

Identification, functional annotation, and stability analysis of miRNA in *Dermatophagoides pteronyssinus*

To the Editor,

House dust mite (HDM) is one of the major allergen sources worldwide which causes up to 85% of respiratory allergic diseases.¹ This species was regarded as the most prevalent source of indoor allergens in the United States and Europe. HDM is a popular name of the taxonomic genus *Dermatophagoides* that include several species, such as *Dermatophagoides pteronyssinus* (DPT) and *Dermatophagoides farinae* (DFA). The sensitization of HDM is strongly associated with asthma.²

MicroRNAs (miRNAs) are small RNAs that play an important role in post-transcriptional regulation of different biological processes, such as cell differentiation and cell death. There is emerging evidence showing a critical role of miRNAs in the regulation of inflammation and allergic airway diseases.³ Previous studies showed that cross-species communication of miRNA was possible. Vaucheret and Chupeau suggested that plant miRNAs could regulate gene expression in mammals.⁴ LaMonte et al reported that human sickle cell erythrocyte miRNAs could be translocated into a malaria parasite *Plasmodium falciparum* and inhibit the growth of the parasite.⁵ Buck et al⁶ also showed that exosomes secreted by nematode parasites could transfer small RNAs to host mice and modulate innate immunity. Hence, we hypothesized that miRNAs in HDM are able to enter human body, affecting lung epithelial cells and modulating immune responses.

Currently, nothing is known about miRNA in dust mites. In our previous studies, we have reported the genomes and transcriptomes of DPT.⁷ In this study, we identified and assigned the putative functions of miRNAs in DPT. We also conducted a comparative study with the miRNA in other arthropods to determine whether our identified miRNAs display critical and specific functions in HDM.

For the analysis of small RNAs in DPT, total RNA was isolated from the homogenate and size-selected by polyacrylamide gel electrophoresis. Libraries were prepared from the 18 to 44nt fraction and sequenced on Illumina HiSeq 2000 in triplicate to produce paired-end reads of 90 bp. The basic information of this sRNA-seq study can be found in Table S1, and the flowchart of our work can be found in Figure S1.

The genome of DPT has been reported previously.⁷ The high-quality raw reads (SRA accession: PRJNA577524) were mapped to the genome (Figure S2). Other types of small noncoding RNAs were removed. The remaining reads were then aligned against the sequence deposited in miRBase database (version 21) via BLAST.

The identified miRNAs were defined as “conserved” if they are identical to known miRNAs or “specific” if different to known miRNAs. After filtering by a series of rigorous criteria (Appendix S1), we have identified 69 conserved and 52 specific miRNAs in DPT (Table S2).

The transcript expression of all identified miRNAs was quantified and normalized by transcript per million (TPM) method. The identified miRNAs and their expression were listed in Figure 1A, and the homologous numbers of DPT miRNA in arthropoda were listed in Table S3. The top 16 identified miRNAs were found with high expression level (≥ 1000 TPM), and the expression of the miRNAs specifically found in HDM is relatively lower than those of conserved group. MiRNA target prediction and functional annotation were performed according to previous method. So far, we identified 65 miRNAs shared in arthropods, including 44 conserved and 21 specific miRNAs in DPT. Through RNA binding tools, we have further predicted targets of the mature miRNAs. For those miRNAs identical to human homologs, their targeted genes were enriched in KEGG pathways. We found that most miRNA target genes could be enriched in immune-related pathways (Table 1 and Table S4), which indicated that these miRNAs could be involved in allergic inflammation triggered by HDM. For example, let-7a-5p was highly expressed in DPT. It has been demonstrated that let-7 miRNAs had a proinflammatory role in allergic mouse lung, and it is confirmed in vitro that interleukin 13 (IL-13) was regulated by mmu-let-7a.⁸

In an attempt to study the stability of mite miRNAs in the environment, we have performed a Taqman miRNA expression assay (Applied Biosystems) to test the quantity of mite-specific miRNA in household dust samples. Seven dust samples were collected from pillows, mattresses, sofa, and chairs from 3 houses in Hong Kong using a vacuum cleaner. Each dust sample collected was separated into two portions: dust samples and mite-free dust samples (Appendix S1). TaqMan miRNA assay was used to detect the expression of miR-5735-3p in the samples. This miRNA was chosen as it is specific in DPT and spider mites *Tetranychus urticae*. The abundance of miR-5735-3p in different samples was plotted in Figure 1B. It validated the presence of miR-5735-3p in DPT and showed that the miRNA was found in household dust samples. The significantly higher level of miR-5735-3p in dust samples compared to mite-free samples indicated the presence of DPT in the dust samples. It also showed the levels of miRNA in mite-free dust samples were significantly higher than those in negative control, indicating that miRNA could be exogenously stable in dust environment outside mite bodies. Hence, due

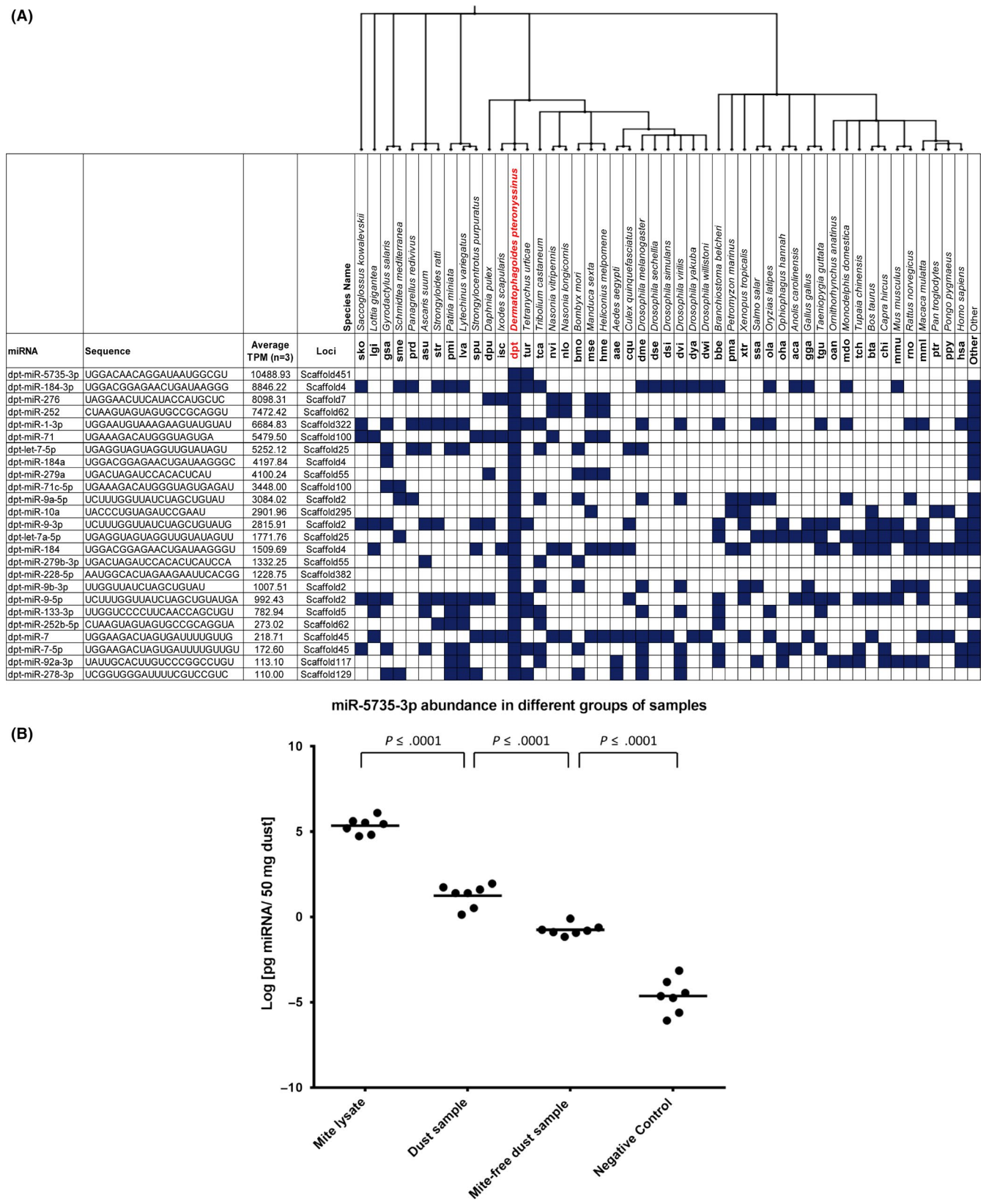


FIGURE 1 A, Conserved and specific miRNAs in DPT. TPMs were calculated by the average TPM from 3 samples. Loci referred to the location of miRNA in DPT genome ⁷. The heatmap plotted the miRNA homologues in different species, the box highlighted in blue means the homologue is presented in the species. We only showed the miRNAs with top expression (TPM>100) in this table. The phylogenetic tree of the species was generated by Interactive tree of life. The full table of identified 69 conserved and 52 specific miRNAs was provided in the Table S2. B, Taqman miRNA assay of miR-5735-3p in different dust samples. There were statistically significant differences in the quantity of miR-5735-3p in different batches of dust samples

TABLE 1 MiRNA target prediction and functional annotation

miRNA	KEGG immune system pathway (human)	Hits	P-value	Targeted genes
dpt-miR-7-5p	B cell receptor signaling pathway	10	.001	AKT3, GRB2, MAP2K2, NRAS, PIK3CB, PIK3CD, PIK3CG, PIK3R3, RAF1, RELA
	Chemokine signaling pathway	7	8.70E-04	PAK1, PIK3CD, PIK3CG, PIK3R3, PTK2, RAF1, RELA
	Fc epsilon RI signaling pathway	10	8.60E-04	AKT3, GRB2, MAP2K2, MAPK9, NRAS, PIK3CB, PIK3CD, PIK3CG, PIK3R3, RAF1
	Fc gamma R-mediated phagocytosis	11	.001	AKT3, CRKL, MARCKSL1, PAK1, PIK3CB, PIK3CD, PIK3CG, PIK3R3, PLPP3, RAF1, WASF2
	Natural killer cell-mediated cytotoxicity	5	.002	PAK1, PIK3CD, PIK3CG, PIK3R3, RAF1
	T cell receptor signaling pathway	12	.001	AKT3, GRB2, MAP2K2, MAPK9, NRAS, PAK1, PIK3CB, PIK3CD, PIK3CG, PIK3R3, RAF1, RELA
	Toll-like receptor signaling pathway	10	.008	AKT3, FADD, MAP2K2, MAPK9, PIK3CB, PIK3CD, PIK3CG, PIK3R3, RELA, TLR4
dpt-let-7a-5p	Natural killer cell-mediated cytotoxicity	6	.001	CASP3, HRAS, KRAS, NRAS, PAK1, TNFRSF10B

Note: This table listed the pathways with P -value $< .01$. The genes listed here belong to humans. Since these miRNAs were found to be identical to their homologous miRNAs in humans, we further checked if their targeted genes could be enriched in allergy-related pathways.

to the stability and small particular size of miRNA, we speculated that miRNA could mix with mite fecal pellets or dust and inhaled by humans. Besides, overexpression of mite-specific miR-5735-3p in immortalized human bronchial epithelial cell line BEAS-2B significantly affected type I interferon signaling pathway (Appendix S1 and Table S5). This pathway is known to induce the immunoregulatory factor IL-17, which contributes to various allergic diseases such as allergic asthma, rhinitis, and dermatitis.⁹

Given that many of the miRNAs identified in DPT are homologous to human miRNAs, and studies are supporting the cross-species transfer and host cell regulation by miRNA, it is tempting to speculate that dust mite miRNA may play a role in human allergic diseases. Experiments in animal models will be required to test the possibility of the entry of mite miRNA into respiratory cells. This is the first systemic investigation of dust mite miRNAs and a comprehensive picture of miRNAs in DPT was shown. In conclusion, our study had laid a foundation for HDM miRNA research and these miRNAs will provide novel targets for further study in allergic diseases.

ACKNOWLEDGMENTS

The authors wish to thank Dr YANG Yi Kevin and Dr KWOK Sui-Lam Jamie for their useful suggestions in miRNA identification, Dr WANG Mingqiang for the genome assembly of dust mite, Dr ZENG Xi for his excellent transcriptome knowledge, and Dr LAU Po-Ying Carol for sample collection. This work is supported by the Hong Kong Research Grants Council (475113; 14175617; T12-402/13-N), NSFC (31328014, 91442118, 91542104), Guangdong Technology Foundation (2014B090901041, 2016A020216029), State Key Research Scheme on Precision Medicine (2016YFC0905802), State Key Laboratory of Respiratory Disease (SKLRD2016ZJ001).

CONFLICT OF INTEREST

The authors have nothing to disclose.

FUNDING INFORMATION

Supported by Hong Kong Research Grants Council (475113; 14175617; T12-402/13-N), NSFC (31328014, 91442118, 91542104), Guangdong Technology Foundation (2014B090901041, 2016A020216029), State Key Research Scheme on Precision Medicine (2016YFC0905802), State Key Laboratory of Respiratory Disease (SKLRD2016ZJ001).

Zhiyuan Yang^{1,2,3}

Angel Tsz-Yau Wan^{2,3}

Xiao-Yu Liu⁴

Qing Xiong^{2,3}

Zhi-Gang Liu⁴

Stephen Kwok-Wing Tsui^{2,3,5} 

¹College of Life Information Science & Instrument Engineering, Hangzhou Dianzi University, Hangzhou, China

²School of Biomedical Sciences, The Chinese University of Hong Kong, Shatin, Hong Kong

³Hong Kong Bioinformatics Centre, The Chinese University of Hong Kong, Shatin, Hong Kong

⁴State Key Laboratory of Respiratory Disease for Allergy at Shenzhen University, School of Medicine, Shenzhen University, Shenzhen, China

⁵Centre for Microbial Genomics and Proteomics, Faculty of Medicine, The Chinese University of Hong Kong, Shatin, Hong Kong

Correspondence

Stephen Kwok-Wing Tsui, School of Biomedical Sciences, the Chinese University of Hong Kong, Shatin, N.T., Hong Kong.

Email: kwtsui@cuhk.edu.hk

ORCID

Stephen Kwok-Wing Tsui  <https://orcid.org/0000-0003-0686-4259>

Yang, Wan, Liu contributed equally to this work.

REFERENCES

1. Calderón MA, Kleine-Tebbe J, Linneberg A, et al. Dust mite respiratory allergy: an overview of current therapeutic strategies. *J Allergy Clin Immunol Pract*. 2015;3:843-855.
2. Lodge CJ, Lowe AJ, Gurrin LC, et al. House dust mite sensitization in toddlers predicts current wheeze at age 12 years. *J Allergy Clin Immunol*. 2011;128:782-788.e9.
3. Pua H, Steiner D, Patel S, et al. MicroRNAs 24 and 27 suppress allergic inflammation and target a network of regulators of T Helper 2 cell-associated cytokine production. *Immunity*. 2016;44:821-832.
4. Vaucheret H, Chupeau Y. Ingested plant MiRNAs regulate gene expression in animals. *Cell Res*. 2012;22:3-5.

5. Lamonte G, Philip N, Reardon J, et al. Translocation of sickle cell erythrocyte MicroRNAs into Plasmodium falciparum inhibits parasite translation and contributes to malaria resistance. *Cell Host Microbe*. 2012;12:187-199.
6. Buck AH, Coakley G, Simbari F, et al. Exosomes secreted by nematode parasites transfer small RNAs to mammalian cells and modulate innate immunity. *Nat Commun*. 2014;5:5488.
7. Liu X-Y, Yang KY, Wang M-Q, et al. High-quality Assembly of dermatophagoides pteronyssinus genome and transcriptome reveals a wide range of novel allergens. *J Allergy Clin Immunol*. 2018;141:2268-2271.e8.
8. Polikepahad S, Knight JM, Naghavi AO, et al. Proinflammatory role for let-7 MicroRNAs in experimental asthma. *J Biol Chem*. 2010;285:30139-30149.
9. Oldford SA, Salsman SP, Portales-Cervantes L, et al. Interferon α 2 and interferon γ induce the degranulation independent production of VEGF-A and IL-1 receptor antagonist and other mediators from human mast cells. *Immunity, Inflamm Dis*. 2018;6:176-189.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

Supplementary documents

Identification, functional annotation and stability analysis of miRNA in *Dermatophagoides pteronyssinus*

Material and methods

Sample collection and small RNA sequencing (sRNA-Seq)

DPT mites were cultured and their RNAs were extracted according to the previous method with some modifications ¹. In brief, DPT mites were cultured at 25 °C in a small air-filtered room at 70-75% relative humidity for 50 days. Total RNA was isolated from the homogenate with Trizol followed by PureLink RNA Mini Kit (Thermo Fisher Scientific). The molecules in the range of 18-44nt were enriched by polyacrylamide gel electrophoresis. These small RNAs were ligated with adapters and were used as templates for cDNA library. The quantity and quality of the RNA sample were measured with a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific) and Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). Next-generation sequencing was performed using Illumina HiSeq 2000 to produce pair-end reads at 90bp per end. . The raw sequencing data had been uploaded to the NCBI database (SRA accession: PRJNA577524).

Quality control and mapping to the genome

Before subsequent analysis, data preprocessing procedures should be performed to extract the relevant information from the raw data. The sequenced reads usually contain 3' adapters and these adapters were trimmed. Mature miRNA is usually longer than 18bp, thus the short reads less than 18bp were also removed ². The lengths of these clean reads were calculated, and the length distribution was plotted.

The genome of DPT has been assembled and reported in our previous work ³. The raw reads that passed the quality control were mapped to the genome by the tool Bowtie ⁴. Both the positive and negative strands were considered, and the tolerance is set by mismatch=1 in the parameter of Bowtie. Also, the loci of the mapped reads in DPT genome have been recorded. To consolidate our results, the reads that failed to map to the genome were discarded in the subsequent study. The remaining reads were denoted as Reads Subset S1 (**Figure S2**).

Identify conserved miRNAs in DPT

As the miRNA profile of *D. pteronyssinus* is currently unknown, we try to infer *D. pteronyssinus* miRNAs from their homologs in other species. The identified miRNAs were categorized into two groups (conserved and specific miRNAs) based on whether they were exactly matched to currently known miRNAs in miRBase database. If a DPT miRNA is identical to at least one of the miRNAs in the database, the miRNA was defined as “conserved miRNA”. If the DPT miRNA showed some mismatches with known miRNAs, this type of miRNA was denoted as “specific miRNA”. Different methods were applied to identify these two kinds of miRNAs in DPT.

All sequences of the mature miRNAs in metazoan clade were downloaded from miRBase (version 21) ⁵. We aligned the Reads Subset S1 against all known metazoan miRNAs in miRBase by BLAST ⁶. The table format of BLAST alignment results was stored.

The alignment satisfying the following criteria were selected by Perl script:

- (1) The length of query sequence (reads) $\geq$ alignment length
- (2) The length of subject sequence (known miRNA) $\geq$ alignment length
- (3) The alignment identity is 100%;

These criteria could ensure an exact-matched alignment and a full identical read sequence of known miRNA. Then, the query reads sequence could be annotated by the

conserved miRNA sequence. The reads that did not satisfy these criteria were denoted as Reads Subset S2.

Identify specific miRNAs in DPT

The methods for identification of specific miRNAs in DPT were based on previous similar studies in other species ⁷⁻⁹. The dominated types of small non-coding RNAs (sncRNA) includes microRNAs (miRNAs), PIWI-interacting RNAs (piRNAs), short-interfering RNAs (siRNAs), small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs) ¹⁰. To consolidate our results, we have identified other types of small ncRNA and removed them in the process.

All of the available non-coding RNAs sequences were downloaded from NONCODE database (version 3) ¹¹. The Reads Subset S2 were aligned against these 177,863 small non-coding RNAs by BLAST. The same criteria as the identification of conserved miRNAs listed above were applied to identify those reads containing a full identical sequence of small non-coding RNAs. These reads were annotated according to their types and were removed from the following analysis.

The coding DNA sequence (CDS) of *Tetranychus urticae* (Two-spotted spider mite), *Tribolium castaneum* (beetle), *Drosophila melanogaster* (fruitfly), *Apis mellifera* (bee) and *Aedes aegypti* (mosquito) were downloaded from Ensembl database ¹². The miRNAs that could map to the CDS of these species were also removed. The remaining reads were denoted as Reads Subset S3 and were further aligned against sequences of known metazoan miRNAs by BLAST. The following criteria were set to select reads that were highly similar to known metazoan miRNAs.

- (1) The length of query sequence (reads) $\geq$ alignment length
- (2) The length of subject sequence (known miRNA) $\geq$ alignment length
- (3) Alignment mismatch ≤ 3 ;
- (4) Gap opening=0 in the alignment;

The reads that did not satisfy these criteria were discarded and the remaining reads were regarded as potential miRNA candidate in DPT. To ensure the accuracy of the miRNA candidate identified, the structures of precursor miRNAs (pre-miRNAs) generating these miRNA candidates were further checked based on thermodynamics theory. The sequences flanking the miRNA candidates (100bp upstream and 100bp downstream) were captured from the DPT genome. These sequences could be regarded as the transcribed pre-miRNA fragment of correspondent miRNA ¹³. The secondary structures of the captured pre-miRNA fragments were predicted with minimum free energy theory using ViennaRNA ¹⁴. Among these pre-miRNA fragments, those meeting the following rigorous criteria were annotated as DPT specific miRNAs:

- (1) The RNA sequence could fold into an appropriate stem-loop hairpin structure;
- (2) The mature miRNA sequences were in one arm of the hairpin structure;
- (3) Less than or equal to five mismatches were presented in the paired nucleotide structure of mature miRNA sequence.
- (4) No loop or break present in the structure of mature miRNA sequence.

These criteria could ensure the accuracy of the identified miRNAs. Further work was conducted to identify the distribution of these miRNAs in other species. The homologous number of miRNAs in arthropoda was calculated.

MiRNA distribution in Arthropoda clade

All known miRNAs of 32 kinds of Arthropods were extracted from miRBase database. As dust mite belongs to the Arthropoda clade, we further investigated if these miRNAs were presented in other Arthropods. The distribution of homologs in different species was plotted. To avoid redundancy of the miRNAs distribution result, only one species was selected for those species in the same genus clade, such as *Drosophila* (12 members) and *Nasonia* (3 members). The homologs of DPT miRNAs in the rest 19 arthropods were identified (**Table S3**)

Transcript quantification and validation of identified miRNAs

The primary application of small RNA sequencing is to estimate the transcript expression. The common measurements of RNA-seq are expressed in RPKM (Reads Per Kilobase per Million mapped reads), FPKM (fragments per kilobase of exon per million fragments mapped) and TPM (transcripts per million) ¹⁵. For miRNA transcript quantification, TPM is a relatively practical measure to quantify miRNA expression because the difference between miRNA lengths is very small ¹⁶. The normalized TPM values¹⁷ of miRNAs were calculated by the following formula:

$$\text{TPM} = \text{miRNA total reads} / \text{total clean reads} \times 10^6$$

The TPM values were also compared between conserved and specific miRNAs.

Validation of the expression of randomly picked miRNAs was performed by TaqMan miRNA assay (Applied Biosystems) according to manufacturer's protocol. TaqMan miRNA assay was performed on two miRNA with relatively high abundance (miR-5735-3p and miR-184). MiRNAs were extracted from mite lysate using mirVana™ miRNA Isolation Kit (Ambion). The quantity and quality of the RNA sample were measured with a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific). Expression of mature miRNAs was determined by the TaqMan miRNA-assay (Applied Biosystems), and normalized using the relative standard curve method (**Figure S3**). All TaqMan-PCRs were performed in triplicates.

MiRNA target prediction and functional annotation

MiRNA targets were predicted according to the method reported by Jatin Shrinet *et al* ¹⁸. In brief, 3'UTR (three prime untranslated region) sequences of *Tribolium castaneum* and *Drosophila melanogaster* were downloaded from UTRdb database ¹⁹. As there were no available 3'-UTR sequences of other closer Chelicerata species (such as spider mite and deer tick), 3'-UTR sequences of *Drosophila* and flour beetle were used here. The homologous genes between dust mite and *drosophila*, dust mite and flour beetle were also identified to ensure the presence of these target genes in dust mite. There are

662,288 curated 3' UTR sequences available in the database that were experimentally validated in diverse species. These sequences were subjected to target prediction using RNAhybrid tool²⁰. Target predictions were performed based on the following criteria:

- (1) Complete complementarity of the 3' UTR and the DTP miRNA sequence
- (2) $p\text{-value} \leq 0.01$
- (3) The minimal free energy of miRNA-mRNA ≤ -25 Kcal/mol

The target genes of a single miRNA were enriched by DAVID tool²¹.

For those miRNAs identical to human homologs, we have further checked if the targeted genes could be enriched in allergy-related pathways.

Detection of miRNA in dust samples

In an attempt to study the stability of mite miRNAs in the environment, we have performed a Taqman miRNA expression assay to test the quantity of mite-specific miRNA in household dust samples²². Seven dust samples were collected from pillows, mattresses, sofa and chairs from 3 houses in Hong Kong using a vacuum cleaner. Each dust sample collected was separated into two parts and treated differently. The first batch (named as Dust sample) was immersed in pure water and the supernatant was collected to extract miRNA using mirVana™ miRNA Isolation Kit (Ambion). And the second batch (named as Mite-free dust sample) was immersed in water and the supernatant was filtered through 0.45 µm Sartorius Minisart syringe filters (Sartorius Stedim Biotech). In detail, dust samples were collected from the vacuum cleaner using a dust bag, and the dust samples were weighted in a 15 ml falcon tube and then immersed with 5 ml deionized water. The sample was then centrifuged for 1 min at 1000g in room temperature. Supernatants were collected and 1 ml of the supernatant was used for miRNA extraction. The remaining supernatants were filtered through the 0.45 µm syringe filter as mentioned and 1 ml of the flow-through was used for miRNA isolation. TaqMan miRNA-assay (Applied Biosystems) was used to detect the expression of miR-5735-3p in the samples. This miRNA was chosen as it can only be found in DPT and spider mites *Tetranychus*

urticae. All TaqMan-PCRs were performed in triplicates and the statistical analysis was calculated by ANOVA.

Isolation of exosome from dust samples

After collection of dust samples, the supernatant was combined with 1/2 volume of Total exosome isolation (from cell media) reagent (Invitrogen) and mixed well until a homogenous solution was formed. The samples were incubated at 4°C overnight and then centrifuged at 4°C at 10,000 g for 1 hr. The supernatant was aspirated and discarded, and the exosome pellet was resuspended in PBS buffer. Total miRNAs were then extracted from the samples using the mirVana™ miRNA Isolation Kit (Ambion) following manufacturer's protocol and TaqMan miRNA-assay (Applied Biosystems) was used to detect the expression of miR-5735-3p in the samples.

Overexpression of mite specific miRNA in human cell model

BEAS-2B, the SV40 large T antigen immortalized human bronchial epithelial cell line derived from normal human bronchial epithelium, was used in this experiment. BEAS-2B cells were purchased from ATCC (CRL-9609) and maintained with LHC-8 medium (Gibco, Carlsbad, CA). The cells were grown at 37°C and 5% CO₂, and fresh media was added every two to three days. All assays were performed in two separate wells with independent cell population ($n=2$).

MiR-5735-3p lentiviral construct was generated as follows: Two complementary single-strand oligonucleotides (ODN) encoding the sequence of the pre-miRNAs were designed and synthesized. Then the 5' phosphorylated, PAGE purified oligos were synthesized (Integrated DNA Technologies, USA). Here showed the ODN sequences designed:

5'-TGACGCTACTATCCTGTTGTCCATTCAAGAGATGGACAACAGGATAATGGCGTTTTTTC-3'
5'-TCGAGAAAAACGCCATTATCCTGTTGTCCATCTCTTGAATGGACAACAGGATAGTAGCGTCA-3'

After that, the single-stranded ODN were hybridized. In detail, 1 μ l of sense oligo and 1 μ l antisense oligo were mixed with 25 μ l annealing buffer and made up to 50 μ l. The mixture was incubated at 95 °C for 4 mins following by 70 °C for 10 mins. After that, the annealed oligos were slowly cooled down to 4 °C. They were then cloned into the pSico vector and the plasmid was named as pSico-miR-5735-3p. Ligation of the phosphorylated oligo and the HpaI-XhoI digested pSico was performed. This plasmid pSico expressed EGFP and Cre would cause EGFP to be recombined out of the construct, activating miRNA expression in the form of shRNA. The sequences of the clones were confirmed by Sanger sequencing.

The lentivirus particle was produced by calcium phosphate-mediated transfection into 293T cells as described. In brief, the subconfluent 293T cells were transfected with 3.5 μ g packaging-plasmid (pCMV-VSVG/ pMD2.G), 6.5 μ g envelope-plasmid (pCMV-deltaR8.47/ psPAX2) and 10 μ g lentiviral transgene plasmid were diluted in 600 μ L OPTI-MEM medium and transfected into 293T cells by lipofectamine 3000. Medium was collected at 48 hrs and 72 hrs post-transfection, passed through a filter with 40 μ m-pore, and then diluted in 5X PEG-it overnight. In the next day, the solution was spin at 1,500 g for 30 mins at 4 °C and resuspended in PBS. After LV transfection, GFP was observed in cells. Recombinant adenoviral stocks expressing Cre recombinase were used to infect the GFP-positive cells. Five days after adenovirus infections, total RNA was extracted from the GFP-negative cells using Trizol and PureLink RNA Mini Kit (Thermo Fisher Scientific) as described above. RNA-Seq of these samples was performed by Grok Bioscience (Hong Kong) using Illumina HiSeq 4000 platform. Transcriptomic analysis was performed using GRCh38.p12 as reference. Bioinformatics analysis of DEGs was performed by HISAT2²³ and Cufflinks²⁴, and then enriched by DAVID analysis. The enrichment results can be found in **Table S5**.

Table S1 – Basic information about this sRNA-seq study.

Item	DPT1	DPT2	DPT3
Raw reads	25,354,824	29,550,836	29,507,774
Reads total length	102,046,412	132,206,681	132,881,793
Reads minimal length	18	18	18
Reads maximum length	44	44	44
Unique reads count	4,232,816	4,669,139	4,555,659
Average count of unique reads	6	6.4	6.4
reads count N50	24,110	25,313	25,477

Table S2 – The full table of conserve and specific miRNAs in DPT. There are 69 conserved miRNAs and 52 DPT specific miRNAs identified.

No.	miRNA	Sequence	Type	Loci	TPM (average)
1	dpt-miR-5735-3p	UGGACAACAGGAUAAUGGCGU	Conserved	Scaffold451	10488.93
2	dpt-miR-184-3p	UGGACGGAGAACUGAUAAAGGG	Conserved	Scaffold4	8846.22
3	dpt-miR-276	UAGGAACUUCAUACCAUGCUC	Conserved	Scaffold7	8098.31
4	dpt-miR-252	CUAAGUAGUAGUGCCGCAGGU	Conserved	Scaffold62	7472.42
5	dpt-miR-1-3p	UGGAAUGUAAAGAAGUAUGGAG	Conserved	Scaffold322	6684.83
6	dpt-miR-71	UGAAAGACAUGGGUAGUGA	Conserved	Scaffold100	5479.50
7	dpt-let-7-5p	UGAGGUAGUAGGUUGUAUAGU	Conserved	Scaffold25	5252.12
8	dpt-miR-184a	UGGACGGAGAACUGAUAAAGGGC	Conserved	Scaffold4	4197.84
9	dpt-miR-279a	UGACUAGAUAUCCACACUCAU	Conserved	Scaffold55	4100.24
10	dpt-miR-71c-5p	UGAAAGACAUGGGUAGUGAGAU	Conserved	Scaffold100	3448.00
11	dpt-miR-9a-5p	UCUUUGGUUAUCUAGCUGUAU	Conserved	Scaffold2	3084.02
12	dpt-miR-10a	UACCCUGUAGAUAUCCGAAU	Conserved	Scaffold295	2901.96
13	dpt-miR-9-3p	UCUUUGGUUAUCUAGCUGUAUG	Conserved	Scaffold2	2815.91
14	dpt-let-7a-5p	UGAGGUAGUAGGUUGUAUAGUU	Conserved	Scaffold25	1771.76
15	dpt-miR-184	UGGACGGAGAACUGAUAAAGGGU	Conserved	Scaffold4	1509.69
16	dpt-miR-279b-3p	UGACUAGAUAUCCACACUCAUCCA	Conserved	Scaffold55	1332.25
17	dpt-miR-228-5p	AAUGGCACUAGAAGAAUUCACGG	Specific	Scaffold382	1228.75
18	dpt-miR-9b-3p	UUGGUUAUCUAGCUGUAU	Conserved	Scaffold2	1007.51
19	dpt-miR-9-5p	UCUUUGGUUAUCUAGCUGUAUGA	Conserved	Scaffold2	992.43
20	dpt-miR-133-3p	UUGGUCCCCUUAACCAGCUGU	Conserved	Scaffold5	782.94
21	dpt-miR-252b-5p	CUAAGUAGUAGUGCCGCAGGUA	Conserved	Scaffold62	273.02
22	dpt-miR-7	UGGAAGACUAGUGAUUUUUGUUG	Conserved	Scaffold45	218.71
23	dpt-miR-7-5p	UGGAAGACUAGUGAUUUUUGUUGU	Conserved	Scaffold45	172.60
24	dpt-miR-92a-3p	UAUUGCACUUGUCCCGGCCUGU	Conserved	Scaffold117	113.10
25	dpt-miR-278-3p	UCGGUGGGAUUUUCGUCCGUC	Conserved	Scaffold129	110.00
26	dpt-miR-451-5p	AAACCGUUACCAUACUGA	Conserved	Scaffold262	79.88
27	dpt-miR-252a	CUAAGUAGUAGUGCCGCAGGUGU	Specific	Scaffold62	73.02
28	dpt-miR-133a	UUUGGUCCCCUUAACCAG	Conserved	Scaffold322	49.63
29	dpt-miR-133a-3p	UUUGGUCCCCUUAACCAGCUG	Conserved	Scaffold5	48.75
30	dpt-miR-276-3p	UAGGAACUUCAUACCGUGCUC	Conserved	Scaffold7	31.96
31	dpt-miR-263a-3p	UGGCUUCUUUAGUGCUUUGA	Specific	Scaffold382	30.33
32	dpt-miR-183-5p	AUGGCACUGGAAGAAUUCA	Specific	Scaffold382	25.31
33	dpt-miR-1a-5p	UGGAAUGUGAAGAAGUAUGG	Specific	Scaffold5	17.92
34	dpt-miR-71a	UGAAAGACGUGGGUAGUGAGA	Specific	Scaffold100	17.76
35	dpt-miR-71-3p	UCUCACUACCUUGUCUUUGGCG	Specific	Scaffold100	15.00

36	dpt-miR-183	AAUGGCACUGGAAGAAUUCACGG	Specific	Scaffold382	14.93
37	dpt-let-7c	UGAGGUAGUAGGUUGUAUGGU	Conserved	Scaffold25	11.19
38	dpt-miR-276a-3p	UAGGAACUUCAUACCGUGCUCU	Conserved	Scaffold7	10.41
39	dpt-let-7f	UGAGGUAGUAGAUUGUAUAGU	Conserved	Scaffold25	9.41
40	dpt-miR-71a-5p	UGAAAGACACGGGUAGUGAGAU	Conserved	Scaffold100	7.69
41	dpt-miR-234	UUAUUGCUUGAGAAUACA	Conserved	Scaffold292	7.11
42	dpt-miR-137	UUAUUGCUUGAGAAUACACGU	Specific	Scaffold292	5.65
43	dpt-miR-1260a	AUCCCACCUCUGCCACCA	Conserved	Scaffold460	5.59
44	dpt-let-7f-5p	UGAGGUAGUAGAUUGUAUAGUU	Conserved	Scaffold25	5.38
45	dpt-miR-9	UCUUUGGUUAUCUAGCUGUGUG	Conserved	Scaffold2	5.06
46	dpt-miR-71-5p	AGAAAGACAUGGGUAGUGAGAU	Conserved	Scaffold100	4.39
47	dpt-miR-92	UAUUGCACUCGUCCCCGCCU	Conserved	Scaffold117	3.86
48	dpt-miR-124	UAAGGCACGCGGUGAAUGC	Conserved	Scaffold49	3.77
49	dpt-let-7c-5p	UGAGGUAGUAGGUUGUAUGGUU	Conserved	Scaffold25	3.74
50	dpt-miR-9c-5p	UCUUUGGUGAUCUAGCCGUGUG	Specific	Scaffold2	3.62
51	dpt-miR-252a-5p	CUAAGUACUAGUGCCGCAGG	Conserved	Scaffold62	3.11
52	dpt-let-7	UGAGGUAGUUGGUUGUAUAGU	Conserved	Scaffold25	3.03
53	dpt-miR-307	UCACAACCUCCUUGAGUGAG	Specific	Scaffold6	2.90
54	dpt-miR-9e-5p	UCUUUGGUUAUCUAGUUGUAU	Specific	Scaffold2	2.84
55	dpt-miR-81a	UGAGAUCAUUGUGAAAGCUAUU	Conserved	Scaffold82	2.80
56	dpt-miR-124-3p	UAAGGCACGCGGUGAAUGCC	Conserved	Scaffold49	2.72
57	dpt-miR-2b-3p	UAUCACAGCCAGCUUUGGUGAGU	Specific	Scaffold100	2.70
58	dpt-miR-79-5p	UUUGGUUAUCUGGCUGUAUGA	Conserved	Scaffold2	2.19
59	dpt-miR-451a	AAACCGUUAACCAUACUGAGUU	Conserved	Scaffold262	1.57
60	dpt-miR-124e-3p	UAAGGCACGCGGUGAAUGCCA	Specific	Scaffold49	1.55
61	dpt-miR-96	UUUGGCACUAGCACAUUUUUG	Conserved	Scaffold382	1.52
62	dpt-let-7e	UGAGGUAGGAGGUUGUAUAGU	Conserved	Scaffold25	1.36
63	dpt-miR-13a-3p	UAUCACAGCCACUUUGAUUAGC	Specific	Scaffold100	1.36
64	dpt-miR-279-3p	UGACUAGAUCACAUUCAUCCA	Conserved	Scaffold55	0.89
65	dpt-let-7g	UGAGGUAGUAGUUUGUAUAG	Conserved	Scaffold25	0.84
66	dpt-miR-5735-5p	ACCUUUCUCUUGUUGUCCACUG	Conserved	Scaffold451	0.81
67	dpt-let-7e-5p	UGAGGUAGGAGGUUGUAUAGUU	Conserved	Scaffold25	0.70
68	dpt-miR-10b	UACCCUGUAGAACCGAAU	Specific	Scaffold295	0.62
69	dpt-miR-96-5p	UUUGGCACUGGCACAUAUUAU	Specific	Scaffold382	0.59
70	dpt-miR-7b	UGGAAGACUUGUGAUUUUUGUU	Conserved	Scaffold45	0.56
71	dpt-miR-263	AAUGGCACUUGAAGAAUUCACGG	Specific	Scaffold382	0.51
72	dpt-miR-228	AAUGGCACUAGAUGAAUUCACGG	Conserved	Scaffold382	0.51
73	dpt-miR-1422f-3p	AAAGAAUAGUCAUGACUGAU	Specific	Scaffold146	0.50
74	dpt-miR-315	UUUUGAUUGUAGCUCAGAAAGC	Specific	Scaffold167	0.46
75	dpt-miR-79	AUAAAGCUAGGUUACCAAAGUUA	Conserved	Scaffold2	0.40
76	dpt-miR-2444	UUUUUGUUGUUUUCUGUUUU	Specific	Scaffold351	0.38
77	dpt-bantam-3p	UGAGAUCAUUGUGAAAGCUGAUU	Conserved	Scaffold82	0.36
78	dpt-miR-9b	UCUUUGGUUUUCUAGCUGUAUG	Conserved	Scaffold2	0.34

79	dpt-miR-133b	UUUGGUCCCCUUAACCAGCUA	Conserved	Scaffold322	0.33
80	dpt-miR-993a-3p	GAAGCUCGUUUCUACAGGUUUC	Conserved	Scaffold379	0.30
81	dpt-miR-92b-3p	UAUUGCACUCGUCCCGGCCUCC	Conserved	Scaffold117	0.28
82	dpt-miR-4967-5p	UGCAUUGUAUCUCGGUUAAC	Specific	Scaffold296	0.22
83	dpt-miR-2a-3p	UAUCACAGCCAGCUUUGGUGAGC	Specific	Scaffold100	0.22
84	dpt-let-7j-5p	UGAGGUAGUAGGUUGGAUAGUU	Conserved	Scaffold25	0.21
85	dpt-miR-17	CAAAGUGCUUACAGUGCA	Conserved	Scaffold47	0.19
86	dpt-miR-285	UAGCACCAUCAGAAUUCAAUUU	Specific	Scaffold176	0.18
87	dpt-miR-252-5p	CUAAGUACUAGUGCCGCAGGAG	Conserved	Scaffold62	0.13
88	dpt-miR-10b-5p	UACCCUGUAGAACCGAAUUUGUG	Conserved	Scaffold295	0.12
89	dpt-miR-190-5p	AGAUUAUGUUUGAUUAUUCUUGGUUG	Conserved	Scaffold17	0.11
90	dpt-miR-6491	GCGGCAGCCCUGGGAAGA	Specific	Scaffold272	0.10
91	dpt-miR-7964a-3p	GUCAACCACGCUCUGCUACC	Specific	Scaffold82	0.09
92	dpt-miR-278	UCGGUGGGAAUUUUCGUCCGUUU	Specific	Scaffold129	0.08
93	dpt-miR-2293	UGAUUUUGUUGAUUUUGAAUU	Specific	Scaffold2	0.08
94	dpt-miR-3878-3p	AUUUUUAUUGUCGUUGUUG	Specific	Scaffold57	0.08
95	dpt-miR-4206-3p	AUUGGAUGGUUUAGUGAG	Specific	Scaffold272	0.07
96	dpt-miR-4770	UGAGUUGACACUGUCGCU	Specific	Scaffold1	0.06
97	dpt-miR-37b-5p	AAGUACUGGUUUCGAUGUUGAU	Specific	Scaffold337	0.06
98	dpt-miR-190a-5p	UGAUUAUGUUUGAUUUUUUGGUUG	Specific	Scaffold17	0.06
99	dpt-miR-219a-5p	UGAUUGUCCAAACGCAAUUCU	Conserved	Scaffold5	0.06
100	dpt-miR-79-3p	AUAAAGCUAGGUUACCAAAGCU	Conserved	Scaffold2	0.04
101	dpt-miR-3613	UGUUGUUGUUUUUUUUUGU	Specific	Scaffold90	0.04
102	dpt-miR-1c-3p	UGGAAUGUUUUGACAAGUGUC	Specific	Scaffold312	0.04
103	dpt-miR-993-3p	GAAGCUCGUUUCUACAGGUUUCU	Specific	Scaffold379	0.04
104	dpt-miR-36c-3p	UCACCGGAUAGUGAUUCAUU	Specific	Scaffold10	0.03
105	dpt-miR-4198-5p	AAUGAUGUAAUAUUAUUG	Specific	Scaffold52	0.03
106	dpt-miR-124-2-5p	GUGUUCACUGUUUGCCUGCAUG	Specific	Scaffold49	0.02
107	dpt-miR-574-5p	UGAGUGUGUGUGUGUGAGUGUGU	Specific	Scaffold90	0.02
108	dpt-miR-184-5p	CCUUAUCAUUUCCGCUGUCCAGA	Conserved	Scaffold4	0.01
109	dpt-miR-2284a	AAAAAAGUUUGAUGUGGUUUU	Specific	Scaffold8	0.01
110	dpt-miR-2325a	GUUUUGUUUCGGUUGUUUUC	Specific	Scaffold49	0.01
111	dpt-miR-8418-5p	UCAUUUCGUUAAUUCCGA	Specific	Scaffold3	0.01
112	dpt-miR-991-3p	UUAAAGUCGAAUUUUGAAAAAU	Specific	Scaffold106	0.01
113	dpt-miR-124a	UAAGGCGCGCGGUGAAUGCC	Specific	Scaffold49	0.01
114	dpt-miR-4121-3p	GUGUUGGAGUUGAUUUUGU	Specific	Scaffold194	0.01
115	dpt-miR-3486-5p	UCUGUGUGUUGAAAUUGA	Specific	Scaffold50	0.01
116	dpt-miR-1814	GUUUUGUUUUCGUUUGUU	Specific	Scaffold105	0.01
117	dpt-miR-1207-3p	UCAGUUGGCCAUCAUUUC	Specific	Scaffold67	0.01
118	dpt-miR-6002-5p	UCGAAGCAAUGGGUAUUCC	Specific	Scaffold70	0.01
119	dpt-miR-3596a	AACUAUACAACCUACUACCUCA	Conserved	Scaffold25	0.01
120	dpt-miR-574	UGAGUGUGUGUGUGUGAGUGUGUG	Specific	Scaffold376	0.01
121	dpt-miR-466i-5p	UGUGUGUGUGUGUGUGUGUGUG	Specific	Scaffold37	0.01

Table S3 – The homologous numbers of DPT miRNA in Arthropoda. For the species in the same genus, only one member was shown in this table.

Order	Species	Latin Name	Total number	Conserved miRNAs	Specific miRNAs
1	tca	<i>Tribolium castaneum</i>	27	10	17
2	bmo	<i>Bombyx mori</i>	26	9	17
3	dme	<i>Drosophila melanogaster</i>	26	8	18
4	aae	<i>Aedes aegypti</i>	21	7	14
5	tur	<i>Tetranychus urticae</i>	20	3	17
6	cqu	<i>Culex quinquefasciatus</i>	18	5	13
7	ame	<i>Apis mellifera</i>	14	5	9
8	dpu	<i>Daphnia pulex</i>	14	5	9
9	hme	<i>Heliconius melpomene</i>	13	3	10
10	mse	<i>Manduca sexta</i>	13	3	10
11	isc	<i>Ixodes scapularis</i>	12	4	8
12	api	<i>Acyrtosiphon pisum</i>	11	4	7
13	aga	<i>Anopheles gambiae</i>	10	3	7
14	nvi	<i>Nasonia vitripennis</i>	10	3	7
15	pxy	<i>Plutella xylostella</i>	7	2	5
16	lmi	<i>Locusta migratoria</i>	2	0	2
17	mja	<i>Marsupenaeus japonicus</i>	1	1	0
18	rmi	<i>Rhipicephalus microplus</i>	0	0	0
19	smr	<i>Strigamia maritima</i>	0	0	0

Table S4 – miRNA target prediction and functional annotation. The genes listed here belong to humans. Since these miRNAs were found to be identical to their homologous miRNAs in humans, we further checked if their targeted genes could be enriched in allergy-related pathways.

miRNA	KEGG immune system pathway (human)	hits	p-value	targeted genes
dpt-miR-7-5p	B cell receptor signaling pathway	10	0.001	<i>AKT3, GRB2, MAP2K2, NRAS, PIK3CB, PIK3CD, PIK3CG, PIK3R3, RAF1, RELA</i>
	Chemokine signaling pathway	7	8.7E-04	<i>PAK1, PIK3CD, PIK3CG, PIK3R3, PTK2, RAF1, RELA</i>
	Fc epsilon RI signaling pathway	10	8.6E-04	<i>AKT3, GRB2, MAP2K2, MAPK9, NRAS, PIK3CB, PIK3CD, PIK3CG, PIK3R3, RAF1</i>
	Fc gamma R-mediated phagocytosis	11	0.001	<i>AKT3, CRKL, MARCKSL1, PAK1, PIK3CB, PIK3CD, PIK3CG, PIK3R3, PLPP3, RAF1, WASF2</i>
	Leukocyte transendothelial migration	4	0.016	<i>PIK3CD, PIK3CG, PIK3R3, PTK2</i>
	Natural killer cell mediated cytotoxicity	5	0.002	<i>PAK1, PIK3CD, PIK3CG, PIK3R3, RAF1</i>
	T cell receptor signaling pathway	12	0.001	<i>AKT3, GRB2, MAP2K2, MAPK9, NRAS, PAK1, PIK3CB, PIK3CD, PIK3CG, PIK3R3, RAF1, RELA</i>
	Toll-like receptor signaling pathway	10	0.008	<i>AKT3, FADD, MAP2K2, MAPK9, PIK3CB, PIK3CD, PIK3CG, PIK3R3, RELA, TLR4</i>
dpt-let-7a-5p	Natural killer cell mediated cytotoxicity	6	0.001	<i>CASP3, HRAS, KRAS, NRAS, PAK1, TNFRSF10B</i>
	Fc epsilon RI signaling pathway	9	0.013	<i>HRAS, KRAS, LYN, MAP2K7, MAPK11, NRAS, PIK3R1,</i>

				<i>PLCG2, VAV2</i>
dpt-miR-1-3p	Chemokine signaling pathway	19	0.036	<i>CCL14, CCL2, CDC42, CRK, CXCL1, CXCL2, CXCL3, CXCL8, GNAI1, GNAI2, GNAI3, GNB1, GNB2, GNG5, PIK3CA, PLCB3, PREX1, RAP1B, SOS2</i>
	Leukocyte transendothelial migration	13	0.048	<i>ACTB, ACTN1, ACTN4, CDC42, F11R, GNAI1, GNAI2, GNAI3, PIK3CA, RAP1B, RASSF5, THY1, VASP</i>
dpt-miR-9-5p	Leukocyte transendothelial migration	50	0.033	<i>ACTN1, ACTN2, ARHGAP35, ARHGAP5, CDC42, CDH5, CLDN10, CLDN14, CLDN18, CLDN2, CLDN7, CLDN8, CTNNA1, CTNNB1, CTNND1, CXCL12, CXCR4, CYBB, EZR, GNAI1, GNAI2, ITGA4, ITGAL, ITGB1, JAM2, JAM3, MLLT4, MMP2, NOX3, OCLN, PIK3CA, PIK3CB, PIK3CD, PIK3R3, PLCG1, PRKCA, PRKCB, PTK2, PTPN11, PXN, RAP1A, RAP1B, RAPGEF4, RASSF5, RHOA, RHOH, ROCK1, TXK, VAV3, VCL</i>
dpt-miR-133b	Leukocyte transendothelial migration	4	0.017	<i>CDC42, CXCR4, MMP9, RHOA</i>
dpt-miR-184	B cell receptor signaling pathway	3	0.049	<i>AKT2, INPPL1, NFATC2</i>
dpt-miR-451a	B cell receptor signaling pathway	2	0.047	<i>AKT1, IKBKB</i>

Table S5 – Gene ontology enrichment analysis of differentially expressed genes between miR-5735-3p overexpression group and control group (Top ten enriched biological processes are shown).

Gene Ontology Term	Count	P-Value	Genes
GO:0060337~type I interferon signaling pathway	8	1.08E-07	<i>EGR1, IFI27, BST2, IFITM1, IFITM2, XAF1, MX1, IFI6</i>
GO:0035456~response to interferon-beta	4	1.73E-05	<i>BST2, IFITM1, IFITM2, XAF1</i>
GO:0001525~angiogenesis	9	5.98E-05	<i>VEGFB, CAV1, ID1, TNFRSF12A, CTGF, TGFBI, RHOB, JAG1, TGFB2</i>
GO:0010951~negative regulation of endopeptidase activity	7	8.97E-05	<i>BST2, SERPINA5, SERPINA4, SERPINA3, SERPINB10, SERPINB2, SLPI</i>
GO:0045071~negative regulation of viral genome replication	5	9.54E-05	<i>BST2, IFITM1, IFITM2, SLPI, MX1</i>
GO:0043066~negative regulation of apoptotic process	11	4.11E-04	<i>VEGFB, EGFR, DUSP1, ID1, TNFRSF10D, SERPINB10, SERPINB2, TP63, LEF1, ASNS, MYC</i>
GO:0007219~Notch signaling pathway	6	6.48E-04	<i>HES1, KRT19, CHAC1, TP63, JAG1, MYC</i>
GO:0006641~triglyceride metabolic process	4	0.001202	<i>CEL, CAV1, APOE, APOC1</i>
GO:0000122~negative regulation of transcription from RNA polymerase II promoter	13	0.001273	<i>EGR1, CAV1, TP63, LEF1, ANKRD1, NRIP1, HES1, IFI27, DKK1, ID1, BHLHE40, TCF4, MYC</i>
GO:0035455~response to interferon-alpha	3	0.001562	<i>BST2, IFITM1, IFITM2</i>

Table S6 – Expression of genes involved in type I interferon signaling pathway.

Gene	Expression (Control)	Expression (Treatment)	log2(fold_change)	p_value	q_value	CTL1	CTL2	OE1	OE2
<i>IFI27</i>	26.9457	200.204	2.89334	5.00E-05	0.0102	24.18	29.71	241.92	158.49
<i>IFI6</i>	97.0317	393.319	2.01917	5.00E-05	0.0102	104.86	89.21	462.16	324.48
<i>BST2</i>	16.8445	37.9956	1.17356	5.00E-05	0.0102	15.27	18.42	42.87	33.12
<i>MX1</i>	12.0771	25.7932	1.09472	5.00E-05	0.0102	12.95	11.2	28.77	22.81
<i>IFITM1, IFITM2</i>	107.462	201.149	0.90445	5.00E-05	0.0102	108.51	106.42	215.55	186.75
<i>EGR1</i>	86.1418	134.872	0.64681	5.00E-05	0.0102	88.84	83.45	144.16	125.59
<i>XAF1</i>	1.01689	2.45536	1.27177	0.0001	0.0173	1.03	1.01	2.68	2.23

*CTL=Control groups of cells with overexpression of control plasmid, OE= Cells with overexpression of miR-5735-3p.

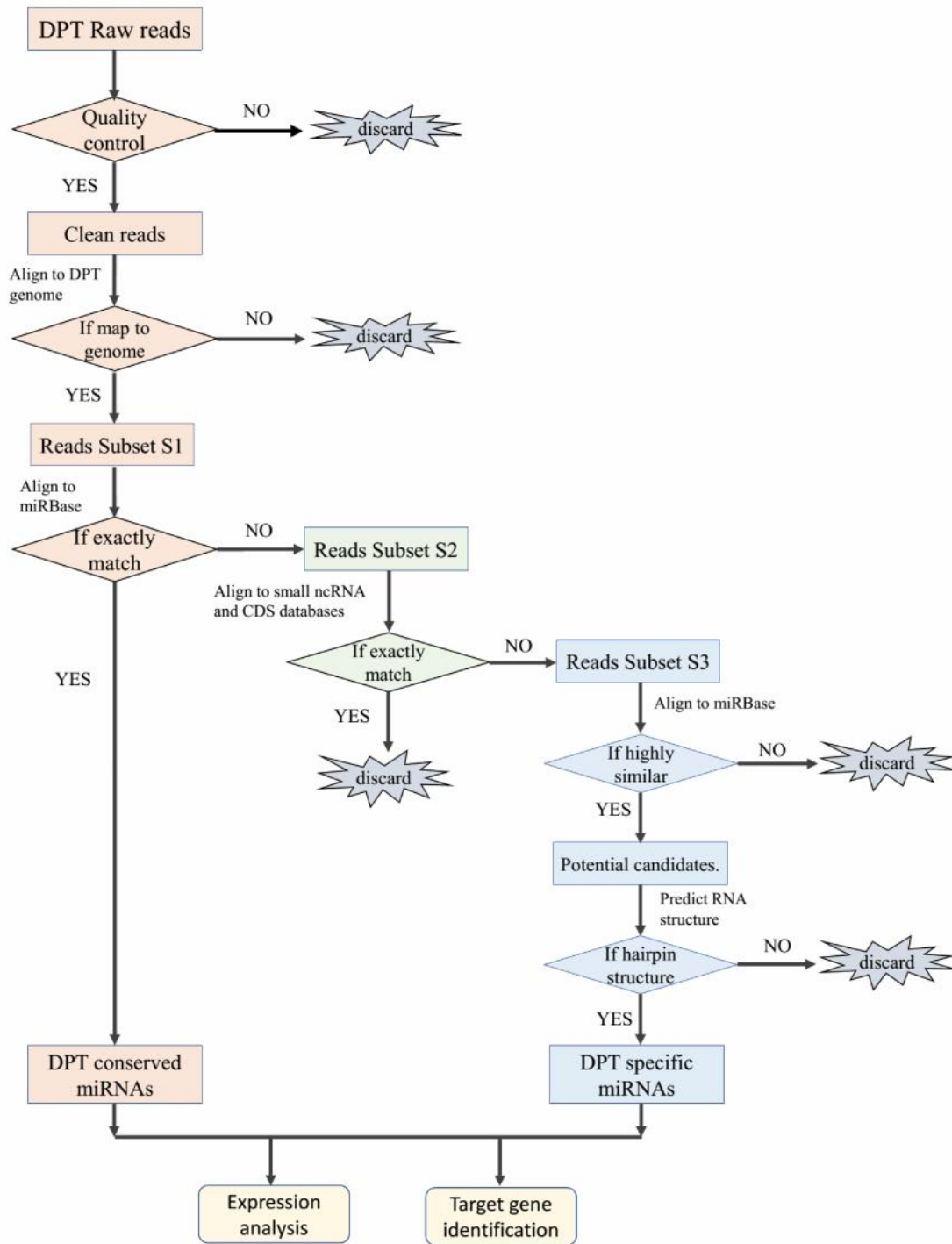


Figure S1 – Flowchart of our work. Our work included two sections of miRNA identification. One section was identification of conserved miRNAs from known homologous sequences, the other section was identification of specific miRNAs from potential candidates.

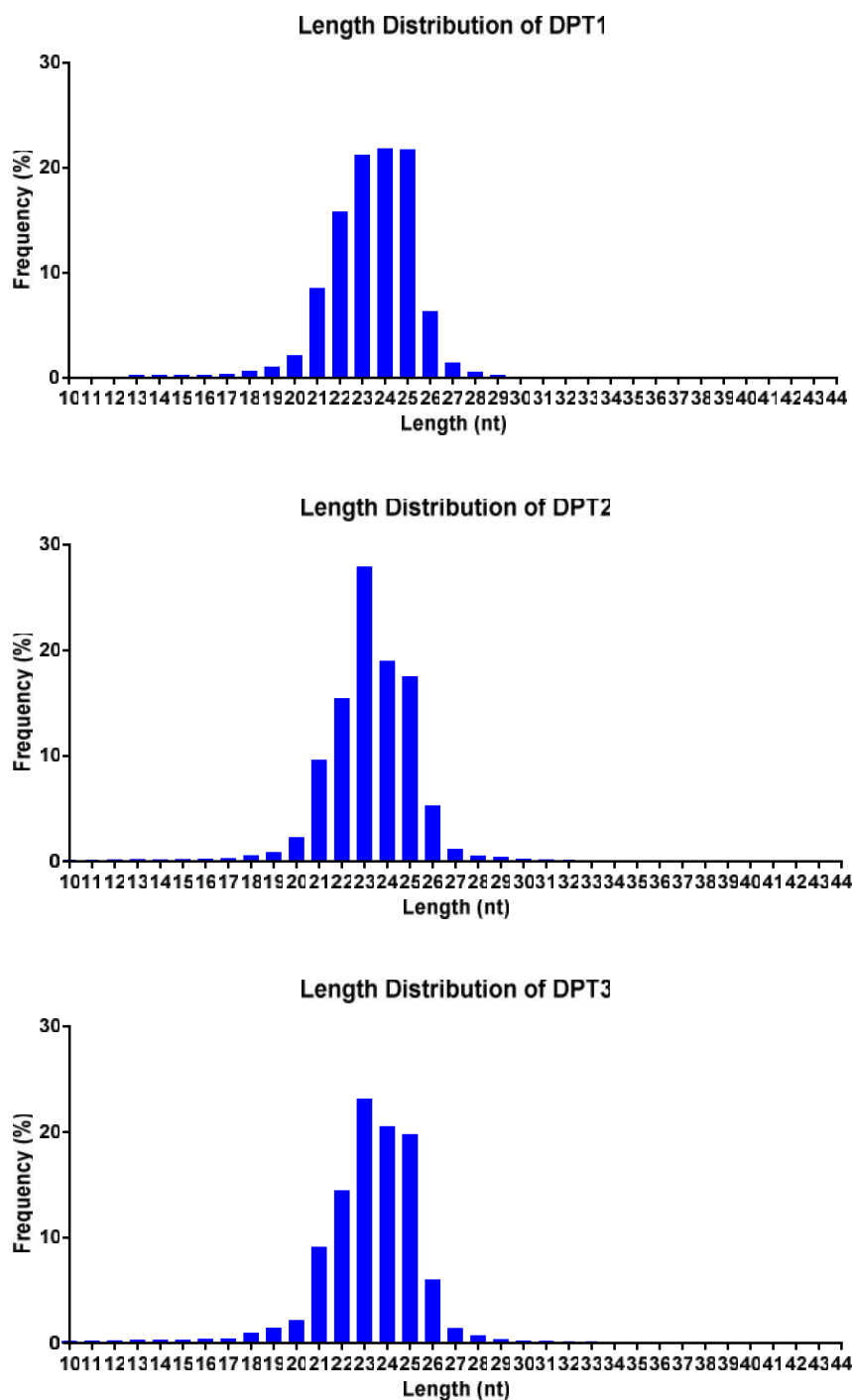


Figure S2 – Length distribution of clean reads. The y-axis indicated that the proportion of the reads in the dataset. The clean reads were range from 18bp to 44bp in three samples. Most of the reads ranged in 22-25bp in length, representing over 80% of the total reads in all samples.

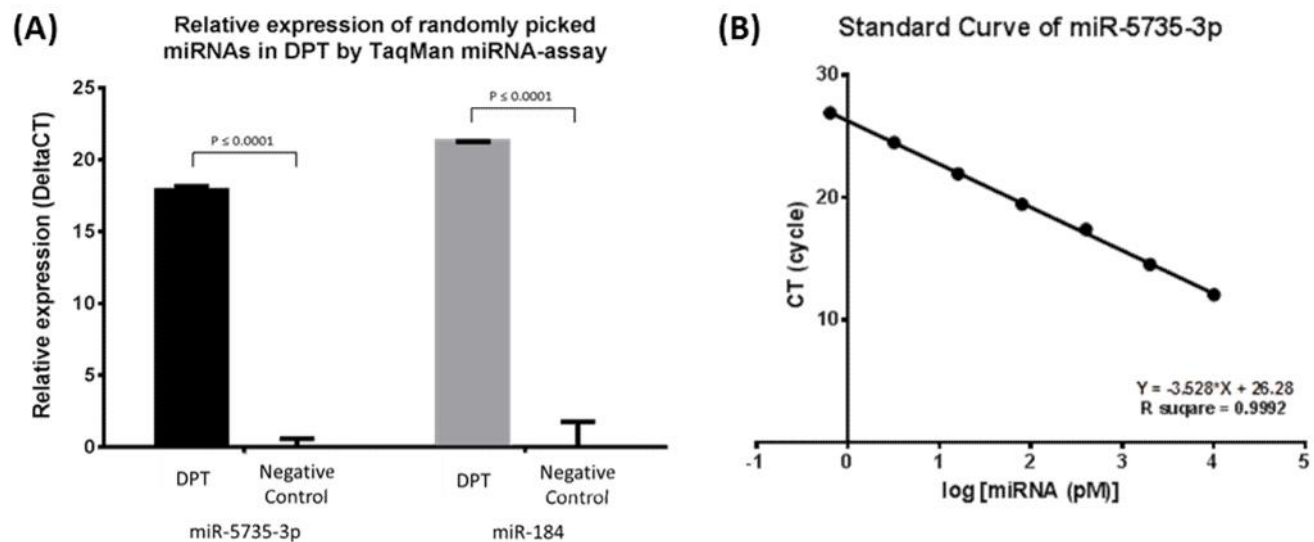


Figure S3 – (A) Validation of the expression of randomly picked miRNAs was performed by TaqMan miRNA-assay (Applied Biosystems). Delta CT was calculated by CT value of mite lysate minus CT value of negative control, and the difference is calculated by Mann-Whitney test. All TaqMan-PCRs were performed in triplicates. It showed that the presence of miR-5735-3p and miR-184 were significant in DPT samples. (B) Standard curve of Taqman miRNA assay of miR-5735-3p. The standard curve was plotted by TaqMan assay results using known concentration of miRNA mimics.

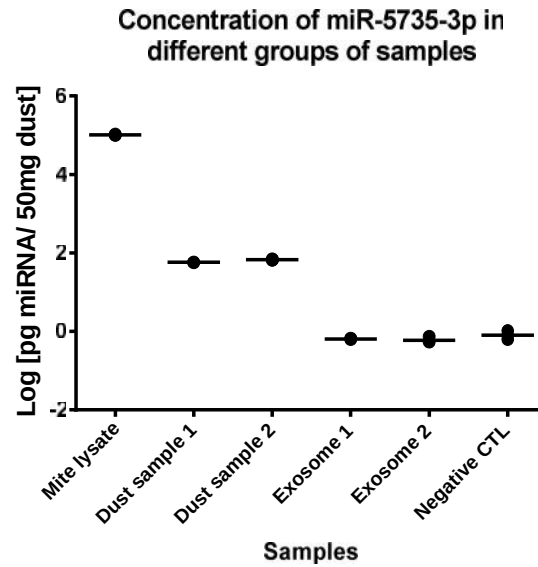


Figure S4 – Concentration of miR-5735-3p in dust and exosome samples. Total miRNAs were extracted from dust samples and isolated exosomes, and TaqMan miRNA assay was performed for miR-5735-3p. All TaqMan-PCRs were performed in triplicates. It showed that the amount of mite specific miR-5735-3p cannot be detected in the exosome samples. Hence, the miR-5735-3p could be a cell-free miRNA. However, as the concentration of exosome isolated was low, it is possible that the amount of miRNA is insufficient for detection by this assay.

References

1. Chan TF, Ji KM, Yim AKY, Liu XY, Zhou JW, Li RQ, et al. The draft genome, transcriptome, and microbiome of *Dermatophagoides farinae* reveal a broad spectrum of dust mite allergens. *J Allergy Clin Immunol*. 2015;135:539–548.
2. Tian T, Wang J, Zhou X. A review: microRNA detection methods. *Org Biomol Chem*. 2015;13:2226–2238.
3. Liu XY, Yang KY, Wang MQ, Kwok JSL, Zeng X, Yang Z, et al. High-quality assembly of *Dermatophagoides pteronyssinus* genome and transcriptome reveals a wide range of novel allergens. *J Allergy Clin Immunol*. 2018;141:2268–2271.e8.
4. Langmead B. Aligning short sequencing reads with Bowtie. *Curr Protoc Bioinforma*. 2010;11:11.7.
5. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res*. 2014;42:D68-73.
6. Altschul SF, Gish W, Miller W, Myers EW, LipmanDJ. Basic local alignment search tool. *J Mol Biol*. 1990;215:403–410.
7. Vishwakarma NP, Jadeja VJ. Identification of miRNA encoded by *Jatropha curcas* from EST and GSS. *Plant Signal Behav*. 2013;8:e23152.
8. Gupta H, Tiwari T, Patel M, Mehta A, Ghosh A. An approach to identify the novel miRNA encoded from H. Annuus EST sequences. *Genomics data*. 2015;6:139–144.
9. Qu B, Qiu Y, Zhen Z, Zhao F, Wang C, Cui Y, et al. Computational identification and characterization of novel microRNA in the mammary gland of dairy goat (*Capra hircus*). *J Genet*. 2016;95:625–637.
10. Choudhuri S. Small noncoding RNAs: Biogenesis, function, and emerging significance in toxicology. *J Biochem Mol Toxicol*. 2010;24:195–216.
11. Zhao Y, Li H, Fang S, Kang Y, Wu W, Hao Y, et al. NONCODE 2016: an

informative and valuable data source of long non-coding RNAs. *Nucleic Acids Res.* 2016;44:D203–208.

12. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, et al. Ensembl 2018. *Nucleic Acids Res.* 2018;46:D754–761.
13. Ding J, Zhou S, Guan J. MiRenSVM: towards better prediction of microRNA precursors using an ensemble SVM classifier with multi-loop features. *BMC Bioinformatics.* 2010;11:S11.
14. Gruber AR, Bernhart SH, Lorenz R. The ViennaRNA web services. *Methods in molecular biology* (Clifton, NJ). 2015;1269: 307–726.
15. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et al. A survey of best practices for RNA-seq data analysis. *Genome Biol.* 2016;17:13.
16. Liu W, Hao Z, Huang L, Chen L, Wei Q, Cai L, et al. Comparative expression profile of microRNAs in *Anopheles anthropophagus* midgut after blood-feeding and *Plasmodium* infection. *Parasit Vectors.* 2017;10:86.
17. Wan SM, Yi SK, Zhong J, Nie CH, Guan NN, Zhang WZ, et al. Dynamic mRNA and miRNA expression analysis in response to intermuscular bone development of blunt snout bream (*Megalobrama amblycephala*). *Sci Rep.* 2016;6:31050.
18. Shrinet J, Jain S, Jain J, Bhatnagar RK, Sunil S. Next generation sequencing reveals regulation of distinct aedes microRNAs during Chikungunya virus development. RibeiroJMC, editor. *PLoS Negl Trop Dis.* 2014;8:e2616.
19. Grillo G, Turi A, Licciulli F, Mignone F, Liuni S, Banfi S, et al. UTRdb and UTRsite (RELEASE 2010): a collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs. *Nucleic Acids Res.* 2010;38:D75–80.
20. Kruger J, Rehmsmeier M. RNAhybrid: microRNA target prediction easy, fast and flexible. *Nucleic Acids Res.* 2006;34:W451–454.

21. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4:44–57.
22. Kramer MF. Stem-loop RT-qPCR for miRNAs. *Curr Protoc Mol Biol.* 2011; 15:10.
23. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods.* 2015;12:357–360.
24. 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;7:562–578.

