

Article

Not peer-reviewed version

Novel Negev-like RNA Viruses Encoding Membrane Proteins of SP24 Family in Hosts Ranging from Algae to Vertebrates

[Sergey Y. Morozov](#) * and [Andrey G. Solovyev](#)

Posted Date: 30 October 2023

doi: 10.20944/preprints202310.1859.v1

Keywords: RNA virus; negeviruses; SP24 protein; taxonomy; membrane-embedded protein; virus evolution



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Novel Negev-like RNA Viruses Encoding Membrane Proteins of SP24 Family in Hosts Ranging from Algae to Vertebrates

Sergey Y. Morozov ^{1,2,*} and Andrey G. Solovyev ^{1,2,3}

¹ A. N. Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow, Russia

² Department of Virology, Biological Faculty, Moscow State University, Moscow, Russia

³ All-Russia Research Institute of Agricultural Biotechnology, Moscow, Russia

* Correspondence: morozov@genebee.msu.ru; Tel.: +7(495)939-31-98

Abstract: This paper describes new horizons in the diversity and taxonomy of negev-like viruses encoding the membrane-embedded SP24 protein. First, our data extend the known host range of SP24-encoding negev-like viruses to include brown algae, fungi, green plants, the phylum *Entoprocta*, the phylum *Mollusca*, and vertebrates. Second, our phylogenetic analysis suggests that the evolution of the SP24 gene family may have involved frequent events of inter-order virus genome shuffling. Third, the identification of 2-3 copies of SP24 protein genes in some virus RNAs shows that virus genomes may have acquired additional SP24 genes during the evolutionary process due to duplications or new acquisition steps. Fourth, the broad host specificity of some SP24-encoded viruses may be related to an important adaptive role of SP24. Fifth, insect and nematode genomes may acquire viral SP24 genes by putative horizontal transfer from negev-like viruses known to infect species of both taxa.

Keywords: RNA virus; negevviruses; SP24 protein; taxonomy; membrane-embedded protein; virus evolution

1. Introduction

Members of the SP24 protein family were first described in chronic bee paralysis virus (CBPV) and anopheline-associated C virus (AACV) (here and below, virus names are given according to the current ICTV rules [1]), which are RNA viruses related to members of the order *Tolivirales* (phylum *Kitrinoviricota*) [2]. This polypeptide contains a poorly conserved positively charged N-terminal region and a conserved domain corresponding to the central region of the protein, including four predicted trans-membrane helical segments [2,3]. The SP24 family protein is one of the major structural components of CBPV virions, together with the RNA2-encoded ORF2 putative glycoprotein, which possesses the N-terminal DiSB-ORF2_chro domain in a short conserved region predicted to form disulfide bridges [4]. CBPV, AACV, and recently described related viruses of the order *Tolivirales*, namely insect tombus bipa-like virus 1, tiger mosquito bi-segmented tombus-like virus, diaphorina citri-associated C virus and linepithema humile C virus 1 as well as the arachnid-specific virus infecting *Tetragnatha maxillosa* (hubei tombus-like virus 28) [5–7] are bipartite RNA viruses, in which RdRp and SP24 are encoded by multicistronic RNAs 1 and 2, respectively.

Another group of multipartite RNA viruses encoding SP24 family proteins are members of the family *Kitaviridae* (order *Martellivirales*, phylum *Kitrinoviricota*). Plant viruses of the genera *Cilevirus* and *Blunervirus* contain the SP24 gene in RNA2 and RNA3 of bipartite and quadripartite genomes, respectively. However, the tripartite genomes of two members of the plant virus genus *Higrevirus* (family *Kitaviridae*) encode SP24 in either RNA2 or RNA3 [8,9].

Several monopartite positive-stranded RNA viruses of insects related to unclassified members of the family *Permutotetraviridae* have also been shown to encode SP24. This virus family is characterized by a unique permuted organization of the replicase RdRp domain and is classified as

an isolated taxon in the kingdom *Orthornavirae* [10]. Some unclassified viruses of this family, in particular shinobi tetravirus, culex permutotetra-like virus, hubei permutotetra-like virus 3 and sarawak virus, contain genomes with three ORFs, where SP24 and DiSB-chro genes represent ORF2 and ORF3 located downstream of the replicase gene [10].

The most abundant groups of monopartite positive-stranded RNA viruses encoding SP24 include negeviruses and negev-like (or nege-like) viruses, which have an obvious evolutionary relationship to the multipartite viruses of the family *Kitaviridae* [9,11]. Initially, viruses of the genus *Negevirus* have been considered to be insect-specific viruses that mainly infect species of the order *Diptera* [11]. In particular, negeviruses are among the most abundant virus taxa in many *Culex* and *Aedes* species, which are the predominant hosts of the eponymous negev virus. Negeviruses are classified into two clades, nelorpiviruses and sandewaviruses, with nelorpiviruses being more closely related to plant kitaviruses [11,12].

Negevirus genomes typically encode three ORFs: ORF1 encodes replicase, while ORFs 2 and 3 encode the predicted DiSB-containing glycoprotein and SP24 family polypeptide, respectively. Interestingly, the DiSB-ORF2_chro domain is part of the virion protein of sogatella furcifera hepe-like virus (order *Hepelivirales*) (see also below, Section 3.5.) [13]. The minority of negeviruses encoding less than three ORFs includes, for example, cordoba virus with a single ORF and buckhurst virus with two ORFs [11,12]. Massive metagenomic searches have significantly increased the number of nege- and negev-like viruses and have shown that the genomic structure of negev-like viruses is quite variable [5,12,14,15]. Moreover, these studies and the in silico studies of public databases have shown that the hosts of negev-like viruses are much more diverse than those of the genus *Negevirus*. Negev-like viruses infect hosts belonging to the class *Insecta* (phylum *Arthropoda*) in addition to the order *Diptera*, namely the orders *Hemiptera*, *Coleoptera*, *Thysanoptera*, *Hymenoptera*, *Odonata*, *Orthoptera*, and *Lepidoptera*. Several host species belonging to the phylum *Arthropoda* (classes *Arachnida*, *Maxillopoda*, *Chilopoda* and *Malacostraca*), as well as to the phyla *Nematoda* and *Cnidaria*, have also been reported [5,12,14,15].

Importantly, the vast majority of negeviruses and negev-like viruses encode the SP24 protein, which is predicted to be a virion component [5,12,14,15]. Indeed, comparative electron microscopy studies of negevirus virions have shown that the particles are pleiomorphic and consist mostly of SP24. These particles have variable sizes, an elliptical shape, and a single projection consisting of glycoprotein with a DiSB signature. Microscopy also reveals stratified layers at the virion surface, probably representing a host-derived lipid bilayer [16]. Although plant-infecting viruses of the family *Kitaviridae* have been shown to encode SP24 and contain lipid membranes, the exact role of SP24 in encapsidation and its function in virus particles remain unclear [9]. Moreover, cileviruses encode a protein (p29) that is found in virions and viroplasms and is capable of forming spherical virus-like particles when expressed as a single engineered protein *in planta*. Strikingly, cileviruses can replicate in insect vector cells in which the accumulation of p29-specific subgenomic RNA is 50 times lower than in plants [9]. It can be hypothesized that p29 is a plant-specific virus movement factor, whereas SP24 is an animal-specific virus movement factor that may directly interact with the virus genomic RNA [3,9], thus facilitating virus transfer in animal (insect) organism.

In this paper, to better understand the taxonomic boundaries of hosts for viruses encoding SP24 and the variability of the negev-like virus genome organizations, we have searched the rapidly increasing amount of new transcriptomic data in the public NCBI database. This search, in our opinion, may open new horizons for planning the future studies on the exact roles of SP24 in virus infection.

2. Materials and Methods

Negev-like nucleotide and protein sequences were collected from the NCBI plant transcriptome database. Sequence comparisons were carried out using the BLAST algorithm (TBLASTn and BLASTp) at the National Center for Biotechnology Information (NCBI). The raw nucleotide sequence reads from each analyzed SRA experiment linked to the virus nucleotide sequence and TSA projects returning SP24 sequence hits were downloaded and subjected to bulk local BLASTX searches (E-

value = $< 1e-105$) against a Refseq virus protein database available at <ftp://ftp.ncbi.nlm.nih.gov/refseq/release/viral/viral.1.protein.faa.gz> (accessed on 20 August 2023). The resulting viral sequence hits from each SRA read were then visually explored. Tentative virus contigs were extended by iterative mapping of the raw reads from each SRA library. This strategy used reiterative BLAST to extract a subset of reads related to the query contig, and these retrieved reads were used to extend the contig, and then the process was repeated iteratively using the extended sequence as the query.

ORFs were identified using the ORF Finder program (<https://www.ncbi.nlm.nih.gov/orffinder/>, accessed on 20 August 2023). Conserved motif searches were conducted in CDD (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>, accessed on 20 August 2023) databases. Membrane-embedded hydrophobic protein regions were predicted using software Membranome (<https://membranome.org>, accessed on 20 August 2023). Phylogenetic analysis was performed with “Phylogeny.fr” (a web service dedicated to reconstruction and analysis of phylogenetic relationships between molecular sequences) by constructing maximum likelihood phylogenetic trees (http://www.phylogeny.fr/simple_phylogeny.cgi, accessed on 20 August 2023). Bootstrap percentages received from 1000 replications were used. Genome sequences of benyviruses viruses were downloaded from the GenBank database.

3. Results

In the last two years, several papers have been published that further encourage us to perform bioinformatics search of SP24 genes in viruses and virus-like RNA assemblies (VLRAs) of order *Martellivirales* infecting all eukaryotes (bikonts and unikonts). First, novel data have shown previously unknown fact that SP24-encoding negev-like viruses also infect flatworms in phylum *Platyhelminthes* [17]. Second, recently obtained transcriptome assemblies from large amount of arthropods, particularly, from 1098 species of spiders (phylum *Arthropoda*, class *Arachnidae*) [18] can provide a good starting basis for further integrated analysis of the variability of genome organizations in SP24-encoding negev-like viruses. Third, novel paper published early 2023 [19] has been reported that a previously undescribed insect virus species—*Recilia dorsalis filamentous virus* (RdFV) encodes SP24 which is included into filamentous virions unlike elliptical negevviruses and CBPV. Moreover, RdFV monopartite genome includes two copies of distantly related SP24 genes (ORFs 4 and 6) (Figure 1). Direct interaction of viral capsid SP24 encoded by ORF 6 with the sperm-specific serpin protein of leafhopper *Recilia dorsalis* mediates invasion of virus into male reproductive organs and binding virions to insect sperm surfaces [19]. As a result, these processes are important for subsequent paternal virus transmission, and SP24, thus, represents proteins factor for vertical virus transmission. One can hypothesized that SP24-mediated paternal virus transmission is a novel type of inter-organismal virus movement in insect populations.

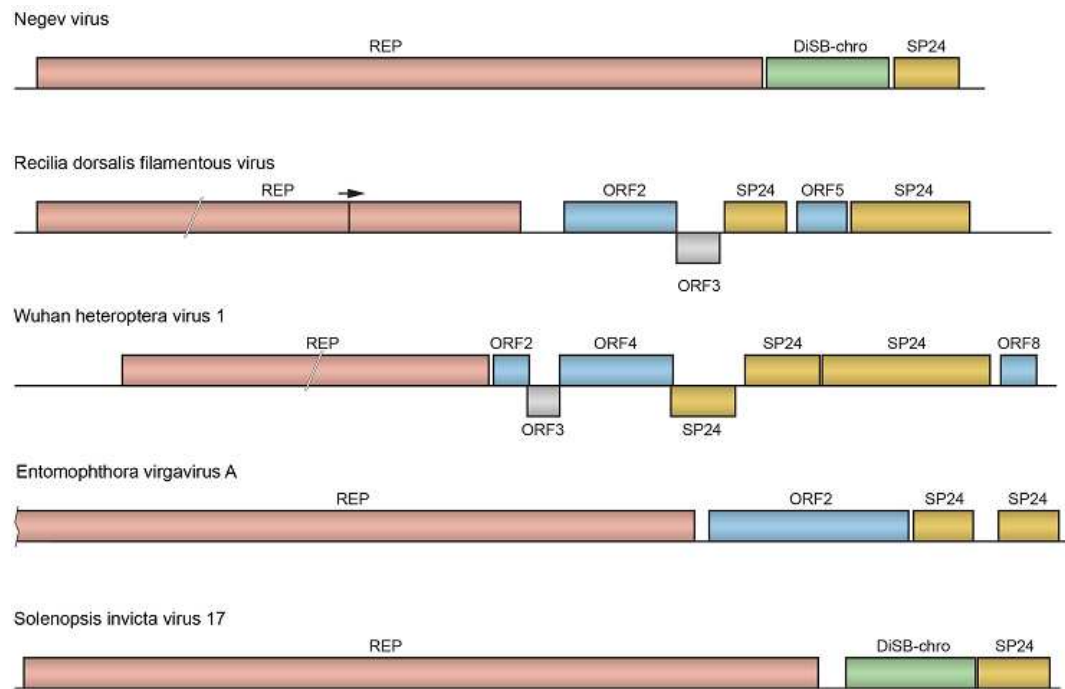


Figure 1. Gene organization of the negevirus and selected negev-like viruses. The replicase protein (REP) is shown in pink (see text for details). SP24 protein is shown in brown. ORFs coding for DiSB-chro domain proteins are in green. Proteins with putative membrane-bound segments are shown in blue. Read-through of leaky terminator codon is shown by arrow above RdFV genome scheme.

3.1. Virus RNA genomes with two or three gene copies encoding SP24

In contrast to many other known negeviruses and negev-like viruses, the genome of a symbiotic virus RdFV infecting rice green zigzag leafhopper *Recilia dorsalis* (class *Insecta*, order *Hemiptera*, family *Cicadellidae*) contains two gene copies encoding SP24 [19] that may correlate with the ability of the virus to form filamentous particles and/or to spread by vertical transmission. In general, this virus encodes seven proteins (Figure 1). Replicase is encoded by ORF1, which has an internal leaky termination codon and may be translated to form two overlapping polypeptides [20]. Among non-replicative proteins of RdFV, only the SP24 polypeptides have significant similarity to proteins of annotated viruses. However, a search of the NCBI TSA database revealed that the transcriptome of insect *Xestocephalus desertorum*, which belongs to the same family *Cicadellidae* as zigzag leafhopper, contains a partial VLRA encoding ORF1 RdRp domain closely related to that of RdFV (e-value $1e-49$), a protein homolog of RdFV ORF5 protein (e-value $1e-19$) and three copies of SP24 (Table 1). Three copies of SP24 genes were also revealed in the annotated genomes of pyrrhocoris apterus virus 1 infecting insect *Pyrrhocoris apterus* (class *Insecta*, order *Hemiptera*, family *Pyrrhocoridae*) and wuhan heteroptera virus 1 isolated from mixed insects of the suborder *Heteroptera* (order *Hemiptera*) [5] (Figure 1) (Table 1). It should be noted that the RdRp domain encoded by ORF1 of RdFV, pyrrhocoris apterus virus 1 and wuhan heteroptera virus 1 are clustered together in the phylogenetic tree of Virgaviridae-like RdRp domains (Supplementary Figure S1).

Another cluster of RdRp domains of negev-like viruses encoding two SP24 gene copies includes, particularly, zeugodacus cucurbitae negev-like virus, broome virga-like virus 1, bactrocera oleae negev-like virus and bactrocera dorsalis negev-like virus isolate Bz infecting mosquitoes and invasive fruit flies from order *Diptera* (class *Insecta*) [21,22] (Table 1) (Supplementary Figure S1). Our search also revealed negev-like VLRA encoding two SP24 gene copies in the transcriptomes of class *Arachnida*, specifically *Scotophaeus scutulatus* (family *Gnaphosidae*) and *Neoscona scylloides* (family *Araneidae*) (Table 1).

Table 1. Examples of the viruses and VLRA encoding two or three SP24 gene copies.

Virus and host species	Host Class/Order	Accession number
recilia dorsalis filamentous virus <i>Recilia dorsalis</i>	<i>Insecta/Hemiptera</i>	OP326514
pyrrhocoris apterus virus 1 <i>Pyrrhocoris apterus</i>	<i>Insecta/Hemiptera</i>	MK024711
VLRA <i>Xestocephalus desertorum</i>	<i>Insecta/Hemiptera</i>	GELC01067077
VLRA <i>Neoscona scylloides</i>	<i>Arachnida/Araneae</i>	ICMG01043884
VLRA <i>Scotophaeus scutulatus</i>	<i>Arachnida/Araneae</i>	ICQK01020075
atrato virga-like virus 6 <i>Psorophora</i> sp.	<i>Insecta/Diptera</i>	MN661118
atrato virga-like virus 7 <i>Psorophora albipes</i>	<i>Insecta/Diptera</i>	MN661129
wuhan heteroptera virus 1 mixed insects	<i>Insecta/Hemiptera</i>	KX883821
dougjudy virga-like virus <i>Calliphora vicina</i>	<i>Insecta/Diptera</i>	MT129773
bactrocera oleae negev-like virus Bactrocera oleae	<i>Insecta/Diptera</i>	MW310375
broome virga-like virus 1 <i>Culex annulirostris</i>	<i>Insecta/Diptera</i>	MT498833
bactrocera dorsalis negev-like virus <i>Bactrocera zonata</i>	<i>Insecta/Diptera</i>	MW310386
zeugodacus tau negev-like virus <i>Zeugodacus tau</i>	<i>Insecta/Diptera</i>	MW310340
zeugodacus cucurbitae negev-like virus <i>Zeugodacus cucurbitae</i>	<i>Insecta/Diptera</i>	MW310350
entomophthora virgavirus A* <i>Entomophthora muscae</i>	<i>Entomophthoromycetes/</i> <i>Entomophthorales</i>	MK231110

* – infects fungi.

3.1.1. Unique fungal virus encoding SP24 gene

Interestingly, the RdRp phylogenetic tree cluster containing Diptera-infecting viruses (in particular, broome virga-like virus 1) also contains the fungal entomophthora virgavirus A (Supplementary Figure S1) with RNA genome encoding two SP24 gene copies (Figure 1) (Table 1). This virus genome has been found in transcriptome of the fungus *Entomophthora muscae*, whose hosts are mainly insects of the order *Diptera* [23]. It can be assumed that this virus has evolutionarily originated due to transfer of an insect virus from a dipteran organism to a fungus and its subsequent adaptation to replicate in fungal hyphal bodies. On the other hand, entomophthora virgavirus A can be a true insect virus, and the finding of virus genome in fungal cells is a consequence of frequent passive transfer of virus RNA from host insects.

3.2. Virus RNA genomes encoding SP24 and infecting arthropods of class Arachnida

The transcriptomes of more than 1000 species of spiders (class *Arachnidae*) [18] recently deposited in the NCBI TSA database have the potential to greatly increase the number of novel RNA viruses encoding SP24. First of all, our search of the NCBI TSA database revealed many spiders infected by solenopsis invicta virus 17 (SINV-17) or its closely related strains (Table 2). This monopartite RNA virus, first described in the red imported fire ant [24], encodes three genes, namely replicase, ORF2 with DiSB-ORF_chro domain and SP24 gene (Figure 1). SINV-17 RdRp domain is phylogenetically most closely related to hubei virga-like virus 8 from Myriapoda sub-philum [5] and

VLRA from the insect *Chrysoperla carnea*, and more distantly related to sandewaviruses (Supplementary Figure S1).

Table 2. Examples of VLRA with encoded proteins having >90% homology to SINV protein SP24.

VLRA origin and accession number	Host Class/Order	Identity to SINV (%) SP24 *
<i>Acheta domesticus</i> , GHUU01023074	<i>Insecta/Orthoptera</i>	100
<i>Orestes mouhotii</i> , GDZN01041390	<i>Insecta/Phasmatodea</i>	95
<i>Polistes metricus</i> , GDHQ01029178	<i>Insecta/Hymenoptera</i>	97
<i>Karoopasma biedouwense</i> , GINP01229017	<i>Insecta/Mantophasmatodea</i>	99
<i>Austrophasmatidae</i> sp., GDYW01015845	<i>Insecta/Mantophasmatodea</i>	98
<i>Gryllus bimaculatus</i> , GISW01177432	<i>Insecta/Orthoptera</i>	99
<i>Nylanderia pubens</i> , JP785474	<i>Insecta/Hymenoptera</i>	96
<i>Bassaniana decorate</i> , ICCJ01037167	<i>Arachnida/Araneae</i>	94
<i>Weintrauboa contortipes</i> , ICJX01013850	<i>Arachnida/Araneae</i>	94
<i>Meotipa spiniventris</i> , IBDD01006961	<i>Arachnida/Araneae</i>	94
<i>Theridion zonulatum</i> , IATV01002901	<i>Arachnida/Araneae</i>	94
<i>Thiania bhamoensis</i> , IAGH01040226	<i>Arachnida/Araneae</i>	94
<i>Oxyopes macilentus</i> , IAPY01004984	<i>Arachnida/Araneae</i>	94
<i>Cyrtophora unicolor</i> , IAYW01028924	<i>Arachnida/Araneae</i>	94
<i>Argyrodes kumadai</i> , IBBR01021510	<i>Arachnida/Araneae</i>	94
<i>Mendoza elongate</i> , ICIJ01021408	<i>Arachnida/Araneae</i>	89
<i>Agelena labyrinthica</i> , IBZQ01034950	<i>Arachnida/Araneae</i>	94
<i>Tegeocelotes corasides</i> , IASA01013670	<i>Arachnida/Araneae</i>	94
<i>Octonoba sybotides</i> , IAZZ01004659	<i>Arachnida/Araneae</i>	94
<i>Leucauge celebesiana</i> , IAWQ01030888	<i>Arachnida/Araneae</i>	94
<i>Cyclosa mulmeinensis</i> , ICEF01010633	<i>Arachnida/Araneae</i>	94
<i>Heptathela higoensis</i> , IBEA01018510	<i>Arachnida/Araneae</i>	94
<i>Oxyopes</i> sp., ICGS01044124	<i>Arachnida/Araneae</i>	94
<i>Argiope minuta</i> , ICGZ01029236	<i>Arachnida/Araneae</i>	94
<i>Uloborus</i> sp., IBJL01027593	<i>Arachnida/Araneae</i>	94
<i>Chikunia albipes</i> , IAUZ01010102	<i>Arachnida/Araneae</i>	94
<i>Ebrechtella tricuspidata</i> , IBYS01040384	<i>Arachnida/Araneae</i>	94
<i>Rhomphaea hyrcana</i> , IAOI01013840	<i>Arachnida/Araneae</i>	94
<i>Latouchia swinhoei</i> , IARH01001083	<i>Arachnida/Araneae</i>	94
<i>Lycosa ishikariana</i> , IBVL01017570	<i>Arachnida/Araneae</i>	94
<i>Dolomedes pegasus</i> , ICEZ01008221	<i>Arachnida/Araneae</i>	94
<i>Araneus seminiger</i> , IAIN01035359	<i>Arachnida/Araneae</i>	94
<i>Thanatus bungei</i> , IALY01037623	<i>Arachnida/Araneae</i>	94
<i>Iwogumoa interuna</i> , IBKG01028989	<i>Arachnida/Araneae</i>	94
<i>Sinopoda okinawana</i> , ICJY01003033	<i>Arachnida/Araneae</i>	94
<i>Thelcticopis severa</i> , ICJF01009625	<i>Arachnida/Araneae</i>	94
<i>Heptathela higoensis</i> , IBNB01017401	<i>Arachnida/Araneae</i>	94
<i>Heptathela yanbaruensis</i> , ICKJ01022680	<i>Arachnida/Araneae</i>	94
<i>Mendoza elongata</i> , ICIJ01011610	<i>Arachnida/Araneae</i>	94
<i>Iwogumoa interuna</i> , IBKG01014765	<i>Arachnida/Araneae</i>	94
<i>Dolomedes triton</i> , GGRN01030862	<i>Arachnida/Araneae</i>	94
<i>Anolis carolinensis</i> , GAJD01024809	<i>Lepidosauria/Iguania</i>	98
<i>Sceloporus jarrovi</i> , GIWZ010003422	<i>Lepidosauria/Iguania</i>	99
<i>Protopterus annectens</i> , GGXP01055242	<i>Euteleostomi/Ceratodontiformes</i>	97

* Percentage identity is shown for SP24 sequence assuming near 100% identity for replicase.

Spiders represent only a part of the arthropod hosts for SINV-17. Our transcriptome search revealed seven insect species that can be infected by SINV-17 (Table 2). Moreover, it is quite unexpected that sequences highly similar to SINV-17 were found in a search of TSA databases of

vertebrate species, namely the West African the lungfish (*Protopterus annectens*), the Yarrow’s spiny lizard (*Sceloporus jarrovii*) and the lizard *Anolis carolinensis* (Table 2).

Additionally, we revealed many novel spider viruses encoding SP24 that are distinct from SINV-17 (Supplementary Figure S1). Previously, several viruses of this type have only been revealed in the golden orb-weaver spider (*Nephila clavipes*) and the spider mite (*Tetranychus urticae*) [25].

3.3. Negev-like virus RNA genomes infecting species of metazoan taxa outside insects

The finding of VLRA highly similar to the negev-like SINV-17 genome in vertebrates (see above) prompted us to perform another comprehensive search of the transcriptomes of non-arthropod metazoan taxa to find new negev-like virus RNA genomes. Our search revealed that viruses only distantly related to SINV-17 infect at least three fish species (Table 3). These species are included in the infraclass *Teleostei* (teleost fishes), clade *Acanthomorphata*.

Table 3. Examples of negev-like virus RNA genomes infecting species of metazoan taxa outside arthropods*.

Host taxonomic position	Name (Virus and VLRA)	Accession number
Teleostei; <i>Anabantiformes</i>	<i>Parargyrops edita</i> VLRA	GIC101264178
Teleostei; Cyprinodontiformes	<i>Aphanius dispar</i> VLRA DN-106	GJEY01002386
Teleostei; <i>Anabantiformes</i>	<i>Channa argus</i> VLRA	GEM101178446
Teleostei; Cyprinodontiformes	<i>Aphanius dispar</i> VLRA DN-113	GJEY01012897
Porifera; <i>Demospongiae</i>	<i>Halisarca dujardini</i> VLRA	GIF101065598
Cnidaria; <i>Actiniaria</i>	beihai anemone virus 1	KX883744
Spiralia; <i>Entoprocta</i>	<i>Loxosomella nordgaardi</i> VLRA	GIMJ01025245
Spiralia; <i>Mollusca</i>	<i>Octopus vulgaris</i> VLRA	GKAX01007523
Spiralia; <i>Mollusca</i>	<i>Charonia lampas</i> VLRA	GIQZ01066080
Spiralia; <i>Mollusca</i>	<i>Potamilus streckersoni</i> VLRA	GJAA01000017
Spiralia; <i>Platyhelminthes</i>	dicrocoelium nege-like virius	OP548619
Spiralia; <i>Platyhelminthes</i>	<i>Girardia</i> sp. VLRA	GHO101019505
Spiralia; <i>Platyhelminthes</i>	provittati virus	BK059742
Spiralia; <i>Platyhelminthes</i>	fasciogiga virus	BK059714
Spiralia; <i>Platyhelminthes</i>	fasciohepa virus 1	BK059715
Spiralia; <i>Platyhelminthes</i>	meterori virus	BK059677
Spiralia; <i>Platyhelminthes</i>	clonorsi virus 1	BK059702
Spiralia; <i>Platyhelminthes</i>	fasciohepa virus 3	BK059717
Spiralia; <i>Platyhelminthes</i>	clonorsi virus 2	BK059703
Spiralia; <i>Platyhelminthes</i>	fasciohepa virus 2	BK059716
Ecdysozoa; <i>Onychophora</i>	<i>Principapillatus hitoyensis</i> VLRA	GJGV01069003
Ecdysozoa; <i>Nematoda</i>	<i>Nippostrongylus brasiliensis</i> VLRA	VDL74358
Ecdysozoa; <i>Nematoda</i>	<i>Anguina tritici</i> VLRA	GKDZ01058772
Ecdysozoa; <i>Nematoda</i>	xingshan nematode virus 1	KX883837
Ecdysozoa; <i>Nematoda</i>	xingshan nematode virus 2	KX883836
Ecdysozoa; <i>Nematoda</i>	xin Zhou nematode virus 1	KX883838
Ecdysozoa; <i>Nematoda</i>	<i>Anisakis pegreffii</i>	HBXC01106013

* Viruses and VLRA do not include those nearly identical to SINV-17.

A VLRA of Crimson seabream (*Parargyrops edita*, order *Anabantiformes*) contains four ORFs (Figure 2). The incomplete ORF1 encodes the C-terminal part of the replicase with RNA helicase and RdRp domains. ORF2 codes for a protein of 224 amino acids with two long hydrophobic segments and no significant similarity to known viral proteins. ORF3 (221 codons) represents a SP24 gene, while ORF4 encodes a protein of 109 amino acids with distant similarity to the VP-2 capsid protein of jingmen tick virus [26]. Another fish virus encoding SP24 is represented by a VLRA from

euryhaline fish (*Aphanius dispar*, order Cyprinodontiformes) (Table 3). This virus genome encodes five proteins (Figure 2). A replicase ORF is followed by overlapping ORF2 and ORF3, which encode hydrophobic proteins of 467 and 268 amino acids, respectively. ORF4 encodes an SP24 protein of 306 amino acid residues, while ORF5 encodes a small protein of 119 residues (Figure 2). This virus is moderately similar to another negev-like virus (Table 3), which also infects the fish *A. dispar* and encodes five proteins, namely replicase, ORF2 protein (492 residues long), ORF3 protein (382 amino acids), ORF4 protein (209 amino acids) representing SP24, and ORF5 (141 residues). It should be noted that RdRp domains encoded by ORF1 of VLRA from *Parargyrops edita* and *Aphanius dispar* are clustered together with negev virus and nelorpiviruses in the phylogenetic tree of *Virgaviridae*-like RdRp domains (Supplementary Figure S1).

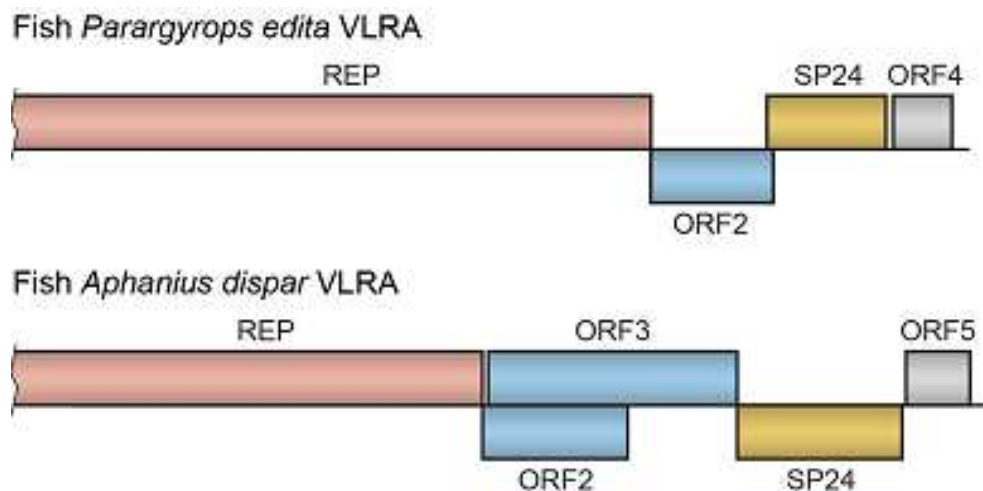


Figure 2. Gene organization of the selected fish VLRA. See Figure 1 for details of ORF naming and shading.

Among metazoan species in the clade *Spiralia*, negev-like viruses encoding SP24 have only been identified in flatworms (phylum *Platyhelminthes*) [17]. Our searches of the NCBI TSA databases revealed additional flatworm viruses that encode SP24-like proteins (Table 3). One more SP24-encoding negev-like virus was found by our search in the goblet worm (*Loxosomella nordgaardi*), which belongs to *Entoprocta*, another phylum of *Spiralia* [27,28] (Table 3). In addition, novel viruses encoding SP24 were found in species of phylum *Mollusca* (classes *Bivalvia*, *Gastropoda* and *Cephalopoda*). Specifically, these animals include *Octopus vulgaris* (class *Cephalopoda*, order *Octopoda*), *Charonia lampas* (triton shell, class *Gastropoda*, order *Littorinimorpha*) and *Potamilus streckersoni* (class *Bivalvia*, order *Unionida*) [28] (Table 3). The octopus SP24-encoding VLRA contains seven ORFs (Figure 3). A partial replicase ORF is followed by ORF2, which encodes a hydrophobic protein of 442 residues with a DiSB-like chro domain. ORF3 encodes a protein of 103 residues with a long predicted transmembrane region. ORF4 encodes a small protein (110 residues) and overlaps with ORF5 of the SP24 protein (304 residues), while ORF6 encodes a highly hydrophobic protein of 254 residues and overlaps ORF7 (285 residues) (Figure 3). In the RdRp phylogenetic tree, RdRp domains of ORF1 in VLRA from *Octopus vulgaris* and *Charonia lampas* are positioned in a large cluster that is quite distant from a subtree containing RdRps from flatworm negev-like viruses (Supplementary Figure S1).

In the animal clade *Ecdysozoa*, species infected by negev-like viruses encoding SP24 were found not only in the phylum *Arthropoda*, but also in the phylum *Nematoda* (Table 3). Furthermore, we found that the velvet worm *Principapillatus hitoyensis*, which belongs to the phylum *Onychophora*, which is phylogenetically closest to *Arthropoda* [28], can be infected by a negev-like virus (Table 3). This velvet worm-specific virus encodes three proteins, the ORF1 replicase with viral methyltransferase domain, macro domain representing an ADP-ribose binding module, OTU domain family (cysteine proteases), superfamily 1 viral RNA helicase domain and the C-terminal RdRp domain; the ORF2 protein with DiSB-like chro domain; the ORF3 protein with SP24 domain (Figure 3).

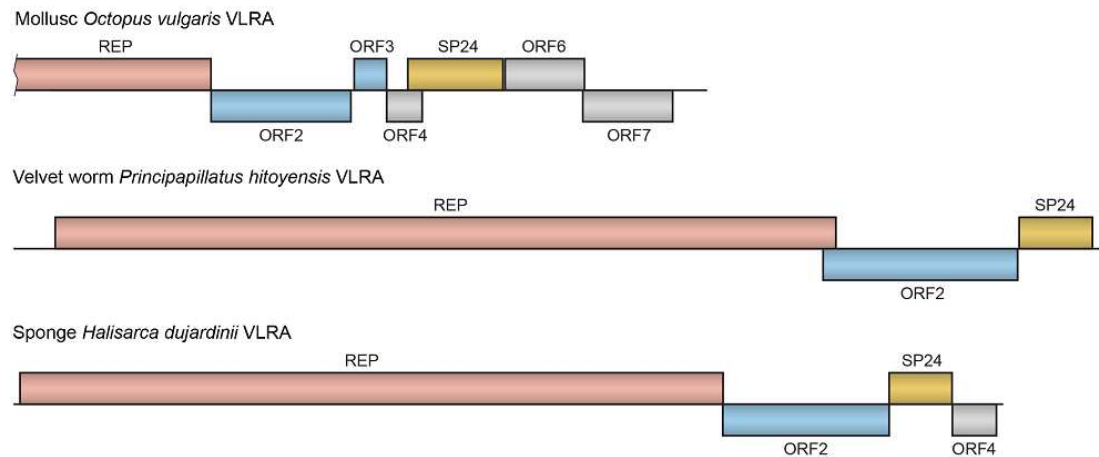


Figure 3. Gene organization of the selected VLRA. See Figure 1 for details of ORF naming and shading.

The phylum *Porifera* (sponges) and the phylum *Cnidaria* contain species that are among the most primitive metazoan organisms [28]. In this study, we report for the first time a sponge VLRA that corresponds to a distant relative of negevviruses (Table 3). This sponge-specific virus infects *Halisarca dujardini* (phylum *Porifera*, class *Demospongiae*, order *Dendroceratida*) and encodes four proteins (Figure 3). The ORF1-encoded replicase has three domains: a replicase with viral methyltransferase domain, a superfamily 1 viral RNA helicase domain, and the C-terminal RdRp domain. ORF1 is followed by ORF2, which encodes a hydrophobic protein of 524 amino acid residues. ORF3 encodes an SP24-like protein (200 residues), while ORF4 encodes a small protein of 136 residues (Figure 3). Previously, a transcriptome search of sea anemones (phylum *Cnidaria*, class *Hexacorallia*, order *Actiniaria*) has revealed a negev-like virus encoding SP24 [5]. This virus, named beihai anemone virus 1 (Table 3), encodes six proteins. ORF1 encodes a replicase with viral methyltransferase domain, macro domain representing an ADP-ribose binding module, superfamily 1 viral RNA helicase domain and the C-terminal RdRp domain. ORF2 encodes a protein (166 residues) that has no counterparts among known viral proteins and overlaps with ORF3, which encodes a highly hydrophobic protein (163 amino acids). ORF4 protein (438 amino acids) contains DiSB-like chro domain, while ORF5 (203 residues) encodes SP24 and ORF6 encodes a small protein (110 amino acids) with no similarity to known viral proteins. Phylogenetic analysis of the RdRp domain of beihai anemone virus 1 showed that the most similar virus is sanxia atyid shrimp virus 1 (SASV1), which infects shrimps of the family *Atyidae* (phylum *Arthropoda*, class *Malacostraca*) (Supplementary Figure S1). This shrimp-specific virus encodes 5 proteins. ORF1 encodes a replicase (2629 amino acids) with viral methyltransferase domain, macro domain, superfamily 1 viral RNA helicase domain and RdRp domain. ORF2 encodes a protein (128 residues) that has no counterparts among known viral proteins. ORF3 encodes a highly hydrophobic protein (164 amino acids). ORF4 protein (570 amino acids) contains the DiSB-like chro domain, while ORF5 (208 residues) encodes SP24. Our data show that several novel species of the class *Malacostraca* can be infected by negev-like viruses encoding SP24. Some of these viruses appear to be related to SASV1 (Supplementary Figure S1). In particular, eriocheir sinensis kita-like virus, which infects Chinese hairy crab (phylum *Arthropoda*, class *Malacostraca*, family *Varunidae*), encodes 3 proteins (Supplementary Figure S1): ORF1 protein is a replicase; ORF2 protein (665 amino acids) is a fusion of SASV1 ORF3 and ORF4 (DiSB) proteins; ORF3 (231 residues) encodes SP24. However, some other negev-like viruses that infect the class *Malacostraca* are only distantly related to SASV1. For example, the RdRp domain of the SP24-encoding VLRA of the amphipod *Gammarus fossarum* (class *Malacostraca*, family *Gammaridae*) is phylogenetically more closely related to a protein encoded by a brown algae VLRA than to other negev-like viruses infecting class *Malacostraca* (Supplementary Figure S1).

3.4. SP24-encoding negev-like viruses infecting species of green plants and brown algae

The discovery of the SP24-encoding kitaviruses in plants [9] suggests that the SP24 gene may also be a genomic component of other plant viruses. Indeed, a novel negev-like RNA virus, *fragaria vesca* associated virus 1 (FVaV-1), has been found to contain the SP24 gene, and an SP-24-encoding gene has been found in the genome of recently annotated chrysanthemum kita-like virus (CKLV) [29]. Our search of plant transcriptomes revealed an additional plant VLRA encoding SP24 (Table 4). The phylogeny of the RdRp domains in FVaV-1, *Triticum polonicum* VLRA and *Glycine dolichocarpa* VLRA (Supplementary Figure S1) shows their rather close relationship to arthropod viruses. We hypothesize that these viruses may have emerged in virus evolution by a host change of an insect virus that was accompanied by an acquired ability to replicate or passively persist in plant cells (see also Section 3.1.1.).

Table 4. Negev-like virus RNA genomes infecting species of green plants and brown algae.

Host taxonomic position	Name (Virus and VLRA)	Accession number
<i>Magnoliopsida, Rosales</i>	<i>fragaria vesca</i> associated virus 1	MN895062
<i>Magnoliopsida, Asterales</i>	chrysanthemum kita like virus	OP807956
<i>Magnoliopsida, Caryophyllales</i>	<i>Amaranthus tuberculatus</i> VLRA	GGGT01091955
<i>Magnoliopsida, Lamiales</i>	<i>Linaria vulgaris</i> VLRA*	HBXK01234259
<i>Magnoliopsida, Caryophyllales</i>	<i>Silene latifolia</i> VLRA	JO777742
<i>Magnoliopsida, Poales</i>	<i>Hordeum vulgare</i> VLRA	GGCQ01066052
<i>Magnoliopsida, Poales</i>	<i>Alloteropsis semialata</i> VLRA	GFYF01029264
<i>Magnoliopsida, Ranunculales</i>	<i>Papaver nudicaule</i> VLRA	GJOR01065662
<i>Magnoliopsida, Ranunculales</i>	<i>Papaver armeniacum</i> VLRA	GJO001068528
<i>Magnoliopsida, Ranunculales</i>	<i>Papaver pavinum</i> VLRA	GJO001016337
<i>Magnoliopsida, Fabales</i>	<i>Glycine dolichocarpa</i> VLRA	GGIW01009300
<i>Magnoliopsida, Oxalidales</i>	<i>Elaeocarpus photiniifolius</i> VLRA	FX134396
<i>Pinopsida, Pinales</i>	<i>Pseudotsuga menziesii</i> VLRA	GFFY01083799
<i>Magnoliopsida, Apiales</i>	<i>Panax ginseng</i> VLRA	GDQW01045297
<i>Phaeophyceae, Fucales</i>	<i>Sargassum vulgare</i> VLRA	GEHA01041094

*—VLRAs of *Linaria vulgaris* and related virus strains are in bold.

VLRA from the common toadflax *Linaria vulgaris* (order *Lamiales*, family *Plantaginaceae*), although encoding SP24, differs significantly from other plant negev-like viruses in terms of RdRp phylogeny (Supplementary Figure S1) and gene organization. Unlike FVaV-1 and CKLV, which have four genes, *Linaria vulgaris* VLRA encodes six potential proteins (Figure 4). Replicase protein of more than 1865 residues is potentially encoded by two overlapping ORFs: ORF1a protein has single annotated domain of superfamily 1 viral RNA helicase, while ORF1b overlaps ORF1 by 555 nucleotides and encodes a protein part with RdRp domain. ORF2 protein (637 amino acids) is distantly related to the predicted DiSB domain-containing glycoprotein of *astegopteryx formosana* negev-like virus 1. ORF3 is completely embedded in ORF2 and encodes a hydrophobic protein with no similarity to known viral proteins. ORF4 protein (329 amino acids) includes a cysteine-rich region, while overlapping ORF5 coding for protein possessing the SP24 domain. ORF6 encodes a small protein of 69 amino acids (Figure 4). Our extensive search of the plant NCBI TSA database revealed that SP24-encoding virus of *Linaria vulgaris* can infect several other monocot and dicot plants, namely, *Hordeum vulgare*, *Alloteropsis semialata*, *Papaver nudicaule*, *Papaver armeniacum*, *Papaver californicum*, *Papaver atlanticum* and *Silene latifolia* (Table 4). The RdRp domain of *Linaria vulgaris* VLRA and its isolates specific for the listed above hosts forms a common cluster with virga-like viruses of brown algae in the RdRp phylogenetic tree (Supplementary Figure S1).

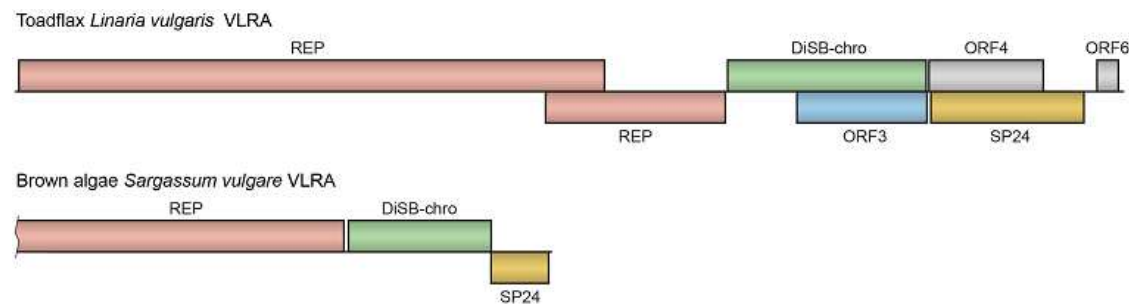


Figure 4. Gene organization of the selected VLRA. See Figure 1 for details of ORF naming and shading.

Strikingly, one of brown algae virga-like viruses infecting *Sargassum vulgare* (class *Phaeophyceae*, order *Fucales*) encodes SP24 (Table 4). This VLRA encodes 3 proteins typical for nege-like viruses (Figure 4). Incomplete ORF1 codes for replicase C-terminal part with superfamily 1 viral RNA helicase domain (pfam01443, e-value—2.74e-38) and RdRp domain (pfam00978, e-value—2.27e-85). ORF2 protein (490 residues) contains DiSB-like domain but has no significant sequence similarity to negevirus proteins. ORF3 encodes protein of 185 amino acids which possesses SP24 domain (e-value—1.06e-07).

3.5. NCBI conserved domain database search reveals horizontal transfer of SP24 protein genes to viruses of order *Hepelivirales* and insect genomes

The NCBI-conserved domain search tool [30] was used to find potential SP24 genes encoded by cellular genomes. Among the four SP24-producing genes, three belong to insects, namely, *Macrosiphum euphorbiae*, order *Hemiptera* (locus CAI6355448, protein size—206 amino acids), *Drosophila melanogaster*, order *Diptera* (locus ABC86319, protein size—202 amino acids), and *Glossina morsitans morsitans*, order *Diptera* (locus ADD20599, protein size—271 amino acids). Another gene is found in the nematode *Nippostrongylus brasiliensis*, order *Strongylida* (locus VDL74358, protein size—184 amino acids). Thus, insect and nematode genomes may acquire viral SP24 genes by putative horizontal transfer from negev-like viruses known to infect species of both taxa (see above).

Another result of our NCBI CDD search is the pioneering discovery of SP24 genes in the genomes of two insect RNA viruses belonging to the order *Hepelivirales*. The genome of hubei hepe-like virus 1 [5] encodes three proteins (accession KX883803). ORF1 encodes a replicase (1661 amino acids) with the C-terminal RdRp domain showing obvious similarity to RdRp of hepeviruses (Table 5). ORF2 encodes a protein (206 amino acids) with SP24 domain, whereas ORF3 encodes a longer protein (560 amino acids) with DiSB-chro domain. Thus, the general genome organization of this virus is different from that of negevirus because of the inverse order of DiSB-chro and SP24 genes. Interestingly, although the replicase of sogatella furcifera hepe-like virus (1844 amino acids) is also related to viruses belonging to the order *Hepelivirales* [13] (Table 5), the genome cistron order in this case is identical to that of negev virus (Figure 1). ORF2 encodes a protein (622 amino acids) with DiSB-chro domain, and ORF3 encodes the SP24 protein (209 amino acids).

Table 5. Similarity of Hubei hepe-like virus 1 RdRp to sequences in NCBI protein database.

Virus and accession number	E-value	Identity to HHLV1 RdRp (%)
<i>Astroviridae</i> sp.; QYJ54474	2e-48	34.54
<i>Hepelivirales</i> sp.; WAY16503	9e-48	34.08
<i>Sogatella furcifera</i> hepe-like virus; YP_009553211	3e-47	34.53
<i>Flumine bastrovirus</i> 3; UQB75993	9e-47	32.84
<i>Hubei sediment bastro-like virus</i> ; QYF50028	2e-45	33.80
<i>Hepelivirales</i> sp.; WAY16393	2e-44	32.34
<i>Bastrovirus-like_virus/VietNam</i> ; YP_009333174	5e-44	33.42
<i>Microbat bastrovirus</i> ; UBK24595	1e-42	31.91

4. Discussion

The present and previously published studies show that there are several important aspects to the evolutionary history of the SP24 protein genes. The high evolutionary impact of the virion SP24 protein is evident because this protein is encoded by many members of at least three orders of the phylum *Kitrinoviricota* (orders *Martellivirales*, *Tolivirales* and *Hepelivirales*) as well as the family *Permutotetraviridae*. Moreover, SP24 appears to be capable of forming virions of different morphologies, from pleiomorphic and elliptical shape to filamentous. In addition, some negev-like viruses of the order *Martellivirales* encode two or even three copies of moderately related SP24 proteins. We propose that these variants of SP24 may have different functions, similar to the tandem duplicated capsid protein genes in members of *Closteriviridae* [31,32] and novel virga-like viruses infecting *Hevea brasiliensis* (rubber tree latent virus 1 and rubber tree latent virus 2) [33]. In this context, it is important to note that one of the SP24 variants in RdFV plays a role in inter-tissue transport through the insect body [19].

We hypothesize that the extremely broad host specificity of SP24-encoded viruses is related to an adaptive role of SP24. These viruses (mainly negev-like viruses) infect most animals (from primitive members of the phyla *Porifera* and *Cnidaria* to arthropods and vertebrates) as well as green plants and brown algae. Such a wide range of host organisms may result from the ability of viruses to rapidly adapt to new hosts upon occasional virus transmission between close and distant eukaryotic taxa. An illustrative example of this phenomenon is SINV-17 and its isolates, which can infect dozens of arthropods (mostly spiders) and some vertebrates. In addition to SINV-17, many insect RNA viruses are also known to be multi-host species [34]. In this paper, similar wide host range covering both monocots and dicots is found for the unique plant virus isolated from *Linaria vulgaris* and showing quite unusual genome organization. Some general properties of RNA viruses that allow high rates of adaptive evolution and facilitate species barriers crossing have been discussed [34]. From an ecological perspective, previous studies have suggested that food-borne transmission is a common route for arthropod predators or scavengers, including spiders and ants. A specific route for arthropods to become infected with viruses is the ingestion of contaminated pollen or feces [35,36].

Our previous data have shown that in virgavirus-like insect viruses, the coding region of replicative RNA helicase is often integrated into insect retrotransposons [37,38]. It is currently clear that horizontal transfer of individual viral genes, including SP24 gene or genome fragments, into insect genomes is a fairly common phenomenon. Such endogenous viral elements (EVE) have been annotated in at least eight insect orders and can be assigned to at least 22 families of positive- and negative-sense RNA viruses [14,39–41]. Studies of EVE have already provided important insights into insect-virus interactions, including the discovery of novel forms of adaptive antiviral immunity [38–42]. Thus, future studies of SP24 genes may shed a light on the functions of the protein in the context of viral genomes and genomes of insects and other eukaryotes.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org. Figure S1: Phylogenetic tree of the RdRp domains in kitaviruses and virgaviridae-like viruses. Whole tree was generated using the maximum likelihood method. Branch lengths correspond to the branch line's genetic distances. Brown alga endornavirus RdRp is used as outgroup. SP24 means encoding of SP24 protein. Viruses and VLRAs are shaded in colors according taxonomic positions of their hosts: clade *Spiralia* is shown in pink; class *Nematoda* is shown in violet; brown Algae—orange; plants—green; class *Malacostraca*—dark blue; class *Arachnida*—dark grey; class *Insecta*—yellow; vertebrates—light blue; fungi and class *Hexacorallia* are unmarked.

Author Contributions: Conceptualization, S.M. and A.S.; methodology, S.M.; software, A.S.; validation, S.M. and A.S.; formal analysis, A.S.; data curation, A.S.; writing—original draft preparation, S.M.; writing—review and editing, A.S. and S.M.; supervision, S.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data are available upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zerbini, F.M.; Siddell, S.G.; Mushegian, A.R.; Walker, P.J.; Lefkowitz, E.J.; Adriaenssens, E.M.; Alfenas-Zerbini, P.; et al. Differentiating between viruses and virus species by writing their names correctly. *Arch. Virol.* **2022**, *167*, 1231-1234.
2. Kuchibhatla, D.B.; Sherman, W.A.; Chung, B.Y.; Cook, S.; Schneider, G.; Eisenhaber, B.; Karlin, D.G. Powerful sequence similarity search methods and in-depth manual analyses can identify remote homologs in many apparently “orphan” viral proteins. *J. Virol.* **2014**, *88*, 10-20.
3. Solovyev, A.G.; Morozov, S.Y. Non-replicative integral membrane proteins encoded by plant alpha-like viruses: emergence of diverse orphan ORFs and movement protein genes. *Front. Plant. Sci.* **2017**, *8*, 1820.
4. Chevin, A.; Coutard, B.; Blanchard, P.; Dabert-Gay, A.S.; Ribiere-Chabert, M.; Thiery, R. Characterization of structural proteins from Chronic Bee Paralysis Virus (CBPV) using mass spectrometry. *Viruses* **2015**, *7*, 3329-3344.
5. Shi, M.; Lin, X.-D.; Tian, J.-H.; et al. Redefining the invertebrate RNA virosphere. *Nature* **2016**, *540*, 539–543.
6. Chiapello, M.; Bosco, L.; Ottati, S.; et al. Complexity and local specificity of the virome associated with tospovirus-transmitting thrips species. *J. Virol.* **2021**, *95*, e00597.
7. Parry, R.; James, M.E.; Asgari, S. Uncovering the worldwide diversity and evolution of the virome of the mosquitoes *Aedes aegypti* and *Aedes albopictus*. *Microorganisms* **2021**, *9*, 1653.
8. Morozov, S.Y.; Lazareva, E.A.; Solovyev, A.G. Sequence relationships of RNA helicases and other proteins encoded by blunervirus RNAs highlight recombinant evolutionary origin of kitaviral genomes. *Front. Microbiol.* **2020**, *11*, 561092.
9. Ramos-González, P. L.; Arena, G.D.; Tassi, A.D.; et al. Kitaviruses: a window to atypical plant viruses causing nonsystemic diseases. *Ann. Rev. Phytopathol.* **2023**, *61*, 97-118.
10. Dorrington, R.A.; Jiwaji, M.; Awando, J.A.; de Bruyn, M.M. Advances of tetravirus research: new insights into the infectious virus lifecycle and an expanding host range. *Curr. Issues Mol. Biol.* **2020**, *34*, 145-162.
11. Kallies, R.; Kopp, A.; Zirkel, F.; et al. Genetic characterization of goutanap virus, a novel virus related to negevirus, cileviruses and higreviruses. *Viruses* **2014**, *6*, 4346-4357.
12. Nunes, M. R.; Contreras-Gutierrez, M. A.; Guzman, H.; et al. Genetic characterization, molecular epidemiology, and phylogenetic relationships of insect-specific viruses in the taxon Negevirus. *Virology* **2017**, *504*, 152-167.
13. Wu, N.; Zhang, P.; Liu, W.; Wang, X. Sogatella furcifera hepe-like virus: first member of a novel Hepeviridae clade identified in an insect. *Virus Res.* **2018**, *250*, 81-86.
14. Kondo, H.; Chiba, S.; Maruyama, K.; Andika, I. B.; Suzuki, N. A novel insect-infecting virga/nege-like virus group and its pervasive endogenization into insect genomes. *Virus Res.* **2019**, *262*, 37–47.
15. Kondo, H.; Fujita, M.; Hisano, H.; Hyodo, K.; Andika, I. B.; Suzuki, N. Virome Analysis of Aphid Populations That Infest the Barley Field: The Discovery of Two Novel Groups of Nege/Kita-Like Viruses and Other Novel RNA Viruses. *Front. Microbiol.* **2020**, *11*, 509.
16. O'Brien, C.A.; McLean, B.J.; Colmant, A.M.G.; et al. Discovery and characterization of Castlereas virus, a new species of negevirus isolated in Australia. *Evol. Bioinform. Online* **2017**, *13*, 1176934317691269.
17. Dheilly, N.M.; Lucas, P.; Blanchard, Y.; Rosario, K. A world of viruses nested within parasites: Unraveling viral diversity within parasitic flatworms (Platyhelminthes). *Microbiol. Spectr.* **2022**, *10*, e0013822.
18. Arakawa, K.; Kono, N.; Malay, A.D.; et al. 1000 spider silkomes: Linking sequences to silk physical properties. *Sci. Adv.* **2022**, *8*, eabo6043.
19. Wan, J.; Liang, Q.; Zhang, R.; Chen, Y.; et al. Arboviruses and symbiotic viruses cooperatively hijack insect sperm-specific proteins for paternal transmission. *Nat. Commun.* **2023**, *14*, 1289.
20. Agranovsky, A. A. Where to stop: Occurrence and evolution of translational recording signals in RNA viruses of eukaryotes. *Gene Expression* **2023**, *22*, 240-249.
21. Williams, S.H.; Levy, A.; Yates, R.A.; et al. The Diversity and Distribution of Viruses Associated with Culex annulirostris Mosquitoes from the Kimberley Region of Western Australia. *Viruses* **2020**, *12*, E717.
22. Zhang, W.; Zhang, Y.C.; Wang, Z.G.; Gu, Q.Y.; Niu, J.Z.; Wang, J.J. The Diversity of Viral Community in Invasive Fruit Flies (Bactrocera and Zeugodacus) Revealed by Meta-transcriptomics. *Microb. Ecol.* **2022**, *83*, 739-752.
23. Linnakoski, R.; Sutela, S.; Coetzee, M.P.A.; et al. Armillaria root rot fungi host single-stranded RNA viruses. *Sci. Rep.* **2021**, *11*, 7336.
24. Xavier, C.A.D.; Allen, M.L.; Whitfield, A.E. Ever-increasing viral diversity associated with the red imported fire ant Solenopsis invicta (Formicidae: Hymenoptera). *Virol. J.* **2021**, *18*, 5.

25. Niu, J.; Zhang, W.; Sun, Q.-Z.; Wang, J.-J. Three novel RNA viruses in the spider mite *Tetranychus urticae* and their possible interactions with the host RNA interference response. *J. Invertebr. Pathol.* **2019**, *166*, 107228.
26. Qin, X.C.; Shi, M.; Tian, J.H.; et al. A tick-borne segmented RNA virus contains genome segments derived from unsegmented virus ancestors. *Proc. Nat. Acad. Sci. USA* **2014**, *111*, 6744-6749.
27. Khalturin, K.; Shunatova, N.; Shchenkov, S.; Sasakura, Y.; Kawamitsu, M.; Satoh, N. Polyzoa is back: the effect of complete gene sets on the placement of Ectoprocta and Entoprocta. *Sci. Adv.* **2022**, *8*, 26, eabo4400.
28. Dunn, C.W.; Giribet, G.; Edgecombe, G.D.; Hejnol, A. Animal phylogeny and its evolutionary implications. *Ann. Rev. Ecol. Evol. Syst.* **2014**, *45*, 371-395.
29. Lenz, O.; Přibylková, J.; Fránová, J.; Koloniuk, I. *Fragaria vesca*-associated virus 1: a new virus related to negeviruses. *Arch. Virol.* **2020**, *165*, 1249-1252.
30. Marchler-Bauer, A.; Bo, Y.; Han, L.; et al. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.* **2017**, *45*, D200-D203.
31. Tatineni, S.; Gowda, S.; Dawson, W.O. Heterologous minor coat proteins of Citrus tristeza virus strains affect encapsidation, but the coexpression of HSP70h and p61 restores encapsidation to wild-type levels. *Virology* **2010**, *402*, 262-270.
32. Mostert, I.; Bester, R.; Burger, J.T.; Maree, H.J. Identification of Interactions between Proteins Encoded by Grapevine Leafroll-Associated Virus 3. *Viruses* **2023**, *15*, 208.
33. Chen, S.; Yu, F.; Li, Z.; et al. Complete genome sequence of a novel virga-like virus infecting *Hevea brasiliensis*. *Arch. Virol.* **2022**, *167*, 965-968.
34. Longdon, B.; Brockhurst, M.A.; Russell, C.A.; Welch, J.J.; Jiggins, F.M. The evolution and genetics of virus host shifts. *PLoS Pathog.* **2014**, *10*, e1004395.
35. Schlappi, D.; Chejanovsky, N.; Yañez, O.; Neumann P. Virus transmission via honey bee prey and potential impact on cocoon-building in labyrinth spiders (*Agelena labyrinthica*). *PLoS ONE* **2023**, *18*, e0282353.
36. Shelomi, M.; Lin, W.; Johnson, B. R.; Furlong, M.J.; Etebari, K. Detection of deformed wing virus (DWV) in the Vietnamese walking stick *Medauroidea extradentata* (Phasmatodea). *Virus Res.* **2021**, *293*, 198263.
37. Lazareva, E.; Lezzhov, A.; Vassetzky, N.; Solovyev, A.; Morozov S. Acquisition of full-length viral helicase domains by insect retrotransposon-encoded polypeptides. *Front. Microbiol.* **2015**, *6*, 1447.
38. Morozov, S. Y.; Lazareva, E. A.; Solovyev, A. G. RNA helicase domains of viral origin in proteins of insect retrotransposons: possible source for evolutionary advantages. *PeerJ.* **2017**, *5*, e3673.
39. Gilbert, C.; Belliardo, C. The diversity of endogenous viral elements in insects. *Curr. Opin. Insect Sci.* **2022**, *49*, 48-55.
40. Blair, C.D.; Olson, K.E.; Bonizzoni M. The Widespread Occurrence and Potential Biological Roles of Endogenous Viral Elements in Insect Genomes. *Curr. Issues Mol. Biol.* **2020**, *34*, 13-30.
41. Wallau, G.L. RNA virus EVEs in insect genomes. *Curr. Opin. Insect Sci.* **2022**, *49*, 42-47.
42. Morozov, S. Y.; Lezzhov, A.A.; Lazareva, E. A.; Erokhina, T.N.; Solovyev, A. G. Potential role of accessory domains in polyproteins encoded by retrotransposons in anti-viral defense of host cells. *Front. Microbiol.* **2019**, *9*, 3193.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.