Genome-Wide Identification of Long Non-coding RNAs in the Gravid Ectoparasite *Varroa destructor*

Zheguang Lin[1], Yibing Liu[1], Xiaomei Chen[1], Cong Han[1], Wei Wang[2], Yalu Ke[2], Xiaoling Su[3], Yujiao Li[4], Heng Chen[1], Hao Xu[1], Guohong Chen[1] and Ting Ji[1]*

[1] Apicultural Research Institute, College of Animal Science and Technology, Yangzhou University, Yangzhou, China,
[2] Wuzhong Animal Health Supervision Institute, Suzhou, China, [3] Jinhua Academy of Agricultural Sciences, Jinhua, China,
[4] Shandong Apiculture Breeding of Improved Varieties and Extension Center, Tai'an, China

Long non-coding RNAs (lncRNAs) emerge as critical regulators across a wide variety of biological functions in living organisms. However, to date, no systematic characterization of lncRNAs has been investigated in the ectoparasitic mite *Varroa destructor*, the most severe biotic threat to honey bees worldwide. Here, we performed an initial genome-wide identification of lncRNAs in *V. destructor* via high-throughput sequencing technology and reported, for the first time, the transcriptomic landscape of lncRNAs in the devastating parasite. By means of a lncRNA identification pipeline, 6,645 novel lncRNA transcripts, encoded by 3,897 gene loci, were identified, including 2,066 sense lncRNAs, 2,772 lincRNAs, and 1,807 lncNATs. Compared with protein-coding mRNAs, *V. destructor* lncRNAs are shorter in terms of full length, as well as of the ORF length, contain less exons, and express at lower level. GO term and KEGG pathway enrichment analyses of the lncRNA target genes demonstrated that these predicted lncRNAs may be potentially responsible for the regulatory functions of cellular and biological progresses in the reproductive phase of *V. destructor*. To our knowledge, this is the first catalog of lncRNA profile in the parasitiformes species, providing a valuable resource for genetic and genomic studies. Understanding the characteristics and features of lncRNAs in *V. destructor* would promote sustainable parasite control.

Keywords: *Varroa destructor*, long non-coding RNA, honey bee, ectoparasitic mite, oogenesis

INTRODUCTION

Colony losses of the western honey bee *Apis mellifera* in the Western world are a serious issue due to the critical role of honey bees in the balance of the ecosystem, sustainable agriculture, and food security (Neumann and Carreck, 2010; Potts et al., 2016; Steinhauer et al., 2018). There has been a consensus that an emerging ectoparasitic mite, *Varroa destructor*, is the principal threatening factor (Neumann and Carreck, 2010; Nazzi and Conte, 2016; Evans, 2019). *V. destructor* feasts on the fat body (Ramsey et al., 2018, 2019) of honey bees, transmits viruses, and affects host immunity (Yang and Cox-Foster, 2005; Rosenkranz et al., 2010; Di Prisco et al., 2016), severely interrupting the social organization and demographic continuity in *A. mellifera* colonies. Without treatment against this mite, infested *A. mellifera* colonies usually die within 6 months to 2 years (Kraus and Page, 1995; Le Conte et al., 2010). In the wake of the occurrence of *V. destructor* in New Zealand in 2000

(Todd et al., 2007; Mondet et al., 2014) and in Hawaii in 2007 (Martin et al., 2012), this mite has been distributed globally except for Australia and a few remote islands (Muli et al., 2018; Noël et al., 2020; Traynor et al., 2020).

V. destructor lives entirely on its host and cannot survive independently (Traynor et al., 2020) with two life cycles: the phoretic (non-reproductive) phase on the body surface of adult bees and the reproductive phase in the sealed brood cells with immature bees (Martin, 1994). Reproduction of the parasite starts from the oogenesis process, which occurs since approximately 6 h later after the invaded cell is capped and is crucial for understanding the reproductive biology of the parasite (Garrido et al., 2000; Häußermann et al., 2019). A great number of studies have been performed on different life phases of this obligate bee parasite, however, molecular studies are very limited as a result of the lack of genomic information. Cornman et al. (2010) analyzed *V. destructor* genome sequence for the first time, which largely facilitated Mondet et al. (2018) to investigate a full life cycle transcriptomic profiling in adult *V. destructor*. Nevertheless, the emerging functional elements of non-coding RNAs (ncRNAs) have yet to report with this parasitic mite.

Numerous genome-wide transcriptome has observed that the majority of transcripts do not code for proteins, and these transcripts are referred to as ncRNAs (Lee et al., 2012). NcRNA is a generic term for all functional RNAs that are transcribed from DNA but not translated into proteins, and the widespread expression of ncRNAs in eukaryotic cells has been demonstrated to be of importance in the processes of development, disease resistance, etc. (Esteller, 2011). Long non-coding RNAs (lncRNAs), most of which are located in the nucleus of eukaryotes, are a cluster of ncRNAs with a length of more than 200 nt, with cap-structure and ploy (A)-tail but usually without a long reading frame (Guttman et al., 2009). LncRNAs can be classified into four groups based on their positional information on genomes, i.e., sense lncRNAs, intergenic lncRNAs (lincRNAs), intronic lncRNAs (ilncRNAs), and antisense lncRNAs (lncNATs) (Harrow et al., 2012; Ma et al., 2013). As functional elements, lncRNAs have been proved to exert their bioactivities by regulating gene expression at the level of epigenetic inheritance, transcription, and post transcription, as well as by affecting protein localization and telomere replication (Furuno et al., 2006; Mercer et al., 2009). Currently, however, studies of lncRNAs in the field of honey bee science are still in its infancy.

The present a few studies have demonstrated that these functional elements of lncRNAs participate in the regulation of the physiological processes of honey bees, such as labor division, ovary development, neural networks, pesticide metabolism, and pathogen resistance (Kiya et al., 2008; Humann et al., 2013; Jayakodi et al., 2015; Chen et al., 2017, 2019; Liu et al., 2019; Fent et al., 2020). Guo et al. (2018a,b) screened the lncRNAs in two honey bee fungi, *Ascospheara apis* and *Nosema ceranae*, providing the first two lncRNA profiles in honey bee pathogenic agents. In the present study, we deeply sequenced the ubiquitous ectoparasite *V. destructor* on the Illumina platform during the oogenesis stage. We identified 6,645 novel lncRNA transcripts corresponding to 3,897 lncRNA genes in the detrimental mite. The subcellular localization patterns of the lncRNAs were predicted, most of which were located in nucleus. The structural features and the subcellular localization of the lncRNAs identified in this study showed consistent with their counterparts in the mammals. In order to comprehensively investigate the main biological properties and functions of the target genes of the putative novel lncRNAs, we performed Gene Ontology (GO) and pathway enrichment analyses. By comprehensively identifying lncRNAs in the gravid *V. destructor*, we aimed to offer novel insights into understanding the basic molecular biology of this ubiquitous ectoparasitic mite of honey bees.

MATERIALS AND METHODS

Gravid Adult Female *Varroa destructor* Mite Collection

V. destructor mites were collected from experimental *A. mellifera* colonies untreated for half a year, which were located in Yangzhou University of China. We employed the mites to mimic phoretic phase before experimental infestation to obtain similar physiological status (see details in Lin et al., 2018). Each freshly capped worker cell was introduced with a mite, and gravid adult mites were collected from sealed worker brood with the tweezers and the paint brushes 2 days later after infestation (**Supplementary Figure 1**). Three samples, each of which contained 15 randomly gathered mites, were prepared for RNA-seq (i.e., Vd-1, Vd-2, and Vd-3; **Table 1**). All the mites were frozen in liquid nitrogen after bathing in the phosphate buffered saline, pH 7.4 (Sigma, MO, United States), twice. The samples were stored in −80°C until RNA extraction.

Library Preparation for lncRNA Sequencing

Total RNA was extracted using TRIzol reagent kit (Invitrogen, CA, United States) according to the manufacturer's protocol. Purity and quantity of total RNA were measured using NanoDrop^{TM} 2000 (Thermo Fisher Scientific, DE, United States). RNA quality was assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, United States) and was checked on RNase free agarose gel electrophoresis. For the sample preparation, 3 μg RNA was used for each of the three *V. destructor* samples, and NEBNext® UltraTM RNA Library Prep Kit for Illumina® (NEB, MA, United States) was used to generate sequencing libraries. After removing rRNAs, the enriched mRNAs and ncRNAs were fragmented into short fragments by using fragmentation buffer and reverse transcribed into cDNA with random primers. Second-strand cDNA were subsequently synthesized by using DNA polymerase I, RNase H, dNTP, and buffer. The cDNA fragments were then purified with QiaQuick PCR extraction kit (Qiagen, Venlo, The Netherlands), end repaired, poly(A) added, and ligated to Illumina sequencing adapters. Uracil-N-Glycosylase was used to digest the second-strand cDNA and the digested products were size selected by agarose gel electrophoresis, PCR amplified, and sequenced on an Illumina HiSeq^{TM} 2500 platform at the Novogene Bioinformatics Institute

TABLE 1 | Throughput and quality of RNA-seq of the three libraries.

Sample name	Total nucleotides (G)	Raw reads	Clean reads	Q20 (%)	Q30 (%)	GC content (%)
Vd-1	13.3	88,932,536	88,072,510	97.3	92.2	43.6
Vd-2	12.6	84,224,988	83,475,384	97.3	92.3	43.7
Vd-3	16.8	111,930,228	108,997,328	97.7	93.5	43.0

FIGURE 1 | Prediction of novel lncRNA of *V. destructor*. **(A)** Venn analysis of the putative novel lncRNAs by using CPC, CNCI, and PFAM. **(B)** The source and distribution percentage of the novel lncRNAs based on the intersection of Venn diagram.
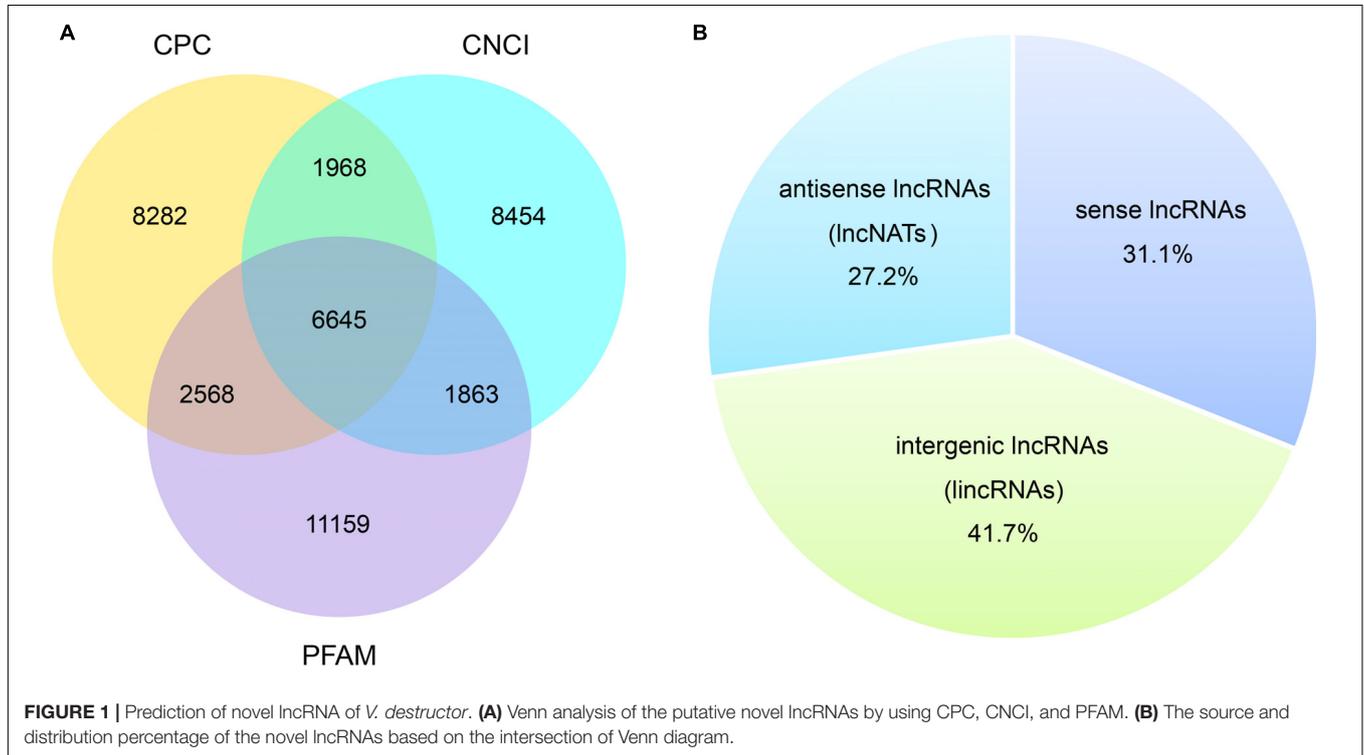
(Beijing, China). The raw sequencing data were uploaded to the National Centre for Biotechnology Information (SRA accession: SRP258850).

LncRNA Identification and Analyses

Clean reads were obtained by removing reads containing adapter, or ploy-N and low-quality reads from the raw data. Meanwhile, we calculated the Q20, Q30, and GC content. The paired-end clean reads were mapped to the reference genome of *V. destructor*[1] (Techer et al., 2019) using HISAT2 (Kim et al., 2015). We then reconstruct the transcripts with StringTie and HISAT2. To identify new transcripts, all the reconstructed transcripts were aligned to the reference genome and were divided into 12 categories by employing Cuffcompare, and transcripts with one of the five class codes "I," "j," "o," "u," and "x" were potentially recognized as novel ones. The putative novel transcripts were further eliminated by removing the ones with length ≤ 200 nt or with exon number <2. Coding Potential Calculator (CPC), Coding-Non-Coding Index (CNCI), and Pfam-scan (PFAM) were jointly used to assess the protein-coding potential of the selected novel transcript candidates and

the intersection was considered as the candidate set of lncRNAs. We used StringTie again to quantify transcripts abundances by calculating the FPKM (expected fragments per kilobase of transcript per million fragments mapped) values. All the putative novel lncRNAs were computed their subcellular localization by means of lncLocator, an online software for lncRNA location prediction based on a stacked ensemble classifier (Cao et al., 2018). LncLocator can predict five subcellular localizations of lncRNAs, including nucleus, cytoplasm (part of cytoplasm except for cytosol, ribosome, and exosome), ribosome, cytosol, and exosome. Then, we searched coding genes 100 kb upstream and downstream of the predicted novel lncRNAs as *cis* target genes (Guil and Esteller, 2012; Feng et al., 2019), which were subjected to enrichment analysis of GO functions and kyoto encyclopedia of genes and genomes (KEGG) pathways.

The statistical analyses (Student's *t*-test) of the characteristic differences between lncRNAs and mRNAs were performed with SPSS Statistics 25.

RT-PCR Validation

Removal of gDNA and synthesis of cDNA was performed with RNA products following the manufacturer's instructions of ReverTra Ace qPCR RT Master Mix (Tiangen, Beijing, China). To

[1]https://www.ncbi.nlm.nih.gov/assembly/GCF_002443255.1

FIGURE 2 | Genomic features of the predicted lncRNAs. Exon number distribution **(A)**, length distribution **(B)**, ORF length distribution **(C)**, and expression level indicated by log10 (FPKM + 1) **(D)** of 32,415 coding transcripts (mRNAs) and 6,645 novel lncRNAs were plotted.
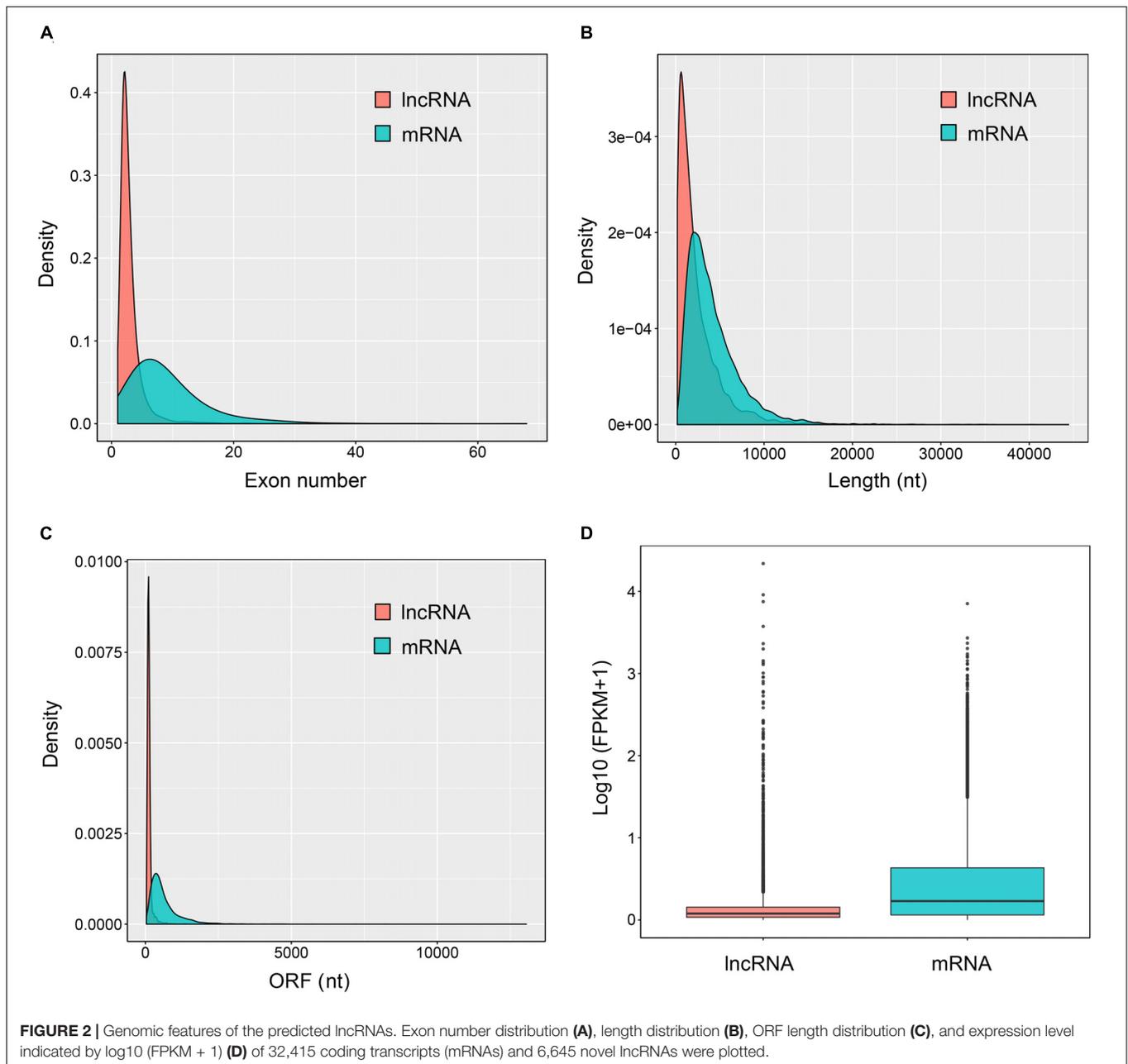
validate the putative lncRNAs in *V. destructor* mites, 16 lncRNAs were randomly selected to determine with PCR amplification, which was carried out with the obtained cDNA in a 20 μL reaction volume mixture (2 × Taq PCR StarMix; GenStar, Beijing, China) on an Eppendorf cycler. PCR profile consists of a pre-denaturation at 94°C for 5 min; followed by 30 cycles including 94°C for 50 s, 55°C for 30 s, and 72°C for 1 min; and a final elongation step at 72°C for 10 min (Guo et al., 2018a). The tested lncRNAs with their forward and reverse primers were presented in **Supplementary Table 1**. PCR products were electrophoresed in 2.5% Tris acetate-EDTA-agarose gel containing 0.01% Gelview (BioTeke, Beijing, China) and visualized under ultraviolet light (Peiqing, Shanghai, China).

RESULTS

Total RNA of the three *V. destructor* samples (Vd-1, Vd-2, and Vd-3) were isolated and sequenced. Overall, 42.8 G sequencing data were generated, corresponding to 285.1 million raw reads and 280.5 million quality filtered (clean) reads were generated from the three cDNA libraries (**Table 1**). G20, G30, and GC content were also shown in **Table 1** with the mean values of 97.4, 92.7, and 43.5%. For the three *V. destructor* samples, 93.2, 93.6, and 90.0% obtained reads were mapped to the reference genome sequence, and the mapped regions of each sample on the genome were shown in **Supplementary Figure 2**. Then 6,645 putative novel non-coding transcripts were predicted using CPC,

FIGURE 3 | Subcellular localization prediction of the sense lncRNAs **(A)**, the intergenic lncRNAs (lincRNAs, **B**), and the antisense lncRNAs (lncNATs, **C**). An online prediction program, lncLocator, which can predict five subcellular localizations of lncRNAs, was used for this analysis. The term "cytoplasm" here includes the part of cytoplasm except for cytosol, ribosome, and exosome.

CNCI, and PFAM (**Figure 1A** and **Supplementary Table 2**), of which 2,066 (31.1%) were sense lncRNAs, 2,772 (41.7%) were lincRNAs, and 1,807 (27.2%) were lncNATs (**Figure 1B**). IlncRNAs were not observed in this mite. In addition, 32,415 protein coding transcripts were obtained, with 32,331 mapped to the reference genome, and the remaining 84 ones were not annotated (**Supplementary Table 2**).

Most lncRNAs contained two exons (55.8%), followed by three (23.8%), four (9.6%), five (4.2%), six (2.2%), and seven (1.1%) (**Figure 2A** and **Supplementary Table 2**). The ratio of lncRNAs is less than one when the number of exons is greater than seven, and they contained at most 27 exons (**Figure 2A** and **Supplementary Table 2**). This was significantly different from the coding transcripts (Student's t-test, $p < 0.001$), which peaked at six exons (9.5%) and were up to 68 exons. Meanwhile, the ratio of mRNAs with three to eight exons were respectively more than seven (**Figure 2A** and **Supplementary Table 2**). Most of both lncRNAs (55.4%) and mRNAs (65.5%) ranged from 1,000 to 5,000 nt in length. But then, 33.2% lncRNAs were less than 1,000 nt and 24.7% mRNAs were between 5,001 nt and 10,000 nt. For the long sequence ($>$10,000 nt), 2.1% lncRNAs and 4.7% mRNAs were occupied. As a result, lncRNAs averaged 2,435 nt in length, which was significantly shorter than protein-coding genes (4,187 nt; Student's t-test, $p < 0.001$; **Figure 2B** and **Supplementary Table 2**). Regarding the length of open reading frames (ORFs) in lncRNAs and mRNAs, we got a similar trend with above. Most of the ORFs of both lncRNAs (58.5%) and mRNAs (51.4%) were in the middle range 100–500 nt in length, followed by 39.9% lncRNAs $\leq$100 nt and 31.9% mRNAs ranging from 501 to 1,000 nt. Consistently, 0.2% lncRNAs and 16.3% mRNAs were respectively greater than 1,000 nt, and the mean length of ORFs in lncRNAs was significantly shorter than that of mRNAs (132 vs. 649 nt; Student's t-test, $p < 0.001$; **Figure 2C** and **Supplementary Table 2**). Further, the expression level of lncRNAs showed significantly lower compared to mRNAs

(**Figure 2D**). Additionally, 16 lncRNAs were randomly chosen to be validated with RT-PCR with 15 (93.8%) successful amplification (**Supplementary Figure 3**). Sanger sequencing confirmed most of the validated lncRNAs (**Supplementary Table 5**) although two of the 15 failed due to the low expression, which were reflected by the dimmed electrophoretic band in **Supplementary Figure 3** (band 2 and 4).

As shown in **Figure 3**, most of the novel lncRNAs were predicted to be localized in the nucleus, followed by in cytoplasm, irrespective of their different sources. In total, only 31, 5, and 5 lncRNAs were predicted in exosome, cytosol, and ribosome, respectively. We then obtained 9,500 target genes of novel lncRNAs in *cis* regulation. GO term analysis indicated the target genes were in the ontology class of cellular component, molecular function, and biological process (**Figure 4A** and **Supplementary Table 3**). The top five enriched GO terms with extremely small padj values participated in protein binding, catalytic activity, and transferase activity, indicating that lncRNAs may play roles in the important process of enzyme-related catalytic activities in *V. destructor* (**Figure 4A** and **Supplementary Table 3**). The KEGG pathway enrichment denoted that the target genes mostly participated in 104 pathways (**Supplementary Table 4**). The top 20 enriched pathways, which was divided into five classes based on their major functions, i.e., organismal systems, cellular processes, metabolism, environmental information processing and genetic information processing, were displayed in **Figure 4B**. The top five enriched pathways were endocytosis, RNA transport, phagosome, ubiquitin mediated proteolysis, and protein processing in endoplasmic reticulum (**Supplementary Table 4**), involving the critical processes of endocytosis, and protein processing and degradation, which belong to the cellular processes and the genetic information processing (**Figure 4B**). Notably, 115 genes are enriched in the RNA transport, which is a fundamental process for gene expression, signifying lncRNAs may play key parts in the transcriptional regulation during the oogenesis of *V. destructor*.
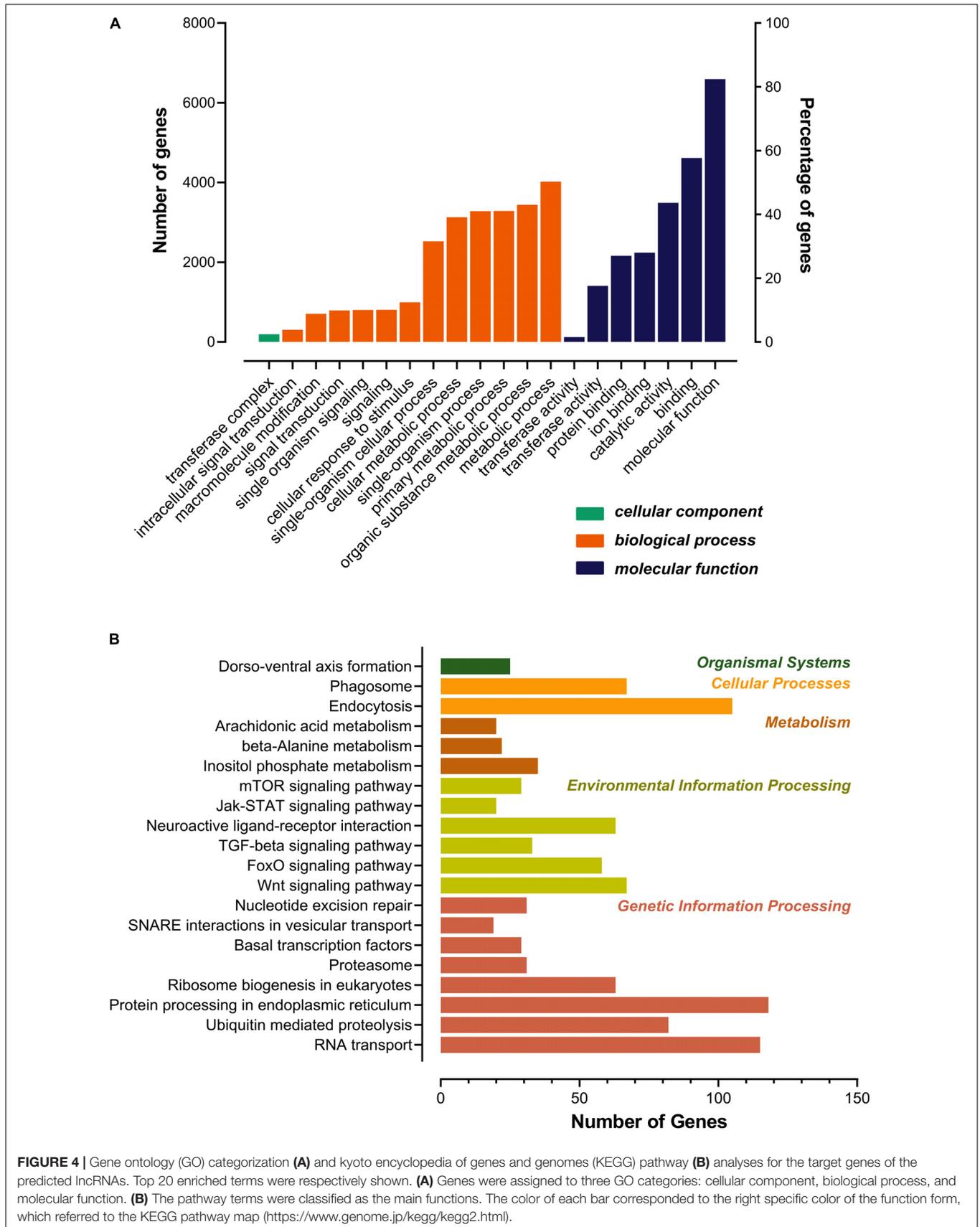
FIGURE 4 | Gene ontology (GO) categorization **(A)** and kyoto encyclopedia of genes and genomes (KEGG) pathway **(B)** analyses for the target genes of the predicted lncRNAs. Top 20 enriched terms were respectively shown. **(A)** Genes were assigned to three GO categories: cellular component, biological process, and molecular function. **(B)** The pathway terms were classified as the main functions. The color of each bar corresponded to the right specific color of the function form, which referred to the KEGG pathway map (https://www.genome.jp/kegg/kegg2.html).

DISCUSSION

LncRNAs have emerged as critical participators in a variety of cellular activities, ranging from simple housekeeping to complex regulatory functions. However, till now, the studies of lncRNAs are mainly conducted in the field of humans, mammal, and crops. In contrast, the research on invertebrates is still at the early stage. Here, we identified the sequences and expression features of the lncRNAs via high-throughput sequencing technology, for the first time, in a devastating ectoparasite, *V. destructor*, of the chief pollinator, honey bee. We obtained 6,645 putative novel lncRNAs from 3,897 gene loci in the *V. destructor*, including 2,066 sense lncRNAs, 2,772 lincRNAs, and 1,807 lncNATs. The characteristics of the lncRNAs and the comparison with mRNAs were also reported. To verify the reliability of the RNA-seq results and the predicted lncRNAs, 16 non-coding transcripts were randomly selected to do RT-PCR validation, and 93.8% lncRNAs amplified signal bands denoting a reliable output of lncRNAs identified in this study. To the best of our knowledge, lncRNA studies have not performed in the parasitiformes species before.

V. destructor reproduction is limited to a short window when the immature honey bee host is concealed in wax cells. Mated female mites target to meal on the larva host 5 h after the invaded cell was capped (Ifantidis, 1988; Donzé and Guerin, 1994) and initiates oogenesis about another hour later (Garrido et al., 2000). The mite then lays the first male egg approximately 70 h after the cell capping (Donzé and Guerin, 1994; Martin, 1994). Considering these timing, we collected the mites 48 h later after the mite was introduced into the freshly capped cells. The artificial infestation has been established to be a suitable method for *V. destructor* research (Dietemann et al., 2013; Lin et al., 2018; Häußermann et al., 2019), and mites parasitized in the capped brood cells are less variable in physiology and fitness than on the adult bee bodies (Milani, 1995). The mites collected were further confirmed gravid with dilated post-abdomen, from which we can even see the eggs inside under microscope (**Supplementary Figure 1B**).

Just as lncRNA screening in other species, lincRNAs, the most extensively studied category of lncRNAs, usually account for the largest proportion (Sun et al., 2013; Wu et al., 2016; Zhan et al., 2016; Guo et al., 2018a), although some studies show otherwise (Guo et al., 2018b). We did not detect any ilncRNAs, which were regarded as the lowest conservative class of lncRNAs (Wang et al., 2016), in *V. destructor*. Intriguingly, this is also the case of *N. ceranae*, another pathogenic agent of honey bees (Guo et al., 2018b). Most of the *V. destructor* lncRNAs contained two exons with the average of 3.0 exons, significantly less than mRNAs (**Figure 2** and **Supplementary Table 2**), which was also in line with the related studies of lncRNAs on other invertebrate counterparts, even mammalian and plant species (Trapnell et al., 2010; Zhang et al., 2014; Wu et al., 2016; Zhan et al., 2016; Guo et al., 2018a,b; Liu et al., 2019). Besides, the shorter length of the putative lncRNAs and the ORFs compared to mRNA also shared similar features with other well-studied species (Trapnell et al., 2010; Zhang et al., 2014; Wang et al., 2016; Wu et al., 2016; Zhan et al., 2016; Guo et al., 2018a,b; Liu et al., 2019). Although lncRNAs always performed lower expression level than the protein-coding genes (this study; Guo et al., 2018a,b; Chen et al., 2019; Liu et al., 2019), their role in the functional activities of organisms has been widely proved to be of significance (Furuno et al., 2006; Mercer et al., 2009; Zhang et al., 2014; Wu et al., 2016; Chen et al., 2019; Liu et al., 2019).

LncRNAs can operate in *cis* to regulate the transcriptional expression of neighboring genes on the same allele (Robinson et al., 2020). The upstream lncRNAs with intersection of promoter or other *cis* elements may regulate gene expression at the level of transcription or post-transcription, and lncRNAs in the downstream or 3'UTR region may have other regulatory functions. LncRNAs in less than 100 kb up/down stream of a gene may serve as *cis* regulatory factors (Guil and Esteller, 2012; Feng et al., 2019). The *cis* target genes were engaged in various molecular functions and biological processes. GO and KEGG analyses revealed the target genes were mainly included in protein binding, enzyme activities, metabolism, signaling molecules and interaction, and so forth (**Figure 4** and **Supplementary Tables 3, 4**). Intriguingly, 25 genes were enriched in a pathway of dorso-ventral axis formation (**Figure 4B** and **Supplementary Table 4**), suggesting that they may play a crucial role in the development and regeneration during the process of oogenesis of the mite.

Similar to proteins, lncRNAs are of importance to be localized in specific cellular compartments, which provides insights for understanding their complex biological functions (Chen, 2016). We predicted the subcellular localizations of the identified sense lncRNAs, lincRNAs, and lncNATs, of which 95.6, 91.7, and 92.4% were respectively accumulated in nucleus (**Figure 3**). These lncRNAs have been proposed to play strong roles in nuclear architecture and gene expression regulation. They are associated with chromatin-modifying complexes, directly influence transcription, act as precursors for small RNAs, participate in stem cell pluripotency and differentiation, and so forth (Chen and Carmichael, 2010). Cytoplasmic lncRNAs, the second most popularly located lncRNAs, have been evidenced to impact gene expression in a variety of ways, such as interfering with protein post-translational modifications with a result of aberrant signal transduction (Lin et al., 2016), acting as decoys for miRNAs and proteins (Cesana et al., 2011; Lee et al., 2016) and affecting mRNA translation in the cytoplasm (Gong and Maquat, 2011). Hence, the lncRNAs in distinct subcellular compartments are of great interest to decipher their diverse functional significance.

CONCLUSION

V. destructor presents tremendous threat to apiculture worldwide, and the in-depth molecular studies on the parasite will facilitate to control this unpopular pest. We reported the lncRNA profile of *V. destructor* by genome-wide RNA-seq in this study, and the genomic and structural features of the lncRNAs showed consistent with their counterparts in other species. Evidence is becoming increasingly clear that the function of lncRNAs is associated with their unique subcellular localization (Chen, 2016), and most of the lncRNAs detected in *V. destructor* were accumulated in nucleus. The target genes of lncRNAs

were inferred to participate in diverse regulatory functions via *cis* regulation by GO term and KEGG pathway enrichment analyses. Our data provide genetic resources for exploration of the functional roles of lncRNAs involved in the ectoparasite *V. destructor*. Further studies would be of interest and value to characterize the expression profile of lncRNAs in the different life stages of the ubiquitous mite.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/genbank/, SRA accession: SRP258850.

AUTHOR CONTRIBUTIONS

TJ, GC, and ZL conceived the ideas and designed the study. YBL, XC, and CH carried out the experimental infestations. ZL, WW, YK, XS, and YJL performed the data analyses. YBL, HC, and HX performed the validation experiment. ZL led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.575680/full#supplementary-material

Supplementary Figure 1 | The dorsal **(A)** and post-abdomen **(B)** appearances of the gravid adult female *Varroa destructor* we collected.

Supplementary Figure 2 | Distribution percentage of the clean reads from Vd-1 **(A)**, Vd-2 **(B)**, and Vd-3 **(C)** mapped to the genome regions of *Varroa destructor*.

Supplementary Figure 3 | RT-PCR validation of the randomly selected 16 lncRNAs. DNA marker (M) were used to indicate the product size in the left lane. Lane 1 to lane 16 were 16 lncRNAs orderly listed in the **Supplementary Table S1**.

Supplementary Table 1 | Primer sets used for RT-PCR validation of the identified lncRNAs.

Supplementary Table 2 | Information of all the transcripts.

Supplementary Table 3 | Gene ontology (GO) output for the *cis* target genes of the identified lncRNAs.

Supplementary Table 4 | Kyoto encyclopedia of genes and genomes (KEGG) pathway output for the *cis* target genes of the identified lncRNAs.

Supplementary Table 5 | Sanger sequencing results of the 13 lncRNAs.

REFERENCES

Cao, Z., Pan, X., Yang, Y., Huang, Y., and Shen, H.-B. (2018). The lncLocator: a subcellular localization predictor for long non-coding RNAs based on a stacked ensemble classifier. *Bioinformatics* 34, 2185–2194. doi: 10.1093/bioinformatics/bty085

Cesana, M., Cacchiarelli, D., Legnini, I., Santini, T., Sthandier, O., Chinappi, M., et al. (2011). A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* 147, 358–369. doi: 10.1016/j.cell.2011.09.028

Chen, D., Chen, H., Du, Y., Zhou, D., Geng, S., Wang, H., et al. (2019). Genome-wide identification of long non-coding RNAs and their regulatory networks lnvolved in *Apis mellifera ligustica* response to *Nosema ceranae* infection. *Insects* 10:245. doi: 10.3390/INSECTS10080245

Chen, L.-L. (2016). Linking long noncoding RNA localization and function. *Trends Biochem. Sci.* 41, 761–772. doi: 10.1016/j.tibs.2016.07.003

Chen, L.-L., and Carmichael, G. G. (2010). Decoding the function of nuclear long non-coding RNAs. *Curr. Opin. Cell Biol.* 22, 357–364. doi: 10.1016/j.ceb.2010.03.003

Chen, X., Ma, C., Chen, C., Lu, Q., Shi, W., Liu, Z., et al. (2017). Integration of lncRNA–miRNA–mRNA reveals novel insights into oviposition regulation in honey bees. *PeerJ* 5:e3881. doi: 10.7717/peerj.3881

Cornman, S. R., Schatz, M. C., Johnston, S. J., Chen, Y. P., Pettis, J., Hunt, G., et al. (2010). Genomic survey of the ectoparasitic mite *Varroa destructor*, a major pest of the honey bee *Apis mellifera*. *BMC Genomics* 11:602–602. doi: 10.1186/1471-2164-11-602

Di Prisco, G., Annoscia, D., Margiotta, M., Ferrara, R., Varricchio, P., Zanni, V., et al. (2016). A mutualistic symbiosis between a parasitic mite and a pathogenic virus undermines honey bee immunity and health. *Proc. Natl. Acad. Sci. U.S.A.* 113, 3203–3208. doi: 10.1073/pnas.1523515113

Dietemann, V., Nazzi, F., Martin, S. J., Anderson, D. L., Locke, B., Delaplane, K. S., et al. (2013). Standard methods for varroa research. *J. Apic. Res.* 52, 1–54. doi: 10.3896/IBRA.1.52.1.09

Donzé, G., and Guerin, P. M. (1994). Behavioral attributes and parental care of *Varroa* mites parasitizing honeybee brood. *Behav. Ecol. Sociobiol.* 34, 305–319. doi: 10.1007/BF00197001

Esteller, M. (2011). Non-coding RNAs in human disease. *Nat. Rev. Genet.* 12, 861–874. doi: 10.1038/nrg3074

Evans, J. D. (2019). Selection for barriers between honey bees and a devastating parasite. *Mol. Ecol.* 28, 2955–2957. doi: 10.1111/mec.15142

Feng, M., Ren, F., Zhou, Y., Zhang, N., Lu, Q., Swevers, L., et al. (2019). Correlation in expression between LTR retrotransposons and potential host *cis*-targets during infection of *Antherea pernyi* with ApNPV baculovirus. *Viruses* 11:421. doi: 10.3390/v11050421

Fent, K., Schmid, M., Hettich, T., and Schmid, S. (2020). The neonicotinoid thiacloprid causes transcriptional alteration of genes associated with mitochondria at environmental concentrations in honey bees. *Environ. Pollut.* 266:115297. doi: 10.1016/j.envpol.2020.115297

Furuno, M., Pang, K. C., Ninomiya, N., Fukuda, S., Frith, M. C., Bult, C., et al. (2006). Clusters of internally primed transcripts reveal novel long noncoding RNAs. *PLoS Genet.* 2:537–553. doi: 10.1371/journal.pgen.0020037

Garrido, C., Rosenkranz, P., Stürmer, M., Rübsam, R., and Büning, J. (2000). Toluidine blue staining as a rapid measure for initiation of oocyte growth and fertility in *Varroa jacobsoni* Oud. *Apidologie* 31, 559–566. doi: 10.1051/APIDO: 2000146

Gong, C., and Maquat, L. E. (2011). lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* 470, 284–288. doi: 10.1038/nature09701

Guil, S., and Esteller, M. (2012). *Cis*-acting noncoding RNAs: friends and foes. *Nat. Struct. Mol. Biol.* 19, 1068–1075. doi: 10.1038/nsmb.2428

Guo, R., Chen, D., Xiong, C., Hou, C., Zheng, Y., Fu, Z., et al. (2018a). Identification of long non-coding RNAs in the chalkbrood disease pathogen *Ascosphaera apis*. *J. Invertebr. Pathol.* 156, 1–5. doi: 10.1016/J.JIP.2018.06.001

Guo, R., Chen, D., Xiong, C., Hou, C., Zheng, Y., Fu, Z., et al. (2018b). First identification of long non-coding RNAs in fungal parasite *Nosema ceranae*. *Apidologie* 49, 660–670. doi: 10.1007/S13592-018-0593-Z

Guttman, M., Amit, I., Garber, M., French, C., Lin, M. F., Feldser, D., et al. (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* 458, 223–227. doi: 10.1038/nature 07672

Harrow, J., Frankish, A., Gonzalez, J. M., Tapanari, E., Diekhans, M., Kokocinski, F., et al. (2012). GENCODE: the reference human genome annotation for the ENCODE Project. *Genome Res.* 22, 1760–1774. doi: 10.1101/GR.1353 50.111

Häußermann, C. K., Giacobino, A., Munz, R., Ziegelmann, B., Palacio, M. A., and Rosenkranz, P. (2019). Reproductive parameters of female *Varroa destructor* and the impact of mating in worker brood of *Apis mellifera*. *Apidologie* 19, 1–14. doi: 10.1007/S13592-019-00713-9

Humann, F. C., Tiberio, G. J., and Hartfelder, K. (2013). Sequence and expression characteristics of long noncoding RNAs in honey bee caste development – potential novel regulators for transgressive ovary size. *PLoS One* 8:e78915. doi: 10.1371/journal.pone.0078915

Ifantidis, M. D. (1988). Some aspects of the process of *Varroa jacobsoni* mite entrance into honey bee (*Apis mellifera*) brood cells. *Apidologie* 19, 387–396. doi: 10.1051/APIDO:19880406

Jayakodi, M., Jung, J. W., Park, D., Ahn, Y.-J., Lee, S.-C., Shin, S.-Y., et al. (2015). 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* 16:680. doi: 10.1186/s12864-015-1868-7

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/NMETH. 3317

Kiya, T., Kunieda, T., and Kubo, T. (2008). Inducible- and constitutive-type transcript variants of kakusei, a novel non-coding immediate early gene, in the honeybee brain. *Insect Mol. Biol.* 17, 531–536. doi: 10.1111/j.1365-2583.2008. 00821.x

Kraus, B., and Page, R. E. (1995). Effect of *Varroa jacobsoni* (Mesostigmata: Varroidae) on feral *Apis mellifera* (Hymenoptera: Apidae) in California. *Environ. Entomol.* 24, 1473–1480. doi: 10.1093/EE/24.6.1473

Le Conte, Y., Ellis, M., and Ritter, W. (2010). *Varroa* mites and honey bee health: can *Varroa* explain part of the colony losses? *Apidologie* 41, 353–363. doi: 10.1051/apido/2010017

Lee, S., Kopp, F., Chang, T.-C., Sataluri, A., Chen, B., Sivakumar, S., et al. (2016). Noncoding RNA NORAD regulates genomic stability by sequestering PUMILIO proteins. *Cell* 164, 69–80. doi: 10.1016/j.cell.2015.12.017

Lee, T.-L., Xiao, A., and Rennert, O. M. (2012). "Identification of novel long noncoding RNA transcripts in male germ cells," in *Germline Development: Methods and Protocols*, eds W.-Y. Chan and L. A. Blomberg (New York, NY: Springer), 105–114. doi: 10.1007/978-1-61779-436-0_9

Lin, A., Li, C., Xing, Z., Hu, Q., Liang, K., Han, L., et al. (2016). The LINK-A lncRNA activates normoxic HIF1α signalling in triple-negative breast cancer. *Nat. Cell Biol.* 18, 213–224. doi: 10.1038/ncb3295

Lin, Z., Liu, Y., Chen, X., Han, C., Wang, W., Ke, Y., et al. (2020). Genome-wide identification of long non-coding RNAs in the gravid ectoparasite *Varroa destructor*. *bioRxiv* [Preprint]. doi: 10.1101/2020.06.14.151340

Lin, Z., Qin, Y., Page, P., Wang, S., Li, L., Wen, Z., et al. (2018). Reproduction of parasitic mites *Varroa destructor* in original and new honeybee hosts. *Ecol. Evol.* 8, 2135–2145. doi: 10.1002/ece3.3802

Liu, F., Shi, T., Qi, L., Su, X., Wang, D., Dong, J., et al. (2019). lncRNA profile of *Apis mellifera* and its possible role in behavioural transition from nurses to foragers. *BMC Genomics* 20:393. doi: 10.1186/s12864-019-5664-7

Ma, L., Bajic, V. B., and Zhang, Z. (2013). On the classification of long non-coding RNAs. *RNA Biol.* 10, 924–933. doi: 10.4161/RNA.24604

Martin, S. J. (1994). Ontogenesis of the mite *Varroa jacobsoni* Oud. in worker brood of the honeybee *Apis mellifera* L. under natural conditions. *Exp. Appl. Acarol.* 19, 199–210. doi: 10.1007/BF00130823

Martin, S. J., Highfield, A. C., Brettell, L., Villalobos, E. M., Budge, G. E., Powell, M., et al. (2012). Global honey bee viral landscape altered by a parasitic mite. *Science* 336, 1304–1306. doi: 10.1126/science.1220941

Mercer, T. R., Dinger, M. E., and Mattick, J. S. (2009). Long non-coding RNAs: insights into functions. *Nat. Rev. Genet.* 10, 155–159. doi: 10.1038/NRG2521

Milani, N. (1995). The resistance of *Varroa jacobsoni* Oud to pyrethroids: a laboratory assay. *Apidologie* 26, 415–429. doi: 10.1051/APIDO:19950507

Mondet, F., De Miranda, J. R., Kretzschmar, A., Le Conte, Y., and Mercer, A. R. (2014). On the front line: quantitative virus dynamics in honeybee (*Apis mellifera* L.) colonies along a new expansion front of the parasite *Varroa destructor*. *PLoS Pathog.* 10:e1004323. doi: 10.1371/journal.ppat.1004323

Mondet, F., Rau, A., Klopp, C., Rohmer, M., Severac, D., Le Conte, Y., et al. (2018). Transcriptome profiling of the honeybee parasite *Varroa destructor* provides new biological insights into the mite adult life cycle. *BMC Genomics* 19:328. doi: 10.1186/s12864-018-4668-z

Muli, E., Okwaro, L. A., Kilonzo, J., Ali, N., and Monthy, G. T. (2018). *Varroa destructor* – free islands in the south west Indian ocean. *Bee World* 95, 122–123. doi: 10.1080/0005772X.2018.1522835

Nazzi, F., and Conte, Y. L. (2016). Ecology of *Varroa destructor*, the major ectoparasite of the western honey bee, *Apis mellifera*. *Annu. Rev. Entomol.* 61, 417–432. doi: 10.1146/ANNUREV-ENTO-010715-023731

Neumann, P., and Carreck, N. L. (2010). Honey bee colony losses. *J. Apic. Res.* 49, 1–6. doi: 10.3896/IBRA.1.49.1.01

Noël, A., Le Conte, Y., and Mondet, F. (2020). *Varroa destructor*: how does it harm *Apis mellifera* honey bees and what can be done about it? *Emerg. Top. Life Sci.* 4, 45–57. doi: 10.1042/ETLS20190125

Potts, S. G., Imperatriz-Fonseca, V., Ngo, H. T., Aizen, M. A., Biesmeijer, J. C., Breeze, T. D., et al. (2016). Safeguarding pollinators and their values to human well-being. *Nature* 540, 220–229. doi: 10.1038/nature20588

Ramsey, S., Gulbronson, C., Mowery, J., Ochoa, R., VanEngelsdorp, D., and Bauchan, G. (2018). A multi-microscopy approach to discover the feeding site and host tissue consumed by *Varroa destructor* on host honey bees. *Microsc. Microanal.* 24, 1258–1259. doi: 10.1017/S1431927618006773

Ramsey, S. D., Ochoa, R., Bauchan, G., Gulbronson, C., Mowery, J. D., Cohen, A., et al. (2019). *Varroa destructor* feeds primarily on honey bee fat body tissue and not hemolymph. *Proc. Natl. Acad. Sci. U.S.A.* 116, 1792–1801. doi: 10.1073/pnas.1818371116

Robinson, E. K., Covarrubias, S., and Carpenter, S. (2020). The how and why of lncRNA function: an innate immune perspective. *Biochim. Biophys. Acta, Gene Regul. Mech.* 1863:194419. doi: 10.1016/j.bbagrm.2019.194419

Rosenkranz, P., Aumeier, P., and Ziegelmann, B. (2010). Biology and control of *Varroa destructor*. *J. Invertebr. Pathol.* 103, S96–S119. doi: 10.1016/J.JIP.2009. 07.016

Steinhauer, N., Kulhanek, K., Antúnez, K., Human, H., Chantawannakul, P., Chauzat, M. P., et al. (2018). Drivers of colony losses. *Curr. Opin. Insect Sci.* 26, 142–148. doi: 10.1016/j.cois.2018.02.004

Sun, J., Lin, Y., and Wu, J. (2013). Long non-coding RNA expression profiling of mouse testis during postnatal development. *PLoS One* 8:e75750. doi: 10.1371/ journal.pone.0075750

Techer, M. A., Rane, R. V., Grau, M. L., Roberts, J. M. K., Sullivan, S. T., Liachko, I., et al. (2019). Divergent evolutionary trajectories following speciation in two ectoparasitic honey bee mites. *Commun. Biol.* 2:357. doi: 10.1038/s42003-019-0606-0

Todd, J. H., Miranda, J. R. D., and Ball, B. V. (2007). Incidence and molecular characterization of viruses found in dying New Zealand honey bee (*Apis mellifera*) colonies infested with *Varroa destructor*. *Apidologie* 38, 354–367. doi: 10.1051/APIDO:2007021

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals

unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. doi: 10.1038/nbt.1621

Traynor, K., Mondet, F., Miranda, J. D., Techer, M., Kowallik, V., Oddie, M., et al. (2020). *Varroa destructor*: a complex parasite, crippling honeybees worldwide. *Trends Parasitol.* 36, 592–606. doi: 10.1016/j.pt.2020.04.004

Wang, Y., Xue, S., Liu, X., Liu, H., Hu, T., Qiu, X., et al. (2016). Analyses of long non-coding RNA and mRNA profiling using RNA sequencing during the pre-implantation phases in pig endometrium. *Sci. Rep.* 6:20238. doi: 10.1038/srep20238

Wu, Y., Cheng, T., Liu, C., Liu, D., Zhang, Q., Long, R., et al. (2016). Systematic Identification and characterization of long non-coding RNAs in the silkworm, *Bombyx mori. PLoS One* 11:e0147147. doi: 10.1371/journal.pone.0147147

Yang, X., and Cox-Foster, D. L. (2005). Impact of an ectoparasite on the immunity and pathology of an invertebrate: evidence for host immunosuppression and viral amplification. *Proc. Natl. Acad. Sci. U.S.A.* 102, 7470–7475. doi: 10.1073/pnas.0501860102

Zhan, S., Dong, Y., Zhao, W., Guo, J., Zhong, T., Wang, L., et al. (2016). Genome-wide identification and characterization of long non-coding RNAs

in developmental skeletal muscle of fetal goat. *BMC Genomics* 17:666. doi: 10.1186/s12864-016-3009-3

Zhang, Y.-C., Liao, J.-Y., Li, Z.-Y., Yu, Y., Zhang, J.-P., Li, Q.-F., et al. (2014). Genome-wide screening and functional analysis identify a large number of long noncoding RNAs involved in the sexual reproduction of rice. *Genome Biol.* 15:512. doi: 10.1186/s13059-014-0512-1