RESEARCH ARTICLE

Insect Molecular Biology

Royal Entomological Society

The RNA interference response to alphanodavirus replication in *Phlebotomus papatasi* sand fly cells

Akira J. T. Alexander[1] | Rhys H. Parry[2] | Maxime Ratinier[3] |
Frédérick Arnaud[3] | Alain Kohl[1,4]

[1]MRC-University of Glasgow Centre for Virus Research, Glasgow, UK

[2]School of Chemistry and Molecular Biosciences, The University of Queensland, St. Lucia, Australia

[3]IVPC UMR754, INRAE, Université Claude Bernard Lyon 1, EPHE, PSL Research University, Lyon, France

[4]Centre for Neglected Tropical Diseases, Department of Tropical Disease Biology and Vector Biology, Liverpool School of Tropical Medicine, Liverpool, UK

Correspondence

Akira J. T. Alexander, MRC-University of Glasgow Centre for Virus Research, Glasgow, UK.
Email: akira.alexander@glasgow.ac.uk

Rhys H. Parry, School of Chemistry and Molecular Biosciences, The University of Queensland, St. Lucia, Australia.
Email: r.parry@uq.edu.au

Alain Kohl, Centre for Neglected Tropical Diseases, Department of Tropical Disease Biology and Vector Biology, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, UK.
Email: alain.kohl@lstmed.ac.uk

Abstract

In this study, we identified and assembled a strain of American nodavirus (ANV) in the *Phlebotomus papatasi*-derived PP9ad cell line. This strain most closely resembles Flock House virus and ANV identified in the *Drosophila melanogaster* S2/S2R cell line. Through small RNA sequencing and analysis, we demonstrate that ANV replication in PP9ad cells is primarily targeted by the exogenous small interfering RNA (exo-siRNA) pathway, with minimal engagement from the PIWI-interacting RNA (piRNA) pathway. In mosquitoes such as *Aedes* and *Culex*, the PIWI pathway is expanded and specialised, which actively limits virus replication. This is unlike in *Drosophila* spp., where the piRNA pathway does not restrict viral replication. In *Lutzomyia* sandflies (family *Psychodidae*), close relatives of *Phlebotomus* species and *Drosophila*, there appears to be an absence of virus-derived piRNAs. To investigate whether this absence is due to a lack of PIWI pathway proteins, we analysed the piRNA and siRNA diversity and repertoire in PP9ad cells. Previous assemblies of *P. papatasi* genome (Ppap_1.0) have revealed a patchy repertoire of the siRNA and piRNA pathways. Our analysis of the updated *P. papatasi* genome (Ppap_2.1) has shown no PIWI protein expansion in sandflies. We found that both siRNA and piRNA pathways are transcriptionally active in PP9ad cells, with genomic mapping of small RNAs generating typical piRNA signatures. Our results suggest that the piRNA pathway may not respond to virus replication in these cells, but an antiviral response is mounted via the exo-siRNA pathway.

KEYWORDS

nodavirus, *Phlebotomus papatasi*, RNAi, siRNA piRNA, vpiRNA, vsiRNA

INTRODUCTION

Phlebotomine sand flies (Diptera: *Psychodidae*) are not only vectors of bunyaviruses of the *Phenuiviridae* family, such as Toscana virus (TOSV) and related phleboviruses, but also members of the *Peribunyaviridae*, *Rhabdoviridae*, *Flaviviridae* and *Reoviridae* families, which have been suggested to be transmitted, or potentially being transmitted by these insects (Jancarova et al., 2023). In Europe, these insects are particularly relevant to human and animal health in regions around the Mediterranean Basin. Still, with climate change, they are increasingly spreading to other regions, thus increasing the geographical risk area (Maroli et al., 2013; Oerther et al., 2020; Schaffner et al., 2024). Moreover, sand flies can transmit *Leishmania* parasites and *Bartonella* bacteria, highlighting their role as vectors requiring urgent research (Alkan et al., 2013; Ayhan & Charrel, 2020; Charrel et al., 2018; Jancarova et al., 2023; Moriconi et al., 2017).

Novel control measures that target vectors may be important in controlling sand fly populations, but this requires a better understanding of their ecology, development, physiology and immunity. Indeed, host responses to viruses in sand flies and sand fly cells remain poorly understood, with very few studies investigating this topic (Telleria et al., 2018). Previous studies in *Lutzomyia longipalpis*-derived LL5 cells described that transfected double-stranded RNA (dsRNA) can induce an antiviral state, and secretion of proteins linked to antiviral immunity in other organisms such as a phospholipid scramblase was described (Martins-da-Silva et al., 2018; Pitaluga et al., 2008). Better understood in insects are RNA interference (RNAi)-based antiviral responses that are generally well conserved, and with many findings from the model insect *Drosophila melanogaster* translatable to important arbovirus vectors such as mosquitoes. Indeed, the mosquito exogenous small interfering RNA (exo-siRNA) pathway is particularly relevant in controlling arbovirus replication across viral families. It is induced by virus-derived dsRNA, which is cleaved by Dicer-2 (Dcr2) protein into predominantly 21 nucleotide (nt) in length, virus-derived small interfering RNAs, vsiRNAs. Importantly, these vsiRNAs are then used by a further effector protein, Argonaute-2 (Ago2) (as part of the RNA Induced Silencing Complex [RISC]), to target complementary viral RNA (Blair, 2023; Blair & Olson, 2015; Olson & Blair, 2015; Prince et al., 2023; Samuel et al., 2018). The P-element Induced WImpy testis (PIWI)-interacting RNA (piRNA) pathway is a second pathway that responds to viral infection and is poorly understood. Importantly, piRNAs in insects are longer, around 24–30 nt. The pathway relies on Argonaute-3 (Ago3) and PIWI proteins, the latter family being expanded in *Aedes aegypti* and *Aedes albopictus* compared with *D. melanogaster* for piRNA production via a so-called ping-pong production mechanism. This results in a predominant positional nt signature, U_1 for primary piRNAs or A_{10} for secondary piRNAs, with a 10 nt overlap between secondary and primary piRNAs. Silencing of piRNA pathway effectors has demonstrated the involvement of this pathway in antiviral responses in mosquitoes, such as Piwi4. Indeed, canonical virus-derived piRNAs, vpiRNAs, are frequently observed in arbovirus-mosquito interactions (Blair, 2023; Santos et al., 2023; Varjak et al., 2018). There are exceptions—for example, in the case of Zika virus infection of *Ae. aegypti*-derived cells, where no canonical vpiRNAs were identified (Varjak et al., 2017). Interestingly, although vpiRNAs have been detected in *D. melanogaster* ovary somatic sheet cells, the pathway may not play a role in antiviral defences in that insect (Petit et al., 2016; Wu et al., 2010). This asks the question of how widespread and functionally relevant antiviral piRNA responses are. Infection of *L. longipalpis* sandflies or *Phlebotomus papatasi* cells with the negative strand RNA viruses (vesicular stomatitis virus, VSV [*Rhabdoviridae*], or TOSV [*Phenuiviridae*], respectively) resulted in the production of 21 nt vsiRNAs but not vpiRNAs (Alexander et al., 2023; Ferreira et al., 2018). Indeed, silencing of Ago2 in *P. papatasi*-derived PP9ad cells increased TOSV replication (Alexander et al., 2023). Similarly, a mitovirus infecting *L. longipalpis* induced the production of 21 nt vsiRNAs, and again, no obvious vpiRNAs were described (Fonseca et al., 2020). An absence of piRNAs was also observed for various viruses of the *Reoviridae* and *Nodaviridae* families infecting *L. longipalpis* – which surprisingly do not show a predominance of 21 nt vsiRNAs either (Aguiar et al., 2015; Ferreira et al., 2018). There are, thus, question marks on whether the induction of canonical exo-siRNA response is virus dependent and the general absence of canonical antiviral piRNA responses.

Here, we showed the detection within previously described small RNA sequencing data from PP9ad cells (Alexander et al., 2023), of a novel strain of the previously identified *D. melanogaster* American nodavirus (ANV) (*Nodaviridae*, *Alphanodavirus*, Flock House virus [FHV]) (Wu et al., 2010), a positive-stranded RNA virus whose genome consists of two RNA strands, named RNA1 and RNA2, in the previously described PP9ad cell line (Alexander et al., 2023). Replication of this nodavirus in *P. papatasi* cells produces 21 nt vsiRNAs but not vpiRNAs. However, these vsiRNAs did not show classical Dcr2 production signature patterns, suggesting that the majority of vsiRNA duplexes were sequestered by the RNAi inhibitory viral B2 protein (which binds as a dimer to small RNA duplexes between 17 and 25 base pairs, preventing Ago2 loading (Chao et al., 2005; Lingel et al., 2005), or to longer dsRNA sequences, thus inhibiting Dcr2 cleavage (Lu et al., 2005)) to counteract the exo-siRNA pathway. Moreover, we re-assessed the *P. papatasi* siRNA and piRNA pathway components and demonstrated their presence in PP9ad cells. These data add to understanding small RNA pathway responses in sand fly cells, suggesting that antiviral exo-siRNA pathway responses are generally active. Still, canonical antiviral piRNA responses may not respond to this viral infection.

RESULTS AND DISCUSSION

Both the siRNA and piRNA pathways are active in PP9ad cells

The initial genome assembly and annotation of the *P. papatasi* genome (Ppap_1.0, RefSeq: GCA_000262795.1) revealed the existence of most core siRNA and piRNA pathway genes (Labbe et al., 2023). However, certain genes, such as *ago2*—which had

RNA INTERFERENCE RESPONSE TO AN ALPHANODAVIRUS

Insect Molecular
Biology

Royal
Entomological
Society

3

been identified from PP9ad cells (Alexander et al., 2023)—were either missing or requiring revision as the predicted genes were shortened or truncated. To address this, we examined the updated *P. papatasi* genome (Ppap_2.1; RefSeq: GCF_024763615.1) to identify the completeness of siRNA and piRNA pathway genes so they could be used to assess transcriptional activity and function in the PP9ad cell line.

In *Aedes* and *Culex* mosquitoes, there is a notable expansion in the repertoire of the piRNA pathway, with expression being prevalent in both somatic and germline tissues and with specialisation of one of these proteins, Piwi4, as antiviral (Varjak et al., 2018). Employing *Ae. aegypti* homologues as references, we identified a single copy homologue of Dcr2 endoribonuclease (LOC129803306, XM_055849757.1) and Ago2 (LOC129803094, XM_055849433.1). Additionally, for the piRNA machinery, we identified two Aubergine-like or PIWI-like genes (LOC129803499, XM_055850102.1 and XM_055850099.1, LOC129803498), as well as one copy of Ago3 (LOC129806092, XM_055854434.1), a Zucchini homologue (LOC129807453, XM_055856739.1) and a Nibbler homologue (LOC129806408, XM_055854972.1). Upon mapping small RNA reads from the PP9ad cell line to these genes and quantifying transcript abundance, we confirmed the transcriptional presence of all identified siRNA and piRNA genes (Figure 1a).

Notably, siRNA machinery was highly transcriptionally active with Dcr2 and Ago2 expression with an average of 143,783 transcripts per million (TPM) and 769,990 TPM, respectively. Comparatively the piRNA machinery genes were less abundantly expressed (between 6082 and 43,282 TPM). Analysis of the small RNA read length distribution showed that most (53%) of small RNAs in these cells originated from the siRNA and miRNA pathways with reads between 21 and 23 nt in length (Figure 1b). Despite this, transcriptional activity of the piRNA pathway was evident, with a significant peak at 28 nt and a bias towards a U_1 position, indicating potential functional activity in these cells (Figure 1b). We investigated the piRNA pathway activity by looking for a 'ping-pong' cycle signature within small RNA populations. We subsetted the small RNAs into the si/miRNA population (Figure 1c, reads 21–24 nt in length), which did not show typical piRNA features, except for a bias towards U at position 1. Additionally, the overlapping pairs of small RNAs of 21–24 nt in length were enriched (z-scores ≥ 1) towards overlap sizes of 17–21 nt in length, suggesting the si/miRNA machinery predominantly processes these read sizes (Antoniewski, 2014). This was in contrast to the piRNA-size populations (Figure 1d, 25–30 nt) containing strong evidence of piRNA activity with 10 nt bp overlap of sRNAs (z-score, 4), a bias for U_1 in piRNAs while only a weak preference in position A_{10} in piRNAs. Given that all reads are analysed as one pool without considering 'strandedness', to manually validate we identified a 1500 nt genomic region in chromosome 1 between positions 60546551and 60544952 with a significant portion of read lengths corresponding to piRNAs (23–29 nt) (Figure 1e), validated the ping-pong overlapping signature of 10 nt (Figure 1f) and demonstrated that reads 23–29 nt originating from the forward strand had a bias towards U1 and reverse strand A_{10} (Figure 1g). This unambiguously demonstrates that the piRNA pathway is active against genomic loci in *P. papatasi* PP9ad cells.

Identification of adventitious nodavirus in *P. papatasi* PP9ad cells

Adventitious viral agents of insect cell lines are common and have been identified by accident or through metagenomics studies (Bishop et al., 2020; Franzke et al., 2018; Parry et al., 2021; Parry & Asgari, 2018; Weger-Lucarelli et al., 2018). Previously, it has been demonstrated that as the siRNA and piRNA pathway target viral genomes, viral-derived small RNA reads are enriched in total small RNA reads, allowing viral contigs to be assembled (Aguiar et al., 2016; Vodovar et al., 2011; Wu et al., 2010). To identify viral sequences in PP9ad cells, we subjected small RNA sequencing from three libraries of PP9ad cells to *de novo* assembly and analysed through a viral metagenomics pipeline (Parry et al., 2021). In total, 313 contigs were assembled ranging from 151 to 4529 nt and were queried to the NCBI non-redundant viral database using BLASTn and BLASTx. Three contigs between 1356 and 1763 in length showed high nt identity (between 93.49% and 96.01%) to RNA 1 and RNA 2 (GenBank ID GQ342965.1, GQ342966.1) from ANV SW-2009a, sometimes called FHV, a common infectious agent of *D. melanogaster* S2 cells (Wu et al., 2010). Manual examination of the ANV-like contigs revealed the two RNA 1 contigs overlapped perfectly with 34 nt, creating a 3107 nt coding complete segment for RNA 1, which shared 96.00% identity to RNA 1 of the SW-2009a strain (GenbankID: GQ342965.1). A complete RNA 2 segment was manually identified and trimmed to 1372 nt with 96.31% nt identity to the IP-VIA-022011 strain of FHV (Genbank ID: JF461542.1), which was also described from a derivative of S2 cells, S2R (Vodovar et al., 2011). Small RNA reads were then remapped to the ANV-PP9ad strain sequence, with 213,989/174,238,062 reads mapping to the reference with a mapping rate of 0.12% and an average coverage of 786× for RNA 1 and 1516× for RNA 2 (Figure 2a).

Examination of the RNAi response to alphanodavirus replication

To further examine if ANV is a bona fide infectious agent of the PP9ad cell line, small RNA libraries from these cells were mapped to the ANV genome (Figure 2b,c).

Examination of the read lengths of small RNAs that mapped to both segments of the ANV genome indicated an overwhelming bias towards small RNA reads of 21 nt size to both RNA 1 and RNA 2 (Figure 2b), 69.6% of total reads ±4.3% for RNA 1 and 74.9% ± 2.1% for RNA 2, with limited small RNAs that corresponded to potential viral piRNA pathway-derived reads (24–29 nt), 8.5% ± 3.2% for RNA 1 and 5.0% ± 1.7% for RNA 2. The production of antigenome mapping small RNA reads at almost equal proportion of the genome strand for this ssRNA virus indicates active replication in these cells (RNA 1 1.05 ± 0.02 sense: 1 antisense and RNA 2 1.02 ± 0.05 sense: 1 antisense).
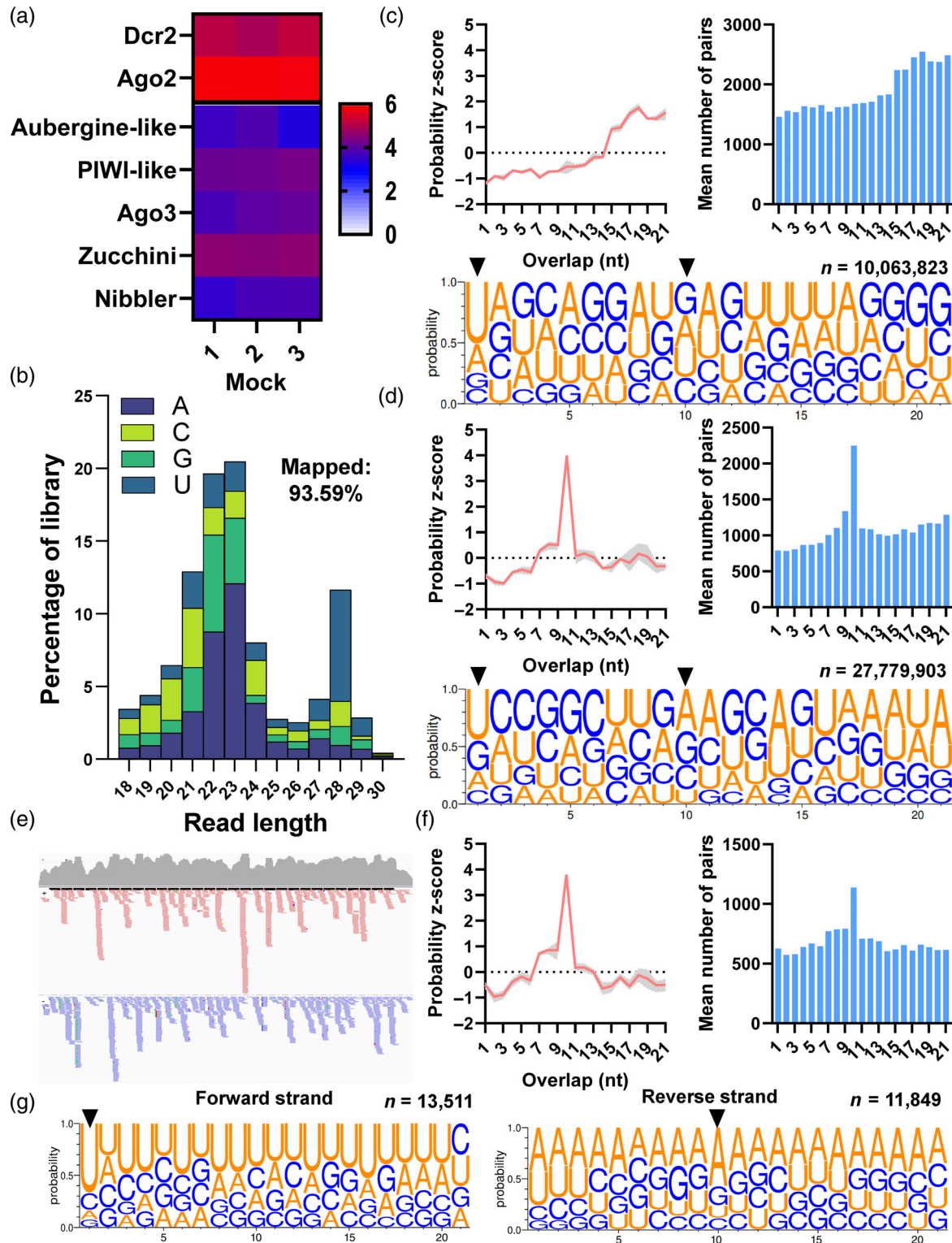
FIGURE 1 Legend on next page.

Further examination of the genomic loci of exclusively 21 nt reads (Figure 2c) indicated an equal coverage of vsiRNAs, with no indication of hotspot regions from which vsiRNAs originated and no bias towards the B1/B2 region of RNA 1 where subgenomic transcription occurs. This is in stark contrast to the mapping profile of the 24–29 nt viral small RNA reads, which were five to tenfold higher in the genome sense (RNA 1, 9.32 ± 3.82 sense: 1 antisense and RNA 2, 5.56 ± 3.69 sense: 1 antisense), suggesting that these

RNA INTERFERENCE RESPONSE TO AN ALPHANODAVIRUS

Insect Molecular
Biology

Royal
Entomological
Society

5

reads are more likely to have derived from the genomic ANV RNA (Figure 2d).

Next, we examined the features of overlapping pairs of viral small RNA reads for both the prototypical vsiRNA (21 nt) and vpiRNA (24–29 nt) read lengths. While both segments produced adequate magnitudes of overlapping vsiRNAs (Figure 3a,b), there was no evidence of a Dcr2-like signature, which is characterised by overlapping read biases of vsiRNAs by 18–20 nt with a peak at 19 nt (Antoniewski, 2014). Additionally, there was no evidence that the small 24–29 nt reads derived from the ANV genome overlapped by 10 nt, which is the prototypical 'ping-pong' signature of viral piRNA pathway processing, or demonstrated the prototypical U_1 and A_{10} bias of the PIWI pathway (Figure 3c,d) (Varjak et al., 2018). As almost all of the reads of this length came from the genomic sense, there was an 80-fold reduction in overlapping pairs of piRNAs compared to the siRNAs.

The B2 protein of *Nodaviridae*, and by extension that of ANV, exhibits two distinct mechanisms of RNAi antagonism. First, it has been shown to bind as a dimer to viral dsRNA duplexes between 17 and 25 base pairs (Chao et al., 2005; Lingel et al., 2005), sequestering away these vsiRNA duplexes before loading them into the holo-RISC complex. Additionally, B2 protein exhibits an even higher affinity for longer dsRNA sequences inhibiting Dcr2 cleavage into the vsiRNA duplexes (Lu et al., 2005). Here, we see no bias in the overlapping pair absolute abundance or *z*-score probability for viral 21 nt vsiRNAs that overlap by 19 nt with two-base 3′ overhangs, a classical feature of Dcr2 cleavage and seen previously in these cells infected with TOSV (Alexander et al., 2023). This suggests that most vsiRNA duplexes are sequestered by the B2 protein and no longer present in these data sets as a consequence of viral counteraction of the exo-siRNA response. It cannot be excluded that the binding of longer dsRNA also comes into play. Indeed, both mechanisms are possible, either separately or alongside each other, and it is not possible to discern between the two. This will require further investigation.

In summary, our bioinformatic evidence suggests that PP9ad cells harbour transcriptionally active and functional siRNA and piRNA pathways, indicating that the absence of virus-derived vpiRNAs in these cells is not attributable to a lack of the piRNA machinery and further investigations are required to see whether the absence of canonical piRNA response as observed here, and previously (Alexander et al., 2023), are general features in sand flies. The absence of a typical

Dcr2 signature in the vsiRNAs from our data set may be explained by sequestering of viral dsRNAs or vsiRNAs duplexes by B2. We can only speculate on this point presently, and for example, mutagenesis of B2 in the PP9ad ANV genome would be required to assess this. Nonetheless, the presence of 21 nt vsiRNAs as observed here and in other viral infections of sand fly cells or sand flies (Alexander et al., 2023; Ferreira et al., 2018; Fonseca et al., 2020) are hallmarks of exo-siRNA pathway induction. Analysis of existing data sets for the presence or absence of Dcr2 cleavage patterns may shed further light on the processes involved in infection of such cells with viruses other than the PP9ad ANV described here. This will require further investigation.

EXPERIMENTAL PROCEDURES

Cell culture

Reads were derived from *P. papatasi*-derived PP9ad cells, which are an adherent form of PP9 cells, and previously described (Alexander et al., 2023) and it is derived from parental PP9 cells obtained from the World Reference Center for Emerging Viruses and Arboviruses (University of Texas Medical Branch, Galveston, Texas, USA).

Small RNA sequencing

Small RNA reads were from previously obtained libraries (Alexander et al., 2023). A summary of the sequencing procedure is described here for completeness. Triplicate wells containing 5×10^6 PP9ad cells were incubated for 3 days at 28°C. RNA was then extracted using Trizol following the manufacturer's protocol, and each triplicate sample was pooled. This was then repeated two more times to generate three independent replicates. Small RNA sequencing was carried out by the Beijing Genomics Institute (BGI Group, Shenzhen, China). Extracted RNA was size separated using polyacrylamide gel electrophoresis, and bands in the region of 18–35 nt were extracted. These had a 3′ and 5′ adaptor added before amplification by reverse transcriptase polymerase chain reaction. After purification of the PCR products, they were heat denatured and circularised using a splint oligo. The single-stranded circular DNA became the library. Library QC was carried out

FIGURE 1 Both the siRNA and piRNA pathways are active in *P. papatasi* PP9ad cells. (a) Transcript abundance of small RNAs deriving from the mRNA of siRNA and piRNA pathway homologues represented as $\log_{10}$ reads per transcript million (TPM) in the PP9ad cell line ($n = 3$). (b) Read profile of small RNAs derived from PP9ad cells mapping to the Ppap_2.1 genome and nucleotide bias of the first position ($n = 3$). (c) Probability *z*-scores (left) and the mean number of overlapping pairs (right) of small RNA length reads (21 nt) mapping to the Ppap_2.1 genome. The mean of $n = 3$ independent repeats. Representative sequence logos of 21 nt small RNA length reads mapping to the Ppap_2.1 genome. The probability of each nucleotide at each position with piRNA pathway bias positions (U_1 and A_{10}) is indicated with arrowheads and trimmed at 21 nt for clarity. (d) Probability *z*-scores (left) and the mean number of overlapping pairs (right) of piRNA length reads (24–29 nt) mapping to the Ppap_2.1 genome. The mean of $n = 3$ independent repeats. Representative sequence logos of all piRNA length reads mapping to the Ppap_2.1 genome. The probability of each nt at each position with piRNA pathway bias positions (U_1 and A_{10}) is indicated with arrowheads, and reads trimmed at 21 nt for clarity. (e) Mapping coverage of piRNA length reads 23–29 nt of a 1500 nt genomic region of chromosome 1 of the *P. papatasi* genome (GenbankID: JANPWV010000001; 60546551-60544952). (f) Probability *z*-scores (left) and the mean number of overlapping pairs (right) of piRNA length reads. (g) Sequence logos of all piRNA length reads originating from the forward strand and reverse strand mapping, trimmed to 21 nt for clarity U_1 in forward strands and A_{10} in reverse strands are indicated with arrowheads.
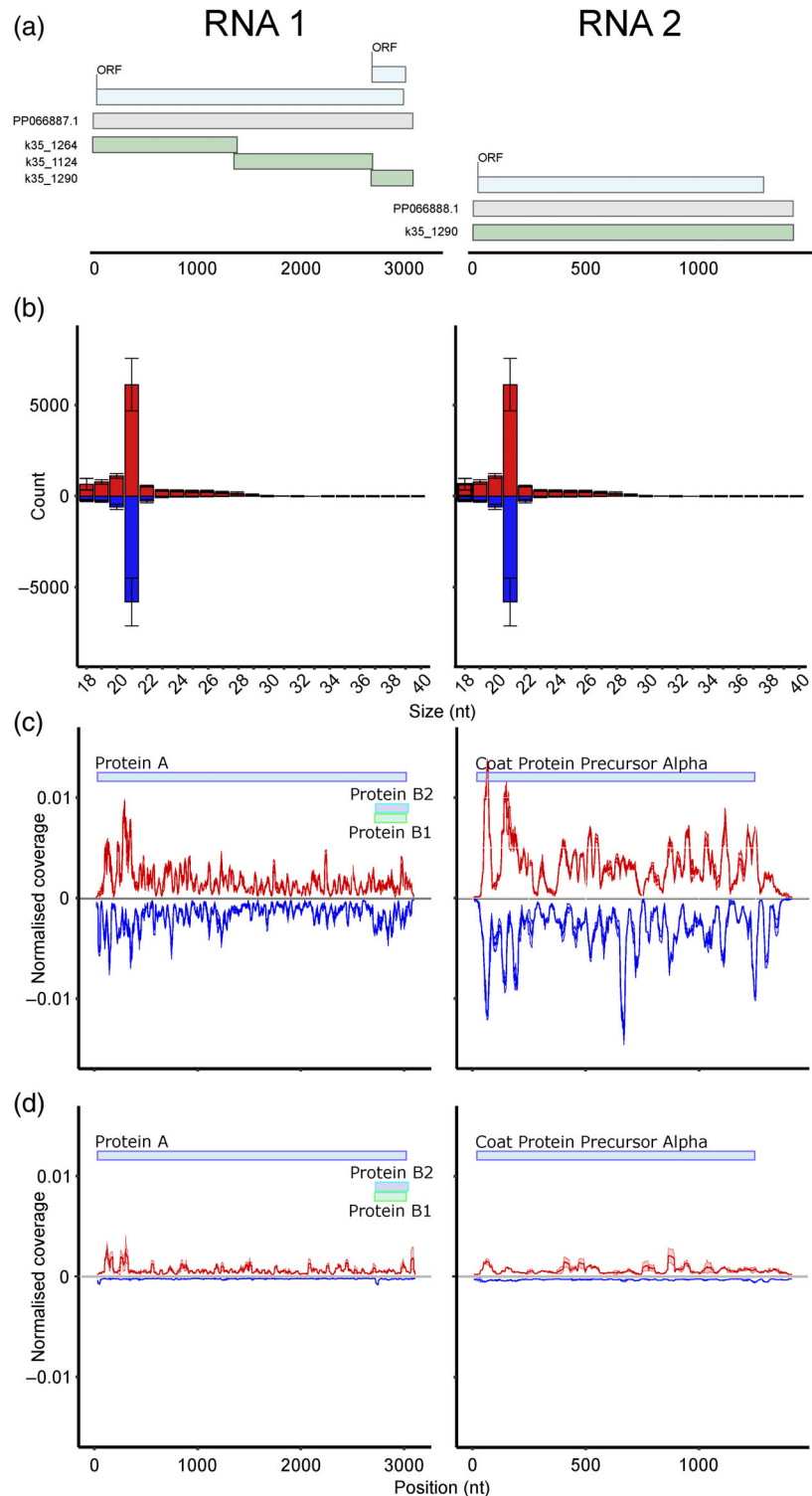
FIGURE 2 ANV is actively replicating in PP9ad cells and is targeted by the exo-siRNA pathway. (a) Alignment of the generated contigs for RNA 1 and RNA 2 of the novel PP9ad ANV. (b) Histogram of small RNA reads from the PP9ad cells 18–40 nt in length, mapping to the RNA 1 and RNA 2 of the PP9ad strain of ANV (PP066887 and PP066888) genome sense (red) and antigenome (blue) with Y-axis showing normalised read count from three independent experiments, error bars indicate $+/-$ SD. (c) Graphs show the distribution and genomic coverage of 21 nt vsiRNAs and (d) the 24–29 nt vpiRNAs along each PP9ad ANV segment length, normalised as a percentage of total reads, the SD between replicates is indicated by the lighter colour. Reads mapping genome sense are shown in red and antigenome in blue. The protein-coding regions are mapped above. The read count per sample was feature-scaled to account for unequal count totals between samples.

RNA INTERFERENCE RESPONSE TO AN ALPHANODAVIRUS

Insect Molecular
Biology

Royal
Entomological
Society

7
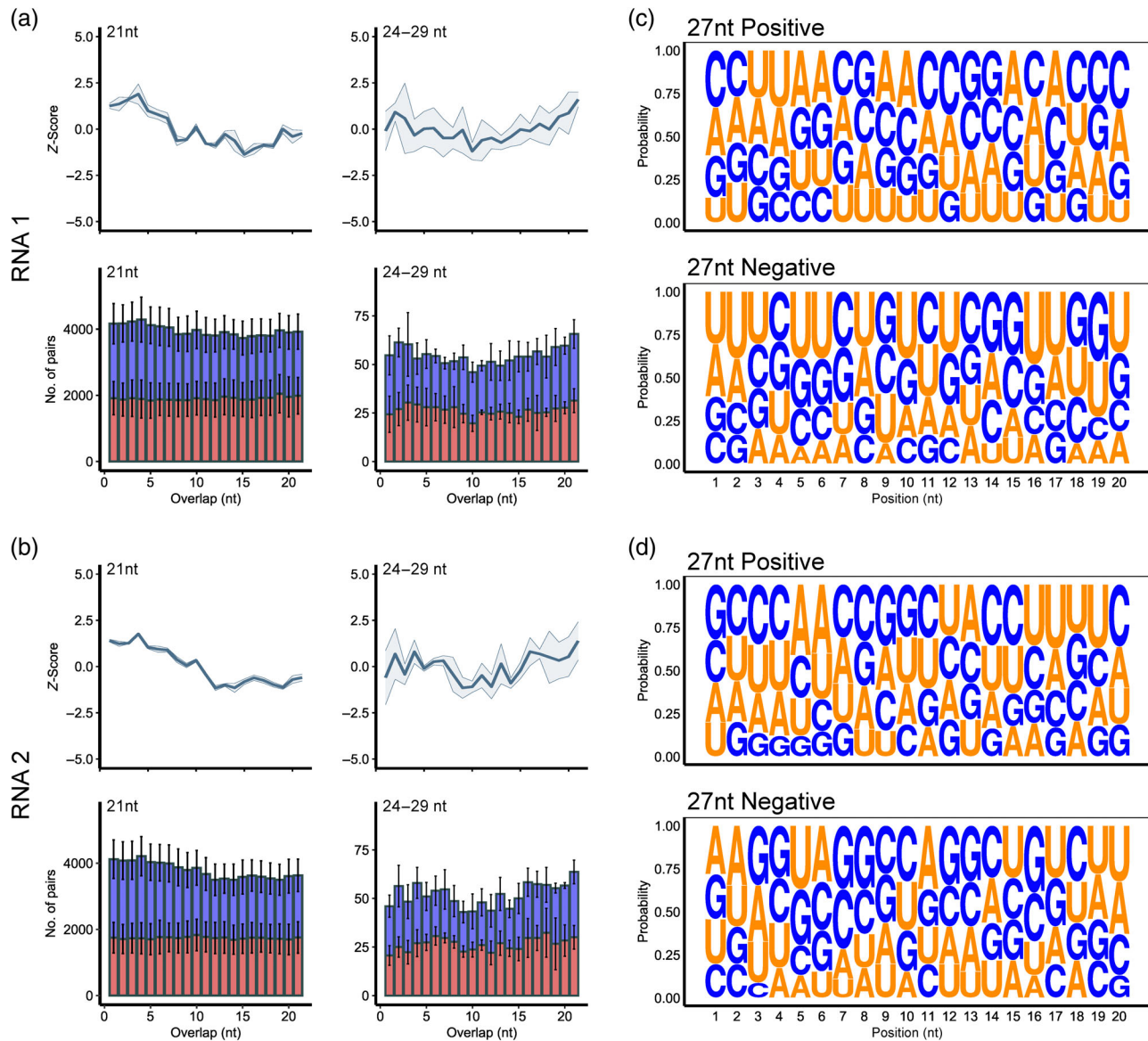
FIGURE 3 No evidence of a prototypical antiviral piRNA response against ANV. Overlapping pairs z-score probability and raw number of pairs of 21 nt vsiRNAs and potential vpiRNAs (24–29 nt) of (a) RNA1 and (b) RNA2 of the PP9ad ANV. Read pairs corresponding to the positive sense (red) genome and negative sense antigenome (blue). Sequence logo analysis of (c) RNA 1 and (d) RNA 2, as determined by seqLogo of nt predominance of the first 20 nt of each 27 nt small RNAs mapping to RNA 1 and RNA 2 genomic or antigenomic segments, as indicated by all replicates combined. The SD is indicated by error bars on the column plot or lighter colour on the line plot.

using an Agilent Technologies 2100 bioanalyser (Agilent Technologies, Santa Clara, CA, USA). The circular library was rolling circle amplified using phi29 polymerase to form DNA nanoballs (DNB). The DNBs were loaded onto a patterned nanoarray, and single-end 50 base reads were generated by combinatorial probe anchor synthesis using a DNBSeq-G400 sequencer (MGI Tech. Co., Shenzhen, China).

Small RNA datasets

Data used in this analysis came from unifected PP9ad cell samples MOCK_1 to MOCK_3 as described by Alexander et al. (2023). These

are available as accessions SRX17532612: SRX17532614 within Bio-Project PRJNA879406 at the NCBI SRA archive.

Small RNA pre-processing and analysis

High throughput RNA sequencing data was analysed, as reported previously (Alexander et al., 2023; Gestuveo et al., 2022). Briefly, fastp was used to remove low-quality (Phred score < 30) and short (<18 nt) reads from sequencing data. To identify potential viral agents, processed sRNA from all samples were assembled using MEGAHIT (v1.2.9) with the following parameters –k-min 15 and –k-max 51 with

a –k-step of 2 (Li et al., 2015) with minimum contig length of 150 nt. Viral contigs were identified using BLASTx against a previously established virus database (Parry et al., 2021). To validate the ANV PP9ad genome, all clean reads were mapped to the RNA 1 and RNA 2 segments (Genbank ID: PP066887 and PP066888) using Bowtie2 (v2.4.5) (Langmead & Salzberg, 2012) and manually inspected with Integrated Genomics Viewer (v 2.7.0) (Thorvaldsdottir et al., 2013).

For the siRNA and piRNA analysis, alignments against the ANV PP9ad genome were carried out using Bowtie2 and unmapped reads were removed using SAMtools (v 1.16.1) (Danecek et al., 2021) and all further analyses were carried out using R 4.3 and RStudio 2023.

For calculation of the coverage of each genome segment, the BAM file was filtered by genome segment, strand and read length. If the strand was negative, the start position was calculated by start plus length, as all mappings are given in the 5′–3′ orientation. Each read in the filtered list was then turned into a matrix of the nts between the start position and the end, start + length. Then, the total of occurrences of each nt from 1 to the length of the genome was counted to give the read coverage for each genome position. The number was then normalised to the total number of mapped reads for the sample to compensate for differences in sample size.

To calculate the number of overlapping pairs, a matrix of all the reads of a given length(s) starting at each genomic nt position was created. As before, if the mapping was on the negative strand, the start position was adjusted to take this into account. The matrix was then split according to the strand mapping, and one of the strands was assigned as the pattern strand and the opposite as the target strand. For each nt in the length of the genome, the pattern—number of reads of the given length(s) starting at that nt position was recalled from the pattern matrix. The opposing target matrix—reads starting at that position on the opposite strand was called from the target matrix. This was repeated for each opposing target nt until the specified overlap length had been recovered; in this case, a 21 nt overlap was investigated, so a read starting at pattern nt—21 on the target strand would overlap by 21 nt. As there cannot be more overlapping reads than reads at a position, if the number of overlapping reads exceeded the input, the number was set to equal the input. The full analysis script is available at https://github.com/AK1RAJ/piMaker/blob/main/piMaker.R.

Reannotation of siRNA and piRNA pathway genes in *P. papatasi*

To identify homologues of siRNA and piRNA machinery within the *P. papatasi* genome, we queried previously validated homologues from *Ae. aegypti*. For the siRNA pathway, Dcr2 (AAEL006794) and Ago2 (AAEL017251) were used, as well as elements of the piRNA machinery such as PIWI proteins (PIWI1-7, AAEL008076 AAEL008098, AAEL013692, AAEL007698, AAEL013233, AAEL013227 and AAEL006287), Ago3 (AAEL007823), Zucchini (AAEL011385) and Nibbler (AAEL005527) (Campbell, Black, et al., 2008; Campbell, Keene,

et al., 2008). These homologues were used as query sequences against Ppap_2.1 using BLASTp and tBLASTn (Camacho et al., 2009).

To analyse the transcriptional activity of identified siRNA and piRNA genes in PP9ad cells, we employed a dual mapping-based and estimation of transcript abundance of cleaned fastq sRNA libraries using salmon (v1.9.0), under default conditions with a kmer size of 17 (Patro et al., 2017).

For examination of the piRNA response, clean reads were aligned to the *P. papatasi* genome (Ppap_2.1; RefSeq: GCF_024763615.1) using the sensitive mapping flag (—sensitive) of Bowtie2 (v2.4.5) (Langmead & Salzberg, 2012). We examined small RNA responses by creating binary alignment files for each small RNA size class relevant to the si/miRNA (21–24 nt) and piRNA (25–30 nt) pathways.

To obtain nt position biases in small RNAs, we extracted all aligned reads irrespective of strand orientation. Reads were then trimmed to a uniform 21 nt from the 5′ end, converted to fasta format and visualised using WebLogo3 (Crooks et al., 2004). Finally, to assess the presence of overlapping si/piRNA molecules, overlapping pairs and overlap probabilities were calculated from alignment files using the 'signature.py' small RNA signatures Python script (Antoniewski, 2014).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Small RNA used in this paper are accessible through the Sequence Read Archive hosted by the National Center for Biotechnology Information under the BioProject accession: (PRJNA879406) (Alexander et al., 2023). The PP9ad strain of the *D. melanogaster* S2 ANV genome has been deposited to Genbank under accession numbers PP066887 and PP066888.

ORCID

Akira J. T. Alexander https://orcid.org/0000-0003-1601-619X
Rhys H. Parry https://orcid.org/0000-0001-9238-1952

REFERENCES

Aguiar, E.R., Olmo, R.P., Paro, S., Ferreira, F.V., de Faria, I.J., Todjro, Y.M. et al. (2015) Sequence-independent characterization of viruses based on the pattern of viral small RNAs produced by the host. *Nucleic Acids Research*, 43, 6191–6206.

Aguiar, E.R., Olmo, R.P., Paro, S., Ferreira, F.V., de Faria, I.J., Todjro, Y.M. et al. (2016) Sequence-independent characterization of viruses based on the pattern of viral small RNAs produced by the host. *Nucleic Acids Research*, 44, 3477–3478.

Alexander, A.J.T., Salvemini, M., Sreenu, V.B., Hughes, J., Telleria, E.L., Ratinier, M. et al. (2023) Characterisation of the antiviral RNA interference response to Toscana virus in sand fly cells. *PLoS Pathogens*, 19, e1011283.

Alkan, C., Bichaud, L., de Lamballerie, X., Alten, B., Gould, E.A. & Charrel, R.N. (2013) Sandfly-borne phleboviruses of Eurasia and Africa: epidemiology, genetic diversity, geographic range, control measures. *Antiviral Research*, 100, 54–74.

Antoniewski, C. (2014) Computing siRNA and piRNA overlap signatures. *Methods in Molecular Biology*, 1173, 135–146.

Ayhan, N. & Charrel, R.N. (2020) An update on Toscana virus distribution, genetics, medical and diagnostic aspects. *Clinical Microbiology and Infection*, 26, 1017–1023.

Bishop, C., Parry, R. & Asgari, S. (2020) Effect of *Wolbachia* wAlbB on a positive-sense RNA negev-like virus: a novel virus persistently infecting *Aedes albopictus* mosquitoes and cells. *The Journal of General Virology*, 101, 216–225.

Blair, C.D. (2023) A brief history of the discovery of RNA-mediated antiviral immune defenses in vector mosquitos. *Microbiology and Molecular Biology Reviews*, 87, e0019121.

Blair, C.D. & Olson, K.E. (2015) The role of RNA interference (RNAi) in arbovirus-vector interactions. *Viruses*, 7, 820–843.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. et al. (2009) BLAST+: architecture and applications. *BMC Bioinformatics*, 10, 421.

Campbell, C.L., Black, W.C., Hess, A.M. & Foy, B.D. (2008) Comparative genomics of small RNA regulatory pathway components in vector mosquitoes. *BMC Genomics*, 9, 425.

Campbell, C.L., Keene, K.M., Brackney, D.E., Olson, K.E., Blair, C.D., Wilusz, J. et al. (2008) *Aedes aegypti* uses RNA interference in defense against Sindbis virus infection. *BMC Microbiology*, 8, 47.

Chao, J.A., Lee, J.H., Chapados, B.R., Debler, E.W., Schneemann, A. & Williamson, J.R. (2005) Dual modes of RNA-silencing suppression by Flock House virus protein B2. *Nature Structural & Molecular Biology*, 12, 952–957.

Charrel, R.N., Berenger, J.M., Laroche, M., Ayhan, N., Bitam, I., Delaunay, P. et al. (2018) Neglected vector-borne bacterial diseases and arboviruses in the Mediterranean area. *New Microbes New Infections*, 26, S31–s36.

Crooks, G.E., Hon, G., Chandonia, J.M. & Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Research*, 14, 1188–1190.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O. et al. (2021) Twelve years of SAMtools and BCFtools. *GigaScience*, 10, giab008.

Ferreira, F.V., Aguiar, E., Olmo, R.P., de Oliveira, K.P.V., Silva, E.G., Sant'Anna, M.R.V. et al. (2018) The small non-coding RNA response to virus infection in the *Leishmania* vector *Lutzomyia longipalpis*. *PLoS Neglected Tropical Diseases*, 12, e0006569.

Fonseca, P., Ferreira, F., da Silva, F., Oliveira, L.S., Marques, J.T., Goes-Neto, A. et al. (2020) Characterization of a novel mitovirus of the sand fly *Lutzomyia longipalpis* using genomic and virus-host interaction signatures. *Viruses*, 13(1), 9.

Franzke, K., Leggewie, M., Sreenu, V.B., Jansen, S., Heitmann, A., Welch, S.R. et al. (2018) Detection, infection dynamics and small RNA response against Culex Y virus in mosquito-derived cells. *The Journal of General Virology*, 99, 1739–1745.

Gestuveo, R.J., Parry, R., Dickson, L.B., Lequime, S., Sreenu, V.B., Arnold, M.J. et al. (2022) Mutational analysis of *Aedes aegypti* Dicer 2 provides insights into the biogenesis of antiviral exogenous small interfering RNAs. *PLoS Pathogens*, 18, e1010202.

Jancarova, M., Polanska, N., Volf, P. & Dvorak, V. (2023) The role of sand flies as vectors of viruses other than phleboviruses. *The Journal of General Virology*, 104(4).

Labbe, F., Abdeladhim, M., Abrudan, J., Araki, A.S., Araujo, R.N., Arensburger, P. et al. (2023) Genomic analysis of two phlebotomine sand fly vectors of *Leishmania* from the New and Old World. *PLoS Neglected Tropical Diseases*, 17, e0010862.

Langmead, B. & Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9, 357–359.

Li, D., Liu, C.M., Luo, R., Sadakane, K. & Lam, T.W. (2015) MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, 31, 1674–1676.

Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. (2005) The structure of the flock house virus B2 protein, a viral suppressor of RNA interference, shows a novel mode of double-stranded RNA recognition. *EMBO Reports*, 6, 1149–1155.

Lu, R., Maduro, M., Li, F., Li, H.W., Broitman-Maduro, G., Li, W.X. et al. (2005) Animal virus replication and RNAi-mediated antiviral silencing in *Caenorhabditis elegans*. *Nature*, 436, 1040–1043.

Maroli, M., Feliciangeli, M.D., Bichaud, L., Charrel, R.N. & Gradoni, L. (2013) Phlebotomine sandflies and the spreading of leishmaniases and other diseases of public health concern. *Medical and Veterinary Entomology*, 27, 123–147.

Martins-da-Silva, A., Telleria, E.L., Batista, M., Marchini, F.K., Traub-Cseko, Y.M. & Tempone, A.J. (2018) Identification of secreted proteins involved in nonspecific dsRNA-mediated *Lutzomyia longipalpis* LL5 cell antiviral response. *Viruses*, 10(1), 43.

Moriconi, M., Rugna, G., Calzolari, M., Bellini, R., Albieri, A., Angelini, P. et al. (2017) Phlebotomine sand fly-borne pathogens in the

Mediterranean Basin: human leishmaniasis and phlebovirus infections. *PLoS Neglected Tropical Diseases*, 11, e0005660.

Oerther, S., Jost, H., Heitmann, A., Luhken, R., Kruger, A., Steinhausen, I. et al. (2020) Phlebotomine sand flies in Southwest Germany: an update with records in new locations. *Parasites & Vectors*, 13, 173.

Olson, K.E. & Blair, C.D. (2015) Arbovirus-mosquito interactions: RNAi pathway. *Current Opinion in Virology*, 15, 119–126.

Parry, R. & Asgari, S. (2018) Aedes Anphevirus: an insect-specific virus distributed worldwide in *Aedes aegypti* mosquitoes that has complex interplays with *Wolbachia* and dengue virus infection in cells. *Journal of Virology*, 92(17), e00224-18.

Parry, R., James, M.E. & Asgari, S. (2021) Uncovering the worldwide diversity and evolution of the virome of the mosquitoes *Aedes aegypti* and *Aedes albopictus*. *Microorganisms*, 9(8), 1653.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A. & Kingsford, C. (2017) Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, 14, 417–419.

Petit, M., Mongelli, V., Frangeul, L., Blanc, H., Jiggins, F. & Saleh, M.C. (2016) piRNA pathway is not required for antiviral defense in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America*, 113, E4218–E4227.

Pitaluga, A.N., Mason, P.W. & Traub-Cseko, Y.M. (2008) Non-specific antiviral response detected in RNA-treated cultured cells of the sandfly, *Lutzomyia longipalpis*. *Developmental and Comparative Immunology*, 32, 191–197.

Prince, B.C., Walsh, E., Torres, T.Z.B. & Ruckert, C. (2023) Recognition of arboviruses by the mosquito immune system. *Biomolecules*, 13(7), 1159.

Samuel, G.H., Adelman, Z.N. & Myles, K.M. (2018) Antiviral immunity and virus-mediated antagonism in disease vector mosquitoes. *Trends in Microbiology*, 26, 447–461.

Santos, D., Feng, M., Kolliopoulou, A., Taning, C.N.T., Sun, J. & Swevers, L. (2023) What are the functional roles of Piwi proteins and piRNAs in insects? *Insects*, 14(2), 187.

Schaffner, F., Silaghi, C., Verhulst, N.O., Depaquit, J. & Mathis, A. (2024) The Phlebotomine sand fly fauna of Switzerland revisited. *Medical and Veterinary Entomology*, 38, 13–22.

Telleria, E.L., Martins-da-Silva, A., Tempone, A.J. & Traub-Cseko, Y.M. (2018) *Leishmania*, microbiota and sand fly immunity. *Parasitology*, 145, 1336–1353.

Thorvaldsdottir, H., Robinson, J.T. & Mesirov, J.P. (2013) Integrative genomics viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, 14, 178–192.

Varjak, M., Donald, C.L., Mottram, T.J., Sreenu, V.B., Merits, A., Maringer, K. et al. (2017) Characterization of the Zika virus induced small RNA response in *Aedes aegypti* cells. *PLoS Neglected Tropical Diseases*, 11, e0006010.

Varjak, M., Leggewie, M. & Schnettler, E. (2018) The antiviral piRNA response in mosquitoes? *The Journal of General Virology*, 99, 1551–1562.

Vodovar, N., Goic, B., Blanc, H. & Saleh, M.C. (2011) In silico reconstruction of viral genomes from small RNAs improves virus-derived small interfering RNA profiling. *Journal of Virology*, 85, 11016–11021.

Weger-Lucarelli, J., Ruckert, C., Grubaugh, N.D., Misencik, M.J., Armstrong, P.M., Stenglein, M.D. et al. (2018) Adventitious viruses persistently infect three commonly used mosquito cell lines. *Virology*, 521, 175–180.

Wu, Q., Luo, Y., Lu, R., Lau, N., Lai, E.C., Li, W.X. et al. (2010) Virus discovery by deep sequencing and assembly of virus-derived small silencing RNAs. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 1606–1611.