been studied in *Anopheles gambiae*, *Drosophila melanogaster*, *Nilaparvata lugens*, and *Apis mellifera* (Mack et al., 2006; Rogers et al., 2008; Zhai et al., 2013; Niu et al., 2014). Current results indicated that these genes and pathways play major roles in insect reproduction but greatly vary among different species (Attardo et al., 2005).

The advent of high throughput sequencing technologies has provided an opportunity to comprehensively examine all genes or a subset of functionally active genes from the genome of a specific species or tissue (Wang et al., 2009). These techniques, especially transcriptomic analyses, have become important to better understand the mechanisms behind insect fecundity. For example, transcriptome sequencing of reproductive tissues has been performed in insects, such as accessory glands and testis of *Bactrocera dorsalis* (Wei et al., 2015; Wei et al., 2016) and *A. gambiae* (Dottorini et al., 2013), and the ovary of *N. lugens*, *A. mellifera* and *Venturia canescens* (Leach et al., 2009; Zhai et al., 2013; Niu et al., 2014). These studies have been useful to identify genes related to sexual gland development, spermatogenesis, and oogenesis.

The diamondback moth (DBM), *Plutella xylostella* (L.) (Lepidoptera, Plutellidae), is a Lepidoptera pest that mainly attacks cruciferous plants (Thorsteinson, 1953). It has a very high reproductive capacity allowing it to invade all regions where cruciferous plants, mainly crops, grow (Zalucki et al., 2012; Furlong et al., 2013). The fundamental reproductive biology of DBM has been widely investigated (Peng et al., 2015), but the molecular mechanisms of reproduction remain unclear. It is therefore important to explore genes and major signal pathways involved in reproduction, including ovary development and oogenesis, which is of significance for screening potential molecular targets to better control *P. xylostella*.

In this study, we used deep sequencing to examine *P. xylostella* genes involved in ovary development and oogenesis. More specifically, we aimed to get a comprehensive view of the transcriptional profiles and putative roles of genes referred to as several crucial reproductive processes of embryogenesis, vitellogenesis, and choriogenesis. This global approach provided valuable insights into the molecular mechanisms related to female reproduction of this species.

2. Materials and methods

2.1. P. xylostella rearing and sample preparations

A susceptible *P. xylostella* strain (i.e. not resistant to pesticides) collected from a Fuzhou (province of Fujian, China) cabbage (*Brassica oleracea* var. *capitata*) field (26.08°N, 119.28°E) in 2004 has been maintained since in laboratory at the Fujian Agriculture and Forestry University. The colony has been on 3–5 weeks old radish seedlings (*Raphanus sativus*) at 25 ± 1 °C, 65 ± 5% RH and L:D = 16:8 h without exposure to insecticides.

Based on a preliminary experiment, we found that the *P. xylostella* ovary began to develop at the pupa stage. The organ was fully formed on the second day of pupation and could be dissect as an independent organ. Therefore, for this study, we collected ovaries from 2-day old pupae to the 5-day old virgin female adults at every 24 h interval (n = 30 per age) as a replication. There were three biological replications. Dissections were performed in sterile PBS-DEPC. The dissected materials were immediately soaked in the 1.5 mL centrifuge tube (Axygen) with RNA-later (Qiagen) solution on ice, and subsequently stored at − 80 °C for RNA extraction.

2.2. RNA extraction

Total RNA extraction of individuals at each age was conducted with RNeasy mini kit (Qiagen, Valencia, CA, USA) as described in the manufacturer's manual. RNA contamination and degradation were analyzed using 1% agarose gel electrophoresis. RNA samples were assessed for purity at absorbance ratios of OD260/280 and OD260/230 using the NanoDrop2000® spectrophotometer (Thermo, USA). To prepare the library, we pooled an equal quantity of total RNA from samples of the different ages.

2.3. cDNA library construction and Illumina sequencing

Each sample with an amount of 3 μg RNA was used for the sample preparation. The library was developed using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) according to manufacturer's instructions. Its quality was then assessed on the Agilent Bioanalyzer 2100 system. The sequences of each replicate were marked by the index codes, and the index-coded samples were clustered on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina) following the manufacturer's instructions. After clustering, the libraries were sequenced on an Illumina Hiseq platform to obtain the 125 bp/150 bp paired-end reads (Fig. S1).

2.4. Transcriptome data analysis

2.4.1. Quality control

Raw data (raw reads) in FASTQ format were first modified into clean data (clean reads) through Perl scripts. This was done by (1) filtering out adapter-only reads, (2) removing reads containing > 10% poly-N, and (3) removing low quality reads to ensure PHRED quality scores ≤ 20. The methods used in our experiment were also consistent with those reported elsewhere, such as in Shi et al. (2016) and Huang et al. (2015). Meanwhile, Q20, Q30, and GC contents were calculated. All of the analyses were carried out using the high-quality clean data.

2.5. Mapping and assembly of clean reads

Sequence files used for reference and annotation were downloaded from the diamondback moth Genome Database (DBM-DB; http://iae.fafu.edu.cn/DBM/index.php). Bowtie v2.2.3 software was used to build the reference genome index (Langmead and Salzberg, 2012), and the clean paired-end reads obtained from RNA-seq were aligned to the reference genome with TopHat v2.0.12 (Trapnell et al., 2009). After TopHat alignment, known and novel transcripts were constructed and identified based on Reference Annotation Based Transcript (RABT) assembly method of Cufflinks v2.1.1 (Trapnell et al., 2010; Trapnell et al., 2012).

2.6. Transcripts expression profiling

The calculations of read numbers mapped for each transcript were performed using HTSeq v0.6.1 and the transcript quantification was calculated with FPKM (expected number of Fragments Per Kilobase of transcript sequence per Million base pairs sequenced) (Trapnell et al., 2010). FPKM value '1' was defined as the threshold for the transcript expression. We took the average values of the three biological replicates as the actual expression value for each of the transcripts. Additionally, based on our RNA-seq data, stage-specific expression of the oogenesis genes were profiled.

2.7. GO and COG annotation

The GOseq R package was applied for Gene ontology (GO) annotation of all expressed genes (Young et al., 2010) and the GO terms were categorized by WEGO. For Clusters of Orthologous Groups (COG)