






Article

Identification and Characterization of Three Novel Iflaviruses in the Cabbage Whitefly *Aleyrodes proletella*

Zhuang-Xin Ye ^{1,2,†} , Guo-Wei Gu ^{3,†}, Peng-Peng Ren ², Chuan-Xi Zhang ² , Jun-Min Li ² , Yan Zhang ^{2,*} 
and Jian-Ping Chen ^{1,2,*} 

¹ College of Forestry and Grassland, Nanjing Forestry University, Nanjing 210037, China; yzx244522794@163.com

² State Key Laboratory for Quality and Safety of Agro-Products, Key Laboratory of Biotechnology in Plant Protection of MARA, Zhejiang Key Laboratory of Green Plant Protection, Institute of Plant Virology, Ningbo University, Ningbo 315211, China; 2011074040@nbu.edu.cn (P.-P.R.); chxzhang@zju.edu.cn (C.-X.Z.); lijunmin@nbu.edu.cn (J.-M.L.)

³ Popularization of Agricultural Technical Service Station of Yuyao, Yuyao 315400, China; ggw_1983@126.com

* Correspondence: zhangxiaoli1028@163.com (Y.Z.); jianpingchen@nbu.edu.cn (J.-P.C.)

† These authors contributed equally to this work.

Simple Summary: Cabbage whitefly (*Aleyrodes proletella*) is a major agricultural pest that damages crops, such as cabbage, broccoli, and kale, by feeding on plant sap. However, the virome of the *A. proletella* remains largely unexplored. In this study, we employed deep sequencing technology to investigate the virome of *A. proletella* and successfully identified the genomes of three novel viruses. Phylogenetic analysis confirmed that the three viruses represent three new species within the genus *Iflavirus* and were assigned to different branches on the phylogenetic tree. These findings enhance our understanding of viral diversity in agricultural pests and provide a foundation for future research and contribute valuable viral resources for the development of biopesticides based on viruses.

Abstract: The cabbage whitefly (*A. proletella*) (Hemiptera: Aleyrodidae) is a major agricultural pest that primarily targets cruciferous crops, such as cabbage, broccoli, and kale, causing extensive damage through feeding and honeydew. However, its associated virome has received limited research attention. In this study, we collected cabbage whiteflies in Xinjiang Agricultural University (43.80833 N, 87.56778 E, 882.3 m), systematically identified the RNA virome of the *A. proletella* and successfully identified three novel iflaviruses (*Aleyrodes proletella* iflavirus 1 (APIV1), *Aleyrodes proletella* iflavirus 2 (APIV2) and *Aleyrodes proletella* iflavirus 3 (APIV3)). APIV1–3 all have a 5' structural protein region and a 3' non-structural protein region. Phylogenetic and sequence identity analyses suggest that APIV1–3 are novel members of the family *Iflaviridae*. Structural modeling using AlphaFold3 revealed a conserved protein core region and a variable outer loop region. This study provides valuable insights into the virome diversity of *A. proletella*, establishing a foundation for future research on virus–host interactions and the potential for biocontrol applications in sustainable agriculture.

Keywords: *Aleyrodes proletella*; novel viruses; *Iflavirus*; meta-transcriptomics



Academic Editor: Zhen Li

Received: 14 February 2025

Revised: 15 March 2025

Accepted: 20 March 2025

Published: 22 March 2025

Citation: Ye, Z.-X.; Gu, G.-W.; Ren, P.-P.; Zhang, C.-X.; Li, J.-M.; Zhang, Y.; Chen, J.-P. Identification and Characterization of Three Novel Iflaviruses in the Cabbage Whitefly *Aleyrodes proletella*. *Insects* **2025**, *16*, 335. <https://doi.org/10.3390/insects16040335>

Copyright: © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the Creative Commons

Attribution (CC BY) license

(<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The cabbage whitefly (*Aleyrodes proletella*), a member of the family Aleyrodidae in the order Hemiptera, is native to East, West, South and North Africa and has since spread to other regions, including South America, North America, New Zealand and Australia [1–3].

In New Zealand, it is widely distributed among native plants, such as New Zealand spurge (*Euphorbia glauca*) and New Zealand celery (*Apium prostratum*), as well as Brassicaceae crops (e.g., kale, cabbage, and Brussels sprouts), ornamental flowers (e.g., plants of the *Aquilegia* genus), and weeds (e.g., chicory) [4]. In addition, *A. proletella* also infests plants from the Asteraceae, Fabaceae, and Papaveraceae families in other countries [5]. *A. proletella* is a sap-sucking insect that penetrates the phloem of plants to feed on sap. The deleterious effects of *A. proletella* on crop growth primarily arises from its feeding activities, wax deposits, and the development of sooty mould due to honeydew excretion. Additionally, its egg-laying behaviour further contributes to plant stress and reduced crop yield [4,6].

With the continuous progress of high-throughput sequencing technologies, meta-transcriptomics has become an essential tool for virus discovery and identification [6–8]. This technology has shown great potential in studying the virome of agricultural insects, facilitating the rapid identification of novel RNA viruses in diverse insect species. For example, multiple novel RNA viruses have been identified in locusts [9], leafhoppers [10,11], planthoppers [12,13] and whitefly (*Bemisia tabaci*) [14]. Although *A. proletella* causes significant agricultural losses in many regions worldwide, its virome is still vastly understudied compared to that of *B. tabaci*. This knowledge gap limits a comprehensive understanding of viral diversity within the family Aleyrodidae. Artificial Intelligence (AI) methods and deep learning algorithms have had a significant impact on various research fields in life sciences, such as virus discovery and protein structure prediction. Recently, AlphaFold3 has been developed and applied to virus identification and viral protein structure prediction to better understand viral evolution [15].

Iflaviridae is a family within the order *Picornavirales* that possesses a positive-sense, single-stranded, non-segmented RNA genome (monocistronic genome) that ranges in length from approximately 9 to 11 kb. The genome encodes a single polyprotein, which is autocatalytically cleaved into structural proteins (SPs) and non-structural proteins (NSPs) involved in replication [16]. Typically, the SPs are encoded in the 5' region of the genome, while the NSPs are encoded in the 3' region. However, some studies have shown gene rearrangement of structural and non-structural proteins within the family *Iflaviridae*, suggesting that it might be more complex than previously understood [17]. Currently, the family *Iflaviridae* contains only one genus, *Iflavirus*, with iflaviruses identified exclusively in arthropods, primarily insects. Once the virus enters the host cell, the infection process occurs rapidly, with progeny viruses produced within hours.

In this study, we collected over 100 cabbage whiteflies from fresh leaf lettuce in Xinjiang Agricultural University and performed high-throughput sequencing. Three novel iflaviruses were identified in *A. proletella* using meta-transcriptomics for the first time, named Aleyrodes prolellus iflavirus 1 (APIV1), Aleyrodes prolellus iflavirus 2 (APIV2) and Aleyrodes prolellus iflavirus 3 (APIV3). Furthermore, the genomic structure, evolutionary relationships, and 3D structural comparisons were analyzed in detail. These findings provide a crucial foundation for understanding the composition of the virome in the *A. proletella* and exploring viral diversity within the family Aleyrodidae.

2. Materials and Methods

2.1. Sample Preparation and RNA Extraction

Cabbage whiteflies were collected from vegetable gardens of Xinjiang Agricultural University (43.80833 N, 87.56778 E, 882.3 m) on fresh leaf lettuce (*Lactuca sativa*) in June 2024 [18]. Then they were maintained in insect-proof cages on fresh romaine lettuce at 25 ± 1 °C, 50–70% relative humidity, and a 14 h light/10 h dark cycle in the laboratory of Ningbo University. Total RNA was extracted from approximately 100 whiteflies using TRIzol reagent (Takara, Beijing, China) according to the manufacturer's protocol. The RNA

quality was assessed using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA).

2.2. Transcriptome Sequencing

The RNA libraries were constructed using the NEBNext Ultra RNA Library Prep Kit (New England Biolabs, Ipswich, MA, USA) and sequenced on an Illumina NovaSeq platform with paired-end 150 bp reads. Raw reads were filtered to remove adaptor sequences and low-quality reads using Trimmomatic (version 0.39) [19]. De novo assembly was performed using Trinity (version 2.8.5) with default parameters. Contigs longer than 200 bp were selected for further analysis. The flited contigs were compared to the iBOL cytochrome oxidase subunit 1 (COI) database [20] using blastn to identify and filter COI sequences. The filtered COI sequence was then subjected to further analysis using the NCBI blastn tool to determine species identity.

2.3. Virus Discovery and Confirmation

The assembled contigs were subjected to DIAMOND (version 0.9.28.129) searches against a local viral sequence database to identify potential viral-like contigs, which consisted of representative viruses downloaded from NCBI Virus on 1 January 2025 (<https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/>) with an E-value cutoff of 1×10^{-5} . To eliminate false positives, clean reads were mapped back to the viral genome to assess the sequencing depth of the viral sequence; the candidate viral contigs were further compared with the NCBI complete nucleotide (NT) and non-redundant (NR) protein databases [21].

2.4. Structural and Genomic Analysis

Open reading frames (ORFs) of the candidate viruses were comprehensively analysed using ORFfinder online server (<https://www.ncbi.nlm.nih.gov/orffinder/>, accessed on 1 January 2025). Conserved protein domains were predicted using a combination of databases, including the Protein Data Bank (PDB) (<https://www.rcsb.org/>, accessed on 1 January 2025) and the Pfam database (<https://www.ebi.ac.uk/interpro/>, accessed on 1 January 2025). Specifically, the pfam_search.pl script was employed to scan for known domain motifs within the viral sequences, and HHblits (version 3.3.0) was used for more sensitive and in-depth sequence alignments, improving the accuracy of domain prediction by Hidden Markov models (HMMs). The predicted ORFs and protein domains were then visualized using the R package gggenes (version 0.5.1).

2.5. Amino Acid Similarity Comparison

Reference sequences of classified iflaviruses and highly homologous sequences were retrieved from GenBank and the ICTV database [22]. Pairwise sequence alignment and identity calculation [23] were constructed based on the amino acid (aa) sequences of the predicted conserved capsid protein (CP) region, RNA-dependent RNA polymerase (RdRp), and complete polyprotein of the selected iflaviruses. Sequence alignment was conducted using MAFFT (version 7.525) [24], and sequence similarity was calculated based on the percentage of amino acid identity.

2.6. 3D Structure Prediction

The polyproteins of iflaviruses were submitted to AlphaFold (<https://alphafoldserver.com/>, accessed on 1 January 2025) for 3D structure prediction. The reliability of the model was assessed using pLDDT (predicted Local Distance Difference Test) scores, with all predictions performed using default settings. pLDDT scores > 70 indicate high confidence in the predicted structure, while scores < 50 suggest low reliability. To evaluate the structural

similarity and variability among the predicted protein structures, pairwise alignments were conducted using the TM-align tool [25].

2.7. Phylogenetic Analysis

Phylogenetic trees were constructed based on the amino acid sequences of RdRp, CP, and complete polyprotein, along with corresponding regions of representative viruses from various families within the order *Picornavirales*. Multiple sequence alignment was performed using MAFFT and retained conserved regions of longer fragments using Gblocks with the parameters (-b2 = 0.55 -b4 = 5 -b5 = h) (version 0.91b). The best amino acid substitution model was selected using ModelTest, and phylogenetic trees were constructed using the Maximum Likelihood (ML) algorithm by RAxML-NG (version 0.9.0) with 1000 bootstrap replicates [26]. The genus *Triatovirus* was used as an outgroup to root the trees.

3. Results

3.1. Three Iflaviruses Identified in *Aleyrodes proletellus*

A total of 97,954 contigs longer than 200 bp were obtained from the de novo assembly of clean reads. The cabbage whitefly species was identified as *A. proletellus*, with a 99.83% COI sequence similarity to the previously reported *A. proletellus* sequence (Accession Number: MK168804.1). Through a homology search against the local virus database, 16 contigs (comprising a total of 17,890 reads) from various viral families were identified. Among these, three nearly complete genomes were successfully identified and designated as *Aleyrodes proletellus* iflavirus 1 (APIV1, 9665 nt, 76.2× coverage) (Figure 1A), *Aleyrodes proletellus* iflavirus 2 (APIV2, 8780 nt, 81.7× coverage) (Figure 1B), and *Aleyrodes proletellus* 3 (APIV3, 8608 nt, 25.5× coverage) (Figure 1C). Comparison with the NCBI NR database revealed the highest sequence similarity of 38.71% between APIV1 and Bat faecal-associated iflavirus 2 (ON872543.1), while APIV2 and APIV3 exhibited highest sequence similarities of 47.08% (Novo Mesto iflavirus 1, OL472192.1) and 41.72% (Myrmica rubra picorna-like virus 8, MW314655.1), respectively (Table 1). These results indicate that the three viruses are likely new species. The three viral contigs were further verified by RT-PCR, followed by Sanger sequencing. There is an overlapping area between every two adjacent fragments for the three viruses to eliminate false positives (such as misassembled contigs, recombination, and others) (Supplementary Table S1). The genome sequences of APIV1 (9665 nt), APIV2 (8780 nt) and APIV3 (8608 nt) were subsequently submitted to GenBank, with the accession numbers PQ888994, PQ888995 and PQ888996, respectively.

Table 1. The detailed information about the three novel viruses identified in *Aleyrodes proletella*.

| Virus Names | Accession | Length (nt) | Homologous Virus Accession | Closest Homologous Virus | Cover (%) | Identity (%) | Aligned Length | E-Value |
|--|-----------|-------------|----------------------------|------------------------------------|-----------|--------------|----------------|---------|
| <i>Aleyrodes proletellus</i> iflavirus 1 (APIV1) | PQ888994 | 9665 | ON872543.1 | Bat faecal-associated iflavirus 2 | 63% | 38.71% | 1873 | 0 |
| <i>Aleyrodes proletellus</i> iflavirus 2 (APIV2) | PQ888995 | 8780 | OL472192.1 | Novo Mesto iflavirus 1 | 98% | 47.08% | 2849 | 0 |
| <i>Aleyrodes proletellus</i> iflavirus 3 (APIV3) | PQ888996 | 8608 | MW314655.1 | Myrmica rubra picorna-like virus 8 | 54% | 41.72% | 1806 | 0 |

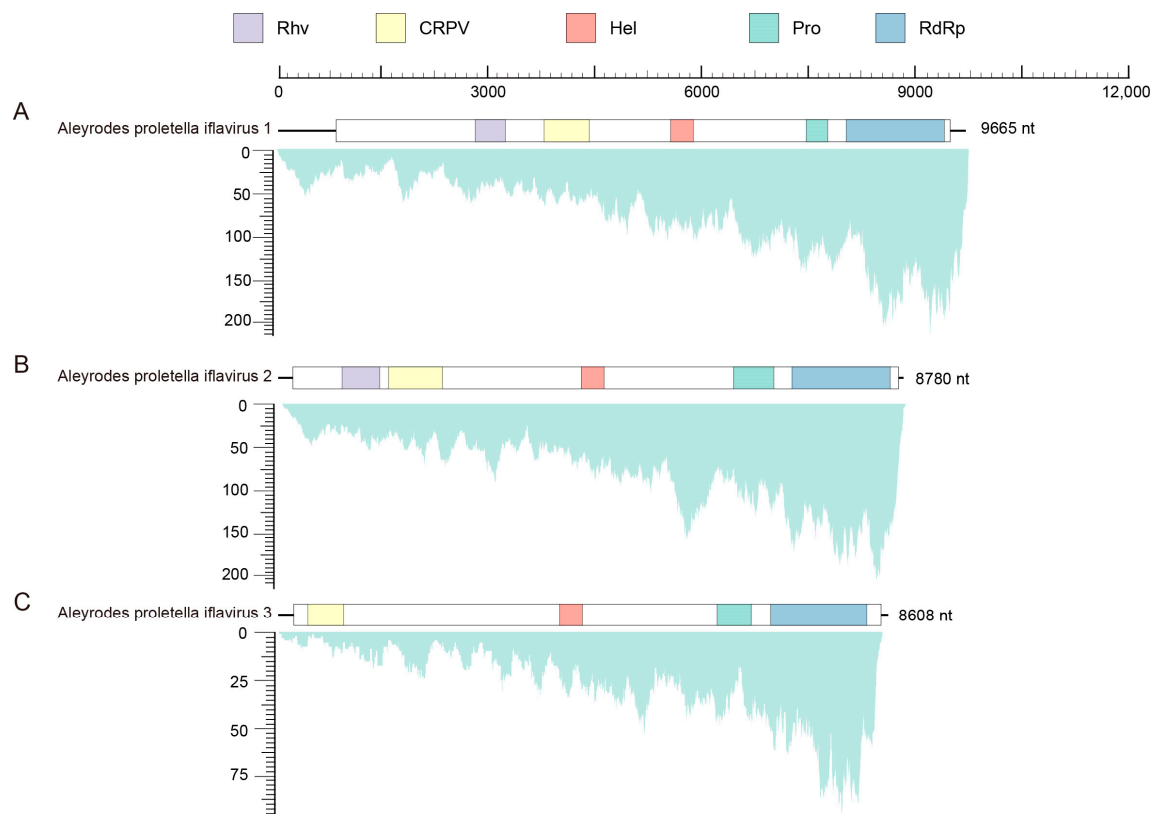


Figure 1. Genome structure and transcriptome raw read coverage of AIPV1 (A), AIPV2 (B) and AIPV3 (C). The white frame represents ORFs. Blocks of different colours represent different domains of the polyprotein. Purple: Rhv, picornavirus-like capsid domain; Yellow: CRPV, CRPV_capsid superfamily domain; Red: Hel, RNA helicase domain; Green: Pro, 3-chymotrypsin-like protease domain; Blue: RdRp, RNA-dependent RNA polymerase domain.

3.2. Genome Structure of Three Novel Viruses

AIPV1, AIPV2 and AIPV3 are predicted to contain a single ORF encoding polyproteins of 2891, 2851, and 2764 amino acids, respectively. InterProScan analysis revealed that both AIPV1 and AIPV2 harbor five conserved domains: a picornavirus-like capsid domain (Rhv), a CRPV_capsid superfamily domain (CRPV), an RNA helicase domain (Hel), a 3-chymotrypsin-like protease domain (Pro), and an RNA-dependent RNA polymerase domain (RdRp). In AIPV1, the SPs comprise an Rhv domain (2737–3163 nt) and a CRPV domain (3709–4348 nt), while the NSPs include a Hel domain (5497–5818 nt), a Pro domain (nucleotides 7414–7717 nt), and an RdRp domain (7978–9364 nt) (Figure 1A). In AIPV2, the structural proteins consist of an Rhv domain (857–1388 nt) and a CRPV domain (1511–2276 nt), whereas the non-structural proteins are represented by a Hel domain (4235–4559 nt), a Pro domain (6386–6956 nt), and an RdRp domain (7211–8600 nt) (Figure 1B). AIPV3 is characterized by a single CRPV domain (372–879 nt) as its structural component, with non-structural proteins including a Hel domain (3057–3381 nt), a Pro domain (5283–5766 nt), and an RdRp domain (6039–7398 nt) (Figure 1C). Notably, the NSPs are arranged in the order Hel–Pro–RdRp, from the N- to the C-terminus. The Hel domain, belonging to superfamily 3 and associated with NTP binding, is characterized by a conserved motif A (2185–2192 nt: GxxGxGKS) [27]; the chymotrypsin-like cysteine protease domain, which resembles the 3C protease of picornaviruses, contains a cysteine protease motif (3495–3498 nt: GxCG) [28] and a substrate-binding motif (3515–3520 nt: GxHxxG) [27]; and the RdRp domain, a member of RNA polymerase superfamily I, har-

bors conserved motifs including motif V (3932–3943 nt: PSG_{x3}TX₃N(S/T)₂) and motif VI (3984–3987 nt: YGDD) [29] (Figure S1).

3.3. The Evolutionary Distances of APIV1–3 and Other Iflaviruses

To determine whether the three iflaviruses are independent virus species or merely different isolates of the same virus, their CP and RdRp were compared with ICTV-confirmed Iflaviruses using the SDT. The results revealed that these three iflaviruses species have consistent similarities in the CP (Figure 2A) and RdRp (Figure 2B), with similarities of 43% (ON872543, Bat faecal-associated iflavirus 2), 40% (NC_002066 Sacbrood virus), and 37% (NC_002066, Sacbrood virus) compared to ICTV-recognized iflaviruses species. Based on ICTV criteria for new Iflaviruse species (intraspecies differentiation requiring >90% capsid protein amino acid sequence similarity) (<https://ictv.global/report/chapter/iflaviridae/iflaviridae/iflavirus>, accessed on 1 January 2025), APIV1, APIV2, and APIV3 were confirmed as three distinct novel Iflaviruses species (Figure 2).

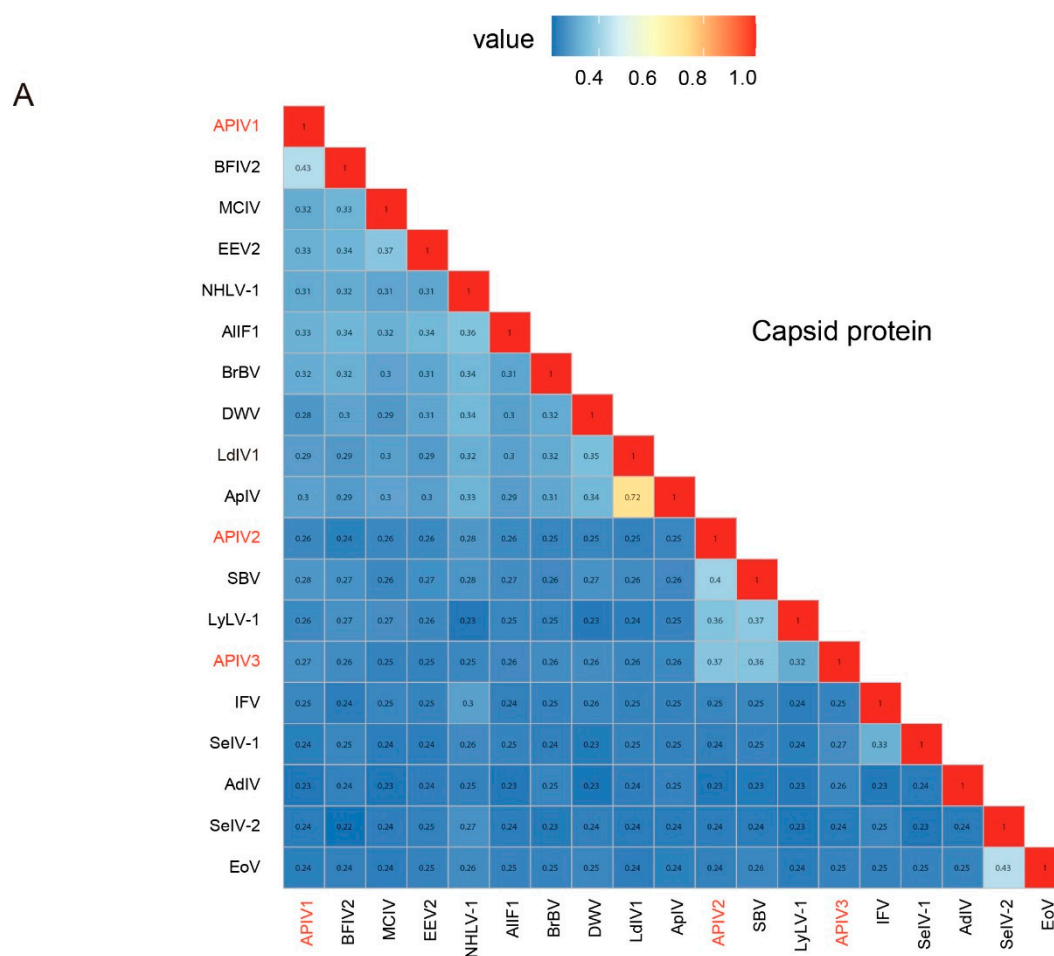


Figure 2. Cont.

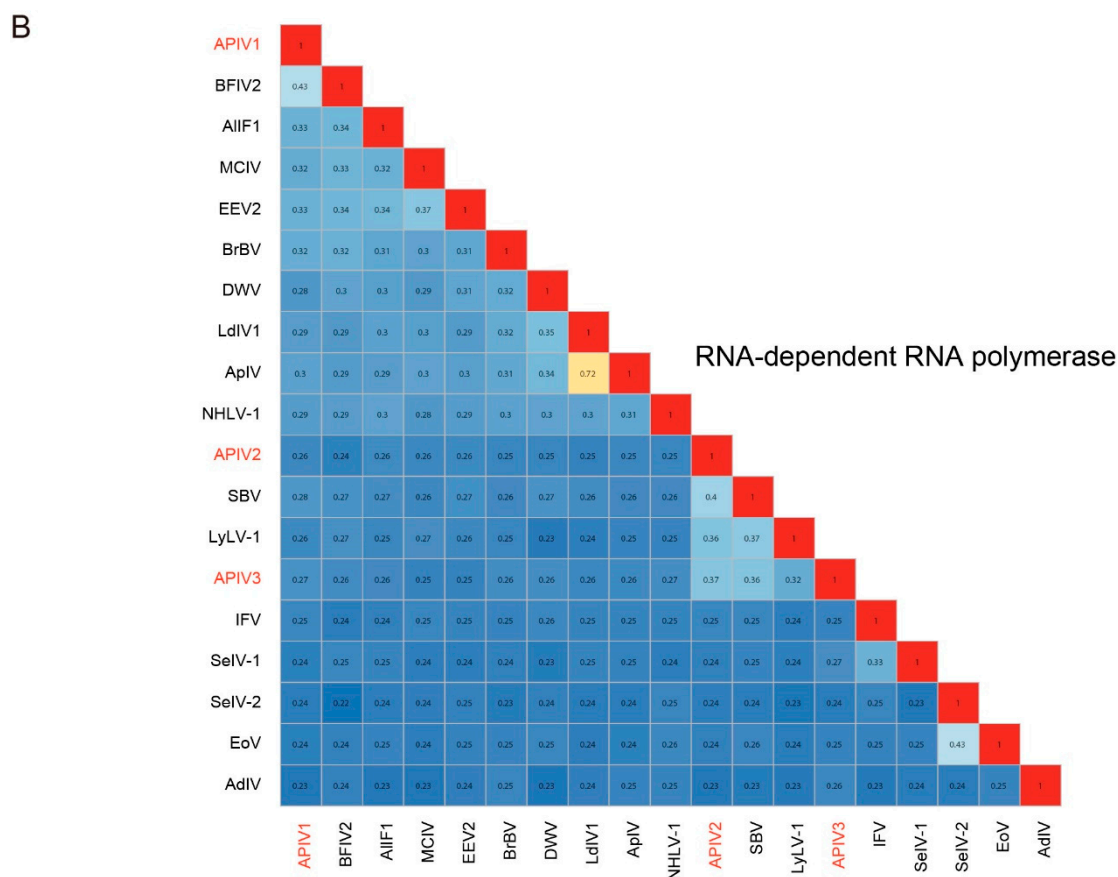


Figure 2. Pairwise distances between CP (A) and RdRp (B) sequences among the iflaviruses identified in cabbage whiteflies. Through the predicted CP and RdRp domains, a pairwise distance matrix was constructed using MAFFT for the three newly discovered iflaviruses and homologous iflaviruses sequences included in ICTV and NCBI GenBank using ggplot2. The new virus sequences identified in this study were highlighted in red. EoV: *Ectropis obliqua* picorna-like virus, SeIV-2: *Spodoptera exigua* iflaviruses 2, AdIV: *Acheta domesticus* iflaviruses, SeIV-1: *Spodoptera exigua* iflaviruses 1, IFV: Infectious flacherie virus, LyLV-1: *Lygus lineolaris* virus 1, SBV: *Sacbrood virus*, ApIV: *Antheraea pernyi* iflaviruses, LdIV1: *Lymantria dispar* iflaviruses 1, DWV: *Deformed wing virus*, BrBV: *Brevicoryne brassicae* virus—UK, NHLV-1: *Nilaparvata lugens* honeydew virus 1, DcPV: *Dinocampus coccinellae* paralysis virus, VDV-2: *Varroa destructor* virus 2, SBPV: *Slow bee paralysis virus*, EEV2: *Exitianus exitiosus* virus 2, MCIV: *Medvezhye Chrysops Iffa*—like virus, BFIV2: *Bat faecal-associated iflaviruses* 2, AIIF1: *Aulacophora lewisii* iflaviruses 1. The viral Accession numbers and name are listed in Supplementary Table S2.

3.4. Phylogenetic Analysis of Three Iflaviruses

To confirm the authentic taxonomic status of the three iflaviruses, phylogenetic trees based on the CP (Figure 3A), RdRp (Figure 3B) and Polyprotein (Figure S2) sequences of representative members of the family *Iflaviridae* were constructed. Both the polyprotein and CP-based phylogenetic trees exhibited highly consistent evolutionary relationships, indicating similar conservation patterns in Iflaviruses evolution. APIV1, APIV2, and APIV3 clustered in three distinct evolutionary branches. APIV1 exhibited the closest phylogenetic relationship with *Bat faecal-associated iflaviruses* 2, while APIV2 and APIV3 grouped with evolutionary branches that include *Hubei picorna-like virus* 40 (NC 032756.1). The high-confidence phylogenetic tree supports the presence of potentially undiscovered Iflaviruses diversity (Figure 3 and Figure S2).

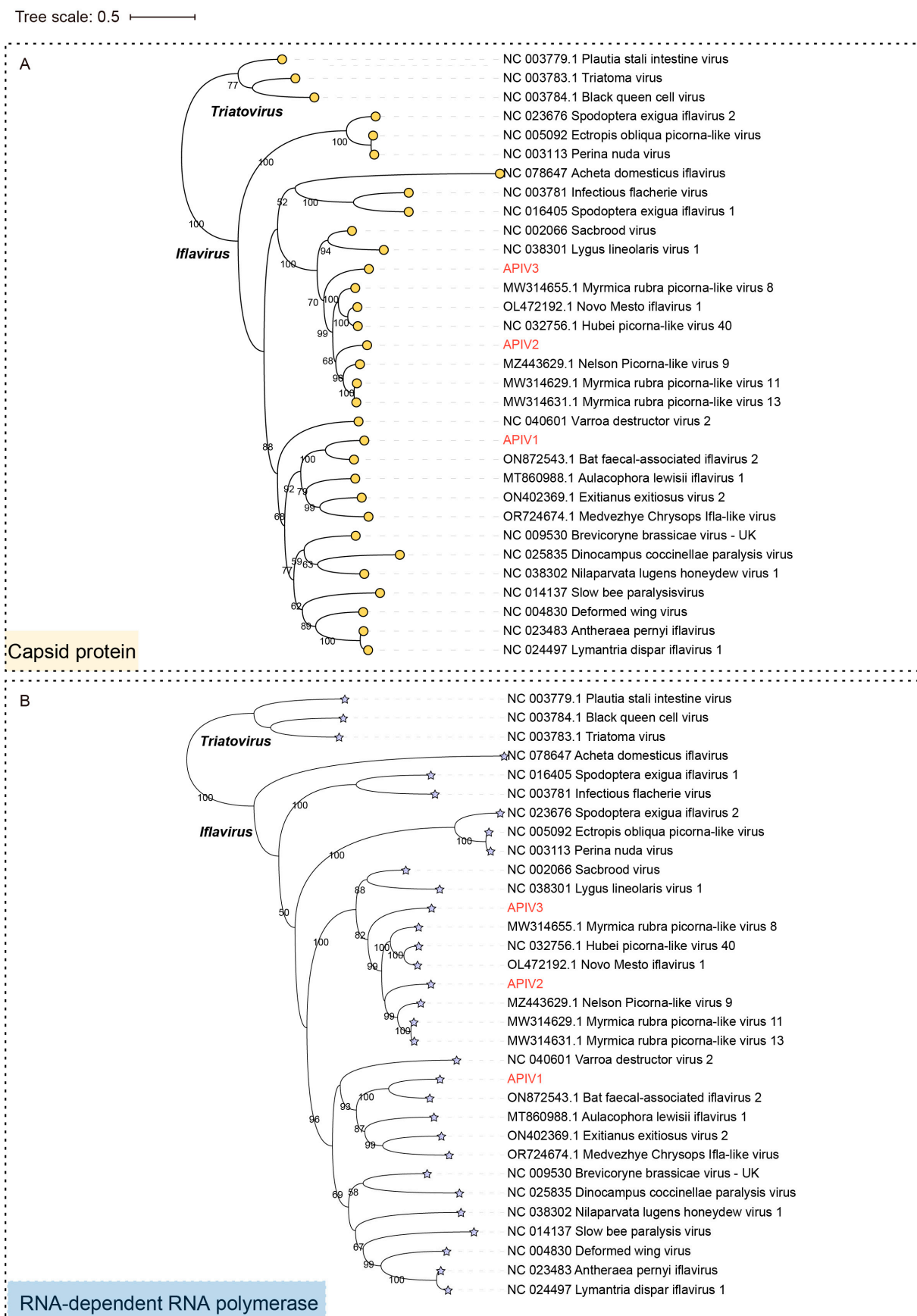


Figure 3. Phylogenetic trees based on the CP (indicated by circular markers) (A) and RdRp (indicated by asterisks) (B) of the three iflaviruses identified in cabbage whiteflies. APIV1, APIV2 and APIV3 are highlighted in red font. Nodes with bootstrap values > 50% are marked.

3.5. Viral 3D Structure Comparison

Phylogenetic analysis revealed that the three viruses did not form a single cluster, suggesting that they may have distinct origins. To further explore the differences among them, we predicted the 3D structures of APIV1–3 and performed a comparative analysis. AlphaFold3 predictions of polyprotein structures yielded average pLDDT scores of 64.93, 70.69, and 61.64 for APIV1, APIV2, and APIV3, respectively. TM-align analysis of the 3D structures revealed highly conserved core regions but significant variations in the external loop regions. The comparison between APIV1 and APIV2 showed the most pronounced variations, with a root-mean-square deviation (RMSD) of 8.60 Å and a TM-score of 0.33784 (Figure 4A). The comparison between APIV1 and APIV3 exhibited slightly less variation, with an RMSD of 7.72 Å and a TM-score of 0.34955 (Figure 4B). The comparison between APIV2 and APIV3 demonstrated relatively higher structural similarity, with an RMSD of 7.64 Å and a TM-score of 0.36407 (Figure 4C).

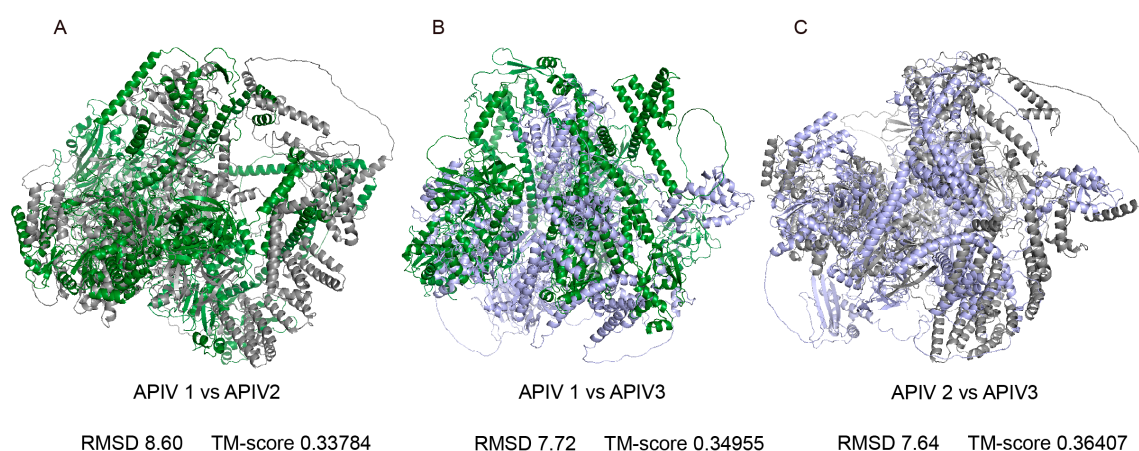


Figure 4. Structural comparison of APIV1, APIV2 and APIV3 proteins and highlighted conserved cores and divergent external loops. Superimposed structures of the three polyproteins: APIV1 (green), APIV2 (grey), and APIV3 (purple). ALPV1 is aligned with ALPV2 (A) ALPV1 is aligned with ALPV3 (B) ALPV2 is aligned with ALPV3 (C).

4. Discussion

Hemipteran insects are major agricultural pests that can transmit a variety of plant viruses, leading to significant economic losses in crop production [30]. Recent studies have reported the identification and discovery, through high-throughput sequencing, of many new viruses (insect-specific viruses, ISVs) in vectors, such as leafhoppers, aphids, and whiteflies, that do not cause disease in hosts and usually have no economic impact on plants [14]. However, despite the close taxonomic status of cabbage whiteflies, the viruses they carry have not yet been effectively investigated. Here, three novel iflaviruses were identified in *A. proletella* using meta-transcriptomics technology for the first time: APIV1, APIV2, and APIV3.

Viruses in the family *Iflaviridae* are positive-sense, single-stranded RNA viruses that are widely distributed across arthropods [31]. According to the International Committee on Taxonomy of Viruses (ICTV), the virus consists of 16 species (<https://ictv.global/report/chapter/iflaviridae/iflaviridae/iflavirus>, accessed on 10 February 2025). Known hosts now include insects belonging to the orders Orthoptera [32], Hymenoptera [33,34], Lepidoptera [35,36], Hemiptera [37], Odonata [17], and Diptera [38]. *Iflavivirus* is primarily transmitted horizontally (such as food contamination) and vertically (from mother to offspring); host infection may be asymptomatic or lead to severe pathology. Deformed wing virus (DWV)-affected bees emerged twice as slowly and had a 30% higher mortality rate

compared to clinically normal bees [39]. *Bombyx mori* Infectious flacherie virus (BmIFV) causes flacherie-like symptoms and is known to result in crop losses of up to 40% [40]. However, *Nilaparvata lugens* honeydew virus 1 (NLHV1) does not induce symptoms in the host [16]. Here, we identified three novel members, APIV1–3, in *A. proletella*, a species whose virome has not been fully explored before, suggesting that the diversity of cabbage whitefly viruses may be significantly underestimated. We used InterProScan to search for conserved domains in the amino acid sequence and identified multiple viral coat domains at the N-terminal end, as well as a helicase and RdRP domain at the C-terminal end (Figure 1). The closest relative to APIV1 in the phylogenetic tree is Bat faecal-associated iflavirus 2 (ON872543.1), which is found in bat faeces and may be associated with the bat's diet [41]. APIV2 was clustered with Novo Mesto iflavirus 1 (OL472192.1), the virus reported in the tomato and weed virome project [42], and APIV3 related to *Myrmica rubra* picorna-like virus 8 (MW314655.1), a virus found in ants [43]. Currently, the transmission pathway, infection mechanisms, and effects on the host species of APIV1–3 remain unclear and require further investigation.

This study employed AlphaFold3 to gain deeper insights into the structural characteristics of these viruses. Structural modelling revealed conserved core regions and highly variable external loops, with RMSD values ranging from 7.64 to 8.60 Å (Figure 4). These findings show some structural flexibility, which may play a crucial role in host adaptation and interaction [44]. The conserved core regions, characterized by high-confidence pLDDT scores and essential viral functions, such as replication and assembly, ensure the fundamental viability of the viruses [45]. In contrast, the variable regions, which exhibit greater structural flexibility, likely mediate host-specific interactions and immune evasion, enabling the viruses to spread and adapt to new hosts [46].

In summary, this study identified three new viruses in *A. proletella*. Based on phylogenetic analysis, they were classified into the family *Iflaviridae*. Our research contributes to a better understanding of the viral diversity in *A. proletella*, offering valuable insights into the coevolution between iflaviruses and their host. Further studies are essential to understanding the possible functions of these viruses and their impact on agricultural productivity.

5. Conclusions

This study is the first to use meta-transcriptomics to identify three novel viruses in the cabbage whitefly (*A. proletella*). By analysing the genomic structure and the evolutionary relationships of these viruses, we determined that the three viruses all belong to the family *Iflaviridae* and designated them as APIV1, APIV2, and APIV3. Further structural predictions of the polyprotein using AlphaFold3 revealed that while the overall structural framework of these viruses is conserved, significant variations exist, particularly in the external loop regions. Currently, the understanding of structural variability in *Iflavirus* remains limited. These findings provide valuable insights into the composition and function of the virome in the cabbage whitefly, as well as virus–host interactions.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/insects16040335/s1>, Figure S1: Multiple sequence alignment showing the conserved motifs in representative iflaviruses and APIV1, APIV 2 and APIV3. The viral Accession numbers and name are listed in Supplementary Table S2; Figure S2: Phylogenetic trees based on the polyprotein of the three iflaviruses identified in cabbage whiteflies. APIV1, APIV2 and APIV3 are highlighted in red font. Nodes with bootstrap values >50% are marked; Table S1: Primers used in this study; Table S2: Viral sequences selected for phylogenetic analysis and sequence alignments in this study; Table S3: Viral sequences identities in this study.

Author Contributions: Y.Z., J.-M.L., C.-X.Z. and J.-P.C. planned and designed the research. Z.-X.Y. and G.-W.G. analysed the data. Z.-X.Y. and P.-P.R. performed the experiments. Y.Z., Z.-X.Y., J.-M.L. and J.-P.C. drafted the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: We would like to express our gratitude to De-Yin Ma for generously providing the samples. This research was funded by the Ningbo Commonweal Project (2023S029), National Natural Science Foundation of China (U23A6006), Postdoctoral Fellowship Program of CPSF (GZC20231186).

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Materials. Further inquiries can be directed to the corresponding authors.

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

References

1. Ludwig, M.; Ludwig, H.; Conrad, C.; Dahms, T.; Meyhöfer, R. Cabbage Whiteflies Colonise *Brassica* Vegetables Primarily from Distant, Upwind Source Habitats. *Entomol. Exp. Appl.* **2019**, *167*, 713–721. [\[CrossRef\]](#)
2. De Barro, P.J.; Carver, M. Cabbage Whitefly, *Aleyrodes proletella* (L.) (Hemiptera: Aleyrodidae), Newly Discovered in Australia. *Aust. J. Entomol.* **1997**, *36*, 255–256. [\[CrossRef\]](#)
3. Martin, J.H.; Mifsud, D.; Rapisarda, C. The Whiteflies (Hemiptera: Aleyrodidae) of Europe and the Mediterranean Basin. *Bull. Entomol. Res.* **2000**, *90*, 407–448. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Springate, S. The Cabbage Whitefly, *Aleyrodes proletella*: Causes of Outbreaks and Potential Solutions. Ph.D. Thesis, University of Greenwich, London, UK, 2016.
5. Koca, A.S.; Kütük, H. Distribution, Host Plants and Natural Enemies of *Aleyrodes proletella* L. (Hemiptera: Aleyrodidae) on Collard (*Brassica oleracea* L. Var. Acephala) in Düzce Province of Turkey. *Bitki Koruma Bül.* **2020**, *60*, 17–24. [\[CrossRef\]](#)
6. Munang’andu, H.M.; Mugimba, K.K.; Byarugaba, D.K.; Mutoloki, S.; Evensen, Ø. Current Advances on Virus Discovery and Diagnostic Role of Viral Metagenomics in Aquatic Organisms. *Front. Microbiol.* **2018**, *8*, 406. [\[CrossRef\]](#)
7. Blomström, A.-L. Viral Metagenomics as an Emerging and Powerful Tool in Veterinary Medicine. *Vet. Q.* **2011**, *31*, 107–114. [\[CrossRef\]](#)
8. Shi, M.; Zhang, Y.-Z.; Holmes, E.C. Meta-Transcriptomics and the Evolutionary Biology of RNA Viruses. *Virus Res.* **2018**, *243*, 83–90. [\[CrossRef\]](#)
9. He, Y.-J.; Ye, Z.-X.; Zhang, C.-X.; Li, J.-M.; Chen, J.-P.; Lu, G. An RNA Virome Analysis of the Pink-Winged Grasshopper *Atractomorpha sinensis*. *Insects* **2023**, *14*, 9. [\[CrossRef\]](#)
10. Han, X.; Wang, H.; Wu, N.; Liu, W.-W.; Cao, M.-J.; Wang, X.-F. Leafhopper *Psammotettix alienus* Hosts Chuviruses with Different Genomic Structures. *Virus Res.* **2020**, *285*, 197992. [\[CrossRef\]](#)
11. Li, L.-L.; Ye, Z.-X.; Chen, J.-P.; Zhang, C.-X.; Huang, H.-J.; Li, J.-M. Characterization of Two Novel Insect-Specific Viruses Discovered in the Green Leafhopper, *Cicadella viridis*. *Insects* **2022**, *13*, 378. [\[CrossRef\]](#)
12. Mao, Q.-Z.; Ye, Z.-X.; Yuan, J.-N.; Ning, C.; Chen, M.-N.; Xu, Z.-T.; Qi, Y.-H.; Zhang, Y.; Li, T.; He, Y.-J.; et al. Diversity and Transmissibility of RNA Viruses in the Small Brown Planthopper, *Laodelphax striatellus*. *J. Virol.* **2024**, *98*, e00191-24. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Yuan, J.-N.; Ye, Z.-X.; Chen, M.-N.; Ren, P.-P.; Ning, C.; Sun, Z.-T.; Chen, J.-P.; Zhang, C.-X.; Li, J.-M.; Mao, Q.-Z. Identification and Characterization of Three Novel Solemo-like Viruses in the White-Backed Planthopper, *Sogatella furcifera*. *Insects* **2024**, *15*, 394. [\[CrossRef\]](#)
14. Huang, H.-J.; Ye, Z.-X.; Wang, X.; Yan, X.-T.; Zhang, Y.; He, Y.-J.; Qi, Y.-H.; Zhang, X.-D.; Zhuo, J.-C.; Lu, G.; et al. Diversity and Infectivity of the RNA Virome among Different Cryptic Species of an Agriculturally Important Insect Vector: Whitefly *Bemisia tabaci*. *npj Biofilms Microbiomes* **2021**, *7*, 43. [\[CrossRef\]](#)
15. Hou, X.; He, Y.; Fang, P.; Mei, S.-Q.; Xu, Z.; Wu, W.-C.; Tian, J.-H.; Zhang, S.; Zeng, Z.-Y.; Gou, Q.-Y.; et al. Using Artificial Intelligence to Document the Hidden RNA Virosphere. *Cell* **2024**, *187*, 6929–6942.e16. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Valles, S.M.; Chen, Y.; Firth, A.E.; Guérin, D.M.A.; Hashimoto, Y.; Herrero, S.; de Miranda, J.R.; Ryabov, E. ICTV Virus Taxonomy Profile: *Iflaviridae*. *J. Gen. Virol.* **2017**, *98*, 527–528. [\[CrossRef\]](#)
17. Zhang, Y.; Ye, Z.-X.; Feng, X.-X.; Xu, Z.-T.; Chen, J.-P.; Zhang, C.-X.; Li, J.-M. Prevalence of Reversed Genome Organizations for Viruses in the Family *Iflaviridae*, Order *Picornavirales*. *Microbiol. Spectr.* **2023**, *11*, e04738-22. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Kuwan, A.; Zhang, G.; Li, J.; Ma, D. Morphological, Genetic and Biological Differences in Two Populations of *Aleyrodes proletella* from Northwestern China. *Entomol. Gen.* **2021**, *41*, 27–37. [\[CrossRef\]](#)
19. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* **2014**, *30*, 2114–2120. [\[CrossRef\]](#)

20. Ratnasingham, S.; Hebert, P.D.N. Bold: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Mol. Ecol. Notes* **2007**, *7*, 355–364. [[CrossRef](#)]
21. Lu, G.; Ye, Z.-X.; He, Y.-J.; Zhang, Y.; Wang, X.; Huang, H.-J.; Zhuo, J.-C.; Sun, Z.-T.; Yan, F.; Chen, J.-P.; et al. Discovery of Two Novel Negevirus in a Dungfly Collected from the Arctic. *Viruses* **2020**, *12*, 692. [[CrossRef](#)]
22. Simmonds, P.; Adriaenssens, E.M.; Lefkowitz, E.J.; Oksanen, H.M.; Siddell, S.G.; Zerbini, F.M.; Alfenas-Zerbini, P.; Aylward, F.O.; Dempsey, D.M.; Dutilh, B.E.; et al. Changes to Virus Taxonomy and the ICTV Statutes Ratified by the International Committee on Taxonomy of Viruses (2024). *Arch. Virol.* **2024**, *169*, 236. [[CrossRef](#)] [[PubMed](#)]
23. Muhire, B.M.; Varsani, A.; Martin, D.P. SDT: A Virus Classification Tool Based on Pairwise Sequence Alignment and Identity Calculation. *PLoS ONE* **2014**, *9*, e108277. [[CrossRef](#)]
24. Katoh, K.; Standley, D.M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)]
25. Zhang, Y.; Skolnick, J. TM-Align: A Protein Structure Alignment Algorithm Based on the TM-Score. *Nucleic Acids Res.* **2005**, *33*, 2302–2309. [[CrossRef](#)] [[PubMed](#)]
26. Kozlov, A.M.; Darriba, D.; Flouri, T.; Morel, B.; Stamatakis, A. RAxML-NG: A Fast, Scalable, and User-Friendly Tool for Maximum Likelihood Phylogenetic Inference. *Bioinformatics* **2019**, *35*, 4453–4455. [[CrossRef](#)]
27. Gorbalenya, A.E.; Koonin, E.V. Viral Proteins Containing the Purine NTP-Binding Sequence Pattern. *Nucleic Acids Res.* **1989**, *17*, 8413–8438. [[CrossRef](#)]
28. Ryan, M.D.; Flint, M. Virus-Encoded Proteinases of the Picornavirus Super-Group. *J. Gen. Virol.* **1997**, *78*, 699–723. [[CrossRef](#)]
29. Koonin, E.V.; Dolja, V.V.; Morris, T.J. Evolution and Taxonomy of Positive-Strand RNA Viruses: Implications of Comparative Analysis of Amino Acid Sequences. *Crit. Rev. Biochem. Mol. Biol.* **1993**, *28*, 375–430. [[CrossRef](#)]
30. Ren, Y.-P.; Dong, W.-H.; Chen, J.-H.; Xue, H.-J.; Bu, W.-J. Identification and Function of microRNAs in Hemipteran Pests: A Review. *Insect Sci.* **2024**. [[CrossRef](#)]
31. Liu, S.; Chen, Y.; Bonning, B.C. RNA Virus Discovery in Insects. *Viruses* **2015**, *8*, 54–61. [[CrossRef](#)]
32. de Miranda, J.R.; Granberg, F.; Onorati, P.; Jansson, A.; Berggren, Å. Virus Prospecting in Crickets—Discovery and Strain Divergence of a Novel Iflavirus in Wild and Cultivated *Acheta domestica*. *Viruses* **2021**, *13*, 364. [[CrossRef](#)] [[PubMed](#)]
33. Dheilly, N.M.; Maure, F.; Ravallec, M.; Galinier, R.; Doyon, J.; Duval, D.; Leger, L.; Volkoff, A.-N.; Missé, D.; Nidelet, S.; et al. Who Is the Puppet Master? Replication of a Parasitic Wasp-Associated Virus Correlates with Host Behaviour Manipulation. *J. Gen. Virol.* **2015**, *282*, 20142773. [[CrossRef](#)]
34. Lanzi, G.; de Miranda, J.R.; Boniotti, M.B.; Cameron, C.E.; Lavazza, A.; Capucci, L.; Camazine, S.M.; Rossi, C. Molecular and Biological Characterization of Deformed Wing Virus of Honeybees (*Apis mellifera* L.). *J. Virol.* **2006**, *170*, 4998–5009. [[CrossRef](#)]
35. Wang, X.-C.; Zhang, J.-M.; Lu, J.; Yi, F.-M.; Liu, C.-F.; Hu, Y.-Y. Sequence Analysis and Genomic Organization of a New Insect Picorna-like Virus, *Ectropis obliqua* Picorna-like Virus, Isolated from *Ectropis obliqua*. *J. Gen. Virol.* **2004**, *17*, 1145–1151. [[CrossRef](#)]
36. Carballo, A.; Williams, T.; Murillo, R.; Caballero, P. Iflavirus Covert Infection Increases Susceptibility to Nucleopolyhedrovirus Disease in *Spodoptera exigua*. *Viruses* **2016**, *12*, 509. [[CrossRef](#)]
37. Ryabov, E.V. A Novel Virus Isolated from the Aphid *Brevicoryne brassicae* with Similarity to Hymenoptera Picorna-like Viruses. *J. Gen. Virol.* **2007**, *88*, 2590–2595. [[CrossRef](#)] [[PubMed](#)]
38. de Oliveira Guimarães, L.; Bahia, S.L.; de Oliveira Ribeiro, G.; do Socorro Foro Ramos, E.; Villanova, F.; dos Santos Moraes, V.; Telles-de-Deus, J.; Helfstein, V.C.; dos Santos, J.M.; Pandey, R.P.; et al. New Iflavirus Species Characterized from Mosquitoes Captured in the Sao Paulo Zoological Facilities. *Microorganisms* **2024**, *12*, 1749. [[CrossRef](#)]
39. Kozii, R.V.; Wood, S.C.; Kozii, I.V.; van Rensburg, C.J.; Moshynskyy, I.; Dvilyuk, I.; Simko, E. Deformed Wing Virus Infection in Honey Bees (*Apis mellifera* L.). *Virology* **2023**, *56*, 636–641. [[CrossRef](#)]
40. Savanurmath, C.J.; Basavarajappa, S.; Hinchigeri, S.B.; Ingalthalli, S.S.; Singh, K.K.; Sanakal, R.D. Relative Incidence of the Silkworm Viral Diseases in Agroclimatic Zones of Northern Karnataka, India. *Bull. Sericultural Res.* **1994**, *5*, 51–55.
41. Van Brussel, K.; Mahar, J.E.; Ortiz-Baez, A.S.; Carrai, M.; Spielman, D.; Boardman, W.S.J.; Baker, M.L.; Beatty, J.A.; Geoghegan, J.L.; Barrs, V.R.; et al. Faecal Virome of the Australian Grey-Headed Flying Fox from Urban/Suburban Environments Contains Novel Coronaviruses, Retroviruses and Sapoviruses. *Virology* **2022**, *576*, 42–51. [[CrossRef](#)]
42. Rivarez, M.P.S.; Pecman, A.; Bačnik, K.; Maksimović, O.; Vučurović, A.; Seljak, G.; Mehle, N.; Gutiérrez-Aguirre, I.; Ravnikar, M.; Kutnjak, D. In-Depth Study of Tomato and Weed Viromes Reveals Undiscovered Plant Virus Diversity in an Agroecosystem. *Microbiome* **2023**, *11*, 60. [[CrossRef](#)] [[PubMed](#)]
43. Viljakainen, L.; Fürst, M.A.; Grasse, A.V.; Jurvansuu, J.; Oh, J.; Tolonen, L.; Eder, T.; Rattei, T.; Cremer, S. Antiviral Immune Response Reveals Host-Specific Virus Infections in Natural Ant Populations. *Front. Microbiol.* **2023**, *14*, 1119002. [[CrossRef](#)] [[PubMed](#)]
44. Zhao, H.-Y.; Nguyen, A.; Wu, D.; Li, Y.; Hassan, S.A.; Chen, J.-J.; Shroff, H.; Piszczek, G.; Schuck, P. Plasticity in Structure and Assembly of SARS-CoV-2 Nucleocapsid Protein. *PNAS Nexus* **2022**, *1*, pgac049. [[CrossRef](#)]

45. Das, N.K.; Vogt, J.; Patel, A.; Banna, H.A.; Koirala, D. Structural Basis for a Highly Conserved RNA-Mediated Enteroviral Genome Replication. *Nucleic Acids Res.* **2024**, *52*, 11218–11233. [[CrossRef](#)]
46. Montoya, V.; McLaughlin, A.; Mordecai, G.J.; Miller, R.L.; Joy, J.B. Variable Routes to Genomic and Host Adaptation among Coronaviruses. *J. Evol. Biol.* **2021**, *34*, 924–936. [[CrossRef](#)] [[PubMed](#)]

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.