UC Riverside
UC Riverside Previously Published Works

Title

A novel virus from Macrosiphum euphorbiae with similarities to members of the family Flaviviridae

Authors

Teixeira, Marcella
Sela, Noa
Ng, James
et al.

A novel virus from *Macrosiphum euphorbiae* with similarities to members of the family *Flaviviridae*

Marcella Teixeira,[1] Noa Sela,[2] James Ng,[3,4] Clare L. Casteel,[5]
Hsuan-Chieh Peng,[3] Sadia Bekal,[6] Thomas Girke,[4,7]
Murad Ghanim[8] and Isgouhi Kaloshian[1,4]

Correspondence
Isgouhi Kaloshian
isgouhi.kaloshian@ucr.edu

[1]Department of Nematology, University of California, Riverside, California, USA

[2]Department of Plant Pathology and Weed Research, Volcani Center, Bet Dagan, Israel

[3]Plant Pathology and Microbiology, University of California, Riverside, California, USA

[4]Institute of Integrative Genome Biology, University of California, Riverside, California, USA

[5]Department of Plant Pathology, University of California, Davis, California, USA

[6]Department of Agricultural and Biological Engineering, University of Illinois, Urbana, IL, USA

[7]Department of Botany and Plant Sciences, University of California, Riverside, California, USA

[8]Department of Entomology, Volcani Center, Bet Dagan, Israel

A virus with a large genome was identified in the transcriptome of the potato aphid (*Macrosiphum euphorbiae*) and was named *Macrosiphum euphorbiae* virus 1 (MeV-1). The MeV-1 genome is 22 780 nt in size, including 3′ and 5′ non-coding regions, with a single large ORF encoding a putative polyprotein of 7333 aa. The C-terminal region of the predicted MeV-1 polyprotein contained sequences with similarities to helicase, methyltransferase and RNA-dependent RNA polymerase (RdRp) motifs, while the N-terminal region lacked any motifs including structural proteins. Phylogenetic analysis of the helicase placed MeV-1 close to pestiviruses, while the RdRp region placed it close to pestiviruses and flaviviruses, suggesting MeV-1 has a positive-polarity ssRNA genome and is a member of the family *Flaviviridae*. Since the MeV-1 genome is predicted to contain a methyltransferase, a gene present typically in flaviviruses but not pestiviruses, MeV-1 is likely a member of the genus *Flavivirus*. MeV-1 was present in nymphal and adult stages of the aphid, aphid saliva and plant tissues fed upon by aphids. However, the virus was unable to multiply and spread in tomato plants. In addition, dsRNA, the replication intermediate of RNA viruses, was isolated from virus-infected *M. euphorbiae* and not from tomato plants infested with the aphid. Furthermore, nymphs laid without exposure to infected plants harboured the virus, indicating that MeV-1 is an aphid-infecting virus likely transmitted transovarially. The virus was present in *M. euphorbiae* populations from Europe but not from North America and was absent in all other aphid species tested.

INTRODUCTION

Aphids are important agricultural pests that cause direct damage to their hosts through feeding and indirect damage by transmitting viruses. Aphids reproduce parthenogenetically by viviparously delivering first instar nymphs and sexually by laying fertilized eggs mostly for overwintering. The first instar resembles the mother and undergoes four moults to become an adult female within a period of about a week. Large populations of aphids can build up rapidly, as a single female can reproduce several nymphs per day and develop quickly. Adult females can be apterous or develop wings in response to crowding, environmental or host-derived cues (De Barro, 1992). To feed, aphids use a pair of stylets that form two slender canals to deliver saliva into the host tissues and acquire plant sap for their nutrition. During feeding, aphids secrete two types of saliva, a gelling saliva that envelops the stylets and forms a sheath as soon as it exits the stylet tip, and a liquid saliva that is delivered throughout the path of stylet penetration, including intracellularly and in the sieve element where aphids feed (Miles, 1968; Tjallingii, 2006).

The potato aphid (*Macrosiphum euphorbiae*) is a generalist with a wide host range including plants in the Solanaceae such as tomato (*Solanum lycopersicum*) and potato (*Solanum tuberosum*). Recently, with the development of genomics and transcriptomic resources and the availability of highly sensitive proteomics technology, the composition of saliva from a number of aphid species, including the potato aphid, has been investigated (Carolan *et al.*, 2009, 2011; Chaudhary *et al.*, 2014, 2015; Cooper *et al.*, 2010, 2011; Harmel *et al.*, 2008; Nicholson *et al.*, 2012; Rao *et al.*, 2013; Vandermoten *et al.*, 2014). Almost all proteins identified in the potato aphid saliva had homologues in the pea aphid (*Acyrthosiphum pisum*), the only aphid with a published genome sequence (International Aphid Genomics Consortium, 2010). While the potato aphid saliva contained proteins with homology to proteins with known functions, the majority of the salivary proteins did not have known motifs and were classified as unknowns (Chaudhary *et al.*, 2015).

Several viruses that infect aphids have been identified. These include two DNA viruses, *Myzus persicae* densovirus (MpDNV) (van Munster *et al.*, 2003b) and *Dysaphis plantaginea* densovirus (DplDNV) (Ryabov *et al.*, 2009), and five RNA viruses, aphid lethal paralysis virus (ALPV) (van Munster *et al.*, 2002), *Rhopalosiphum padi* virus (Moon *et al.*, 1998), *Acyrthosiphon pisum* virus (van der Wilk *et al.*, 1997), *Brevicoryne brassicae* virus (Ryabov, 2007) and rosy apple aphid virus (RAAV) (Ryabov *et al.*, 2009). While these aphid DNA viruses are around 5 kb in size the RNA viruses are larger, around 10 kb in size. Most of these viruses negatively affect aphid survival and fecundity. The best-studied aphid virus is the picorna-like RNA virus ALPV, which can infect several small-grain aphids, including *Rhopalosiphum padi* (Rybicki & von Wechmar, 1982; van Munster *et al.*, 2002; Williamson *et al.*, 1988). Aphids infected with ALPV display uncoordinated movements, followed by paralysis and rapid population decline (Williamson *et al.*, 1988). Therefore, identification of viruses has been mainly based on altered phenotype of the aphid host.

The recent advent of high-throughput sequencing, such as RNASeq, has accelerated the identification of virus sequence discovery in organisms including arthropods (Chiu, 2013; Junglen & Drosten, 2013; Li *et al.*, 2015; Lipkin & Firth, 2013; Liu *et al.*, 2015). Here we report the identification of a novel virus from *M. euphorbiae* by sequencing and assembly of the aphid transcriptome. Characterization of the virus genome and its presence in the plant host, various aphid species and aphid populations from New and Old World countries are described.

RESULTS

### Identification of a virus genome in the potato aphid transcriptome

To build genomics resources for *M. euphorbiae*, the transcriptome of this aphid was assembled from mixed developmental stages by Illumina sequencing. The predicted *M. euphorbiae* proteome derived from this transcriptome was used to assist in the identification of potato aphid salivary proteins by nanoflow liquid chromatography tandem MS (Chaudhary *et al.*, 2014, 2015). Over 105 proteins were identified in the saliva of this aphid. Among these proteins were peptides matching three contigs that had no matches to sequences in the National Center for Biotechnology Information (NCBI) database or the pea aphid genome. Therefore, these contigs were annotated as unknowns and considered to be potato-aphid-specific. Improvements to the potato aphid transcriptome assembly incorporated these three contigs into a larger transcript of 22 780 nt. The transcript contained a single ORF, with translation initiation and termination sites, encoding a protein of 7333 aa. The identification of a long single ORF suggested proper assembly of this transcript. To confirm this, we subcloned and sequenced the putative coding region. Sequence analysis of the overlapping subclones indicated proper assembly of the transcript (data not shown).

The transcript had a 290 nt 5′ non-coding region (NCR) and a 472 nt 3′ NCR ending with only four adenines, suggesting the likely absence of a poly(A) tract. Since this transcript was discovered by RNA-Seq, which is biased toward identification of poly(A)-containing RNA, we investigated the presence of a longer poly(A) tract at the 3′ end of this transcript using 3′ RACE. The 3′ RACE indicated the presence of a single and complete 3′ NCR sequence and added 16 additional adenines to the existing four-adenine sequence, indicating the presence of a poly(A) tract.

To compare this sequence with known proteins, the full-length sequence was used in BLAST analysis. The full-length transcript (GenBank number KT309079) was too long for this analysis. Therefore, the sequence was divided into two and used in BLASTX searches against NCBI databases. The N-terminal half revealed similarity to short stretches of amino acid sequences. The first region (567–636 aa) had similarity to the recently identified plant virus gentian Kobu-sho-associated virus (GKaV; 30/70 and E value 2^{-10}) (Kobayashi *et al.*, 2013) and to RAAV (21/51 and E value 2^{-4}) (Ryabov *et al.*, 2009). The second region (201–338 aa) had similarity to *O*-acetyl-ADP-ribose deacetylase 1 containing Macro domain Poa1p, a high-affinity ADP ribose-binding module present in a variety of proteins from different organisms, including an *A. pisum* protein (XP001947976.2; 43/129 and E value 1^{-7}). BLASTX analysis of the C-terminal half of the KT309079 transcript revealed top hits to the recently characterized soybean cyst nematode virus 5 (SbCNV-5) (E value 2^{-50}) (Bekal *et al.*, 2014) and to GKaV (E value 4^{-38}). In addition, BLASTX revealed weak similarities to helicases from different organisms, including bacteria (E value 3^{-11}), fungi (E value 5^{-11}) and insects (E values 9^{-11} to 1^{-9}).

Domain search analysis using a protein family database (Pfam) (Finn *et al.*, 2014) confirmed the presence of the
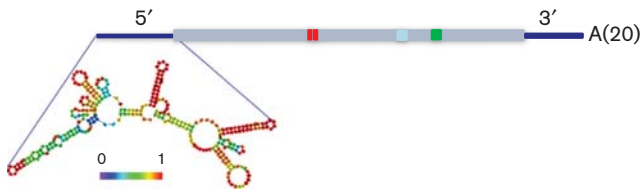
Fig. 1. Predicted RNA structures in the 5′ NCR and conserved amino acid sequence motifs in the *Macrosiphum euphorbiae* virus 1 (MeV-1) polyprotein. Red boxes represent the location of the helicase domain (4066–4415 aa), the pale blue box represents the location of the methyltransferase domain (6201–6419 aa) and the green box represents the location of the RdRp domain (6857–7070 aa). A predicted stable RNA hairpin loop exists in the MeV-1 5′ NCR. The rainbow scale indicates a range of base-pairing probabilities from 0 to 1, violet to red.

helicase (E value 5.24^{-13}) domain and identified two additional domains for RNA-dependent RNA polymerase (RdRp) (E value 2.48^{-06}) and methyltransferase (E value 2.9^{-08}) (Fig. 1). Methyltransferases are involved in virus capping, while RdRp and helicase are essential proteins for RNA virus replication. The similarity of KT309079

transcript to viruses and identification of virus-associated domains indicated that this transcript is of viral origin, seemingly exclusive to *M. euphorbiae*, and was therefore named *Macrosiphum euphorbiae* virus 1 (MeV-1).
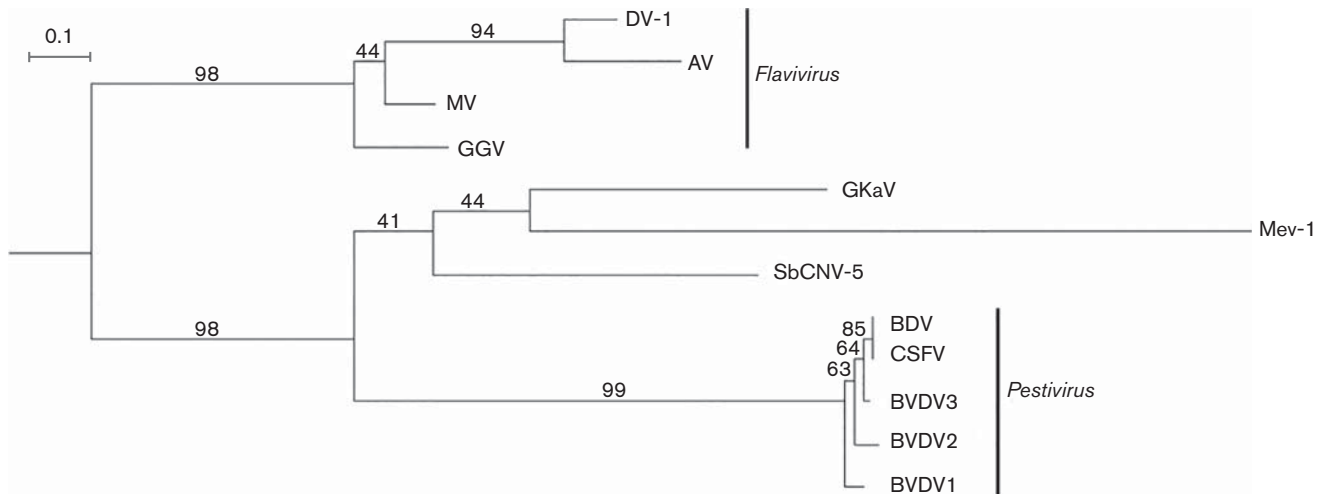
MeV-1 belongs to the family *Flaviviridae*

Both SbCNV-5 and GKaV belong to the family *Flaviviridae*, which contains three genera, *Flavivirus*, *Hepacivirus* and *Pestivirus*. Members of this family are positive-polarity ssRNA viruses that encode a single ORF, which is translated into a polyprotein subsequently cleaved into structural and non-structural proteins (Liu *et al.*, 2014). The N-terminal half of the polyprotein encodes the structural proteins, composed of a small basic capsid protein and two or three envelope proteins (King *et al.*, 2012). The C-terminal half encodes non-structural proteins, including a serine protease, in addition to the RNA helicase, methyltransferase and RdRp (King *et al.*, 2012). Besides the helicase, methyltransferase and RdRp domain, searches did not reveal any additional domains in the MeV-1 ORF, including structural domains. Instead, in the 5′ NCR of MeV-1, multiple stable hairpins were identified (Fig. 1).

Phylogenetic analysis of the conserved RdRp domain of members of the family *Flaviviridae* placed MeV-1 close to the
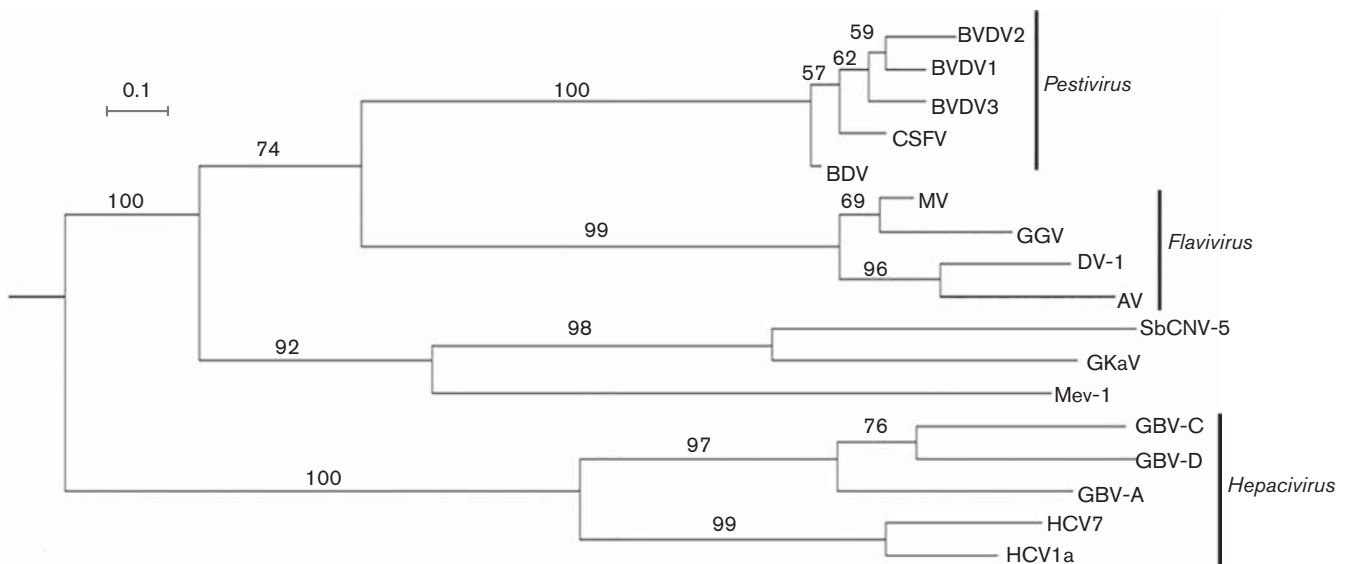
Fig. 2. Phylogenetic analysis of the RdRp domains of MeV-1 and members of the family *Flaviviridae*. The RdRp protein regions were aligned with that of hepatitis C virus (HCV; GenBank accession no. AF011751), aa 2521–2795. The viruses used in the comparison were: Meaban virus (MV; ABBB90668), Gadgets Gully virus (GGV; ABBB90669), dengue virus type 1 (DV1; AAB70695), Aroa virus (AV; AAC58750), SbCNV-5 (YP_009028573), GKaV (YP_007438864), classical swine fever virus (CSFV; CAA61161), border disease virus (BDV; AAB60887), bovine viral diarrhea virus 1 (BVDV1; AAA42854), BVDV2 (NP_044731.1), BVDV3 (ACM79934), HCV1a (AAB67036), HCV7 (ABN05226), GB virus A (GBV-A; AAB71133), GBV-C (AAC55951) and GBV-D (ADK12630). Genus designations are shown to the right of the phylogram. Numbers and nodes represent bootstrap values as percentages. The lengths of branches are proportional to the numbers of changes. Bar, 0.1 aa substitutions per site.

Fig. 3. Phylogenetic analysis of the helicase domains of MeV-1 and members of the family *Flaviviridae*. Helicase domains were aligned with that of BVDV1, aa 2095–2190. See Fig. 2 for abbreviations of the virus names used in the comparison. Genus designations are shown to the right of the phylogram. Numbers and nodes represent bootstrap values as percentages. The lengths of branches are proportional to the numbers of changes. Bar, 0.1 aa substitutions per site.

genera *Flavivirus* and *Pestivirus* (Fig. 2). A second phylogenetic analysis, performed with the helicase domain, placed MeV-1 as a sister clade to the genus *Pestivirus* (Fig. 3). However, the presence of a methyltransferase domain, which is associated with flaviviruses but not pestiviruses (King *et al.*, 2012), indicates that MeV-1 is likely a member of the genus *Flavivirus*.

MeV-1 replicates in *M. euphorbiae* and is present in all aphid developmental stages

Mixed developmental stages of *M. euphorbiae* were used to assemble its transcriptome. To investigate MeV-1 presence in different aphid developmental stages, adults and third and fourth instars were tested. MeV-1 was detected in all these life stages (Fig. 4a), suggesting the virus is vertically transmitted. To confirm the transovarial transmission of the virus, adult aphids collected as first instar nymphs from the posterior ends of the mothers while being delivered, and grown on naive tomato plants, were also tested for the presence of the virus. Of the 20 aphids tested, all were positive for MeV-1 (Fig. 4b; and data not shown).

Some previously identified aphid-infecting viruses were detected in only a proportion of the aphids in a colony (Ryabov, 2007, 2009). To test the frequency of virus-infected aphids in the *M. euphorbiae* colony, randomly selected single adult aphids were tested for the presence of the virus. The virus was present in all adult aphids tested, suggesting the virus is ubiquitously present in our potato aphid colony (Fig. 4c).

The ubiquitous presence of the virus in the potato aphid population suggested the possibility that the virus was integrated into the aphid genome. Previous work has shown that sequences homologous to RNA viruses are integrated

into host genomes and are actively transcribed (Belyi *et al.*, 2010; Taylor *et al.*, 2010). To investigate whether MeV-1 is also integrated into the aphid genome, aphid DNA was used as a template in PCR with the virus primers. The aphid ribosomal gene was amplified as a control. Although a product of the expected size was detected
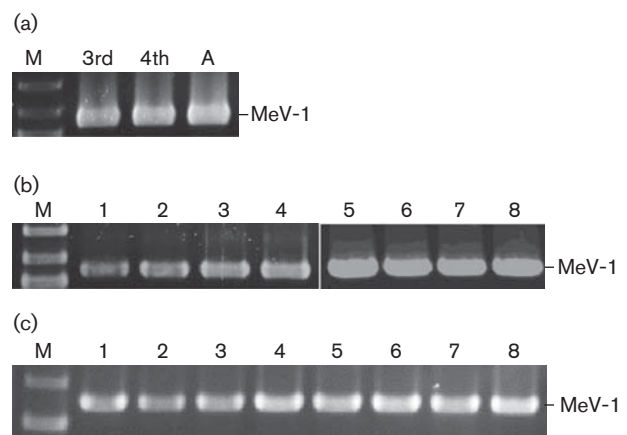
Fig. 4. MeV-1 is ubiquitously present in the *M. euphorbiae* (WU11) colony and is vertically transmitted to progeny. MeV-1 transcripts were detected using reverse transcription PCR (RT-PCR). (a) Third and fourth instar nymphs and young adults were processed for MeV-1. (b) First instar nymphs were collected with a brush from adult aphids while being laid, before touching tomato leaflets, and transferred to a naive tomato plant. One week later, when nymphs had moulted into adults, single aphids were processed for MeV-1 presence. (c) Single adult aphids were randomly collected and tested for MeV-1 presence. M, Molecular mass marker.

with the ribosomal primers, no product was detected with the MeV-1 primers, indicating that MeV-1 is not integrated into the potato aphid genome (data not shown).

To determine whether the virus is able to replicate in the aphid, MeV-1-infected aphids were processed and analysed for the presence of dsRNA, the replication intermediate of RNA viruses. Leaves of cucumber (*Cucumis sativus*) infected with cucurbit yellow stunting disorder virus (CYSDV), a positive-polarity ssRNA virus, were used as positive control for dsRNA isolation. In agarose gel analysis, a major band of over 19 kb was detected in the MeV-1-infected aphid sample treated with DNase I, RNase A and proteinase K (Fig. 5a), indicating the presence of dsRNA and strongly suggesting that MeV-1 is able to multiply in the aphid. In addition to this band, a larger, more intensely stained dsRNA band appeared upon concentrating the aphid sample treated with DNase I, RNase A

and proteinase K (Fig. 5a) and not in the dilution of the concentrates (Fig. 5b). It is likely that this larger band is an artefact of the concentration procedure. The identity of this larger RNA species has not been established and the presence of an additional virus in this aphid colony has not been ruled out.

MeV-1 is specific to *M. euphorbiae* populations from Europe

BLASTX searches indicated that MeV-1 is a unique virus, not yet reported, to our knowledge, from any other aphid species, and seems to be restricted to *M. euphorbiae*. To confirm the virus is exclusive to *M. euphorbiae*, five additional aphid species (*Aphis fabae*, *Aphis gossypii*, *A. pisum*, *Macrosiphum rosae* and *Myzus persicae*) from four countries were tested for the presence of MeV-1 (Fig. 6a). No MeV-1 was detected in any of these aphid
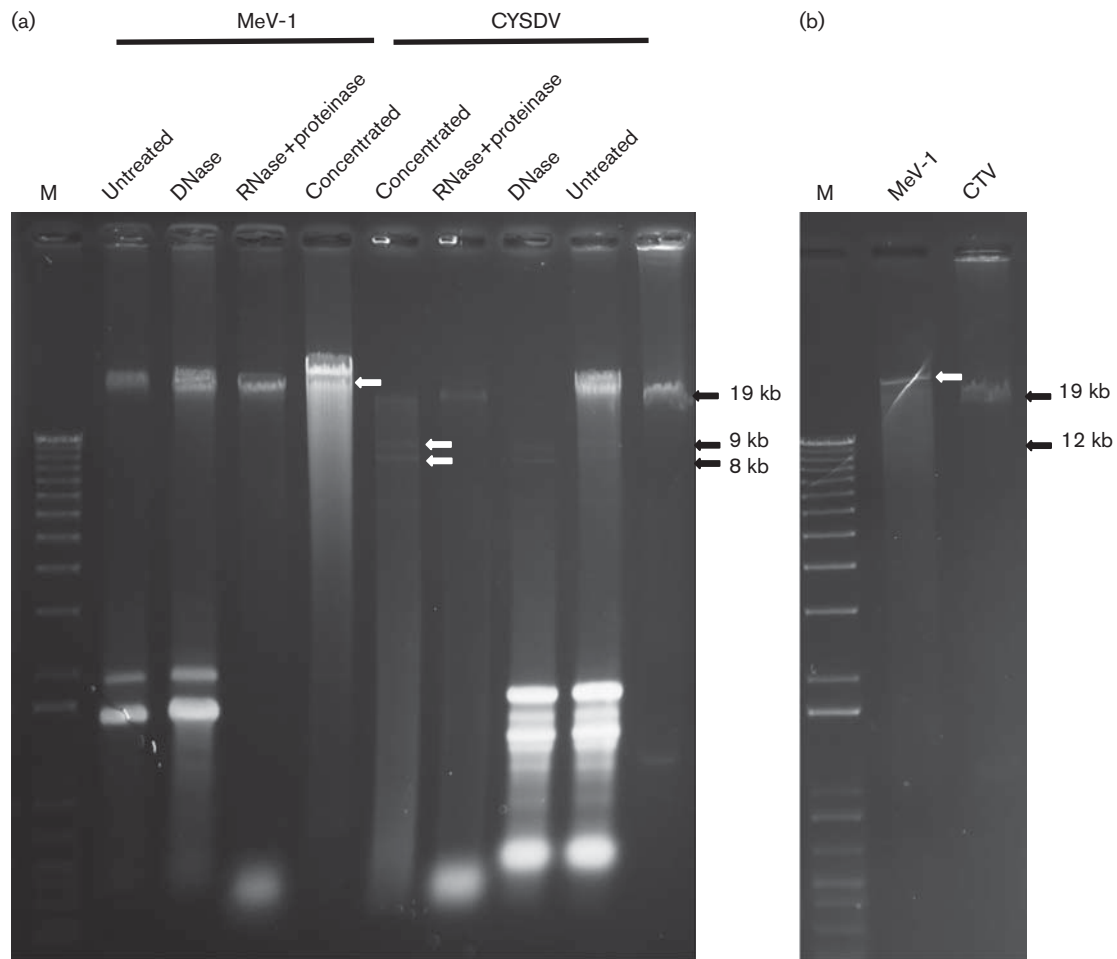
Fig. 5. Analysis of dsRNAs recovered from MeV-1-infected *M. euphorbiae*. The dsRNAs of CYSDV-infected cucumber were used as a positive control. (a) MeV-1 and CYSDV dsRNAs (arrows) were isolated, using CF11 cellulose chromatography, treated with DNase I, RNase A and proteinase K, and concentrated. (b) Concentrated MeV-1 dsRNA was diluted 1:5. (a, b) PCR-amplified full-length cDNA (19 296 bp) of the *Citrus tristeza* virus (CTV; strain T36) and a 1 kb ladder (M) were used as size standards.
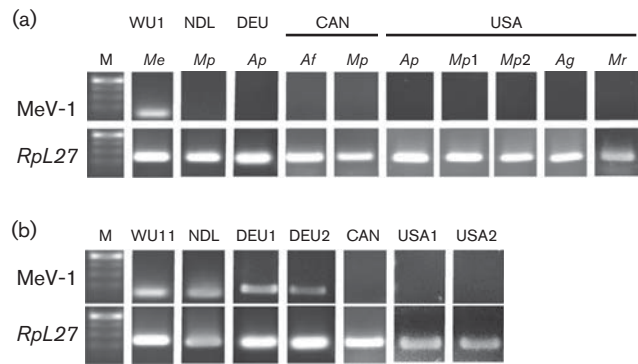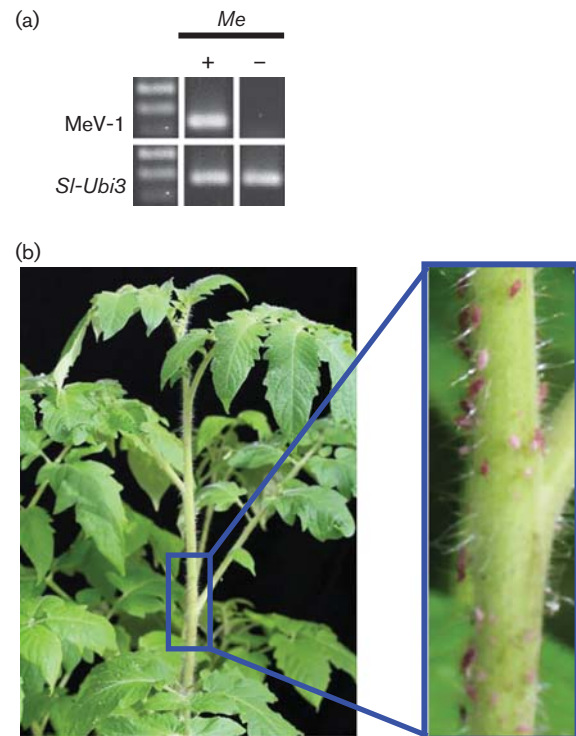
Fig. 6. MeV-1 is present only in European populations of *M. euphorbiae* (*Me*). Aphids were tested for the presence of MeV-1 using RT-PCR. (a) Aphid species: *Myzus persicae* (*Mp*); *A. pisum* (*Ap*); *Aphis fabae* (*Af*), *Aphis gossypii* (*Ag*) and *M. rosae* (*Mr*) from The Netherlands (NDL), Germany (DEU), Canada (CAN) and the USA. (b) *M. euphorbiae* from NDL, DEU, CAN and USA. WU11 is the *M. euphorbiae* colony from which the virus was identified and came originally from France. Arabic numerals indicate different populations of the aphid species. M, Molecular mass marker. Aphid ribosomal gene *RpL27* was used as a positive control.

Fig. 7. MeV-1 is delivered by *M. euphorbiae* (*Me*) into plant tissues during feeding but does not cause symptoms in tomato plants. (a) Leaves from naive tomato plants or plants infested for 2 weeks with *M. euphorbiae* were used in RT-PCR for MeV-1 detection. *Sl-Ubi3* was used as control. (b) Tomato plants infested with *M. euphorbiae* WU11 population for 4 weeks showed no symptoms of viral infection.

species. The primers used for the virus detection annealed to the RdRp and the 3′ NCR. To exclude the possibility that variants of MeV-1 could be present in these aphids and that the primer pair utilized was not specific to the variants, we used a different primer pair (MevirF2 and MevirR1), in the conserved RdRp domain, to detect the virus. No virus was detected with the second set of primers, confirming the previous findings that MeV-1 is not present in these aphid species (data not shown).

The population of *M. euphorbiae* in which MeV-1 was identified was originally acquired from France. To test for the presence of this virus among *M. euphorbiae* populations from different continents, *M. euphorbiae* was obtained from The Netherlands, Germany, Canada and the USA. MeV-1 was detected in the European populations but not in the Canadian or the USA populations (Fig. 6b). Taken together, these results indicate that this virus is likely specific to *M. euphorbiae* and may not be present in North America.

MeV-1 is not capable of independent systemic movement in tomato, but is delivered into tomato leaves by *M. euphorbiae*

Detection of peptides derived from MeV-1 proteins in *M. euphorbiae* saliva suggested that the virus is delivered into plant tissues during aphid feeding. To test for the presence of MeV-1 in plant tissues, tomato leaflets heavily infested with MeV-1-infected *M. euphorbiae* were used. MeV-1 was detected in leaves of aphid-infested plants but not in leaves of control naive plants not exposed to aphids (Fig. 7). To confirm that the amplification products

obtained were from MeV-1, PCR amplicons from aphid-infested tomato leaflets were sequenced. The sequence of the amplified products showed 100% identity to MeV-1 (data not shown).

Considering that *M. euphorbiae* is a pest known to vector plant viruses (Blackman & Eastop, 2000), MeV-1 could be a plant virus transmitted by this aphid. However, no symptoms characteristic of viral infections were observed in aphid-infested tomato plants (Fig. 7b). To investigate whether MeV-1 is a plant virus, the presence of the virus was monitored at 6 h, 24 h, 48 h, 72 h and 2 weeks post-MeV-1-infected aphid infestation. The virus was detected only in the 2-week-infested leaflets, suggesting slow accumulation of the virus in tomato leaves (data not shown).

To determine whether MeV-1 was able to multiply and spread in plants, virus presence was evaluated systemically. MeV-1-infected aphids were caged onto tomato leaflets for 2 weeks and the upper uninfested leaflets, without contact with aphids, were used for virus detection. No virus was detected in the upper uninfested leaflets (data not shown), indicating MeV-1 does not move systemically.

(a)

(b)
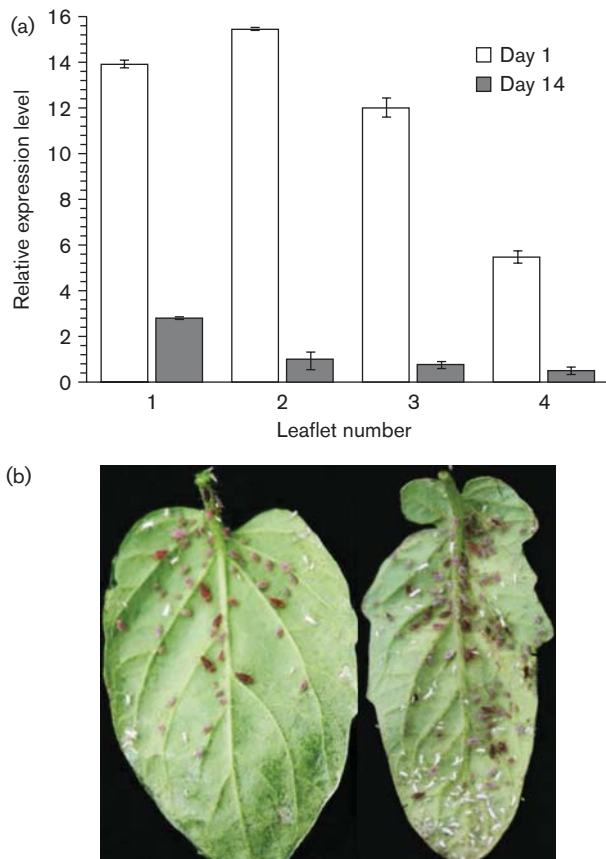
Fig. 8. MeV-1 titre in tomato half-leaflets measured by RT-quantitative PCR (qPCR). (a) Four heavily infested tomato plants were cleared of *M. euphorbiae*. A leaflet on each plant was cut into halves longitudinally through the midrib. The detached half was processed immediately for MeV-1 detection. The second half was left attached to the plant, free of aphids, for 14 additional days before processing for MeV-1 detection. Bars indicate mean of two technical replicates, and error bars show ±SD. (b) Tomato leaflets of similar aphid infestation levels were used for this experiment.

Heavily infested tomato leaves with MeV-1-infected aphids were also analyzed for the presence of dsRNA after clearing of the leaves from the aphids. No dsRNA was observed in these aphid infested leaves (data not shown) suggesting that MeV-1 most likely does not replicate in tomato.

To confirm that MeV-1 is not a plant virus, the dynamics of the virus titre within a leaflet were investigated. Heavily infested tomato plants were cleared of MeV-1-infected aphids, and leaflets were cut through the midvein, collecting half of the leaflet and leaving the second half attached to the plant. Analysing the first halves of the leaflets for the presence of MeV-1, the virus could be detected in all leaflets tested albeit at different titres (Fig. 8a, b). However, 2 weeks later, the virus titre was dramatically reduced in the second halves of these leaflets (Fig. 8a), suggesting

elimination of the viral RNA. Taken together, our data indicate that MeV-1 is not a plant virus.

DISCUSSION

Our work is an example for identification of a novel virus from aphids by means of high-throughput sequencing technology. This virus, MeV-1, has a positive-polarity ssRNA genome with a single ORF encoding a polyprotein, and is the largest aphid virus genome identified to date. BLAST analysis showed that MeV-1 has similarities to members of the family *Flaviviridae*, which is mainly composed of arthropod or arthropod-borne viruses, and is the only aphid virus identified to date that belongs to this family. While phylogenetic analysis of the RdRp domains does not clearly place MeV-1 and its close relatives, SbCNV-5 and GKaV, to a clade within this family, the helicase domains place these three viruses in the *Pestivirus* clade. Although MeV-1, SbCNV-5 and GKaV do not appear to encode the non-structural protein 1, which is highly conserved among flaviviruses, all three viruses have predicted methyltransferase domains, which are present in flaviviruses and not in pestiviruses (King *et al.*, 2012; Vlachakis *et al.*, 2013). Methyltransferases are responsible for genome capping in flaviviruses (Vlachakis *et al.*, 2013). In addition, the 5′ NCRs of both MeV-1 and SbCNV-5 contain stable hairpin structures observed in members of the genus *Flavivirus* and important for translation. Although not reported earlier, close inspection of the 5′ NCR of GKaV also showed hairpin structures. The 5′ NCRs of members of the other two genera in *Flaviviridae*, *Pestivirus* and *Hepacivirus*, have an internal ribosome entry site, a translation initiation site recognized by the host ribosome (Daly & Ward, 2003). Taken together, this information indicates that MeV-1 is closer to members of the genus *Flavivirus* or may constitute a part of a new virus clade that bridges the genera *Flavivirus* and *Pestivirus*. As more viruses are identified through high-throughput sequencing technologies, the diversity of viruses, including arthropod viruses, will increase and will likely result in new phylogenetic clades.

MeV-1 contains a poly(A) tract, which is not common among the members of the family *Flaviviridae*. Although it is not clear whether GKaV has a poly(A) tract, it is likely that SbCNV-5 has a poly(A) tract as it was discovered among the transcriptome sequences of the soybean cyst nematode developed by reverse transcription (RT) using poly(A)-RNA-enriched templates (Bekal *et al.*, 2011). During assembly of transcriptomes developed by high-throughput sequencing, sequence ends are trimmed and it is likely that the poly(A) tract from SbCNV-5 was removed as in the MeV-1 sequences. Our approach, however, did not determine the location of the poly(A) tract within the 3′ NCR and whether the poly(A) is located internally. Internally located poly(A) tracts, though rare, have been previously reported for both human and plant viruses, including the human-pathogenic chikungunya

virus (Khan *et al.*, 2002), the tick-borne encephalitis virus (Asghar *et al.*, 2014), the plant-pathogenic barley strip mosaic virus (Agranovsky *et al.*, 1978) and the hibiscus latent Singapore virus (Srinivasan *et al.*, 2005).

Considering that members of the *Flaviviridae* are exclusive to insects that feed on animals or are transmitted by insects to animals, the discovery of SbCNV-5 and MeV-1 in plant-feeding nematodes and aphids adds to the diversity of this family of viruses. Although GKaV was detected from symptomatic gentian plants, no vector transmitting this virus was linked with the disease. In contrast, both MeV-1 and SbCNV-5 are present in plant pests, with no visible virus symptoms associated with their respective infested plant hosts.

The detection by MS of viral peptides in *M. euphorbiae* saliva suggests that the MeV-1 virion is able to cross salivary gland membrane barriers and is delivered into tomato leaves via the saliva (Chaudhary *et al.*, 2014). Indeed, RT-PCR confirmed the presence of viral RNA in tomato leaflets. No viral dsRNA was detected in aphid-cleared plant tissues previously heavily infested with MeV-1-infected potato aphids, and the virus titre decreased dramatically in leaves after removal of MeV-1-infected aphids. This indicates that MeV-1 is not able to replicate in tomato and suggests plant immune responses are involved in eliminating the virus. Alternatively, host factors required for MeV-1 replication are missing in tomato. Most aphid viruses reported to date are detected in the aphid saliva or plant host tissues (Ban *et al.*, 2007; Ryabov *et al.*, 2009; van Munster *et al.*, 2003a). Although viral multiplication in plant tissues has not been investigated for all these viruses, aphid viruses do not seem to multiply in the plants (Ryabov *et al.*, 2009). Since MeV-1 is delivered into plants and can be detected in plants, the virus could be transmitted horizontally from plants to aphids.

In our colony, potato aphids seem to be universally infected with MeV-1, as the virus was detected from every single aphid tested as well as from numerous small populations of the aphid started from a single female (data not shown). In addition, all adult aphids collected as first instar nymphs, while being delivered by their mothers and grown on naive plants, were also infected with MeV-1, suggesting that the virus is transmitted vertically and likely transovarially. Vertical transmission from mother to nymphs has been demonstrated for DplDNV in the rosy apple aphid, and for MpDNV in *Myzus persicae* (Ryabov *et al.*, 2009; van Munster *et al.*, 2003a). However, the vertical transmission in both cases did not seem to be absolute as not all the progeny of virus-infected females were infected with the virus. The reason for this selective transmission remains unknown.

The detection of a large molecular mass dsRNA of similar size to the MeV-1 transcript in MeV-1-infected *M. euphorbiae* indicates that this virus is able to multiply in the aphid. In addition, the absence of dsRNA in the tomato plants

infected with the virus indicates that MeV-1 is likely an aphid virus and not a plant virus. The initial *M. euphorbiae* population (WU11) from which MeV-1 was identified originated from France and had been reared under laboratory conditions for at least 15 years. Therefore, the extent of MeV-1 infection among *M. euphorbiae* populations was unknown. Moreover, it was not clear whether MeV-1 was also present in additional aphid species. Interestingly, surveying European and North American *M. euphorbiae* populations showed that the virus is only present in *M. euphorbiae* from European countries and not in North American populations, suggesting movement of aphid populations in these neighbouring European countries. It is likely, therefore, that this virus is present in *M. euphorbiae* from additional countries in Europe. A survey of the presence of MeV-1 among additional aphid species did not detect the virus in any of the five aphid species tested from either North America or Europe, suggesting MeV-1 is limited to *M. euphorbiae*. Whether the virus can be transmitted vertically through the plant host to these additional aphid species, or to additional insect species such as whiteflies, remains to be investigated.

The effect of MeV-1 on the *M. euphorbiae* population strain WU11 is unclear since aphids from this virus-infected population do not show obvious pathology or unusual phenotypic characteristics. The identification of MeV-1-free *M. euphorbiae* populations will allow investigation of whether the virus confers subtle phenotypes to the aphid. Horizontal transfer of MeV-1 from host to virus-free *M. euphorbiae* populations will allow investigation of subtle differences in aphid biology and host preference in the same aphid population with an identical genetic background.

METHODS

Tomato growth and aphid infestation. Tomato cv. UC82b was used. Germinated seedlings, planted into Edna's Best Potting Soil (E.B. Stone Organics, San Jose, CA) in 946 ml styrofoam cups, were maintained in a clean plant growth room with 16 h light photoperiod at 24 °C. Plants were fertilized weekly with MiracleGro (18-18-21; Stern's MiracleGro Products).

For time-course aphid infestation experiments, 10 aphids were caged onto each tomato leaflet. After 6 h, 24 h, 42 h, 72 h or 2 weeks of aphid infestation, aphids were removed and leaflets collected and frozen for further processing. For systemic MeV-1 movement, lower tomato leaflets were caged with 10 aphids for 2 weeks.

For dsRNA detection, heavily aphid-infested tomato plants were used. Leaflets were completely cleaned of aphids and frozen at −80 °C. A few leaflets were processed for virus presence before use in dsRNA detection.

***M. euphorbiae* colony and aphid samples.** The colony of *M. euphorbiae* WU11, imported from France, was maintained on tomato cv. UC82b. The colony was kept inside cages in a pesticide-free greenhouse supplemented with light for a 16 h light photoperiod. Aphid samples originating from other research laboratories were received in RNAlater or Trizol.

Transcriptome assembly. The potato aphid transcriptome was reassembled by processing the sequence reads using the clustering algorithm SEED and assembled with Velvet/Oases (1.0.15/0.1.18) (Schulz *et al.*, 2012) as described by Bao *et al.* (2011).

Bioinformatic and phylogenetic analyses. BLAST analysis was performed by searching the non-redundant NCBI protein database. Domain search was performed using the Pfam database. Both searches were completed in June 2015.

Stable hairpins were predicted using the RNAfold webserver (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) by selecting basic options for the fold algorithms.

Protein sequence alignments and phylogenetic trees were performed using MAFFT sequence alignment server. Conserved regions of RdRp and helicase were chosen as described by Longdon *et al.* (2010) to build alignments. Phylogenetic trees were constructed using the neighbour-joining method with 1000 bootstrap replications.

RNA extraction and virus detection. Total RNA was extracted using Trizol and treated with DNase, and 3 μg RNA was used for cDNA synthesis using Superscript III reverse transcriptase (Invitrogen) and oligo-dT primers according to the manufacturer's recommendations. For single-aphid RNA extraction, acrylamide (Fisher) was added as a carrier before precipitation. PCR was performed in 25 μl using 1 μl cDNA as template and one of the following sets of primers (MevirF1, 5′-GTACACTTGCCTTACCTTACTGT-3′ and MevirR1b, 5′-AACACGGGTCACGACCTTAG-3′; MevirF2, 5′-CAA-ATGGGACAAGACCAAGACC-3′ and MevirR1, 5′-GATATTCCCG-CAGACTTATGTCG-3′). PCR conditions were 94 °C for 5 min, followed by 35 cycles of 94 °C for 30 s, 58 °C for 30 s, 72 °C for 30 s, and a final cycle of 72 °C for 3 min. As controls for RT, aphid ribosomal protein L27 (*RpL27*) (Atamian *et al.*, 2013) and tomato ubiquitin (*Sl-Ubi3*) (Bhattarai *et al.*, 2010) were used. Products were separated by electrophoresis in 1.5% agarose gels and visualized by ethidium bromide staining.

Quantitative PCR (qPCR) was performed in an iCycler5 IQ (Bio-Rad) using MevirF2 and MevirR2 (5′-GCACAGTAAGGTAAGGCAAGTG-3′) primers and iQ SYBR Green Supermix (Bio-Rad) in 15 μl with the same programme as for PCR. Two technical replicates were performed and the generated threshold cycle (C_T) was used to calculate transcript abundance relative to *Sl-Ubi3*.

dsRNA isolation and analysis. Mixed developmental stages of aphids or plant samples were subjected to dsRNA isolation by CF-11 cellulose (Sigma) column chromatography, and the resulting products were treated with DNase I, RNase A and proteinase K, and concentrated by Zymoclean gel RNA recovery kit (Zymo Research). Samples were separated by electrophoresis in 0.8% agarose gels and visualized by ethidium bromide staining (Salem *et al.*, 2009; Valverde *et al.*, 1990).

3′ RACE and sequencing. Aphid cDNA was synthesized using Superscript III reverse transcriptase (Invitrogen) and primer B26 (5′-GACTCGAGTCGACATCGATTTTTTTTTT-3′). The 3′ end of the transcript was amplified using a virus-specific forward primer (MevirF1) and a RACE-specific reverse primer (B25, 5′-GACTCGA-GTCGACATCGAT-3′). A second PCR was performed using nested virus-specific primer MevirF3 (5′-CTAAGGTCGTGACCCGTGTT-3′) and primer B25. The amplicon was separated by electrophoresis and gel-purified using the MiniElute purification kit (Qiagen), cloned into pCRII-TOPO (Life Technologies) and transformed into *Escherichia coli* strain DH5α. Plasmids were extracted from six transformed colonies using the Zyppy plasmid prep kit (Zymo Research) for sequencing. Sequences were analysed using Geneious.

MeV-1 subcloning and sequencing. Nine overlapping MeV-1 subclones were amplified using aphid cDNA as template and virus-specific primers (Table S1, available in the online Supplementary Material). Amplicons were separated by electrophoresis, gel-purified and cloned as described earlier. Four clones from each subclone were sequenced, using primer walking when necessary, and analysed.

ACKNOWLEDGEMENTS

REFERENCES

Agranovsky, A. A., Dolja, V. V., Kavsan, V. M. & Atabekov, J. G. (1978). Detection of polyadenylate sequences in RNA components of barley stripe mosaic virus. *Virology* **91**, 95–105.

Asghar, N., Lindblom, P., Melik, W., Lindqvist, R., Haglund, M., Forsberg, P., Överby, A. K., Andreassen, Å., Lindgren, P. E. & Johansson, M. (2014). Tick-borne encephalitis virus sequenced directly from questing and blood-feeding ticks reveals quasispecies variance. *PLoS One* **9**, e103264.

Atamian, H. S., Chaudhary, R., Cin, V. D., Bao, E., Girke, T. & Kaloshian, I. (2013). In planta expression or delivery of potato aphid *Macrosiphum euphorbiae* effectors *Me10* and *Me23* enhances aphid fecundity. *Mol Plant Microbe Interact* **26**, 67–74.

Ban, L., Didon, A., Jonsson, L. M., Glinwood, R. & Delp, G. (2007). An improved detection method for the *Rhopalosiphum padi* virus (RhPV) allows monitoring of its presence in aphids and movement within plants. *J Virol Methods* **142**, 136–142.

Bao, E., Jiang, T., Kaloshian, I. & Girke, T. (2011). SEED: efficient clustering of next-generation sequences. *Bioinformatics* **27**, 2502–2509.

Bekal, S., Domier, L. L., Niblack, T. L. & Lambert, K. N. (2011). Discovery and initial analysis of novel viral genomes in the soybean cyst nematode. *J Gen Virol* **92**, 1870–1879.

Bekal, S., Domier, L. L., Gonfa, B., McCoppin, N. K., Lambert, K. N. & Bhalerao, K. (2014). A novel flavivirus in the soybean cyst nematode. *J Gen Virol* **95**, 1272–1280.

Belyi, V. A., Levine, A. J. & Skalka, A. M. (2010). Unexpected inheritance: multiple integrations of ancient bornavirus and ebolavirus/marburgvirus sequences in vertebrate genomes. *PLoS Pathog* **6**, e1001030.

Bhattarai, K. K., Atamian, H. S., Kaloshian, I. & Eulgem, T. (2010). WRKY72-type transcription factors contribute to basal immunity in tomato and *Arabidopsis* as well as gene-for-gene resistance mediated by the tomato R gene *Mi-1*. *Plant J* **63**, 229–240.

Blackman, R. L. & Eastop, V. F. (2000). *Aphids on the World's Crops*. John Wiley & Sons, Chichester, UK.

Carolan, J. C., Fitzroy, C. I. J., Ashton, P. D., Douglas, A. E. & Wilkinson, T. L. (2009). The secreted salivary proteome of the pea

aphid *Acyrthosiphon pisum* characterised by mass spectrometry. *Proteomics* 9, 2457–2467.

Carolan, J. C., Caragea, D., Reardon, K. T., Mutti, N. S., Dittmer, N., Pappan, K., Cui, F., Castaneto, M., Poulain, J. & other authors (2011). Predicted effector molecules in the salivary secretome of the pea aphid (*Acyrthosiphon pisum*): a dual transcriptomic/proteomic approach. *J Proteome Res* 10, 1505–1518.

Chaudhary, R., Atamian, H. S., Shen, Z., Briggs, S. P. & Kaloshian, I. (2014). GroEL from the endosymbiont *Buchnera aphidicola* betrays the aphid by triggering plant defense. *Proc Natl Acad Sci U S A* 111, 8919–8924.

Chaudhary, R., Atamian, H. S., Shen, Z., Briggs, S. P. & Kaloshian, I. (2015). Potato aphid salivary proteome: enhanced salivation using resorcinol and identification of aphid phosphoproteins. *J Proteome Res* 14, 1762–1778.

Chiu, C. Y. (2013). Viral pathogen discovery. *Curr Opin Microbiol* 16, 468–478.

Cooper, W. R., Dillwith, J. W. & Puterka, G. J. (2010). Salivary proteins of Russian wheat aphid (Hemiptera: Aphididae). *Environ Entomol* 39, 223–231.

Cooper, W. R., Dillwith, J. W. & Puterka, G. J. (2011). Comparisons of salivary proteins from five aphid (Hemiptera: Aphididae) species. *Environ Entomol* 40, 151–156.

Daly, M. & Ward, F. (2003). Liver disease. In *Clinical Pharmacy and Therapeutics*, 3rd edn, pp. 209–228. Edited by R. Walker & C. Edwards. London: Churchill Livingstone.

De Barro, P. J. (1992). The role of temperature, photoperiod, crowding and plant quality on the production of alate viviparous females of the bird cherry-oat aphid, *Rhopalosiphum padi*. *Entomol Exp Appl* 65, 205–214.

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L. & other authors (2014). Pfam: the protein families database. *Nucleic Acids Res* 42, D222–D230.

Harmel, N., Létocart, E., Cherqui, A., Giordanengo, P., Mazzucchelli, G., Guillonneau, F., De Pauw, E., Haubruge, E. & Francis, F. (2008). Identification of aphid salivary proteins: a proteomic investigation of *Myzus persicae*. *Insect Mol Biol* 17, 165–174.

International Aphid Genomics Consortium (2010). Genome sequence of the pea aphid *Acyrthosiphon pisum*. *PLoS Biol* 8, e1000313.

Junglen, S. & Drosten, C. (2013). Virus discovery and recent insights into virus diversity in arthropods. *Curr Opin Microbiol* 16, 507–513.

Khan, A. H., Morita, K., del Carmen Parquet, M., Hasebe, F., Mathenge, E. G. M. & Igarashi, A. (2002). Complete nucleotide sequence of chikungunya virus and evidence for an internal polyadenylation site. *J Gen Virol* 83, 3075–3084.

King, A. M. Q., Adams, M. J., Carstens, E. B. & Lefkowitz, E. J. (editors) (2012). Ninth Report of the International Committee on Taxonomy of Viruses. In *Virus Taxonomy*, p. ii. San Diego: Elsevier.

Kobayashi, K., Atsumi, G., Iwadate, Y., Tomita, R., Chiba, K.-i., Akasaka, S., Nishihara, M., Takahashi, H., Yamaoka, N. & other authors (2013). Gentian Kobu-sho-associated virus: a tentative, novel double-stranded RNA virus that is relevant to gentian Kobu-sho syndrome. *J Gen Plant Pathol* 79, 56–63.

Li, C. X., Shi, M., Tian, J. H., Lin, X. D., Kang, Y. J., Chen, L. J., Qin, X. C., Xu, J., Holmes, E. C. & Zhang, Y. Z. (2015). Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *eLife* 4, 4.

Lipkin, W. I. & Firth, C. (2013). Viral surveillance and discovery. *Curr Opin Virol* 3, 199–204.

Liu, Y., Liu, H., Zou, J., Zhang, B. & Yuan, Z. (2014). Dengue virus subgenomic RNA induces apoptosis through the Bcl-2-mediated PI3k/Akt signaling pathway. *Virology* 448, 15–25.

Liu, S., Chen, Y. & Bonning, B. C. (2015). RNA virus discovery in insects. *Curr Opin Insect Sci* 8, 54–61.

Longdon, B., Obbard, D. J. & Jiggins, F. M. (2010). Sigma viruses from three species of *Drosophila* form a major new clade in the rhabdovirus phylogeny. *Proc Biol Sci* 277, 35–44.

Miles, P. W. (1968). Insect secretions in plants. *Annu Rev Phytopathol* 6, 137–164.

Moon, J. S., Domier, L. L., McCoppin, N. K., D'Arcy, C. J. & Jin, H. (1998). Nucleotide sequence analysis shows that *Rhopalosiphum padi* virus is a member of a novel group of insect-infecting RNA viruses. *Virology* 243, 54–65.

Nicholson, S. J., Hartson, S. D. & Puterka, G. J. (2012). Proteomic analysis of secreted saliva from Russian wheat aphid (*Diuraphis noxia* Kurd.) biotypes that differ in virulence to wheat. *J Proteomics* 75, 2252–2268.

Rao, S. A. K., Carolan, J. C. & Wilkinson, T. L. (2013). Proteomic profiling of cereal aphid saliva reveals both ubiquitous and adaptive secreted proteins. *PLoS One* 8, e57413.

Ryabov, E. V. (2007). A novel virus isolated from the aphid *Brevicoryne brassicae* with similarity to Hymenoptera picorna-like viruses. *J Gen Virol* 88, 2590–2595.

Ryabov, E. V., Keane, G., Naish, N., Evered, C. & Winstanley, D. (2009). Densovirus induces winged morphs in asexual clones of the rosy apple aphid, *Dysaphis plantaginea*. *Proc Natl Acad Sci U S A* 106, 8465–8470.

Rybicki, E. P. & von Wechmar, M. B. (1982). Characterisation of an aphid-transmitted virus disease of small grains. Isolation and partial characterization of three viruses. *J Phytopathol* 103, 306–322.

Salem, N. M., Chen, A. Y., Tzanetakis, I. E., Mongkolsiriwattana, C. & Ng, J. C. (2009). Further complexity of the genus *Crinivirus* revealed by the complete genome sequence of Lettuce chlorosis virus (LCV) and the similar temporal accumulation of LCV genomic RNAs 1 and 2. *Virology* 390, 45–55.

Schulz, M. H., Zerbino, D. R., Vingron, M. & Birney, E. (2012). *Oases*: robust *de novo* RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* 28, 1086–1092.

Srinivasan, K. G., Min, B. E., Ryu, K. H., Adkins, S. & Wong, S. M. (2005). Determination of complete nucleotide sequence of *Hibiscus latent Singapore virus*: evidence for the presence of an internal poly(A) tract. *Arch Virol* 150, 153–166.

Taylor, D. J., Leach, R. W. & Bruenn, J. (2010). Filoviruses are ancient and integrated into mammalian genomes. *BMC Evol Biol* 10, 193.

Tjallingii, W. F. (2006). Salivary secretions by aphids interacting with proteins of phloem wound responses. *J Exp Bot* 57, 739–745.

Valverde, R. A., Nameth, S. T. & Jordan, R. L. (1990). Analysis of double-stranded RNA for plant virus diagnosis. *Plant Dis* 74, 255–258.

van der Wilk, F., Dullemans, A. M., Verbeek, M. & van den Heuvel, J. F. J. M. (1997). Nucleotide sequence and genomic organization of *Acyrthosiphon pisum* virus. *Virology* 238, 353–362.

van Munster, M., Dullemans, A. M., Verbeek, M., Van Den Heuvel, J. F. J. M., Clérivet, A. & van der Wilk, F. (2002). Sequence analysis and genomic organization of Aphid lethal paralysis virus: a new member of the family *Dicistroviridae*. *J Gen Virol* 83, 3131–3138.

van Munster, M., Dullemans, A. M., Verbeek, M., van den Heuvel, J. F., Reinbold, C., Brault, V., Clérivet, A. & van der Wilk, F. (2003a). Characterization of a new densovirus infecting the green peach aphid *Myzus persicae*. *J Invertebr Pathol* 84, 6–14.

van Munster, M., Dullemans, A. M., Verbeek, M., van den Heuvel, J. F. J. M., Reinbold, C., Brault, V., Clérivet, A. & van der Wilk, F. (2003b). A new virus infecting *Myzus persicae* has a genome organization similar to the species of the genus *Densovirus*. *J Gen Virol* **84**, 165–172.

Vandermoten, S., Harmel, N., Mazzucchelli, G., De Pauw, E., Haubruge, E. & Francis, F. (2014). Comparative analyses of salivary proteins from three aphid species. *Insect Mol Biol* **23**, 67–77.

Vlachakis, D., Koumandou, V. L. & Kossida, S. (2013). A holistic evolutionary and structural study of *Flaviviridae* provides insights into the function and inhibition of HCV helicase. *PeerJ* **1**, e74.

Williamson, C., Rybicki, E. P., Kasdorf, G. G. F. & Von Wechmar, M. B. (1988). Characterization of a new picorna-like virus isolated from aphids. *J Gen Virol* **69**, 787–795.