Agua Salud alphavirus defines a novel lineage of insect-specific alphaviruses discovered in the New World

Kyra Hermanns[1], Marco Marklewitz[1,2,3], Florian Zirkel[4]†, Gijs J. Overheul[5], Rachel A. Page[3], Jose R. Loaiza[3,6,7], Christian Drosten[1,2], Ronald P. van Rij[5] and Sandra Junglen[1,2,*]

Abstract

The genus *Alphavirus* harbours mostly insect-transmitted viruses that cause severe disease in humans, livestock and wildlife. Thus far, only three alphaviruses with a host range restricted to insects have been found in mosquitoes from the Old World, namely Eilat virus (EILV), Taï Forest alphavirus (TALV) and Mwinilunga alphavirus (MWAV). In this study, we found a novel alphavirus in one *Culex declarator* mosquito sampled in Panama. The virus was isolated in C6/36 mosquito cells, and full genome sequencing revealed an 11 468 nt long genome with maximum pairwise nucleotide identity of 62.7 % to Sindbis virus. Phylogenetic analyses placed the virus as a solitary deep rooting lineage in a basal relationship to the Western equine encephalitis antigenic complex and to the clade comprising EILV, TALV and MWAV, indicating the detection of a novel alphavirus, tentatively named Agua Salud alphavirus (ASALV). No growth of ASALV was detected in vertebrate cell lines, including cell lines derived from ectothermic animals, and replication of ASALV was strongly impaired above 31 °C, suggesting that ASALV represents the first insect-restricted alphavirus of the New World.

INTRODUCTION

The genus *Alphavirus* (family *Togaviridae*) includes 31 approved virus species [1] and two putative additional species that have been recently described and await ratification by the International Committee on Virus Taxonomy (ICTV) [2, 3]. Alphaviruses are mostly mosquito-borne viruses that can cause severe diseases in humans, livestock and wildlife. Their host range includes mammals, birds, reptiles, amphibians and fish. Old World alphaviruses, such as chikungunya virus (CHIKV), can cause acute febrile illness with arthralgia, while New World alphaviruses, such as Western equine encephalitis virus (WEEV), are neuroinvasive and can cause encephalitis [4, 5]. The currently approved alphaviruses form 11 serological complexes that group accordingly in phylogenetic analyses [1].

The alphavirus genome consists of single-stranded, positive-sense RNA encoding two ORFs for non-structural proteins (NSPs) and structural proteins (SPs), respectively [1]. The ORFs are framed by untranslated regions (UTRs). The 3′-UTR of most alphaviruses contains repeated sequence elements (RSEs) that seem to be important for the successful infection of insect vectors [6–8]. An exception is the group of salmonid alphaviruses that have short 3′-UTRs without RSEs [9]. These viruses have no known insect vector and were placed in a basal phylogenetic position to all known alphaviruses, suggesting an aquatic origin of the genus [10].

In addition to the vertebrate-infecting groups of mosquito-borne alphaviruses and aquatic alphaviruses, another group of alphaviruses, which is restricted in its host range

to mosquitoes, was discovered in recent years. The group contains three insect-restricted alphaviruses that have been detected in mosquitoes from the Old World [2, 3, 11]. These viruses, Eilat virus (EILV), Taï Forest alphavirus (TALV) and Mwinilunga alphavirus (MWAV), form a monophyletic sister clade to the WEE complex [3, 11]. TALV and MWAV could not be isolated in cell culture. However, EILV was isolated in cell culture from infected mosquitoes but an insect-restricted virus (Negev virus) of the unclassified group of negeviruses was also present in the same cell culture sample [11]. A reverse genetic system of EILV was established to separate EILV from the negevirus and was used for studies on host range restriction [11].

In contrast to insect-restricted viruses of other families, insect-restricted alphaviruses are rarely detected in mosquitoes and show very low prevalence rates in mosquito populations [2, 12]. In addition, screenings of diverse invertebrates for RNA viruses revealed no novel alphaviruses [13, 14]. The objective of this study was to assess the diversity of alphaviruses in mosquitoes of the New World using mosquitoes collected in forested areas of the Panama Canal Zone, Panama.

METHODS

Mosquito collection

In total, 13806 mosquitoes were sampled in the area of the Panama Canal Zone, Central Panama, in 2013 and 2014. Details on mosquito sampling and identification will be described in another publication.

RT-PCR screening

Mosquitoes were homogenized individually in 500 µl Leibovitz's L-15 medium (Gibco – Thermo Fisher Scientific) using ceramic beads and a SpeedMill Plus (Analytik Jena). Stainless steel beads were used for the homogenization of *Culex* mosquitoes. In total, 1414 pools containing five to 11 mosquitoes were generated according to species and sampling location by combining 100 µl cleared supernatant of individual mosquito homogenates. RNA was extracted from pooled supernatants using the MagNA Pure 96 Instrument with the MagNA Pure 96 DNA and Viral NA Small Volume Kit (Roche Diagnostics). SuperScript III reverse transcriptase (Invitrogen – Thermo Fisher Scientific) was used for cDNA synthesis according to the manufacturer's instructions. All pools were screened for alphaviruses with a generic RT-PCR as previously described [2]. For PCR-positive pools, RNA was extracted from homogenates of individual mosquitoes using the QIAamp Viral RNA Mini Kit (Qiagen). cDNA was synthesized and samples were tested for viruses by PCR as described above.

Virus isolation and plaque purification

The homogenate of mosquito pool MP416 was used for primary virus isolation in C6/36 and Vero cells [15]. Briefly, the supernatant of the mosquito suspension was filtrated through a 0.45 µm filter and cells seeded in 24-well plates were infected with 100 µl (F) and 10 µl (F10) of the suspension.

Seven days post-infection (dpi) 100 µl of the supernatant was passaged on fresh cells. This procedure was repeated four times. Cells were observed daily for signs of cytopathic effects (CPEs). A virus stock was generated from the second passage of MP416-F10 and harvested 6 dpi. Agua Salud alphavirus (ASALV) was plaque-purified from this stock using a plaque assay in C6/36 cells as described previously [16]. 1 ml of the serially diluted virus stock was used to infect the cells and after 1 h the inoculum was replaced by the overlay. Plaques were picked 7 dpi using a pipette tip, transferred to fresh C6/36 cells and incubated for 3 days. This was repeated twice with a reduction of the plaque assay incubation time to 6 days. A plaque-purified stock was generated (hereafter named ASALV-PP) and the virus titre was determined by a plaque assay as described above. Deep sequencing was performed to verify the purity of the ASALV-PP stock.

Cell lines

The two *Aedes albopictus* cell lines C6/36 (ECACC 89051705) and U4.4 (obtained from the Radboud Institute for Molecular Life Sciences, Nijmegen, The Netherlands) were used for growth kinetics. C6/36 cells were cultivated in L-15 medium with 5% FCS (Biochrom – Merck KGaA) and 1% L-glutamine (Gibco – Thermo Fisher Scientific). U4.4 cells were cultivated in L-15 medium with 20% FCS, 2% Tryptose Phosphate Broth (Gibco – Thermo Fisher Scientific), 1% Non-Essential Amino Acids (NEAA) (Gibco – Thermo Fisher Scientific) and 1% L-glutamine. Both cell lines were incubated at 28 °C without CO_2. The three fish cell lines BF-2 (*Lepomis macrochirus* – CCLV-RIE 290), CHSE-214 (*Oncorhynchus tshawytscha* – CCLV-RIE 1104) and FHM (*Pimephales promelas* – CCLV-RIE 57) were obtained from the Friedrich-Loeffler-Institute (Greifswald, Germany). BF-2 cells were cultivated in minimal essential medium (MEM) with Hanks' salt (Gibco – Thermo Fisher Scientific) with 10% FCS and 850 mg $NaHCO_3 l^{-1}$ (Gibco – Thermo Fisher Scientific). CHSE-214 cells were cultivated in MEM with Earle's salt (Gibco – Thermo Fisher Scientific) with 10% FCS, 1% NEAA and 120 mg sodium pyruvate l^{-1} (Sigma-Aldrich – Merck KGaA). FHM cells were cultivated in MEM with Hanks' salt with 10% FCS and 850 mg $NaHCO_3 l^{-1}$. The snake cell line VH2 (*Daboia russelii* – GCC 90102539) was cultivated in MEM with Hanks' salt with 10% heat-inactivated FCS, 1% L-glutamine and 1% NEAA. The frog cell line ICR-2A (*Rana pipiens* – HPACC 89072615) was cultivated in L-15 medium with 40% distilled water, 10% FCS and 1% L-glutamine. VH2 and FHM cells were incubated at 28 °C with and without CO_2, respectively. The ICR-2A, BF-2 and CHSE-214 lines were incubated at room temperature.

Growth kinetics and infection of vertebrate cell lines

All infections were performed with ASALV-PP in duplicate. To measure the amount of viral genome copies, RNA was extracted from cell culture supernatant using the NucleoSpin RNA Virus kit (Macherey-Nagel), cDNA was synthesized as described above and a specific quantitative RT-PCR was established (MP416-ASALV-TM-F, 5′-CCGTACTCGAAA

CAGACATTGC-3′; MP416-ASALV-TM-R, 5′-TCGTCAAC GCCTAGATCCTCTA-3′; MP416-ASALV-TM, 5′–6-FAM/ ACAAATCCC/ZEN/AGGACGACTCG/Iowa Black FQ-3′). One day before infection of C6/36 and U4.4 cells, 2.5×10^5 cells per well were seeded in 24-well plates. Cells were infected at an m.o.i. of 0.1 and 0.01 in 300 µl medium without additives and washed three times with 1 ml PBS and once with 1 ml medium with additives after 1 h of incubation at 28 °C. 1 ml medium with additives was added and 50 µl (C6/36) or 75 µl (U4.4) of supernatant was taken for RNA extraction. Samples were taken every 6 h for 48 h and after 72 h (C6/36) or every 24 h for 3 days (U4.4). C6/36 cells were incubated at 28–32 °C to assess the temperature sensitivity of ASALV as described previously [17]. Vertebrate cell lines were inoculated with ASALV to analyse the *in vitro* host range. BF-2, CHSE-214, FHM, ICR-2A and VH2 cells were seeded 1 day before infection (1×10^5 per 24-well) and infected at an m.o.i. of 1 in 300 µl medium without additives. After 1 h of incubation, 700 µl medium with additives was added and 75 µl supernatant was taken for RNA extraction. The supernatant was passaged weekly for four passages and an aliquot of the supernatant was taken before passaging.

Genome sequencing

For deep sequencing, RNA from infectious cell culture supernatants of the second passage of MP416-F10 (ASALV strain PA-2013-MP416) and of the plaque-purified stock ASALV-PP were used. cDNA synthesis and sequencing with the MiSeq desktop sequencer (Illumina) was performed as described previously [2]. Genome ends of ASALV strain PA-2013-MP416 were amplified using the 3′ and 5′ RACE System for Rapid Amplification of cDNA Ends (ThermoFisher Scientific) and PCR products were sequenced by Sanger sequencing (Microsynth).

Genomic and phylogenetic analyses

All sequences were assembled and analysed in Geneious R9.1.8 [18]. ASALV genome analyses were performed based on the full genome sequence of ASALV strain PA-2013-MP416. For phylogenetic analyses, nucleotide sequences of the structural protein ORF encoding the E2, 6K and E1 proteins and the non-structural protein ORF (nsP1-nsP4) excluding parts of nsP3 of all established alphavirus species, TALV, MWAV and ASALV, were aligned by a MAFFT-E v7.308 [19] translational alignment in Geneious. An optimized maximum-likelihood phylogenetic tree with the GTR substitution model and 1000 bootstrap replicates was calculated using PhyML as implemented in Geneious [20]. The trees were rooted to the midpoint.

Small RNA library preparation and analysis

U4.4 cells were cultured as described previously [21], and were seeded at a density of 2×10^6 cells per well in six-well plates and infected the next day with ASALV at an m.o.i. of 0.1. The cells were harvested for total RNA isolation at 72 hpi in RNA-Solv reagent (Omega Biotek R630-02). Small RNA libraries were prepared from 1 µg of total RNA using the NEBNext Multiplex Small RNA Library Prep Kit for Illumina (NEB E7560S). The amplified libraries were size selected on a 6% acrylamide/1× TBE gel and quantified using the Agilent 2100 Bioanalyzer System, and pooled libraries were sequenced on an Illumina HiSeq4000 machine by Plateforme GenomEast. Small RNA reads were mapped to the ASALV genome (ASALV-PP) using Bowtie (Galaxy Tool Version 1.1.2 [22]) allowing for one mismatch and the genome distribution was obtained by plotting the 5′ end position of the ASALV mapping reads on the viral genome. Reads were normalized to total library size (reads per million).

Accession numbers

The complete genome sequence of ASALV strain PA-2013-MP416 was assigned GenBank accession number MK959114. The complete genome sequence of ASALV-PP was assigned GenBank accession number MK959115. The small RNA sequencing data were deposited at the Sequence Reads Archive with accession number PRJNA559096.

RESULTS

Detection of a novel alphavirus

To analyse the genetic diversity of alphaviruses in mosquitoes from the New World, we tested 13806 mosquitoes originating from sylvatic habitats in Panama. Samples were combined into 1414 pools and tested with a generic RT-PCR [2]. One pool (MP416), consisting of 10 *Culex declarator* mosquitoes captured in a forest fragment surrounded by agriculture, contained a sequence showing 75% nucleotide identity to TALV, suggesting the detection of a novel alphavirus. The virus was tentatively named Agua Salud alphavirus (ASALV). Testing of the individual mosquitoes from pool MP416 revealed only one ASALV-positive mosquito (M9506). This corresponds to a prevalence of 0.00724%.

Virus isolation and purification

ASALV was isolated in C6/36 cells by two approaches, either using undiluted mosquito homogenate (MP416-F) or a 1:10 dilution (MP416-F10). MP416-F induced a strong CPE after the second cell culture passage with dead and aggregated cells 3 dpi (Fig. 1a right panel). MP416-F10 induced a weaker CPE with reduced cell growth and rounded or stretched cells 3 dpi (Fig. 1a middle panel). Both supernatants were positive for ASALV by RT-PCR. Due to the observed CPE differences between the two isolates, the supernatants were further tested for other viruses. A virus with 99% nucleotide identity to the negevirus Wallerfield virus (WALV) [23] was detected in the supernatant of MP416-F but not in MP416-F10. Negeviruses are fast growing in cell culture and induce strong CPEs similar to that observed for MP416-F [24]. ASALV was thus plaque-purified from the third passage of MP416-F10 (hereafter named ASALV-PP). ASALV-PP induced distinct, medium-sized plaques in C6/36 cells 6 dpi (Fig. 1b) and deep sequencing of infectious cell culture supernatant confirmed a pure ASALV stock with no other viruses present. ASALV-PP showed 99.97% pairwise nucleotide identity to the wild-type
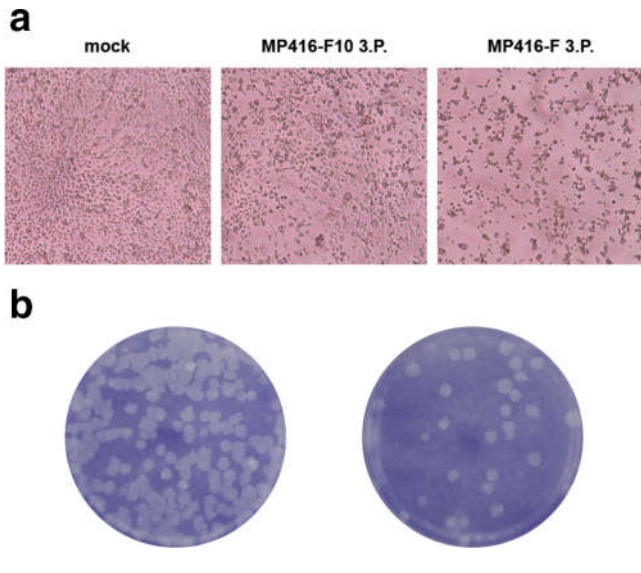
Fig. 1. Virus isolation. (a) Photographs of mock-infected C6/36 cells and cells infected either with filtrated homogenate (MP416-F) or filtrated and 1:10 diluted homogenate (MP416-F10) 3 dpi. (b) Plaque morphology of the plaque-purified strain ASALV-PP in C6/36 cells 6 dpi.

strain, corresponding to four single nucleotide exchanges. Three of these exchanges led to amino acid changes, two in nsP2 (I597V and A627T) and one in nsP4 (V295I).

ASALV genome analysis

The complete genome of ASALV consisted of 11468 nt and showed a typical alphavirus genome organization including conserved sequence elements (CSEs) and cleavage sites (Fig. 2a). ASALV showed a maximal pairwise nucleotide identity of 62.7% to Sindbis virus (SINV). Pairwise comparison of the NSP- and SP-ORFs yielded 66.5 and 52.9 % maximal amino acid identity to Whataroa virus (WHATV), respectively, a mosquito-borne alphavirus from New Zealand. Pairwise protein identities between ASALV and other alphaviruses are shown in Supplementary file 1 (available in the online version of this article). According to the species demarcation criteria for alphaviruses of the ICTV, alphaviruses belonging to different species show in most cases more than 10% divergence in their amino acid sequences of the complete coding regions, although species demarcation is based on a combination of genetic and biological characteristics such as differences in virulence, host or mosquito vector usage [1]. ASALV differed by at least 47.1% in the SP-ORF and by at least 33.5% in the NSP-ORF at the amino acid level from
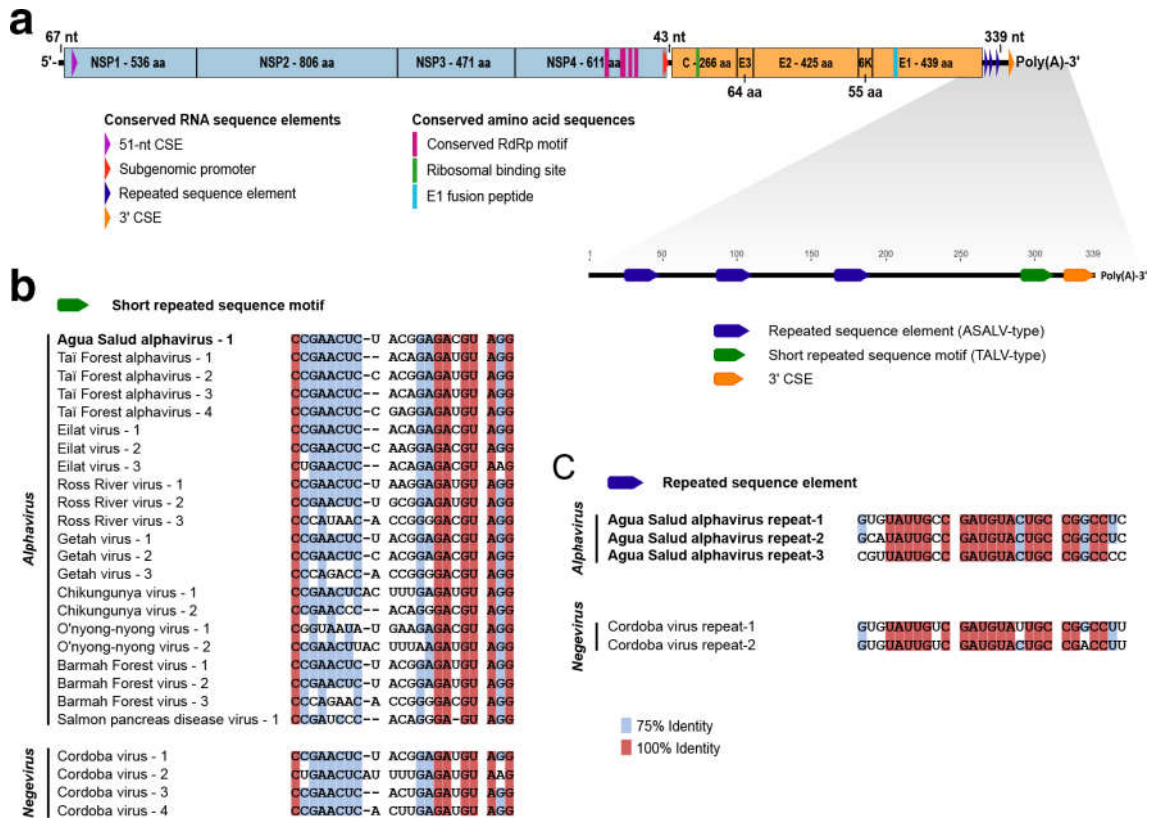

Fig. 2. Genome analyses of ASALV. (a) Schematic illustration of the ASALV genome including motifs and CSEs. The ORFs are indicated in blue (NSP) and orange (SP). Amino acid length of the mature peptides and UTR nucleotide length are displayed. (b) MAFFT-E alignment of the SRS motif of ASALV, representative alphaviruses and Cordoba virus (NC_034156 – unclassified negevirus [25]). For GenBank accession numbers of the alphaviruses see Fig. 3. (c) MAFFT-E alignment of the RSE of ASALV and Cordoba virus.

other alphaviruses, was detected in a different vector species than other alphaviruses and is unlikely to infect mammalian hosts (see below). ASALV is thus proposed to define a novel alphavirus species.

The 3′-UTR of ASALV was 339 nt long and contained one copy of the short repeated sequence (SRS) motif previously described in the 3′-UTRs of TALV and other alphaviruses without U-rich regions [2]. The copy of the SRS motif was 100% identical to the first copy of Getah virus, a mosquito-borne alphavirus that is widespread from Eurasia to Australia (Fig. 2b). In addition, three copies of a novel RSE were detected in the 3′-UTR of ASALV. The novel RSE was not found in other alphaviruses but two copies of it were detected in the putative 3′-UTR of Cordoba virus, an unclassified negevirus [25]. The Cordoba virus RSE copies showed 74–89 % identity to the three copies of ASALV (Fig. 2c). The 3′-UTR of Cordoba virus further contained four copies of the SRS motif present in alphaviruses (Fig. 2b). Similar SRS motifs were also detected in the 3′-UTRs of other negeviruses within the *Nelorpivirus* group and Santana virus.

Phylogenetic analysis

Maximum-likelihood phylogenetic analyses based on the structural polyprotein E2-6K-E1 ORF placed ASALV on a long solitary branch in basal position to the WEE complex and to the clade comprising the insect-restricted viruses TALV, MWAV and EILV from the Old World (Africa and Asia) (Fig. 3). ASALV was in apical position to Trocara virus (TROV), which was isolated from mosquitoes collected in the Amazon Basin in Brazil and Peru, and can infect vertebrate cells, mice and hamsters [26]. ASALV was placed at the same position in phylogenetic analyses based on the non-structural polyprotein ORF (nsP1-nsP4) where ASALV remained basal to SINV, WHATV and Aura virus (AURAV) and to the clade of the insect-restricted viruses, while WEEV and the related recombinant viruses of the WEE complex clustered with Eastern equine encephalitis virus (EEEV) (Supplementary file 2). The observed phylogenetic distance between ASALV and established virus species is in agreement with the genetic distance to other alphaviruses and supports the suggestion that ASALV might define a new alphavirus species.
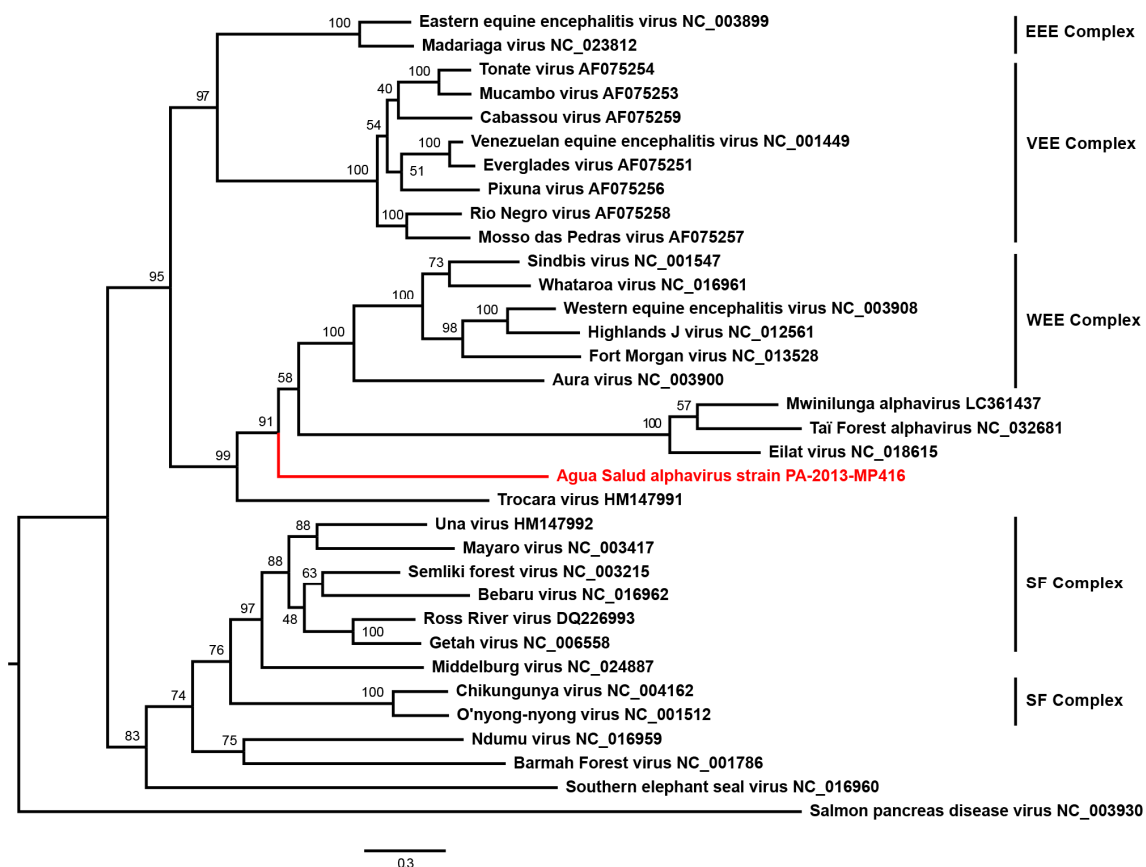

Fig. 3. Phylogenetic relationship of ASALV. The phylogenetic tree was inferred based on a MAFFT-E translational alignment of the nucleotide sequences of the region of the structural protein ORF encoding the E2, 6K and E1 protein of all established alphavirus species, TALV, MWAV and ASALV strain PA-2013-MP416. An optimized maximum-likelihood phylogenetic tree with the GTR substitution model and 1000 bootstrap replicates was calculated using PhyML. The tree was rooted to the midpoint. GenBank accession numbers are shown next to the virus names. EEE, eastern equine encephalitis; VEE, Venezuelan equine encephalitis; WEE, western equine encephalitis; SF, Semliki Forest.

In vitro host range of ASALV

As ASALV branched between the insect-specific and vertebrate-pathogenic alphaviruses, we next analysed the *in vitro* host range. ASALV replicated to high titres 24–48 hpi in the mosquito cells C6/36 (Fig. 4a) and U4.4 (Fig. 4b). Replication kinetics at different temperatures in C6/36 cells revealed a temperature-sensitivity of ASALV with delayed growth at 31 °C and complete block at 32 °C (Fig. 4c). This suggests that ASALV is not able to infect vertebrates with body temperatures over 32 °C, such as mammals and birds. We further tested the ability of ASALV to infect cell lines derived from ectothermic vertebrates (e.g. frogs, snakes and fish) with incubation temperatures between 20 and 30 °C. No replication was detected in the tested vertebrate cells VH2, ICR-2A, FHM, CHSE or BF-2 (Fig. 4d). This was in contrast to previous observations for the vertebrate-infecting alphaviruses WEEV, Fort Morgan virus and Highlands J virus, which could replicate in VH2 and FHM cells [27]. Infection of Vero cells incubated at 30 °C was likewise negative.

Viral small RNA profiles in mosquito cells

To determine if ASALV is targeted by RNAi in mosquito cells, we analysed small RNAs in infected U4.4 cells. As expected, total small RNA reads showed prominent populations of small interfering RNAs (siRNAs) (21 nt), micro RNAs (miRNAs) (~22 nt) and piRNAs (25–30 nt) (Fig. 5a). ASALV derived reads were almost exclusively 21 nt in length, indicative of Dicer-2-dependent biogenesis, and mapped to both the positive and the negative strand in approximately equal numbers (Fig. 5b). These viral siRNAs are distributed across the entire genome, with a slightly higher coverage in the first quarter of the viral genome (Fig. 5c). In contrast to the viral siRNAs, hardly any viral small RNAs in the size range of 25–30 nt were produced, indicating that ASALV is not efficiently processed into piRNAs (Fig. 5b).

DISCUSSION

In this study we discovered and characterized a novel alphavirus defining a deep rooting lineage in a basal phylogenetic relationship to the three insect-restricted alphaviruses EILV, MWAV and TALV, and to the WEE complex. ASALV differs from arboviruses in the WEE complex by its sensitivity to temperatures above 31 °C and the consequential lack of a mammalian host. The other insect-restricted alphaviruses were detected in *Culex pipiens*, *Culex quinquefasciatus*, *Culex decens* and *Anopheles coustani* mosquitoes, while ASALV was found in *Culex declarator* mosquitoes [2, 3, 11, 28]. Based on these genetic and biological characteristics, ASALV probably represents a new alphavirus species and represents the first insect-restricted alphavirus detected in the New World.

Phylogenetic analyses suggest a marine origin of the genus *Alphavirus* [10]. The basal salmonid alphaviruses have no known insect vector and are temperature-sensitive above 15 °C in cell lines derived from fish and mosquitoes [29, 30]. EEEV can infect snakes [31–33] and WEEV and related viruses can replicate in reptilian and fish cells, supporting
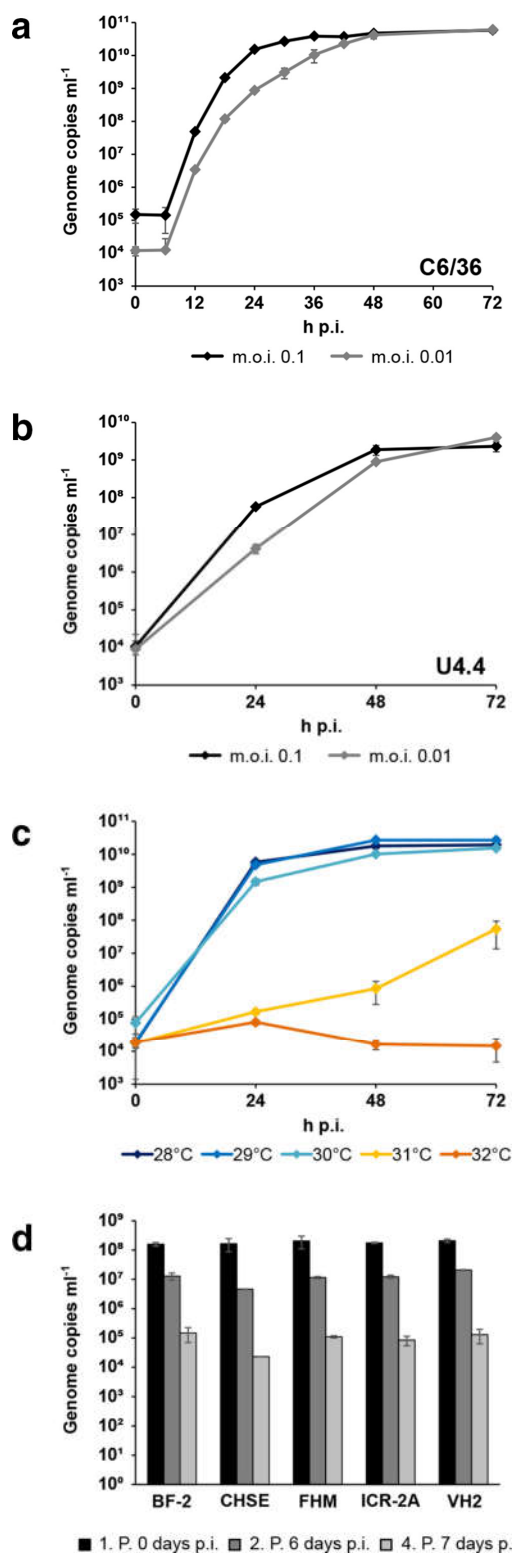
Fig. 4. *In vitro* host range. (a) Growth kinetics of ASALV-PP in C6/36 cells. (b) Growth of ASALV-PP in U4.4 cells. (c) Temperature-dependent replication of ASALV-PP in C6/36 cells infected with an m.o.i. of 0.1. (d) Infection trials with ASALV-PP in cell lines from ectothermic animals with an m.o.i. of 1. Each data point represents the mean of duplicates with standard deviation.
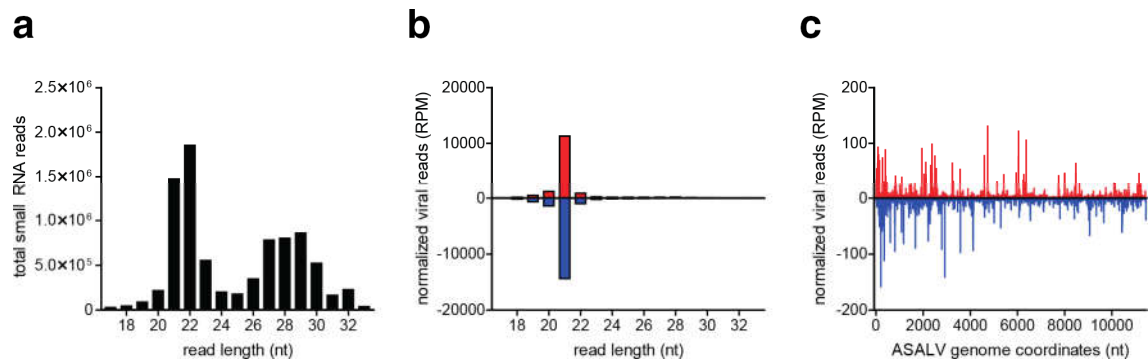
Fig. 5. Small RNA responses targeting ASALV in U4.4 cells. (a) Total number of small RNA reads in libraries from ASALV-infected U4.4 cells. (b) Size distribution of ASALV-derived small RNAs mapping to the viral positive (red) or negative (blue) RNA strand. (c) Location of ASALV-derived 21 nt siRNAs across the genome. Small RNAs were mapped to the viral genome allowing one mismatch and normalized to library size (reads per million, RPM). The 5′ positions of viral siRNAs are plotted.

the theory of a marine origin of alphaviruses [27]. All insect-restricted alphaviruses known to date have been found in *Culex* mosquitoes [2, 3, 28], which are known to feed on ecto-thermic hosts [34, 35]. Because ASALV was isolated from a *Culex declarator* mosquito and was temperature-sensitive, we infected cell lines derived from ectothermic host (amphibians, reptiles and fish) at low temperatures, but in contrast to the related viruses from the WEE complex no virus replication was observed, further confirming that ASALV is an insect-specific virus.

The presence of RSEs in the 3′-UTR of insect-associated alphaviruses and their absence in aquatic alphaviruses may help to understand the evolutionary transition from aquatic to terrestrial alphaviruses. It has been shown that the RSEs in the 3′-UTR of SINV are important for the efficient translation of the viral genome in insect cells [6]. In contrast to the insect-restricted alphaviruses, which contain RSEs, the salmonid alphavirus Sleeping Disease Virus has a short 3′-UTR without RSEs [9] and insertion of SINV RSEs in the 3′-UTR of Sleeping Disease Virus improves the infection of insect cells [6]. Hence, the acquisition of RSEs in the 3′-UTR of aquatic alphaviruses might have enabled the infection and acquisition of insects as hosts [36]. Such a putative host range expansion could have occurred through recombination events with related invertebrate viruses such as negeviruses [13]. Negeviruses are widely distributed in mosquito and sandfly populations from Africa, the Americas, Europe and Asia [25]. They belong to the alphavirus supergroup and fall into two diverse groups in phylogenetic analyses, named *Nelorpivirus* and *Sandewavirus*, in relationship to segmented plant viruses [24, 25, 37]. Inter-estingly, the ASALV 3′-UTR contains RSEs, which are also found in the 3′-UTR of the unclassified negevirus Cordoba virus. ASALV and other alphaviruses further contain an SRS motif, which is also present in the 3′-UTRs of negevi-ruses within the *Nelorpivirus* clade. The presence of similar sequence elements in alpha- and negeviruses may represent putative ancient sites of recombination. RNA recombination occurs in alphaviruses and the first known recombination event was between an EEEV-like and a SINV-like virus from

which WEEV, Highlands J virus and Fort Morgan virus have emerged [27, 38]. In addition, alphaviruses are related to plant viruses of the families *Virgaviridae*, *Bromoviridae* and *Closteroviridae* and it is believed that multiple recombina-tion events took place during the evolution of the alphavirus supergroup [37, 39]. The transfer of functional UTR elements through recombination between different viral families was also previously described for 3′-UTR cap-independent trans-lation enhancers in plant viruses [40].

For the infection of mammalian hosts and the evolution of a dual-host tropism, additional changes, such as adaptation to higher temperatures and immune evasion, were necessary. Adaptions of alphaviruses to higher temperatures seem to have occurred in the past and the temperature sensitivity of alphaviruses is in agreement with the phylogenetic placement. The fish-infecting alphaviruses can replicate at temperatures up to 15 °C [29] whereas the insect-infecting alphaviruses can replicate at temperatures up to 30 °C. ASALV and the other insect-restricted alphaviruses could be intermediates from a transition of alphaviruses with solely aquatic hosts to terrestrial alphaviruses with a dual-host cycle infecting insects and vertebrates.

Further support for this hypothesis and multiple independent adaptation events to vertebrate hosts are peculiarities of AURAV. AURAV is phylogenetically placed apical to the insect-restricted viruses and basal to the arboviruses of the WEE complex. AURAV was isolated from *Culex* sp. and *Aedes serratus* mosquitoes captured in Brazil and Argentina and has no known vertebrate host [41, 42]. Most vertebrate-adapted alphavirus genomes contain an RNA structure, namely downstream loop (DLP), which allows eukaryotic transla-tion initiation factor 2-independent translation initiation if dsRNA-activated protein kinase R (PKR) activation occurs in infected vertebrate cells. In contrast, AURAV has a subop-timal DLP structure and is only able to replicate in BHK cells, which express low levels of PKR, or in PKR knockout cells, indicating an incomplete adaptation to vertebrate hosts [43]. Interestingly, the observed structural differences in the DLP

probably derived from independent introduction events in the clades of SINV, Semliki Forest virus and EEEV [43].

Viruses may be targeted by two small RNA pathways in mosquitoes, the siRNA/RNAi pathway and the piRNA pathway [44], of which the former is considered a major antiviral defence response. We find that ASALV viral RNA is processed into siRNAs, but not piRNAs. This is unexpected as viral piRNAs are produced upon infection of mosquito cells with other alphaviruses such as Chikungunya, Sindbis and Semliki Forest virus [21, 45–47]. The majority of viral piRNAs in these infections derive from the viral positive strand in the region encoding the subgenomic RNA, suggesting that this RNA is a major substrate for viral piRNA biogenesis. While ASALV also produces a subgenomic RNA, our data indicate that this is not sufficient to trigger efficient viral piRNA production. Together, these results suggest that double stranded replication intermediates from ASALV are subject to Dicer-2 cleavage, efficiently producing viral siRNAs that can trigger an RNAi response.

Thus far, only insect-specific alphaviruses in basal phylogenetic relationship to the WEE complex have been detected, but additional insect-specific alphaviruses related to other serological complexes might be discovered in the future. ASALV and additional alphavirus isolates with restricted host range might help to understand the evolution of the dual-host tropism in the genus *Alphavirus*.

Author contributions
Conceptualization: K.H. and S.J.; investigation: K.H., M.M., G.J.O. and F.Z.; formal analysis: K.H. and R.P.V.R.; resources: M.M., R.A.P., J.R.L., S.J. and C.D.; funding acquisition: S.J. and R.P.V.R; writing – original draft preparation, K.H.; writing – review and editing, all authors.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References
1. Chen R, Mukhopadhyay S, Merits A, Bolling B, Nasar F *et al*. ICTV virus taxonomy profile: *Togaviridae*. *J Gen Virol* 2018.

2. Hermanns K, Zirkel F, Kopp A, Marklewitz M, Rwego IB *et al*. Discovery of a novel alphavirus related to Eilat virus. *J Gen Virol* 2017;98:43–49.

3. Torii S, Orba Y, Hang'ombe BM, Mweene AS, Wada Y *et al*. Discovery of *Mwinilunga alphavirus*: a novel alphavirus in Culex mosquitoes in Zambia. *Virus Res* 2018;250:31–36.

4. Chen W, Foo S-S, Sims NA, Herrero LJ, Walsh NC *et al*. Arthritogenic alphaviruses: new insights into arthritis and bone pathology. *Trends Microbiol* 2015;23:35–43.

5. Zacks MA, Paessler S, alphaviruses E. Encephalitic alphaviruses. *Vet Microbiol* 2010;140:281–286.

6. Garcia-Moreno M, Sanz MA, Carrasco L. A viral mRNA motif at the 3'-untranslated region that confers translatability in a cell-specific manner. Implications for virus evolution. *Sci Rep* 2016;6:19217.

7. Ou JH, Trent DW, Strauss JH. The 3'-non-coding regions of alphavirus RNAs contain repeating sequences. *J Mol Biol* 1982;156:719–730.

8. Pfeffer M, Kinney RM, Kaaden OR. The alphavirus 3'-nontranslated region: size heterogeneity and arrangement of repeated sequence elements. *Virology* 1998;240:100–108.

9. Weston J, Villoing S, Bremont M, Castric J, Pfeffer M *et al*. Comparison of two aquatic alphaviruses, salmon pancreas disease virus and sleeping disease virus, by using genome sequence analysis, monoclonal reactivity, and cross-infection. *J Virol* 2002;76:6155–6163.

10. Forrester NL, Palacios G, Tesh RB, Savji N, Guzman H *et al*. Genome-Scale phylogeny of the alphavirus genus suggests a marine origin. *J Virol* 2012;86:2729–2738.

11. Nasar F, Palacios G, Gorchakov RV, Guzman H, Da Rosa APT *et al*. Eilat virus, a unique alphavirus with host range restricted to insects by RNA replication. *Proc Natl Acad Sci U S A* 2012;109:14622–14627.

12. Halbach R, Junglen S, van Rij RP. Mosquito-specific and mosquito-borne viruses: evolution, infection, and host defense. *Curr Opin Insect Sci* 2017;22:16–27.

13. Shi M, Lin X-D, Tian J-H, Chen L-J, Chen X *et al*. Redefining the invertebrate RNA virosphere. *Nature* 2016;540:539–.

14. Chandler JA, Liu RM, Bennett SN. Rna shotgun metagenomic sequencing of northern California (USA) mosquitoes uncovers viruses, bacteria, and fungi. *Front Microbiol* 2015;06:185.

15. Junglen S, Kopp A, Kurth A, Pauli G, Ellerbrok H *et al*. A new flavivirus and a new vector: characterization of a novel flavivirus isolated from uranotaenia mosquitoes from a tropical rain forest. *J Virol* 2009;83:4462–4468.

16. Junglen S, Korries M, Grasse W, Wieseler J, Kopp A *et al*. Host range restriction of insect-specific flaviviruses occurs at several levels of the viral life cycle. *mSphere* 2017;2.

17. Marklewitz M, Zirkel F, Kurth A, Drosten C, Junglen S. Evolutionary and phenotypic analysis of live virus isolates suggests arthropod origin of a pathogenic RNA virus family. *Proc Natl Acad Sci U S A* 2015;112:7536–7541.

18. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M *et al*. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 2012;28:1647–1649.

19. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013;30:772–780.

20. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W *et al*. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 2010;59:307–321.

21. Vodovar N, Bronkhorst AW, van Cleef KWR, Miesen P, Blanc H *et al*. Arbovirus-derived piRNAs exhibit a ping-pong signature in mosquito cells. *PLoS One* 2012;7:e30861.

22. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D *et al*. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res* 2018;46:W537–W544.

23. Auguste AJ, Carrington CVF, Forrester NL, Popov VL, Guzman H *et al*. Characterization of a novel Negevirus and a novel bunyavirus isolated from Culex (Culex) declarator mosquitoes in Trinidad. *J Gen Virol* 2014;95:481–485.

24. Kallies R, Kopp A, Zirkel F, Estrada A, Gillespie T *et al*. Genetic characterization of goutanap virus, a novel virus related to negeviruses, cileviruses and higreviruses. *Viruses* 2014;6:4346–4357.

25. Nunes MRT, Contreras-Gutierrez MA, Guzman H, Martins LC, Barbirato MF *et al*. Genetic characterization, molecular epidemiology, and phylogenetic relationships of insect-specific viruses in the taxon Negevirus. *Virology* 2017;504:152–167.

26. Travassos da Rosa AP, Turell MJ, Watts DM, Powers AM, Vasconcelos PF *et al*. Trocara virus: a newly recognized alphavirus (Togaviridae) isolated from mosquitoes in the Amazon Basin. *Am J Trop Med Hyg* 2001;64:93–97.

27. Allison AB, Stallknecht DE, Holmes EC. Evolutionary genetics and vector adaptation of recombinant viruses of the Western equine encephalitis antigenic complex provides new insights into alphavirus diversity and host switching. *Virology* 2015;474:154–162.

28. Bennouna A, Gil P, El Rhaffouli H, Exbrayat A, Loire E *et al*. Identification of Eilat virus and prevalence of infection among Culex pipiens L. populations, Morocco, 2016. *Virology* 2019;530:85–88.

29. Hikke MC, Verest M, Vlak JM, Pijlman GP. Salmonid alphavirus replication in mosquito cells: towards a novel vaccine production system. *Microb Biotechnol* 2014;7:480–484.

30. McLoughlin MF, Graham DA. Alphavirus infections in salmonids-a review. *J Fish Dis* 2007;30:511–531.

31. Bingham AM, Graham SP, Burkett-Cadena ND, White GS, Hassan HK *et al*. Detection of eastern equine encephalomyelitis virus RNA in North American snakes. *Am J Trop Med Hyg* 2012;87:1140–1144.

32. Graham SP, Hassan HK, Chapman T, White G, Guyer C *et al*. Serosurveillance of eastern equine encephalitis virus in amphibians and reptiles from Alabama, USA. *Am J Trop Med Hyg* 2012;86:540–544.

33. White G, Ottendorfer C, Graham S, Unnasch TR. Competency of reptiles and amphibians for eastern equine encephalitis virus. *Am J Trop Med Hyg* 2011;85:421–425.

34. Burkett-Cadena ND, Eubanks MD, Hassan HK, Unnasch TR, Guyer C *et al*. Blood feeding patterns of potential arbovirus vectors of the genus Culex targeting ectothermic hosts. *Am J Trop Med Hyg* 2008;79:809–815.

35. Cupp EW, Zhang D, Yue X, Cupp MS, Guyer C *et al*. Identification of reptilian and amphibian blood meals from mosquitoes in an eastern equine encephalomyelitis virus focus in central Alabama. *Am J Trop Med Hyg* 2004;71:272–276.

36. Carrasco L, Sanz MA, González-Almela E. The regulation of translation in Alphavirus-Infected cells. *Viruses* 2018;10:70.

37. Wolf YI, Kazlauskas D, Iranzo J, Lucía-Sanz A, Kuhn JH *et al*. Origins and evolution of the global RNA Virome. *mBio* 2018;9.

38. Hahn CS, Lustig S, Strauss EG, Strauss JH. Western equine encephalitis virus is a recombinant virus. *Proc Natl Acad Sci U S A* 1988;85:5997–6001.

39. Strauss JH, Strauss EG, Alphaviruses Rin. Recombination in alphaviruses. *Seminars in Virology* 1997;8:85–94.

40. Miras M, Sempere RN, Kraft JJ, Miller WA, Aranda MA *et al*. Interfamilial recombination between viruses led to acquisition of a novel translation-enhancing RNA element that allows resistance breaking. *New Phytol* 2014;202:233–246.

41. Causey OR, Casals J, Shope RE, UDOMSAKDI S, Udomsakdi S. Aura and Una, two new group A arthropod-borne viruses. *Am J Trop Med Hyg* 1963;12:777–781.

42. Rümenapf T, Strauss EG, Strauss JH. Subgenomic mRNA of aura alphavirus is packaged into virions. *J Virol* 1994;68:56–62.

43. Ventoso I. Adaptive changes in alphavirus mRNA translation allowed colonization of vertebrate hosts. *J Virol* 2012;86:9484–9494.

44. Miesen P, Joosten J, van Rij RP. PIWIs go viral: arbovirus-derived piRNAs in vector mosquitoes. *PLoS Pathog* 2016;12:e1006017.

45. Miesen P, Girardi E, van Rij RP. Distinct sets of PIWI proteins produce arbovirus and transposon-derived piRNAs in *Aedes aegypti* mosquito cells. *Nucleic Acids Res* 2015;43:6545–6556.

46. Morazzani EM, Wiley MR, Murreddu MG, Adelman ZN, Myles KM. Production of virus-derived ping-pong-dependent piRNA-like small RNAs in the mosquito soma. *PLoS Pathog* 2012;8:e1002470.

47. Schnettler E, Donald CL, Human S, Watson M, Siu RWC *et al*. Knockdown of piRNA pathway proteins results in enhanced Semliki Forest virus production in mosquito cells. *J Gen Virol* 2013;94:1680–1689.