

RESEARCH ARTICLE

Open Access



Genome-wide identification of lncRNAs associated with chlorantraniliprole resistance in diamondback moth *Plutella xylostella* (L.)

Bin Zhu, Manyu Xu, Haiyan Shi, Xiwu Gao and Pei Liang*

Abstract

Background: Long noncoding RNAs (lncRNAs) are now considered important regulatory factors, with a variety of biological functions in many species including insects. Some lncRNAs have the ability to show rapid responses to diverse stimuli or stress factors and are involved in responses to insecticide. However, there are no reports to date on the characterization of lncRNAs associated with chlorantraniliprole resistance in *Plutella xylostella*.

Results: Nine RNA libraries constructed from one susceptible (CHS) and two chlorantraniliprole-resistant *P. xylostella* strains (CHR, ZZ) were sequenced, and 1309 lncRNAs were identified, including 877 intergenic lncRNAs, 190 intronic lncRNAs, 76 anti-sense lncRNAs and 166 sense-overlapping lncRNAs. Of the identified lncRNAs, 1059 were novel. Furthermore, we found that 64 lncRNAs were differentially expressed between CHR and CHS and 83 were differentially expressed between ZZ and CHS, of which 22 were differentially expressed in both CHR and ZZ. Most of the differentially expressed lncRNAs were hypothesized to be associated with chlorantraniliprole resistance in *P. xylostella*. The targets of lncRNAs via cis- (<10 kb upstream and downstream) or trans- (Pearson's correlation, $r > 0.9$ or < -0.9 , $P < 0.05$) regulatory effects were also identified; many of the differently expressed lncRNAs were correlated with various important protein-coding genes involved in insecticide resistance, such as the ryanodine receptor, uridine diphosphate glucuronosyltransferase (UGTs), cytochrome P450, esterase and the ATP-binding cassette transporter.

Conclusions: Our results represent the first global identification of lncRNAs associated with chlorantraniliprole resistance in *P. xylostella*. These results will facilitate future studies of the regulatory mechanisms of lncRNAs in chlorantraniliprole and other insecticide resistance and in other biological processes in *P. xylostella*.

Keywords: lncRNA, *Plutella xylostella*, Chlorantraniliprole, RNA-seq

Background

The diamondback moth (DBM), *Plutella xylostella* (L., Lepidoptera: Plutellidae), is a major insect pest of cruciferous vegetables and is considered an especially troublesome pest because of its ability to rapidly develop high resistance to insecticides used for its control [1]. To date, *P. xylostella* has developed resistance to several types of insecticides and has become one of the most resistant pests in the world [2].

Chlorantraniliprole is a new type of anthranilic diamide insecticide with a novel mode of action that activates the muscle ryanodine receptor (RyR), which controls internal calcium release in the sarcoplasmic reticulum. Activation of RyR causes rapid cessation of feeding, lethargy, muscle paralysis and, finally, insect death [3]. Because of this novel mode of action, chlorantraniliprole is very effective in controlling several orders of insects, especially lepidopteran pests. However, in recent years, *P. xylostella* has developed high levels of resistance to chlorantraniliprole in many countries, including China [4–7].

Previous studies indicate that enhanced activity of detoxification enzymes such as cytochrome P450 monooxygenase

* Correspondence: liangcau@cau.edu.cn
Department of Entomology, China Agricultural University, 2
Yuanmingyuan West Road, Beijing 100193, People's Republic of China



(P450), carboxylesterase (CarE) and glutathione *S*-transferases (GSTs) [8, 9] and point mutation of the target (RyR) [10–12] may be associated with chlorantraniliprole resistance in *P. xylostella*.

By using high-throughput RNA sequencing (RNA-seq) technology, Lin et al. identified 1,215 genes that may be involved in chlorantraniliprole resistance in three field-resistant *P. xylostella* strains, of which several genes were associated with calcium signaling, vascular smooth muscle contraction and cardiac muscle contraction pathways, as well as in the metabolism of xenochemicals such as insecticides [13].

Several studies have investigated mechanisms of chlorantraniliprole resistance in the past few years and many protein-coding genes have been proven to be involved in chlorantraniliprole resistance. However, research on regulatory mechanisms of these protein-coding genes remains very limited.

Most recently, microRNAs (miRNAs) have been associated with chlorantraniliprole resistance in *P. xylostella* [14]. MiRNA is a kind of endogenous small non-coding RNA (ncRNA), which regulates the expression of target genes at the transcriptional level; it has gained significant interest and popularity over the last decade [15]. Currently, another type of ncRNA, long non-coding RNA (lncRNA), has gained significant attention from researchers. Previous studies indicate that lncRNAs could show quick response to diverse stimuli or stress factors and might be involved in responses to insecticides [16, 17], so we hypothesized lncRNAs may also be associated with chlorantraniliprole resistance in *P. xylostella*.

lncRNAs are non-protein coding transcripts longer than 200 nucleotides. They were once considered inconsequential transcriptional noise. However, recent studies have shown that lncRNAs play important regulatory roles in many biological processes, including transcriptional regulation, post-transcriptional control and epigenetic processes [18, 19]. According to the position and direction of transcription in relation to protein-coding genes, lncRNAs can be further classified into several categories, such as sense, antisense, intronic and intergenic [20]. Like mRNAs, many identified lncRNAs are transcribed by RNA polymerase II, hence they are presumably capped, polyadenylated and spliced. In addition, there are also a few non-polyadenylated lncRNAs transcribed by RNA polymerase III [21]. Most lncRNAs are located only in the nucleus, but some are cytoplasmic or are in both the nucleus and cytoplasm [22].

Currently, RNA-sequencing (RNA-seq) is a very powerful approach to identify lncRNAs. In the present study, a laboratory susceptible *P. xylostella* strain and two chlorantraniliprole-resistant strains were used, and nine strand-specific RNA-seq libraries that combine rRNA removal were constructed. Four types of lncRNAs

were obtained, and the relative expression of some were found to be significantly altered in chlorantraniliprole-resistant populations. These results lay a solid foundation for further study of the roles of lncRNAs in regulation of insecticide resistance in *P. xylostella*.

Results

Identification and characterization of lncRNAs in *P. xylostella*

High-throughput strand-specific RNA-seq was performed in the three DBM strains (CHS, CHR, ZZ), each with three biological replicates. A total of 1,198,903,526 raw reads were obtained from the nine libraries, with an average of 133 million reads per sample. After the low-quality reads were removed, 1,110,303,222 clean reads with high quality were retained (Additional file 1). Clean reads that could be mapped to the *P. xylostella* genome (GCA_000330985.1) were then used for transcript assembly and annotation. A total of 68,118 transcripts corresponding to 43,041 loci were initially generated. Then, we filtered protein-coding transcripts according to the annotated DBM reference genome (known DBM mRNAs) and transcripts with a single-exon and those that were shorter than 200 nucleotides were removed. The remaining 3,015 transcripts (corresponding to 2,289 loci) were subsequently used for protein-coding capacity prediction by using the Coding Potential Calculator (CPC) [23], Coding-Non-Coding Index (CNCI) [24], Pfam-scan (PFAM) [25] and PLEK [26] (Fig. 1b). Finally, 1,309 reliably expressed lncRNAs corresponding to 1,096 loci were obtained and were classified into four categories including 'u' (intergenic), 'i' (intronic), 'x' (anti-sense) and 'o' (sense-overlapping) according to their genomic location and referring to the neighboring genes. Specifically, the 'u' category contained transcripts falling in the intergenic regions between two protein-coding loci. The 'i' category contained transcripts falling entirely within an intron of a known protein coding gene. The 'x' category contained transcripts that have generic exonic overlap with a known protein coding gene on the opposite strand. The 'o' category contained the transcripts partial overlapping with a coding gene on the same genomic strand (Fig. 1a, b).

Most of the identified lncRNAs fell into class u, with 877 lncRNAs (67.00%), whereas 190 (14.51%), 76 (5.81%) and 166 (12.68%) lncRNAs belonged to classes i, x and o, respectively (Fig. 1c). All lncRNA sequences are listed in Additional file 2.

Previous studies in mammals have shown that the expression of lncRNAs is significantly lower than those of protein-encoding genes [27]. To determine whether *P. xylostella* lncRNAs have similar features, we measured the expression level (fragments per kilobase of exon per million fragments mapped, FPKM) of the identified lncRNAs; they generally showed a lower level of expression compared to protein-coding mRNAs (Fig. 2a).

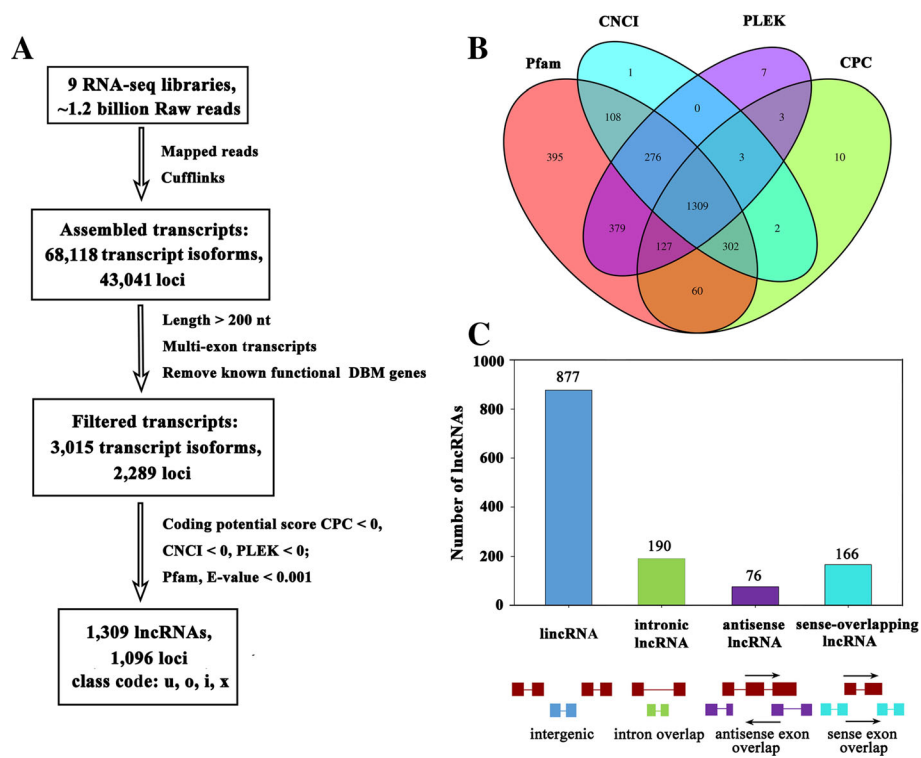


Fig. 1 The computational pipeline for identifying lncRNAs in *P. xylostella* from RNA-seq data and their classification. **a** The lncRNA identification pipeline flowchart; **b** Coding potential analysis using the four methods; **c** The classification of identified lncRNAs, red rectangles or lines represent the exon or intron of protein-coding gene, respectively; Blue, green, purple and light blue rectangles or lines represent the exon or intron of lncRNA, respectively

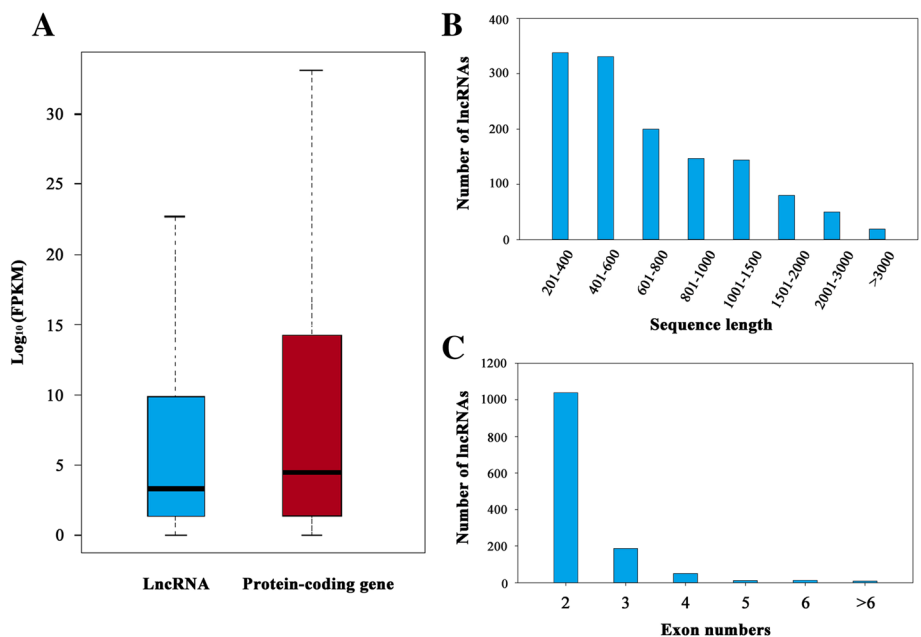


Fig. 2 Characterization of *P. xylostella* lncRNAs. **a** Comparison of expression value (FPKM) in *P. xylostella* lncRNAs and protein coding genes; For the box-plot: midline, median; box limits, 25th percentile (first quartile) and 75th percentile (third quartile); upper whisker, min (max(x)), third quartile + 1.5x interquartile range (IQR; third-quartile minus first-quartile values); lower whisker, max(min(x)), first quartile - 1.5x IQR; **b** Size distribution of *P. xylostella* lncRNAs; **c** The distribution of exon number of lncRNAs

The length and exon number of the identified lncRNAs were also analyzed. The size distribution of these lncRNAs ranged from 200 nucleotides to 7,193 nucleotides, with approximately 78% of lncRNAs shorter than 1000 nucleotides (Fig. 2b). Characterization of the genomic location revealed that the exon number of these lncRNAs ranged from 2 to 13; 1,038 (79.30%) *P. xylostella* lncRNAs had two exons, 186 (14.21%) had three exons, and 22 (1.68%) lncRNAs had more than five exons (Fig. 2c).

Analysis of differentially expressed lncRNAs

To systematically identify chlorantraniliprole resistance-associated lncRNAs, a differential expression analysis was performed among the three strains. In total, 64 lncRNAs (45 lincRNAs, 13 sense-overlapping lncRNAs and 6 intronic lncRNAs) were identified as differentially expressed between CHR and CHS ($P < 0.05$, \log_2 (fold change) > 1), of which 34 were down-regulated and 30 were up-regulated in the CHR strain (Additional file 3, Fig. 3a, c, e). Interestingly, among these differentially expressed lncRNAs, we found 7 lncRNAs that were specifically expressed in CHS and 5 lncRNAs that were specifically expressed in CHR (Additional file 3).

In addition, 83 lncRNAs (57 lincRNAs, 12 sense-overlapping lncRNAs and 14 intronic lncRNAs) were differentially expressed between ZZ and CHS ($P < 0.05$, \log_2 (fold change) > 1), of which 34 were down-regulated and 49 were up-regulated in ZZ (Additional file 3, Fig. 3b, d, f). Among these differentially expressed lncRNAs, 8 lncRNAs were found to be specifically expressed in ZZ and one lncRNA was specifically expressed in CHS (Additional file 3).

Compared to CHS, 22 lncRNAs (15 lincRNAs, 5 sense-overlapping lncRNAs and 2 intronic lncRNAs) were found to be differentially expressed in both CHR and ZZ, of which 9 were down-regulated and 13 were up-regulated in both resistant strains (Fig. 4, additional file 3). Among these lncRNAs, 4 lncRNAs were specifically expressed in both resistant strains (Additional file 3).

To validate the RNA-seq data, three lncRNAs that were differentially expressed in both CHR and ZZ compared to CHS (TCONS_00000650, TCONS_00041352, TCONS_00056025), three lncRNAs that were differentially expressed only between CHS and CHR (TCONS_00000795, TCONS_00001205, TCONS_00033373), and three lncRNAs that were differentially expressed only between CHS and ZZ (TCONS_00011964, TCONS_00017539, TCONS_00017842) were randomly selected and their relative expression levels were quantified by qRT-PCR. The expression patterns of almost all selected lncRNAs showed a similar trend between the results of sequencing and qRT-PCR except for TCONS_00033373, which was significantly up-regulated in both the CHR and ZZ strains by qRT-PCR (Fig. 5). However, according to

the sequencing results, this lncRNA was significantly up-regulated only in the CHR strain. Pearson correlation coefficient between RNA-Seq data and qRT-PCR data was 0.970, which indicates that the RNA-Seq data was highly correlated with that of the qRT-PCR (Fig. 6).

Functional analysis of resistance-associated lncRNAs

Previous studies showed that lncRNAs may play a cis-regulatory role in mediating the expression of neighboring genes [28]. We searched for protein-coding genes 10 kb upstream or downstream of the differentially expressed lncRNAs. 138 protein-coding genes (including 19 overlapped protein-coding genes) were found close to 64 differentially expressed lncRNAs between CHR and CHS, and 177 protein-coding genes (including 26 overlapped protein-coding genes) were found close to 83 differentially expressed lncRNAs between ZZ and CHS (Additional file 4). Among these neighboring protein-coding genes, only the ATP-binding cassette subfamily G member 1 (ABCG1) gene overlapped with TCONS_00022357 (down-regulated in ZZ) was previously associated with insecticide resistance [29] (Additional file 4). When Gene Ontology (GO) analysis of the neighboring genes was performed, we found that most of the differentially expressed lncRNAs between CHS and CHR or between CHS and ZZ may play a role in binding-associated functions in cis mode, because most of the neighboring genes were annotated as binding-related GO terms, such as metal ion binding, ATP binding, DNA binding. Meanwhile, neighboring genes were also enriched in transcription in the Biological Process category (BP) and the nucleus, integral component of membrane, cytoplasm in Cellular Component (CC) categories (Additional file 4).

Many studies have also shown that lncRNAs can function as trans-regulatory elements. Thus, potential targets of the differentially expressed lncRNAs in trans-regulatory relationships were predicted using a co-expression analysis. Pearson's correlation test ($r > 0.9$ or < -0.9 , P -value < 0.05) was used in this study. For differentially expressed lncRNAs and protein-coding genes between CHS and CHR, a total of 802 interaction relationships (580 positive and 222 negative correlations) were detected (Additional file 5). Similarly, for differentially expressed lncRNAs and protein-coding genes between CHS and ZZ, a total of 2420 interaction relationships (1,823 positive and 597 negative correlations) were obtained (Additional file 5), which was approximately 3-fold of that between CHS and CHR. This huge difference may be caused by the obvious differences between CHR and ZZ. The former was a laboratory resistant strain selected from CHS with only chlorantraniliprole, while the later came from the field with quite different genetic background and was exposed to many different insecticides over a long time period.

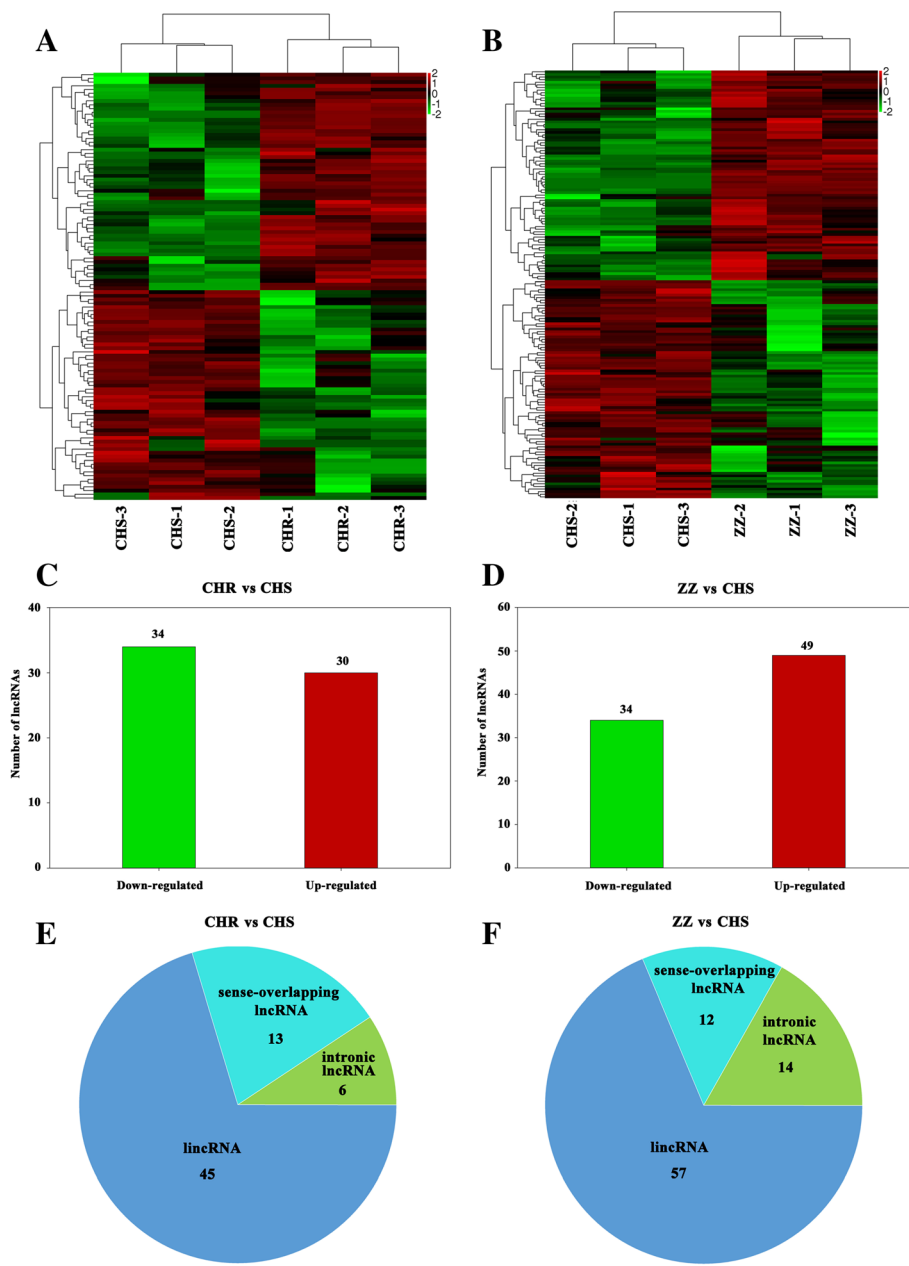
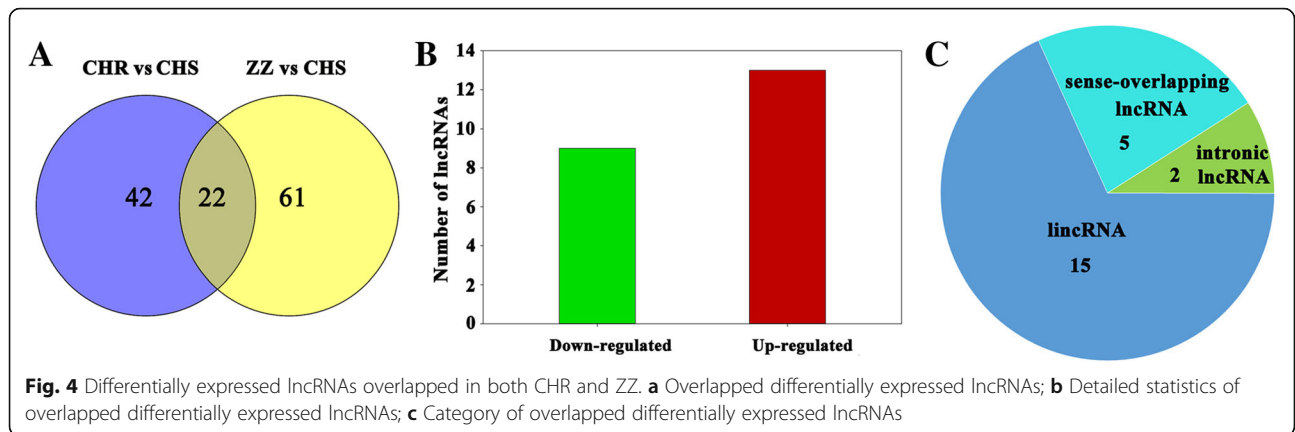


Fig. 3 Differentially expressed lncRNAs identified among CHS, CHR and ZZ. **a** Hierarchical clustering of the differentially expressed lncRNAs between CHS and CHR; **b** Hierarchical clustering of the differentially expressed lncRNAs between CHS and ZZ; **c** Detailed statistics of differentially expressed lncRNAs between CHS and CHR; **d** Detailed statistics of differentially expressed lncRNAs between CHS and ZZ; **e** Category of differentially expressed lncRNAs between CHS and CHR; **f** Category of differentially expressed lncRNAs between CHS and ZZ

Interestingly, 2 lncRNAs, TCONS_00013329 and TCONS_00056155, were found to be co-expressed with the ryanodine receptor (the target of chlorantraniliprole; TCONS_00013329 was significantly up-regulated in both resistant strains and TCONS_00056155 was significantly down-regulated only between CHS and CHR. Many lncRNAs are also co-expressed with various protein coding genes involved in insecticide resistance, such as

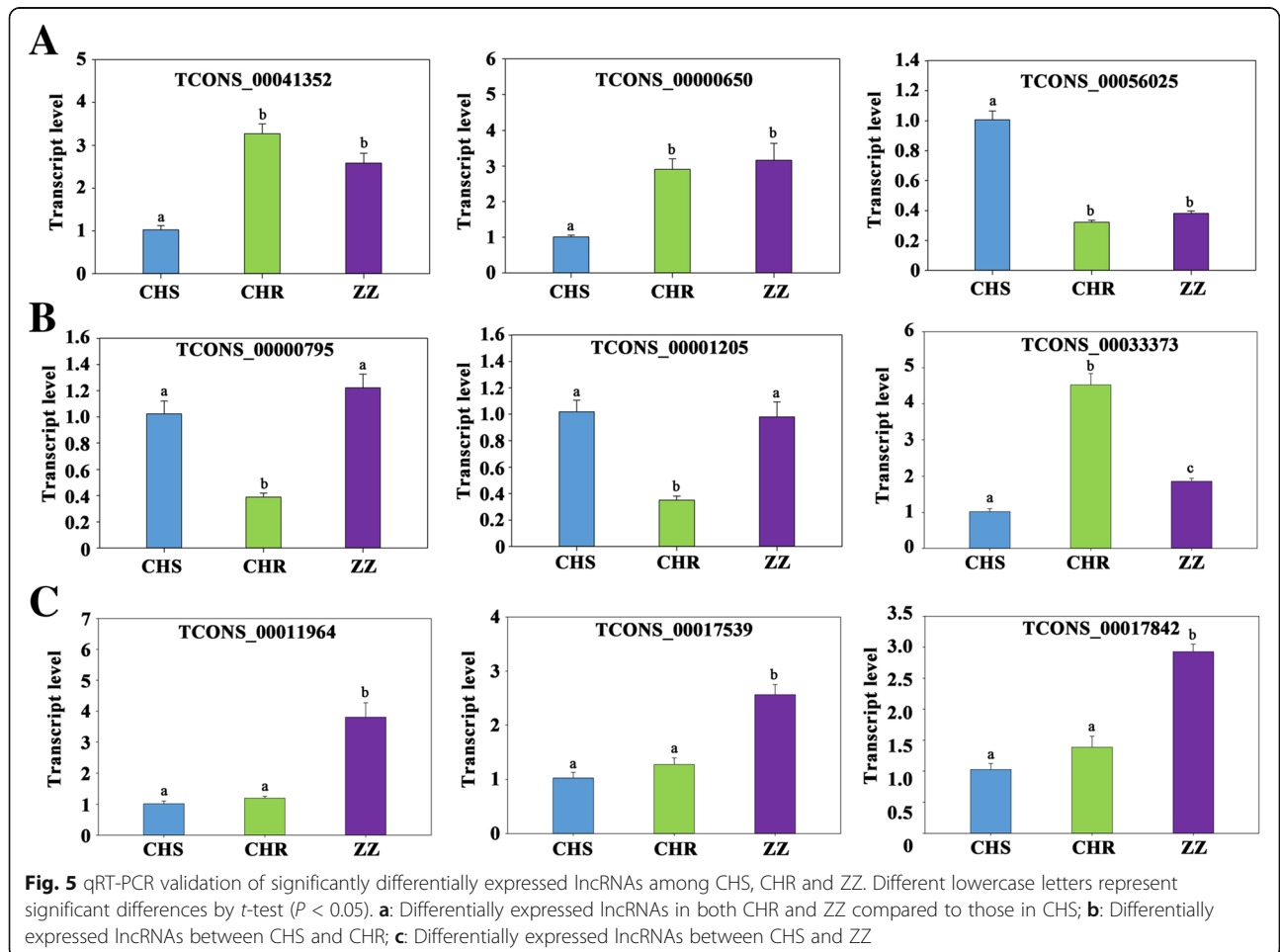
UDP-glucuronosyltransferase (UGTs), cytochrome P450, esterase, glutathione S-transferase (GSTs), ATP-binding cassette transporter (ABC), heat shock protein (HSP) and cuticle protein (Additional file 5). GO enrichment analysis based on the correlated target protein-coding genes was also performed. All correlated protein-coding genes of the differently expressed lncRNAs in the two comparison groups were enriched in similar GO terms, such as metal

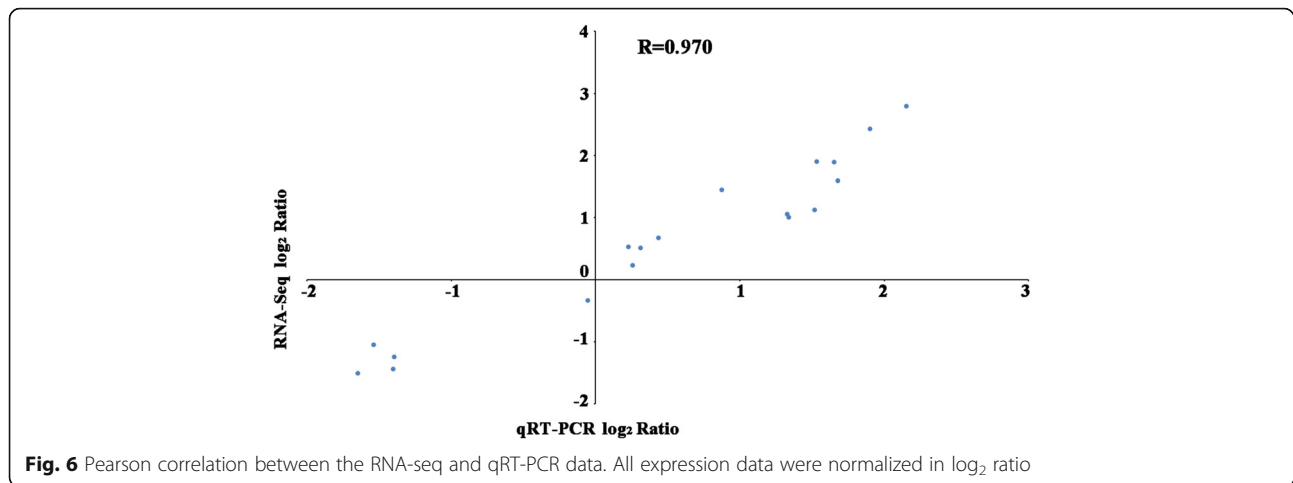


ion binding, ATP binding, DNA binding, zinc ion binding, serine-type endopeptidase activity and RNA binding in Molecular Function (MF); transcription, regulation of transcription, DNA integration, innate immune response and oxidation-reduction process in Biological Process (BP); and nucleus, cytoplasm, integral component of membrane in Cellular Component (CC) (Additional file 5).

Discussion

In recent years, the vast majority of studies on lincRNA have been conducted in mammals, especially in humans. Studies of insect lincRNAs are still in preliminary stages. With the rapid development of high-throughput techniques, a batch of lincRNAs has been identified in several insect species, such as *Drosophila melanogaster* [30, 31],





Apis mellifera [32], *Apis cerana* [33], *Anopheles gambiae* [34], *Aedes aegypti* [35], *Nilaparvata lugens* [36], *Bombyx mori* [37] and *P. xylostella* [16], of which only *N. lugens* and *P. xylostella* are agricultural pests.

The publication of the *P. xylostella* genome in 2013 [38] benefits the identification of lncRNAs considerably. In a recent annotation of the DBM genome in NCBI, 707 transcripts were annotated as ncRNAs. In addition, Etebari et al. [16] also identified 3844 DBM long intergenic non-coding RNAs (lncRNAs) from RNA deep-sequencing data downloaded from the NCBI Sequences Read Archive.

In the current study, a total of 1,309 lncRNAs belonging to four types were identified from 9 strand-specific RNA-seq libraries, of which 77 could be blast-searched in the NCBI database. These were annotated as DBM ncRNAs and shared a similar locus, including 67 lincRNAs, 5 intronic lncRNAs, 1 sense-overlapping lncRNA and 4 antisense lncRNAs (Additional file 2). In addition, 190 lncRNAs partially overlapped (shared similar loci in the same scaffold of the DBM genome) with the lincRNAs identified by Etebari et al. [16] according to the transcript location supplied in their research, but because of differences in the library building method and the RNA-seq data used in our research, some of the overlapped lncRNAs were re-classified to other lncRNA types rather than lincRNA. Specifically, 4 of these overlapped lncRNAs were re-classified as antisense lncRNAs, including TCONS_00003280 (overlapped with lincRNA_210, a lincRNA named by Etebari et al. [16]), TCONS_00006143 (overlapped with lincRNA_379), TCONS_00008192 (overlapped with lincRNA_495) and TCONS_00030950 (overlapped with lincRNA_1763); 18 were re-classified as intronic lncRNAs, such as TCONS_00018632 (overlapped with lincRNA_1075) and TCONS_00019803 (overlapped with lincRNA_1140); and 8 were re-classified as sense-overlapping lncRNAs, such as TCONS_00010817 (overlapped with lincRNA_658) and TCONS_0001

1269 (overlapped with lincRNA_673) (Additional file 2). Moreover, there were 17 overlapping lncRNAs between the 190 lincRNAs identified by Etebari et al. [16] and the 77 lncRNAs annotated in the NCBI (Additional file 2). Therefore, 1,059 novel lncRNAs were identified in this research. The number of putative lncRNAs detected in this study was less than that reported by Etebari et al. [16], mainly because all the transcripts that contained only one exon were retained in their research, but only multiple exon transcripts were used for lncRNA identification in the present study.

In addition, to our knowledge, this is the first application of rRNA removal and strand-specific RNA sequencing to study the DBM transcriptome. This method allows non-polyA transcripts to be obtained, which is an obvious advantage compared to polyA enrichment sequencing [39]. Strand information was also included in our sequencing data, which made it easy to distinguish sense transcripts from antisense transcripts. As a result, anti-sense lncRNAs were identified in *P. xylostella* for the first time. Different types of lncRNAs may play their roles in different way, so a detailed classification of lncRNAs would help us to further understand their various functions [20].

Differently expressed lncRNAs among CHS, CHR and ZZ were analyzed in the present study, most associated with chlorantraniliprole or other insecticide resistance. The CHR strain was established from CHS by successive selection with chlorantraniliprole and has been reared under the same laboratory conditions as CHS; all 64 differentially expressed lncRNAs between them are likely to be associated with chlorantraniliprole resistance. The ZZ strain was collected from the field and had developed middle to high levels of resistance to several commonly used insecticides besides chlorantraniliprole, such as beta-cypermethrin, abamectin, spinosad and indoxacarb (unpublished data from a local plant protection station). Each of these insecticides has a distinctive mode of action. Therefore, the 83 differentially expressed lncRNAs in the

ZZ strain likely result from the comprehensive effects of these different insecticides as well as from other environmental factors.

When the differentially expressed lncRNAs in CHR and ZZ strains were put together, 22 of them overlapped. These common differentially expressed lncRNAs are very likely involved in chlorantraniliprole resistance because they were differentially expressed in both laboratory-selected and field-collected resistant strains, and both of these strains were resistant to chlorantraniliprole. However, due to the complexity of insecticide resistance mechanisms in the ZZ strain, these 22 lncRNAs may also reveal resistance mechanisms common to other insecticides besides chlorantraniliprole. Interestingly, four of these 22 lncRNAs were specifically expressed in the two resistant strains. We speculated that their transcription might be induced by long-term exposure to chlorantraniliprole, and these four lncRNAs may play key roles in chlorantraniliprole resistance in *P. xylostella*. The other 42 lncRNAs differentially expressed in the CHR strain may also be involved in chlorantraniliprole resistance. Though this hypothesis was not supported by data from the field resistant strain, these lncRNAs should not be ignored in the further study of the mechanisms of chlorantraniliprole resistance in *P. xylostella*. We suspected that most of these unique differentially expressed lncRNAs may play specific roles in resistance only to chlorantraniliprole. Meanwhile, the lncRNAs that were differentially expressed only in the ZZ strain are more likely to be involved in resistance to other insecticides besides chlorantraniliprole.

Notably, several of their overlapping lincRNAs for the differently expressed lncRNAs identified in the present study have been found to be involved in insecticide resistance in *P. xylostella* by Etebari et al. [16]. For example, lincRNA_2514 overlapped with TCONS_00044413 is involved in chlorpyrifos, fipronil and Bt resistance; lincRNA_1623 overlapped with TCONS_00028420 in Bt and fipronil resistance; lincRNA_494 overlapped with TCONS_00008143 in Bt resistance; and lincRNA_1624 overlapped with TCONS_00028420 in chlorpyrifos resistance, respectively [16]. Our finding of the involvement of these lncRNAs in chlorantraniliprole resistance increases the possibility that these lncRNAs play some important roles in insecticide resistance regulation.

To further study the roles of lncRNAs possibly associated with chlorantraniliprole resistance, we predicted the potential function of the differently expressed lncRNAs using cis and trans methods. In the cis prediction, numerous protein-coding genes were found within 10 kb upstream or downstream from the concerned lncRNAs, most of which may play a role in binding-associated activity. In the trans prediction, potential targets of the differently expressed lncRNAs were predicted using co-expression analysis and many target protein-coding genes involved in insecticide resistance were

identified, indicating that lncRNAs may regulate insecticide resistance by directly affecting these target genes. For example, XM_011567276, annotated as cytochrome P450 6B6, was co-expressed with 11 lncRNAs (TCONS_00044883, TCONS_00026933, TCONS_00019595, TCONS_00032346, TCONS_00065690, TCONS_00019598, TCONS_00008336, TCONS_00037191, TCONS_00007659, TCONS_00065766 and TCONS_00052631). Previous studies showed that one lncRNA could regulate multiple protein-coding genes and vice versa [40]. Here, these 11 lncRNAs may collectively regulate the expression of cytochrome P450 6B6, thus enhancing the metabolism of insecticides in *P. xylostella*. In fact, Peng et al. [41] have identified a set of lncRNAs highly correlated with the expression of P450 in mouse liver during maturation. Interestingly, the ryanodine receptor, the main target of chlorantraniliprole, was also found to be co-expressed with two lncRNAs (TCONS_00013329 and TCONS_00056155), and these lncRNAs may directly control the expression of the ryanodine receptor to mediate chlorantraniliprole resistance. In addition to this, several binding terms were identified as enriched GO terms for the target mRNAs in both comparison groups. lncRNAs play important roles in regulating biological functions through various mechanisms that are not fully understood; these proposed mechanisms include regulation based on RNA-protein interactions as well as RNA-RNA interactions and RNA-DNA interactions [42]. Here, binding terms were identified as enriched GO terms for the correlated mRNAs in both comparison groups, and it is very likely that lncRNAs may act primarily through these interactions.

Conclusions

In the current study, 1,309 lncRNAs were identified from 9 RNA-seq libraries of *Plutella xylostella*, including 877 intergenic lncRNAs, 190 intronic lncRNAs, 76 antisense lncRNAs and 166 sense-overlapping lncRNAs. In addition, several lncRNAs showed significant expression changes in the two chlorantraniliprole-resistant strains; some were identified as co-expressed with several genes involved in insecticide resistance, especially the ryanodine receptor, the target of chlorantraniliprole. These results provide solid bases for further investigation of the roles of lncRNAs in regulation of chlorantraniliprole and other insecticide resistance and in other biological processes in *P. xylostella*.

Methods

Insects

The susceptible DBM strain (CHS) was collected in the vegetable fields of Beijing and maintained in our laboratory without any insecticide treatments for more than 10 years. The chlorantraniliprole-resistant strain (CHR) was derived from the CHS strain by uninterrupted selection

with chlorantraniliprole for more than 70 generations. The Zhangzhou strain (ZZ) was collected in the vegetable fields of Zhangzhou, Fujian province, southeastern China in 2015; before sequencing, the ZZ strain was selected with chlorantraniliprole for two generations in our laboratory. Moreover, the toxicity of chlorantraniliprole to the CHS, CHR and ZZ populations was tested using a leaf dipping method as described elsewhere [43]; the CHR and the ZZ strains showed 65-fold and 42-fold resistance to chlorantraniliprole, respectively, compared to the susceptible CHS strain [44]. All stages of *P. xylostella* were maintained at 27 ± 1 °C, with an RH of 40–60% for radish seedlings (*Raphanus sativus* L.) and a photoperiod of 16:8 h (L:D). *P. xylostella* adults were provided with 10% (W/V) honey solution and were allowed to lay eggs on radish seedlings.

RNA extraction, library preparation and sequencing

Thirty 3rd instar *P. xylostella* larvae were collected in a PE tube as one sample. Trizol Reagent (Invitrogen, Carlsbad, CA, USA) was used to isolate total RNA according to the manufacturer's instructions. RNA degradation and contamination were assessed on 1% agarose gels and RNA concentration was measured using a NanoDrop 2000 (Thermo Fisher Scientific Inc, USA).

Library construction and RNA-Seq were performed by the OE Biotechnology Corporation (Shanghai, China). Total RNA from 9 samples (three independent biological replicates for each of the CHS, CHR and ZZ strains) with RNA integrity number (RIN) values above 8 were used to construct RNA-Seq libraries using the TruSeq stranded total RNA preparation kit with Ribo-Zero Gold (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. Sequencing was performed on the Illumina HiSeq™2500 and 150 bp paired-end reads were generated.

Bioinformatics analysis

Raw data in FASTQ format were first processed using the NGS QC Toolkit [45]. In this step, clean data (clean reads) were obtained by removing reads containing adapters, reads containing poly-N, low quality reads (lower than 20) and contaminants from the raw data. At the same time, the Q20, Q30 and GC content of the clean data were calculated. All the downstream analyses were based on clean data with high quality.

The latest reference genome and gene model annotation files of *P. xylostella* (GCA_000330985.1) were downloaded from the NCBI FTP site (ftp://ftp.ncbi.nlm.nih.gov/genomes/Plutella_xylostella/). The clean reads from each library were first aligned to the DBM genome using TopHat [46], then the mapped reads were assembled using Cufflinks in a reference-based approach [47].

Identification of lncRNAs

The assembled transcripts were annotated using the Cuffcompare program from the Cufflinks package [46]. According to the annotations of the DBM genome sequence, the known protein-coding transcripts as well as the rRNA, tRNA, snRNA, snoRNA, pre-miRNA and pseudogenes were first removed. Meanwhile, transcripts with single exons and those that were shorter than 200 bps were also excluded from further non-coding analysis. The coding potential for the remaining transcripts was calculated by using CPC [23], CNCI [24], Pfam [25] and PLEK [26]. Transcripts revealing coding potential with a CPC score > 0, CNCI score > 0, PLEK_score > 0 and Pfam-scan > 0.001 were all removed. The identified lncRNAs were finally separated into four types using the class code module in Cuffcompare [46].

Differential expression analysis

The number of reads mapped to each lncRNA and protein-coding transcript was determined using HT-Seq software (<http://www-huber.embl.de/users/anders/HTSeq/doc/index.html>) [48]. The expression level of each transcript was measured by FPKM. Differential expression analysis was performed using the DESeq R package [49]. The DESeq package implements the negative binomial model to compute differentially expressed transcripts. *P*-value < 0.05 and $|\log_2(\text{fold change})| > 1$ were considered as significantly differential expression. Hierarchical Clustering was performed using the Agilent GeneSpring GX software (version 11.5.1).

Quantitative real-time PCR (qRT-PCR)

Quantitative real-time PCR was performed to experimentally validate the relative expression levels of the identified lncRNAs. Total RNA from the same samples used for deep sequencing were used for the first-strand cDNA synthesis using PrimeScript™ RT reagent Kit with gDNA Eraser (Perfect Real Time) (Takara Biotechnology, Dalian, China) per the manufacturer's instructions. qRT-PCR analysis was carried out using SYBR Premix Ex Taq (Takara Biotechnology, Dalian, China). Each reaction was performed on an ABI 7500 Real Time PCR system (Applied Biosystems) with three biological replicates. The relative expression levels of lncRNAs and protein coding genes were calculated using the $2^{-\Delta\Delta C_t}$ method [50]. Ribosomal protein L32 mRNA was used as a reference gene. Pearson correlation coefficient between qRT-PCR data and RNA-Seq data was calculated to validate RNA-Seq experiments. All primers used in this study are listed in Additional file 6.

Functional analysis of resistance-associated lncRNAs

None of the DBM lncRNAs are functionally annotated in current databases. Thus, we estimated their functions

based on the functional annotations of their related protein-coding genes. In cis, we searched for all the protein-coding genes 10 kb upstream or downstream of the differently expressed lncRNAs. In trans, we used the expression levels of the differently expressed lncRNAs and protein-coding genes to analyze their co-expression relationships. Pearson correlation with P -value < 0.05 and Pearson's correlation coefficients ($r > 0.9$ or < -0.9) were considered as correlated expression. All identified neighboring genes and co-expressed genes were used for the GO enrichment analysis, respectively. A $P < 0.05$ was considered statistically significant. The GO analysis was divided into Molecular Function, Biological Process and Cellular Component. The results allowed us to predict the functional classification in which the target genes of the differentially expressed lncRNAs were enriched.

Additional files

- Additional file 1: Table S1.** Summary of RNA-seq data. (DOCX 13 kb)
- Additional file 2: Table S2.** lncRNAs identified in this study. (XLSX 395 kb)
- Additional file 3: Table S3.** Differentially expressed lncRNAs between CHS and CHR as well as CHS and ZZ. (XLSX 26 kb)
- Additional file 4: Table S4.** Protein-coding genes 10 kb upstream or downstream of the differently expressed lncRNAs. (XLSX 45 kb)
- Additional file 5: Table S5.** Protein-coding genes co-expressed with the differently expressed lncRNAs. (XLSX 256 kb)
- Additional file 6: Table S6.** All primers used in this study. (XLSX 9 kb)

Abbreviations

ABC: ATP-binding cassette transporter; CarE: Carboxylesterase; FPKM: Fragments per kilobase of exon per million fragments mapped; GSTs: Glutathione S-transferases; HSP: Heat shock protein; lncRNA: Long noncoding RNA; P450: Cytochrome P450 monooxygenase; qRT-PCR: Quantitative real-time reverse transcription polymerase chain reaction; RNA-seq: RNA sequencing; RyR: Ryanodine receptor; UGTs: Uridine diphosphate glucuronosyltransferase

Acknowledgments

We thank all contributors of the present study. We also thank Dr. Zhenyu Li at the Institute of Plant Protection, Guangdong Academy of Agricultural Sciences (Guangzhou, China) for collection of field populations of *Plutella xylostella*.

Funding

This work was supported by the National Natural Science Foundation of China (31572023), and the funding body has no roles in the design of the study or collection, analysis, and interpretation of data or in writing the manuscript.

Availability of data and materials

The raw data for this article have been deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive under SRX2458890, SRX2458891, SRX2458892, SRX2458893, SRX2458894, SRX2458919, SRX2458920, SRX2458921 and SRX2458922.

Authors' contributions

PL and BZ conceived of and designed the experiments. BZ performed the experiments and data analysis. MYX analyzed the data. HYS conducted the qPCR validation. PL and XWG contributed reagents/materials and provided the experimental environment. BZ and PL wrote the manuscript. All authors have read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

The radish seeds used in this study were purchased from the institute of vegetables and flowers, Chinese academy of agricultural sciences. The CHS strain of *Plutella xylostella* used in this study were initially collected in Beijing in 2000, and have been maintained in our laboratory at the China Agricultural University since then. The CHR strain were derived from the CHS by continuous selection with chlorantraniliprole. No specific permit was required for the field collection of ZZ strain and the location is not privately-owned or protected in any way. The species in the genus of *Plutella* are common agricultural pests and are not included in the "List of Endangered and Protected Animals in China".

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 11 January 2017 Accepted: 2 May 2017

Published online: 15 May 2017

References

- Zalucki MP, Shabbir A, Silva R, Adamson D, Shu-Sheng L, Furlong MJ. Estimating the economic cost of one of the world's major insect pests, *Plutella xylostella* (Lepidoptera: Plutellidae): just how long is a piece of string? *J Econ Entomol.* 2012;105:1115–29.
- Furlong MJ, Wright DJ, Dosdall LM. Diamondback moth ecology and management: problems, progress, and prospects. *Annu Rev Entomol.* 2013;58:517–41.
- Lahm GP, Cordova D, Barry JD. New and selective ryanodine receptor activators for insect control. *Bioorg Med Chem.* 2009;17:4127–33.
- Li ZY, Feng X, Liu SS, You MS, Furlong MJ. Biology, ecology, and management of the diamondback moth in China. *Annu Rev Entomol.* 2016;61:277–96.
- Wang XL, Wu YD. High levels of resistance to chlorantraniliprole evolved in field populations of *Plutella xylostella*. *J Econ Entomol.* 2012;105:1019–23.
- Wang XL, Khakame SK, Ye C, Yang YH, Wu YD. Characterisation of field-evolved resistance to chlorantraniliprole in the diamondback moth, *Plutella xylostella*, from China. *Pest Manag Sci.* 2013;69:661–5.
- Ribeiro LM, Wanderley-Teixeira V, Ferreira HN, Teixeira AA, Siqueira HA. Fitness costs associated with field-evolved resistance to chlorantraniliprole in *Plutella xylostella* (Lepidoptera: Plutellidae). *B Entomol Res.* 2014;104:88–96.
- Liu X, Wang HY, Ning YB, Qiao K, Wang KY. Resistance selection and characterization of chlorantraniliprole resistance in *Plutella xylostella* (Lepidoptera: Plutellidae). *J Econ Entomol.* 2015;108:1978–85.
- Jiang WH, Lu WP, Guo WC, Xia ZH, Fu WJ, Li GQ. Chlorantraniliprole susceptibility in *Leptinotarsa decemlineata* in the North Xinjiang Uygur Autonomous Region in China. *J Econ Entomol.* 2012;105:549–54.
- Trocza B, Zimmer CT, Elias J, Schorn C, Bass C, Davies TG, et al. Resistance to diamide insecticides in diamondback moth, *Plutella xylostella* (Lepidoptera: Plutellidae) is associated with a mutation in the membrane-spanning domain of the ryanodine receptor. *Insect Biochem Molec.* 2012;42: 873–80.
- Trocza BJ, Williams AJ, Williamson MS, Field LM, Lüemmen P, Davies TG. Stable expression and functional characterisation of the diamondback moth ryanodine receptor G4946E variant conferring resistance to diamide insecticides. *Sci Rep.* 2015;5:14680.
- Guo L, Liang P, Zhou X, Gao X. Novel mutations and mutation combinations of ryanodine receptor in a chlorantraniliprole resistant population of *Plutella xylostella* (L). *Sci Rep.* 2014;4:6924.
- Lin Q, Jin F, Hu Z, Chen H, Yin F, Li Z, et al. Transcriptome analysis of chlorantraniliprole resistance development in the diamondback moth *Plutella xylostella*. *PLoS One.* 2013;8, e72314.
- Li X, Guo L, Zhou X, Gao X, Liang P. miRNAs regulated overexpression of ryanodine receptor is involved in chlorantraniliprole resistance in *Plutella xylostella* (L). *Sci Rep.* 2015;5:14095.
- Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell.* 2004;116:281–97.
- Etebari K, Furlong MJ, Asgari S. Genome wide discovery of long intergenic non-coding RNAs in diamondback moth (*Plutella xylostella*) and their expression in insecticide resistant strains. *Sci Rep.* 2014;5:14624.

17. Di C, Yuan JP, Wu Y, Li JR, Lin HX, Hu L, et al. Characterization of stress-responsive lincRNAs in *Arabidopsis thaliana* by integrating expression, epigenetic and structural features. *The Plant J.* 2014;80(5):848–61.
18. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. *Nat Rev Genet.* 2009;10:155–9.
19. Yoon JH, Abdelmohsen K, Gorospe M. Post-transcriptional gene regulation by long noncoding RNA. *J Mol Biol.* 2013;425:3723–30.
20. Ma L, Bajic VB, Zhang Z. On the classification of long non-coding RNAs. *RNA Biol.* 2013;10(6):924–33.
21. Gibb EA, Brown CJ, Lam WL. The functional role of long non-coding RNA in human carcinomas. *Mol Cancer.* 2011;10:1–17.
22. Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science.* 2007;316:1484–8.
23. Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, Gao G. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* 2007;35(Web Server issue):W345–9.
24. Sun L, Luo H, Bu D, Zhao G, Yu K, Zhang C, et al. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res.* 2013;41(17), e166.
25. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. *Nucleic Acids Res.* 2013;42(D1):D222–30.
26. Li A, Zhang JY, Zhou ZY. PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. *BMC Bioinformatics.* 2014;15:311.
27. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 2012;22(9):1775–89.
28. Zhan SY, Dong Y, Zhao W, Guo JZ, Zhong T, Wang LJ, et al. Genome-wide identification and characterization of long non-coding RNAs in developmental skeletal muscle of fetal goat. *BMC Genomics.* 2016;17:66.
29. Guo Z, Kang S, Zhu X, Xia J, Wu Q, Wang S, et al. Down-regulation of a novel ABC transporter gene (*Pxwhite*) is associated with *Cry1Ac* resistance in the diamondback moth, *Plutella xylostella* (L.). *Insect Biochem Mol Biol.* 2015;59:30–40.
30. Young RS, Marques AC, Tibbit C, Haerty W, Bassett AR, Liu JL, Ponting CP. Identification and properties of 1,119 candidate lincRNA loci in the *Drosophila melanogaster* genome. *Genome Biol Evol.* 2012;4(4):427–42.
31. Chen B, Zhang Y, Zhang X, Jia S, Chen S, Kang L. Genome-wide identification and developmental expression profiling of long noncoding RNAs during *Drosophila* metamorphosis. *Sci Rep.* 2016;6:23330.
32. Humann FC, Tiberio GJ, Hartfelder K. Sequence and expression characteristics of long noncoding RNAs in honey bee caste development – potential novel regulators for transgressive ovary size. *PLoS One.* 2013;8(10), e78915.
33. Jayakodi M, Jung JW, Park D, Ahn YJ, Lee SC, Shin SY, et al. Genome-wide characterization of long intergenic non-coding RNAs (lincRNAs) provides new insight into viral diseases in honey bees *Apis cerana* and *Apis mellifera*. *BMC Genomics.* 2015;16(1):680.
34. Jenkins AM, Waterhouse RM, Kopin AS, Marc AT. Long non-coding RNA discovery in *Anopheles gambiae* using deep RNA sequencing. *bioRxiv.* 2014. doi.org/10.1101/007484.
35. Etebari K, Asad S, Zhang GM, Asgari S. Identification of *Aedes aegypti* long intergenic non-coding RNAs and their association with wolbachia and dengue virus infection. *PLoS Negl Trop Dis.* 2016;10(10), e0005069.
36. Xiao HM, Yuan ZT, Guo DH, Hou BF, Yin CL, Zhang WQ, et al. Genome-wide identification of long noncoding RNA genes and their potential association with fecundity and virulence in rice brown planthopper, *Nilaparvata lugens*. *BMC Genomics.* 2015;16:749.
37. Wu Y, Cheng T, Liu C, Liu D, Zhang Q, Long R, et al. Systematic identification and characterization of long non-coding RNAs in the silkworm, *Bombyx mori*. *PLoS One.* 2016;11(1), e0147147.
38. You M, Yue Z, He W, Yang X, Yang G, Xie M, et al. A heterozygous moth genome provides insights into herbivory and detoxification. *Nat Genet.* 2013;45(2):220–5.
39. Zhao W, He X, Hoadley KA, Parker JS, Hayes DN, Perou CM. Comparison of RNA-Seq by poly (a) capture, ribosomal RNA depletion, and DNA microarray for expression profiling. *BMC Genomics.* 2014;15:419.
40. Wilusz JE, Sunwoo H, Spector DL. Long noncoding RNAs: functional surprises from the RNA world. *Gene Dev.* 2009;23(13):1494–504.
41. Peng L, Paulson A, Li H, He X, Lu H, Klaassen CD, et al. Long noncoding RNAs and transcription of cytochrome P450s in mouse liver during maturation. *FASEB J.* 2013;27(1):1102.7.
42. Li Y, Syed J, Sugiyama H. RNA-DNA triplex formation by long noncoding RNAs. *Cell Chem Biol.* 2016;S2451–9456(16):30346–4.
43. Guo L, Wang Y, Zhou XG, Li ZY, Liu SZ, Liang P, et al. Functional analysis of a point mutation in the ryanodine receptor of *Plutella xylostella*(L) associated with resistance to chlorantraniliprole. *Pest Manag Sci.* 2014;70:1083–9.
44. Li XX, Zhu B, Gao XW, Liang P. Over-expression of UDP-glycosyltransferase gene *UGT2B17* is involved in chlorantraniliprole resistance in *Plutella xylostella* (L.). *Pest Manage Sci.* 2016. doi:10.1002/ps.4469.
45. Cock PJ, Fields CJ, Goto N, Heuer ML, Rice PM. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Res.* 2010;38(6):1767–71.
46. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics.* 2009;25(9):1105–11.
47. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc.* 2012;7(3):562–78.
48. Anders S, Huber W. Differential expression of RNA-Seq data at the gene level—the DESeq package. 2012.
49. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol.* 2010;28(5):511–5.
50. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods.* 2001;25:402–8.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

