

Genetic characterisation of an Iflavirus associated with a vomiting disease in the Indian tropical tasar silkworm, *Antheraea mylitta*

Kangayam M. Ponnuel^a, Joachim R. de Miranda^b, Olle Terenius^{c,*}, Wenli Li^d, Katsuhiko Ito^e, Diksha Khajje^a, G Shamitha^f, Anupama Jagadish^a, Himanshu Dubey^a, Rakesh K. Mishra^a

^a Genomic Division, Seri-biotech Research Laboratory, Carmelaram Post, Kodathi, Bengaluru 560035, India

^b Department of Ecology, Swedish University of Agricultural Sciences, Box 7044, Uppsala 750 07, Sweden.

^c Department of Cell and Molecular Biology, Uppsala University, Box 596, Uppsala SE-751 23, Sweden

^d School of Life Science and Medicine, Dalian University of Technology, Dagong Road No.2, Liao Dong Wan New District, Panjin 124221, China

^e Department of Science of Biological Production, Graduate School of Agriculture, Tokyo University of Agriculture and Technology, Fuchu, Tokyo 183-8509, Japan

^f Department of Zoology, Kakatiya University, Warangal 506 009, India

ARTICLE INFO

Keywords:

Viral disease
Iflavirus
Antheraea mylitta
ORF
Genome

ABSTRACT

Antheraea mylitta, the Tropical tasar silkworm, is frequently affected by a vomiting disease called Virosis by sericulturists although not confirmed being of viral origin. Based on the symptoms and the disease pattern, the causal agent is however suspected to be a virus. The condition involves a series of characteristic and progressive symptoms that generally culminates in the death of the larva. The disease is common in autumn season (Sep-Oct), with widespread distribution causing severe damage to the tasar silk industry. The leads for this study were obtained from a transcript identified in the EST database in a different study, which matched the positive strand of Iflavirus, an RNA virus known to infect insects. In the present study the genome of this novel Iflavirus was characterised and the full length of the genome was found to be 9728 nucleotides long encoding for a single large open reading frame (ORF) with flanking NTR regions at 5' and 3' ends and a natural poly A tail at the 3' end. The ORF encoded structural proteins at the N-terminal end and non-structural proteins at the C-terminal end with a predicted 2967 amino acid long polyprotein. The structural proteins consisted of 4 proteins (VP1-VP4) and the non-structural proteins consisted of helicase, RNA-dependent RNA polymerase and 3C-protease. The virus is found in various tissues (midgut, fatbody, trachea, Malpighian tubules and silk gland) and also has a vertical route of transmission, i.e., from gravid females to the offspring. Based on the available data, the virus is a new member of Iflaviridae for which we propose the name *Antheraea mylitta* Iflavirus (AmIV).

1. Introduction

Antheraea mylitta, the tropical tasar silkworm of India, is commonly affected by a disease characterised by a shrinking body of the infected larvae, generally occurring in the 5th instar, close to the cocooning stage of the life cycle (Fig. 1AB). The infected silkworms show signs of poor eating and become sluggish. Due to the disease, they also lose the ability to hold onto twigs and drop down to the ground. The disease occurs throughout the year, especially in rainy seasons (Jun - Sep) and autumn (Sep - Oct) months. The disease was previously suspected to spread mainly through the droppings of other infected silkworms, disintegrated silkworms, and pathogen-contaminated rearing sites and appliances. High temperature, poor leaf quality of the feed, and high humidity is

found to aggravate this disease. As the infection progresses, the silkworms show retarded growth, lethargy and refusal to feed. Eventually the larvae vomit their gastric contents, which is followed by death. Based on the symptoms, the disease is known among sericulturists under different names: Virosis, Vomiting disease and Anal lip sealing disease. The observed symptoms were noted for a long time at the field level; however, the disease etiology and mode of transmission was not known. It was also noted that the dead worms did not putrefy, as is the case with bacterial diseases, which suggests that the disease is most likely caused by a virus (Chen et al., 2012). Similar symptoms have been observed in *Antheraea pernyi* (Chinese oak silk moth), where it was shown that the disease is caused by an Iflavirus (Geng et al., 2014). We therefore hypothesised that the symptoms observed in *A. mylitta* may be similarly

* Corresponding author.

E-mail address: olle.terenius@icm.uu.se (O. Terenius).

<https://doi.org/10.1016/j.virusres.2022.198703>

Received 9 June 2021; Received in revised form 28 August 2021; Accepted 28 January 2022

Available online 30 January 2022

0168-1702/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

associated with a viral infection, possibly also an Iflavirus.

Iflaviruses are small icosahedral viruses with a monopartite RNA genome of about 10 kilobases in size that are classified in Family *Iflaviridae* within the Order Picornavirales (Valles et al., 2017). They are a very common and widely distributed group of insect-specific viruses with often a dominant oral-fecal transmission route and an intestinal infection profile (van Oers 2010; Valles et al. 2017). Iflaviruses are well represented among the known insect viral diseases, particularly gastrointestinal diseases in insects with high local densities that

facilitate transmission, such as social Hymenoptera and lepidopteran insects (van Oers 2010; de Miranda et al. 2021), or insects that are reared commercially, e.g., in apiculture, sericulture or for food-feed (Maciel-Vergara and Ros, 2017; de Miranda et al., 2021). Iflaviruses have a monopartite genome coding for a single large open reading frame (ORF) (Valles et al., 2017). The ORF encodes for a leader protein followed by a cassette of four capsid proteins, followed by several non-structural proteins including (in order) the helicase, 3C-protease and RNA-dependent RNA polymerase (*RdRp*). The order of the various

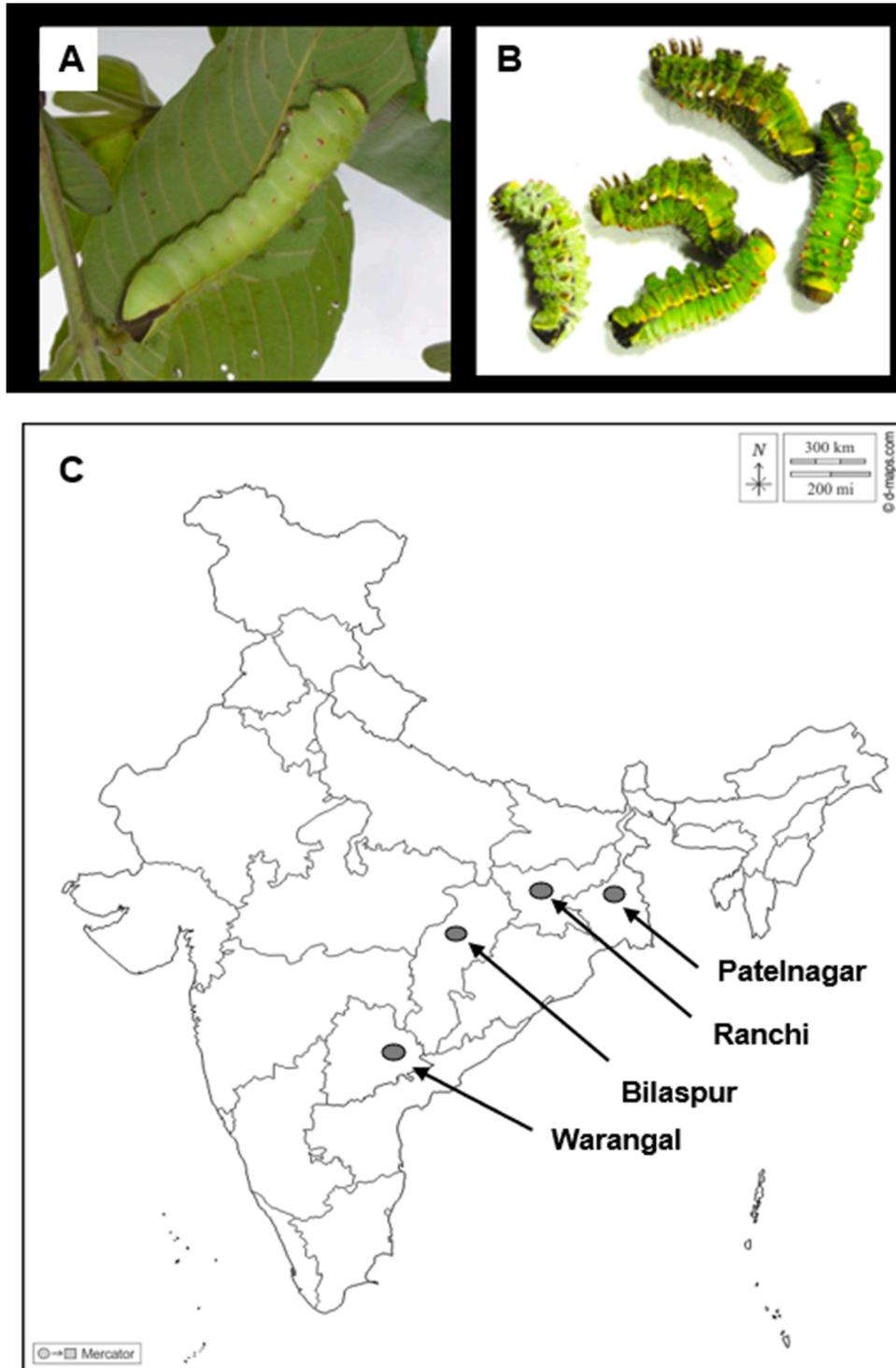


Fig. 1. Healthy and diseased *A. mylitta* larva. **A:** Healthy larva, **B:** Diseased larvae (shortened in body length). **C:** *A. mylitta* silkworm samples obtained from different locations within India. (Map obtained from https://d-maps.com/carte.php?num_car=279&lang=en)

peptides in the polyprotein is highly characteristic for the Iflaviridae (Valles et al., 2017). The genomic RNA of Iflaviruses is translated into a polyprotein that is subsequently processed by autolytic cleavage and by the virus-encoded 3C-protease into functional proteins for virus replication and particle construction (King et al., 2012; Procházková et al., 2020). The capsid proteins combine to form the virus capsid protomer, with 12 pentameric units of five protomers each encapsulating a single copy of genomic RNA to form the icosahedral virus particle (Procházková et al., 2020). The various capsid protein domains have specific roles in the structural integrity of the virus particle and in the infection process, while the conserved domains in the non-structural components of the Iflavirus genome, such as the helicase, 3C-protease and *RdRp* are important in the replication, translation and processing of the virus genome and polyprotein (Quevillon et al., 2005).

In this study, we present the genome sequence of a novel Iflavirus; *Antheraea mylitta* Iflavirus (AmIV) obtained from diseased *A. mylitta* silkworms, and provide evidence of its causative link to the disease symptoms and pathology described above, as well as its genomic and phylogenetic characterization.

2. Materials and methods

2.1. Initial discovery

During the screening of an *A. mylitta* EST database (<http://www.cdfb.org.in/wildsilkbases/home.php>), a singular transcript (Amfb1024) was identified (Arunkumar et al., 2008) with high similarity to Deformed wing virus: a well-known Iflavirus of honeybees and other pollinators (Beaurepaire et al., 2020). The transcript encoded a part of an *RdRp* gene, which is extremely conserved across all viruses with an RNA genome (Venkataraman et al., 2018), and thus highly diagnostic for the putative presence of an RNA virus, and likely an Iflavirus.

2.2. Sample collection

Cocoons were obtained from silkworm seed production centres, which in turn had purchased the cocoons from tribal communities in areas where tasar sericulture is practiced (Fig. 1C). Eggs obtained from mating of eclosed adults were washed thoroughly with 0.2% sodium hypochlorite solution to remove any surface contaminants. After this, the eggs were incubated at $28\text{ }^{\circ}\text{C} \pm 2\text{ }^{\circ}\text{C}$ for the emergence of larvae. The emerged larvae were screened for the presence of AmIV using RT-PCR-based assays based on the AmIV *RdRp* gene identified through screening the *A. mylitta* EST database, as described below.

2.3. RNA extraction and cDNA synthesis

RNA was extracted from experimental and field samples using an RNAiso Plus kit (TaKaRa, India). The tissue samples (100 mg) were crushed in the RNAisoplus mixture using a micro pestle followed by chloroform extraction to remove the cell debris and finally precipitation of RNA using isopropyl alcohol. All the stages involved centrifugation at 17,500 g for 8 min at $4\text{ }^{\circ}\text{C}$. The RNA concentration was checked using a NanoDrop™ 2000/2000c Spectrophotometer (Thermo Fisher Scientific). cDNA was synthesized from the RNA using the PrimeScript 1st strand cDNA synthesis kit (TaKaRa, India) following the manufacturer's recommended protocol. The RNA was treated with DNase prior to synthesis of cDNA using DNase I Amplification grade (Thermo Fischer Scientific, USA) based on the manufacturer's instructions.

2.4. 3' RACE

The transcript sequence retrieved from the EST database was used to design a 3' RACE assay in order to (1) confirm the presence of AmIV in different *A. mylitta* samples, (2) to confirm the sequence and initial identity of the AmIV *RdRp* gene, and (3) to obtain a longer sequence for

designing a diagnostic RT-PCR assay (Supplementary Fig. 1). 3' RACE was chosen since the Iflavirus *RdRp* is located towards the 3' end of the genome (Valles et al., 2017). 3' RACE was performed on RNA extracted from *A. mylitta* samples displaying vomiting disease symptoms, using an anchored oligo-dT primer and a selection of forward primers targeting 3' end of the viral genome (Table 1, Supplementary Fig. 1) using a 3'RACE kit (Invitrogen, Bangalore, India) following the manufacturer's instructions. The amplification products were sequenced by Sanger sequencing (Eurofins India Pvt. Ltd, Bangalore, India) and the sequences were manually assembled.

2.5. RT-PCR assays

A Reverse Transcriptase - Polymerase Chain Reaction (RT-PCR) assay was designed with primers located in the AmIV *RdRp* gene identified through the EST database screening and completed by the 3' RACE analyses. The PCR was conducted in 10 μl volumes containing 1 μl cDNA, 100 nM of each forward and reverse primers (Table 1) and 5 μl of 2X EmeraldAmp® GT PCR master mix (TaKaRa, India). The samples were subjected to the following thermal profile for amplification: 2 min at $94\text{ }^{\circ}\text{C}$, followed by 30 cycles of [30 s at $94\text{ }^{\circ}\text{C}$, 30 s at $57\text{ }^{\circ}\text{C}$ and 1 min at $72\text{ }^{\circ}\text{C}$] in a BioRad PTC 200 Thermal cycler (BioRad, Bangalore, India). The PCR products were resolved on a 1.2% agarose gel stained with ethidium bromide for visualization of the amplified product.

2.6. Real-time quantitative RT-qPCR

The quantitative Real Time PCR reaction mix consisted of 1 μl each of forward and reverse primers (conc. 100 nM), 5 μl of SYBR green master mix (TaKaRa, India), 0.1 μl of ROX (Carboxyrhodamine) and the volume made up to 10 μl with water after accounting for the cDNA addition of 1 μl . The qPCR was performed on a Stratagene Mx3005P thermocycler (Agilent Technologies, Germany) with the following cycling parameters: $95\text{ }^{\circ}\text{C}$ for 2 min followed by 40 cycles of [$95\text{ }^{\circ}\text{C}$ for 30 s, $55\text{ }^{\circ}\text{C}$ for 30 s and $72\text{ }^{\circ}\text{C}$ for 35 s]. Three technical replicates were run for each assay. Non-template control (water) without the template was also included in all the assays. A standard curve was generated by performing serial dilutions of the cloned fragment of the *RdRp* domain to quantify the copy number in the unknown samples. The cycle threshold values obtained from the analysis were used in determination of copy numbers.

2.7. Virus purification

The fatbody of naturally AmIV-infected *A. mylitta* silkworms was homogenised in PBS buffer with 2 g of tissue. The resulting homogenate was filtered through 2 layers of cheese cloth and subjected to centrifugation at 100 g for 2 min. The crude homogenate was used as a virus inoculum for the infection studies. The amount of AmIV in the inoculum was determined by RT-qPCR, using the assays described above.

2.8. Virus inoculation

A crude homogenate of AmIV was prepared in PBS buffer (pH 7.0) obtained from naturally infected, symptomatic (as described in the introduction) *A. mylitta* samples, with a concentration of 1×10^8 AmIV genome equivalents/ml. Thirty newly ecdysed 4th instar *A. mylitta* larvae (five groups of six larvae each for each time point) were injected ventrally with 50 μl crude homogenate (1×10^5 AmIV genome equivalents) using a 25 G needle. Similarly, a control group of 36 4th instar larvae were injected with 50 μl PBS (pH 7.0). At 0 h, six larvae from the control group were checked to ensure that the larvae were free from infection (Eberle et al., 2012). The larvae were reared under natural conditions on an Arjun tree (*Terminalia arjuna*) covered with nylon nets and were checked at every 6 h for growth and symptom development (loss of appetite, shrunken body, losing grip on the twigs, lethargy). Six larvae each from the inoculated and control groups were sacrificed at 6

Table 1
Primers for 3' RACE and PCR.

Target	Primer	Sequence (5'3')	Size(bp)
3' RACE	R3 Forward	CTTGTGGAATCCCTTCAGGA	115
3' RACE	R1 Reverse	GCTGGCGTTCTGGATCGTTA	
3' RACE	R4 Forward	GCCAGCTAATTCGTCTGTGT	767
3' RACE	R2 Reverse	TATCCCATGTTTCAGCATCCA	
3' RACE	R5 (oligo T primer with adapter sequence)	GACCACGCGTATCGATGTCGAACCTTTTITTTTTTTTTTITV	NA
3' RACE	R6 (specific to the adapter sequence in R5)	GACCACGCGTATCGATGTCGAC	
3' RACE	R7 Forward	GAAGCATTGCGAAAAGAAAGG	631
3' RACE	R8 Reverse	ATACACACATCGTCGCCGTA	
AmIV RdRp	R7 Forward	GAAGCATTGCGAAAAGAAAGG	623
	R8 Reverse	ATACACACATCGTCGCCGTA	
AmIV RdRp (qPCR)	qPCR Forward	CTTGTGGAATCCCTTCAGGA	157
	qPCR Reverse	ATACACACATCGTCGCCGTA	
<i>A. mylitta</i> β -actin	Forward	ACCAACTGGGACGACATGGAGAAA	124
	Reverse	TCTCTCTGTTGGCCTTGGGTTGA	
<i>Nosema</i> β -tubulin	Forward	CTTTGGACAATCTGGTGCTG	181
	Reverse	GAGAAGGGTTCCCATTCCTG	
CPV	Forward	GCAAGTACATGCCCGTAGGT	511
	Reverse	TGCCATAAGTGCTTGACAGC	

h, 12 h, 24 h, 36 h, 48 h post inoculation for pathogen analyses. The fatbody tissues from these samples were dissected, weighed and RNA was extracted from 100 mg of pooled fatbody tissue, followed by cDNA synthesis and qPCR analysis.

2.9. Tissue tropism

AmIV is a novel virus whose tissue tropism is not known. We used the RT-PCR assays described above to check for the presence of AmIV in dissected fatbody, midgut, Malpighian tubules, silk glands and trachea from naturally infected *A. mylitta* silkworm field samples. The RNA was extracted from these tissues, cDNA was synthesised and amplified by qPCR as described above.

2.10. Vertical transmission

Vertical transmission has been reported for several Iflaviruses infecting a range of insect hosts, including lepidopterans (Geng et al., 2017; Yuan et al., 2017; Yanez et al., 2020; Valles et al., 2017) and the same was tested in this study. Cocoons were collected from the silkworm egg production centres and male and female adult moths were allowed to emerge from the cocoons followed by mating for 8–12 h. After mating, the female moths were allowed to oviposit for a period of 24–48 h in a dark chamber. The laid eggs were incubated at 25C with 80% relative humidity for a period of 10 days. The mother moth, eggs and newly hatched larvae were screened for the presence of virus. The virus positive moth's egg samples were separated and washed with 0.2% sodium hypochlorite to remove any surface contaminants. The newly hatched larvae from these fertilized eggs along with the eggs and the infected female moth samples were tested for the presence of the virus through RT-PCR. Egg samples from un-infected mother moths were also included in the analysis. The β -actin gene was used as an internal control for confirming the integrity of the extracted RNA samples.

2.11. Co-infection of AmIV and pébrine disease

It has been suggested that *Spodoptera exigua* Iflavirus survives and, in a co-infected state in the host along with the baculovirus *Spodoptera exigua* nucleopolyhedrovirus, establishes a synergistic effect beneficial for the survival of the co-infecting *S. exigua* Iflavirus (Jakubowska et al., 2016). In this study, two sets of samples were used for co-infection analysis. The first set of samples concerned various tissues dissected from covertly infected adult moths collected from Warangal, which were tested by RT-PCR for the presence of AmIV and *Nosema bombycis* (Esvaran et al., 2019).

The second set of samples concerned dissected tissues from diseased

silkworm larvae field samples with clear symptoms of vomiting disease, collected from Karimnagar, Mahadevpur and Warangal. These samples were tested by RT-PCR for the presence of AmIV and *A. mylitta* cytoplasmic polyhedrosis reovirus (targeting segment 9) (Qanungo et al., 2002) and for the presence of *Nosema bombycis*.

2.12. Completing the AmIV genome sequence

Total RNA was isolated from an *A. mylitta* fatbody infected with AmIV using TRizol reagent (Thermo Fisher Scientific, Carlsbad, CA, USA) and RNeasy Mini Kit (Qiagen, Hilden, Germany). Whole-transcriptome shotgun sequencing (RNA-seq) was outsourced to Macrogen Japan (<https://www.macrogen-japan.co.jp/>). RNA-seq libraries were generated from total RNAs using the TruSeq RNA Sample Preparation Kit (version 2) (Illumina, San Diego, CA, USA) and 100 bp paired-end sequencing was performed using the HiSeq 2000 Sequencing System (Illumina) in accordance with the manufacturer's instructions. Quality analysis of the obtained raw reads was performed by the FastQC toolkit (Leggett et al., 2013). Low-quality reads and adapter contamination were removed by Trimmomatic (Bolger et al., 2014). The resultant high-quality reads were assembled by Trinity RNA-seq de-novo assembler (ver. R20140717) into contigs (Grabherr et al., 2011), which were subsequently combined with the EST data, and the 3'RACE sequences into a near-complete nucleotide sequence of the AmIV genome, which was deposited in GenBank under the accession number MW115117.

2.13. Genome organization and analysis

The nucleotide sequence was used in deciphering the information of the genome organisation with regards to the functional domains and putative proteolytic cleavage sites. The nucleotide sequence derived from sequencing was used in the ExPASy translate tool (Gasteiger et al., 2003) to derive the information of the amino acid sequence of the ORF. The obtained sequence data was annotated for structural- and non-structural regions in the genome using conserved domains identified by InterProScan search (Quevillon et al., 2005), 3C-protease cleavage sites through Geng et al. (2014), and conserved residues reported elsewhere (Murakami et al., 2013) within these structural- and non-structural regions.

Conserved residues identified among closely related viruses was used in multiple sequence alignment by ExPASy BoxShade (https://embnet.vital-it.ch/software/BOX_form.html) for both structural- and non-structural protein regions between other viruses from the order Picornavirales (Supplementary Table 1) (Murakami et al., 2013). Conserved motifs between AmIV and *Antheraea pernyi* Iflavirus (ApIV; Geng et al., 2014) were analysed using MEME suite with default parameters (Bailey

et al., 2015).

In order to derive the amino acid sequence identity and similarity between selected viruses in the order Picornavirales (Supplementary Table 1), the Pal2Nal tool (Suyama et al., 2006) was used to derive the identity and similarity index. The amino acid substitution rates for synonymous- and non-synonymous substitution were also calculated between AmIV and ApIV by codeml in the PAML package using the Pal2Nal tool in order to derive the dN/dS ratio to understand the significance of the substitution rates.

2.14. Phylogenetic analysis

The polyprotein amino acid sequences of AmIV were aligned to a selection of 34 related Iflaviruses (Supplementary Table 1) by CLUSTAL-Omega at the EMBL-EBI bioinformatics platform at www.ebi.ac.uk, using the default gap creation and extension penalties. The phylogenetic analyses were conducted in MEGA-X (Kumar et al. 2018). The complete amino acid sequence of the Iflavirus polyprotein was used for phylogenetic analysis. All positions containing gaps or missing data in the multiple alignment were excluded from the phylogenetic analyses, resulting in a total of 1100 characters (amino acids) for the main analysis. A more detailed analysis was also conducted for seven taxa on the internal clade where AmIV is located, using Darwin bee virus 3 as a local outgroup. These analyses involved a total of 2119 characters, separated into the Lp region (117 characters), the VP region (876 characters), the Helicase region (832 characters) and the 3C-protease/RdRp region (294

characters). The phylogenetic relationship between the different taxa was inferred using the Maximum Likelihood method and JTT matrix-based model (Jones et al., 1992). Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Joining and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model, and then selecting the topology with superior log likelihood value. The phylogenetic tree with the highest likelihood was retained. The statistical confidence for each branching node was determined by bootstrapping the alignment 1000 times, to calculate the percentage of bootstrapped trees retaining the taxa clustered by each node in the most likely tree.

3. Results

3.1. AmIV genome organisation

The novel Iflavirus possesses a positive strand RNA of at least 9728 nucleotides that codes for a single 2967 amino acid polyprotein flanked on the 5' and 3' ends by Non-Translated Regions (NTR) and includes a natural poly-A tail at the 3' terminus. The AmIV 5' NTR was 435 nucleotides shorter than the 5' NTR determined for ApIV (Geng et al., 2014), which most likely means that the AmIV sequence is incomplete at the 5' end. The AmIV genome is A/U rich (68.8%), which is comparable to other Iflaviruses (Fujiyuki et al., 2004; Lanzi et al., 2006; Ongus et al., 2004; Ghosh et al., 1999; Isawa et al., 1998). The structural proteins are located towards the N-terminus of the polyprotein, from where they are

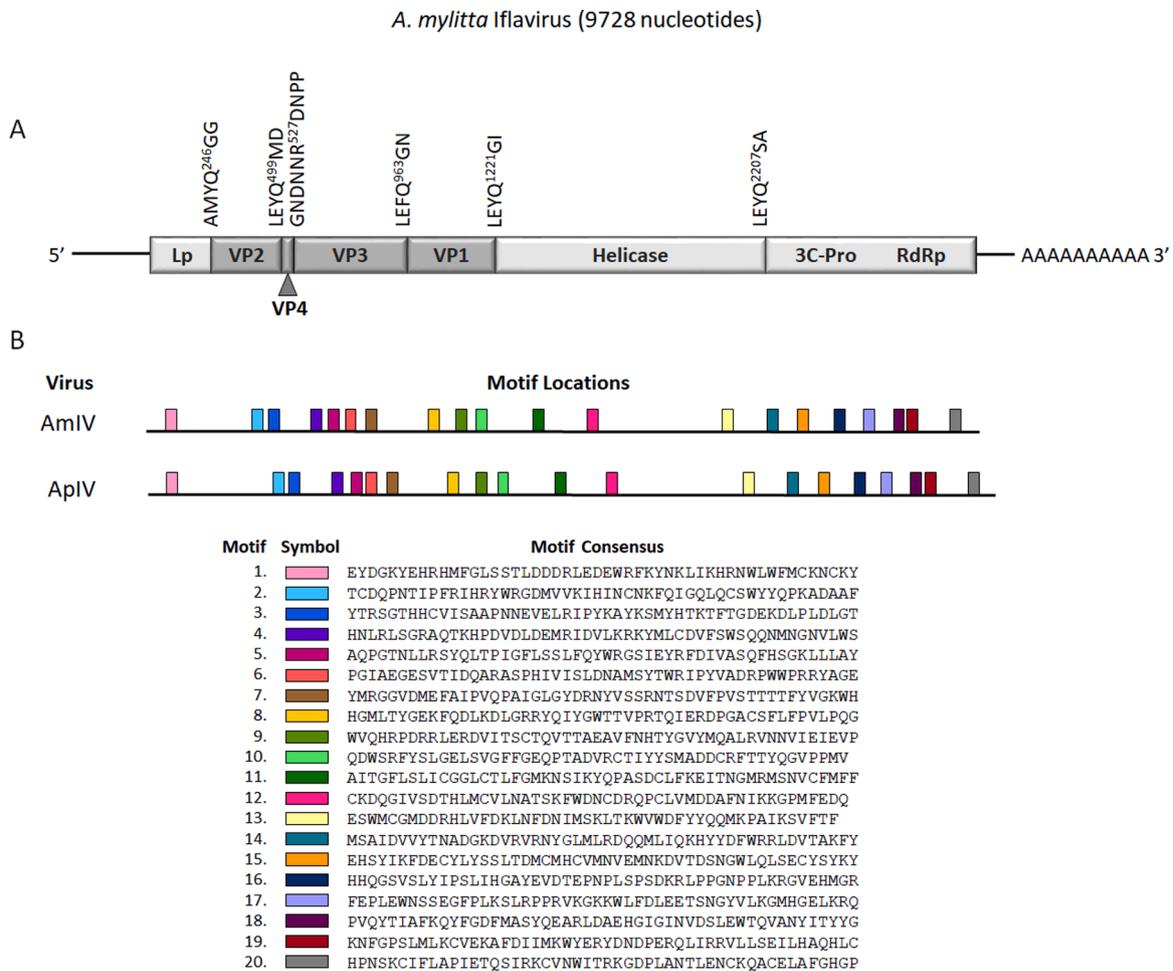


Fig. 2. A: Genome organization of AmIV, showing the 5' and 3' NTR, poly-A tail and the single ORF encoding for a polyprotein. The location of the Leader peptide (Lp), the structural proteins (VP1-VP4), the Helicase, the 3C-protease and the RNA-dependent RNA polymerase are also shown, as well as the putative protease processing sites separating the functional peptides from the polyprotein. B: Location of twenty protein motifs that are highly conserved between AmIV and ApIV.

processed proteolytically into the four capsid proteins (order VP2, VP4, VP3 and VP1; Fig. 2A) that form the basic protein unit for constructing the Iflavirus particle (Procházková et al., 2020; Ghosh et al., 1999; Isawa et al., 1998). The non-structural proteins are located in the C-terminal half of the polyprotein and are necessary for processing the polyprotein into functional units (3C-protease) and replicating the RNA genome (Helicase, RNA-dependent RNA polymerase: Fig. 2A). At the N-terminus of the polyprotein is located a 246 amino acid Leader protein (Lp: Fig. 2A) thought to be involved in infectivity and virulence (de Miranda and Genersch, 2010). The putative autocatalytic and 3C-protease cleavage sites where the AmIV polyprotein is processed into functional units (de Miranda et al., 2010a; de Miranda et al., 2010b; Geng et al., 2014) were also identified (Fig. 2A).

3.2. Conserved domains and amino acid residues

InterProScan revealed five major domains in the polyprotein that are strongly conserved across many RNA virus families whose functions are well supported by biochemical and virus structural studies (Procházková et al., 2020; Guan et al., 2017; Venkataraman et al., 2018). The domains include two rhinovirus-like capsid protein domains between amino acids 300–485 and 580–773, a conserved cricket paralysis virus-like capsid protein domain between amino acids 982–1203, an RNA helicase domain between amino acids 1536–1644, and a series of well-defined RdRp domains between amino acids 2538–2941 (Venkataraman et al., 2018). Additionally, twenty motifs that are highly conserved between AmIV and ApIV were also identified across all the structural- and non-structural regions of the genome (Fig. 2B).

Each of the structural- and non-structural functional proteins of the deduced AmIV amino acid sequence was compared with homologous proteins from related viruses from the order Picornavirales (Supplementary Table 1), in order to identify highly conserved amino acids within the functional domains identified above (Forslund et al., 2011). The conserved NxNxFAQG VP2 domain was identified in all the picornaviruses included in the analysis, with both Ns and the G residues universally conserved (Supplementary Fig. 2A). In the 27-amino-acid long conserved motif WxGxLxxxFxGxxxxxxGxxxxYxP of the VP3 domain, conserved amino acids were found in four positions in the alignment: the 2nd (R), the 3rd (G), the 21st (L) and the 27th (P), thus only two of the eight presumed conserved amino acids (G and P) were found to be conserved (Supplementary Fig. 2A). Similarly, in the FxRG domain of VP1, only the residue G was found to be conserved. The conserved regions from the non-structural protein region also displayed conserved motifs with few conserved amino acid residues for the Helicase A, B and C subdomains, the 3C protease subdomains and the RdRp subdomains (Supplementary Fig. 2B).

3.3. Identity/similarity index and dN/dS ratio

The identity and similarity index between the viruses selected from the Picornavirales is displayed a variable matrix (Supplementary Table 2). ApIV (Geng et al., 2014) is most closely related to AmIV, with an identity index of 80.51% and similarity index of 86.95% (Supplementary Table 2).

The non-synonymous nucleotide substitution rate (dN/dS) between AmIV and ApIV was calculated to be 0.0173. A dN/dS ratio of <1.0 is indicative of stabilizing selection (Kryazhimskiy and Plotkin, 2008), acting against evolutionary change between AmIV and ApIV in the polyprotein amino acid sequence.

3.4. Phylogenetic analysis

The phylogenetic analysis revealed that AmIV is most closely related to ApIV and firmly embedded within the Iflaviridae phylogenetic tree, on a well-separated branch within clade-I together with several other lepidopteran iflaviruses identified by RNA sequencing from the

Mulberry silkworm *Bombyx mori*, the Pine processionary moth *Thaumetopoea pityocampa*, the Gypsy moth *Lymantria dispar*, the Red postman butterfly *Heliconius erato* and the Cotton bollworm *Helicoverpa armigera* (Fig. 3, Supplementary Table 1). There is excellent support for the clade as a whole and the clustering of AmIV with ApIV, but the relationships between the remaining taxa are less clear. The uncertainty concerns primarily whether *Thaumetopoea pityocampa* Iflavirus clusters with AmIV and ApIV (supported by the structural proteins) or with *Bombyx mori* Iflavirus (supported by the non-structural proteins). However, the *Thaumetopoea pityocampa* Iflavirus Lp gene clusters with *Heliconius erato* Iflavirus, well away from AmIV and ApIV, resulting in a very weakly supported short branch clustering *Bombyx mori* Iflavirus with AmIV and ApIV in the global phylogeny. The relative positions of *Lymantria dispar* Iflavirus-1, *Helicoverpa armigera* Iflavirus and *Heliconius erato* Iflavirus within the clade are relatively consistent across the genome (except, again, for the Lp gene) and thus well supported. There are several other weakly supported branches elsewhere in Clade-I, which are also generally the result of different associations between the viruses depending on which genomic region that is analysed.

3.5. AmIV incidence in wild adult *A. mylitta* moths

A total of 50 wild adult *A. mylitta* silkworm samples (25 males + 25 females) were screened for the presence of AmIV by end-point RT-PCR. Fifteen male silkworms (60%) and 17 female silkworms (68%) were found to be positive for AmIV, with little difference in incidence between the four sampling sites (Table 2).

3.6. Inoculation and virus propagation

From the infection of AmIV in the 4th instar larvae, a steady increase of AmIV was observed starting 6 h increasing up to 48 h post injection (Fig. 4). The larvae developed the typical infection symptoms described in the introduction.

3.7. Tissue tropism and vertical transmission

The presence of amplification bands in various tissues obtained from the field samples such as fatbody, midgut, Malpighian tubules, silk glands and trachea indicated the virus presence, displaying ubiquitous presence of the virus in different tissues once the host gets infected (Fig. 5A). An RT-PCR analysis also revealed that the presence of the virus was higher in the fatbody followed by Malpighian tubules, silk gland and midgut. Trachea displayed the lowest copy number as compared with other tissues (Fig. 5B).

Amplification of bands confirmed the presence as well as transfer of the virus from the infected mother moth to offspring (Fig. 5C). The absence of the amplification band in the eggs laid by an uninfected mother moth further confirms that the vertical transmission is of transovarial nature.

3.8. Co-infection

Sixty-four percent (32/50) adult *A. mylitta* silkworm samples were found by RT-PCR to be covertly infected with AmIV with equal infection rates in male (15/25) and female (17/25) silkworms. However, only 9 (18%) of the 32 samples that were positive for AmIV were also positive for *Nosema* sp. (Table 3). On the contrary the diseased larval silkworm samples, with symptoms of vomiting disease, were 100% positive for AmIV. Of these, 51 (68%) were infected with only AmIV, while 23 (30.7%) were positive for both AmIV and *Nosema* sp.. A single sample infected with both AmIV and *Nosema* (i.e. 1/75, or 1.3%) was also positive for CPV (Table 4).

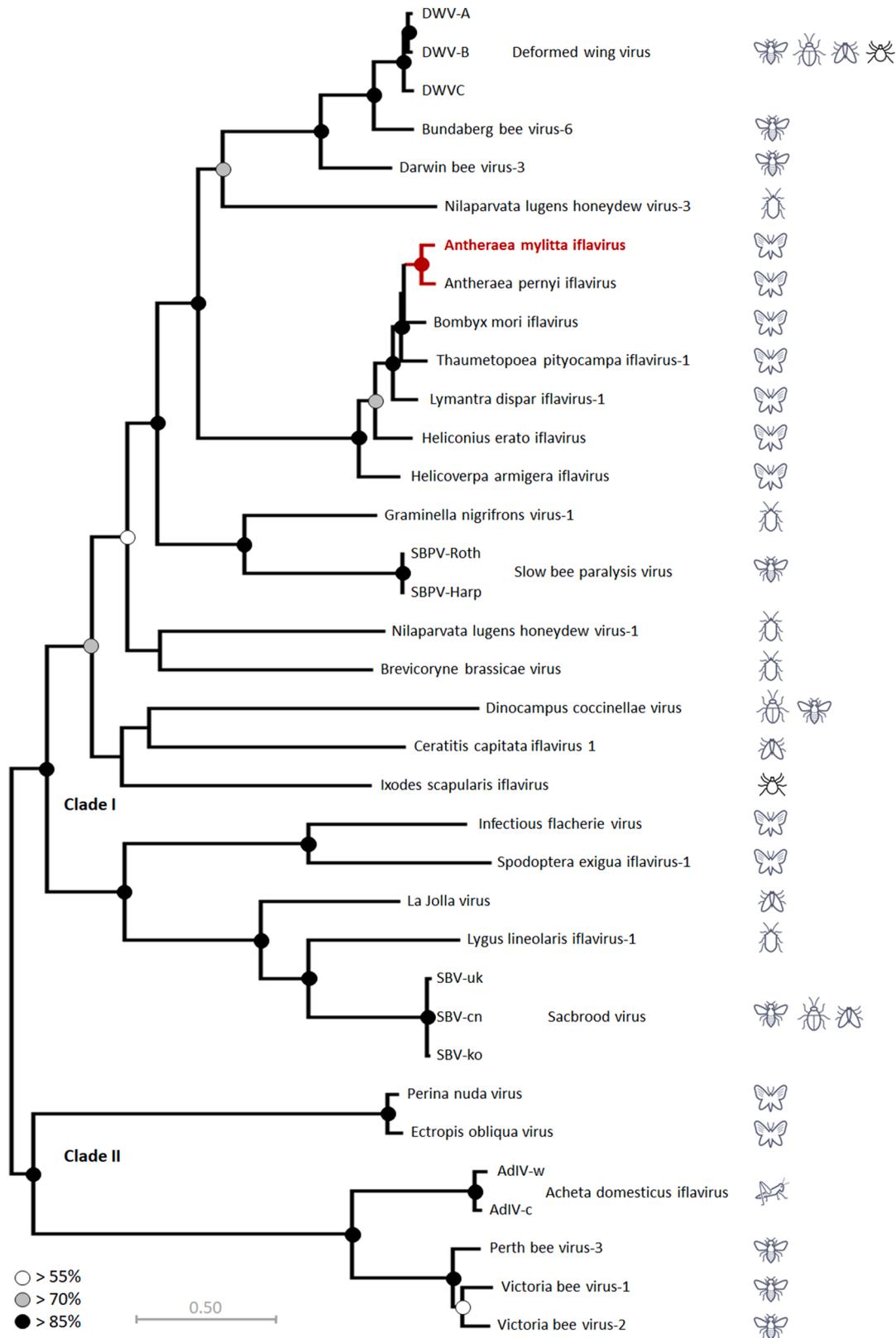


Fig. 3. Location of the AmIV (red branch) in the Iflavirus phylogenetic tree, based on analysis of the full polyprotein amino acid sequences of a representative selection of Iflaviruses (Supplementary Table 1). The evolutionary history was inferred by using the Maximum Likelihood method and JTT matrix-based model. The tree with the highest log likelihood (-51,417.43) is shown. The percentage of trees in which the associated taxa clustered together is represented by circles shown next to the branches. The tree is drawn to scale, with branch lengths measured in the number of amino acid substitutions per site. The insect icons next to the taxa represent the insect families the virus has been proven to infect.

Table 2
Samples of adult *Antheraea mylitta* from sericulture farms screened for covert infection of AmIV. For locations c.f. Fig. 1C.

Sample Origin	Total Samples			Positive for AmIV	
	Total	Females	Males	Females	Males
Patelnagar, West Bengal	13	8	5	4 (50%)	3 (60%)
Ranchi, Jharkhand	7	0	7	0	6 (86%)
Bilaspur, Chhattisgarh	16	8	8	7 (88%)	4 (50%)
Warangal, Telangana	14	9	5	6 (67%)	2 (40%)
Total	50	25	25	17 (68%)	15 (60%)

4. Discussion

In this study, we have analysed several genetic and biological characteristics of a novel Iflavirus identified in *A. mylitta*, the wild Tasar silkworm of India, and present arguments that this is a new virus species, for which we propose the name *Antheraea mylitta* Iflavirus (AmIV). The virus has a single-stranded RNA genome that is at least 9728 nucleotides long and encodes a single 2967 amino acid polyprotein flanked at the 5'

and 3' ends by Non-Translated Regions and terminating in a natural poly-A tail at the 3' end. These non-translated regions contain secondary and tertiary RNA structure that are involved in regulating and coordinating viral replication and translation. The size of these NTRs is variable from species to species (Lu et al., 2006). The AmIV 5' NTR is 435 nucleotides shorter than the 5' NTR region of its closest relative, ApIV, implying that the AmIV genome sequence may well be incomplete at the 5' end. The genome was found to be A/U rich, which is observed to also be the case in other Iflaviruses (Fujiyuki et al., 2004; Lanzi et al., 2006; Ongus et al., 2004). The polyprotein of the new virus contains a short leader peptide at the N-terminus, followed by four viral capsid proteins VP1-VP4 with the helicase, 3C protease and RdRp protein located towards the C-terminal end. This genome organization is highly characteristic of the Iflaviridae, a common and widely distributed insect-specific virus family in the Order Picornavirales (Valles et al. 2017). In the Iflaviridae, the mature proteins are released from the polyprotein through protease activity, primarily through the virus-encoded 3C-protease, which cleaves at conserved protease recognition sites (King et al., 2012). These recognition sites are also quite specific for each virus, such that each virus primarily processes its own

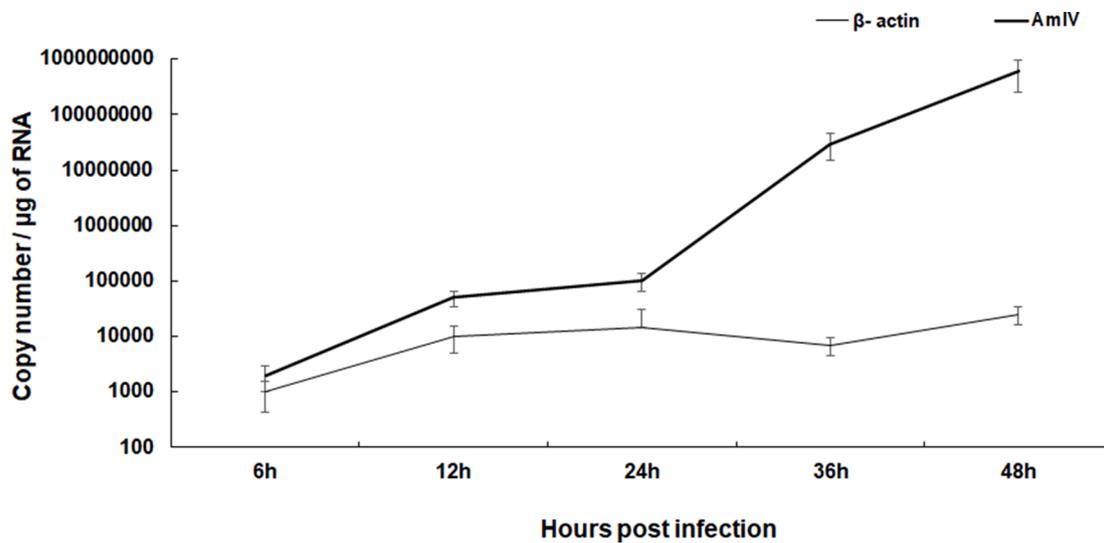


Fig. 4. Inoculation of the Iflavirus into 4th instar larvae to check for infectivity at different time points. The fatbody tissue was collected at different timepoints in order to check for the viral copy number.

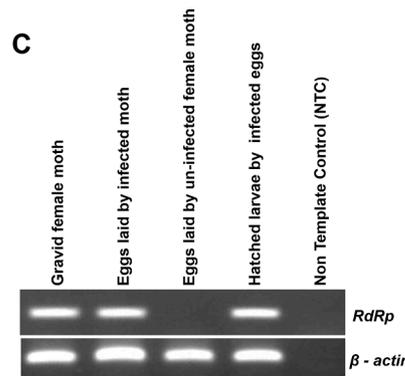
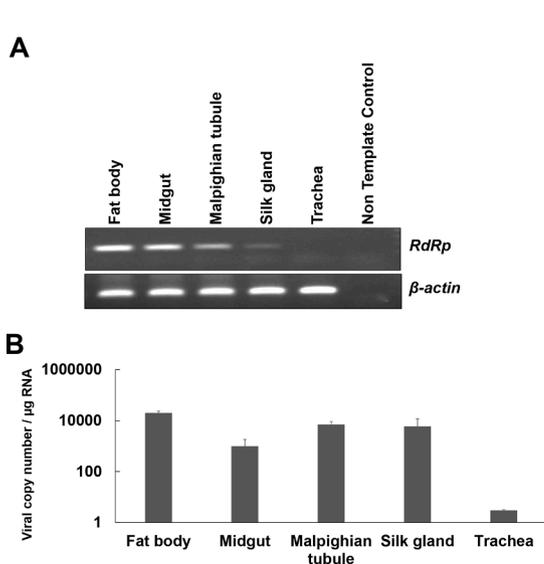


Fig. 5. A: Tissue tropism of AmIV in the infected *A. mylitta* field samples using the AmIV RdRp end-point RT-PCR assay (R7/R8 primers; Table 1). The β -actin end-point RT-PCR assay was used for quality control of the RNA/cDNA samples. **B:** Quantitative RT-qPCR analysis of the AmIV load in different infected silkworm tissues (pool of individual tissue samples), using the AmIV RdRp RT-qPCR assay. The *A. mylitta* β -actin assay was used as an internal reference gene for RNA quality control. **C:** Vertical transmission of the virus from infected mother moth to the offspring. The samples tested included the gravid female, eggs and hatched larvae.

Table 3

Adult *A. mylitta* samples were tested for co-infection of AmIV and *Nosema*. Silk moths without symptoms were collected from Warangal and tested with end-point PCR assays for β -tubulin (*Nosema*; DNA sample), and RdRp (AmIV; cDNA sample) genes.

	Total samples	Positive for AmIV alone	Co-infected with <i>Nosema</i>
Female moth, Warangal	25	17 (68%)	5 (20%)
Male moth, Warangal	25	15 (60%)	4 (16%)
Total	50	32 (64%)	9 (18%)

Table 4

A. mylitta larval samples were tested for co-infection of AmIV, *Nosema* and CPV. All silkworm larvae had symptoms of vomiting disease and were collected from the Warangal and Karimnagar districts of Telangana. They were tested with end-point PCR assays for β -tubulin (*Nosema*; DNA sample), CPV and RdRp (AmIV; cDNA sample) genes.

Silkworm rearing location	Total samples	Presence of AmIV in diseased silkworm		
		Positive for AmIV alone	Co-infected with <i>Nosema</i>	Co-infected with <i>Nosema</i> and CPV
Jakaram, Warangal (Telangana)	35	21 (60%)	14 (40%)	0
Mahadevpur, Karimnagar (Telangana)	40	30 (75%)	9 (22.5%)	1 (2.5%)
Total	75	51 (68%)	23 (30.7%)	1 (1.3%)

polyproteins. Six putative 3C-protease cleavage sites were identified on the novel virus polyprotein, as well as an autocatalytic cleavage site between VP4 and VP3. These protease recognition sites are located in suitable positions to release the leader peptide, the capsid proteins and the various non-structural proteins involved in virus replication and processing (Jiang et al., 2014; van Oers, 2010). A number of peptide motifs were identified across the AmIV genome that are highly conserved across the Picornavirales, such as the NxNxFQxG, WxGxLx₃FxFx₇Gx₅YxP, and FxRG motifs in the structural proteins VP2, VP3, and VP1, respectively (Liljas et al., 2002). Further motifs were identified in the non-structural proteins, such as the helicase, 3C-protease and RdRp. The helicase domain was identified between amino acids 1032 and 2207 and belongs to the helicase superfamily 3. This superfamily is generally found to be associated with NTP binding and contain 3 conserved motifs A-Gx₂GxGKS, motif B-Qx₅DD and motif C-KKx₄Px₅NSN, which was found to be slightly different in AmIV from the consensus sequence KGx₄Sx₅STN (Cheng et al., 2013; Gorbalenya and Koonin, 1989; Koonin and Dolja, 1993). The 3C protease of picornavirus resembles the chymotrypsin-like cysteine protease domain. The conserved cysteine-protease motifs GxCG (Ryan and Flint, 1997) and a substrate-binding motif GxHxx₃G (Gorbalenya et al., 1989) was observed within this region. The RdRp domain in AmIV belongs to RNA polymerase superfamily I (Kamer and Argos, 1984) that contains conserved I-VIII motifs including PSGx₃Tx₃N (S/T and YGDD, which are characteristic motifs of the RdRp domain (Koonin and Dolja, 1993).

Comparative sequence analysis is necessary to derive sequence-structure-function relationships as this forms the first step to understand the way in which the sequence is organised. The similarity index between AmIV and other picornaviruses was calculated and ApIV displayed the highest amino acid identity with 80.51% indicating its polyprotein to be highly similar to the polyprotein of AmIV. The synonymous- and non-synonymous substitution rates of codons (dN/dS) between ApIV and AmIV was low (0.0173) indicating a high degree of codon conservation between these two iflaviruses (Kryazhimskiy and Plotkin, 2008).

The novel AmIV exists as a covert infection in about 68% of female and 60% of male adult *A. mylitta* silk moths collected from different parts of India. However, the AmIV was present in about 100% of *A. mylitta* larval samples with overt symptoms of vomiting disease. The controlled infection studies could replicate the vomiting disease symptoms, with the inoculum from the crude extract obtained from diseased silkworms. The presence of AmIV in all the diseased silkworms with vomiting disease symptoms as well as the controlled infection studies indicate that AmIV is strongly associated with the vomiting disease symptoms. The support for AmIV causing the vomiting disease also comes from a study of the closely related ApIV where injection of ApIV purified with a discontinuous sugar gradient lead to vomiting in *Antheraea pernyi* larvae (Geng et al., 2014). From earlier studies, the silkworms from different seasons were screened for the presence of virus; although the incidence of disease was noticed in all seasons, the disease symptoms were more pronounced and resulted in death and heavy cocoon crop loss only in rainy and autumn seasons (Mahobia and Yadav 2010). We hypothesize that the AmIV is prevalent in these silkworm populations as covert infections, and when environmentally triggered (high humidity, high temperature, poor leaf quality) the covert infection results in overt infection resulting in death of these silkworms (Kumar et al., 1990). Iflaviruses have been observed to exist in natural populations without showing any disease symptoms, hence maintaining their population in the preferred host (Jakubowska et al., 2014).

From the studies on ApIV-associated vomiting disease and ApIV transmission, it was noted that oral administration of ApIV resulted in only local infection, whereas the virus actively replicated and spread to various tissues when it was injected (Geng et al., 2017). The AmIV became active once inside the host and was capable of spreading and propagating in various tissues such as fatbody, midgut, Malpighian tubules, silk glands and trachea. However, the viral load was found to be higher in the fatbody in comparison to other tissues, indicating fatbody to be the preferred tissue for propagation. Also, when the *A. mylitta* were tested for infectivity after injection, the virus was capable of multiplication in the host fatbody tissues starting at 6 h and increasing up to 48 h post infection. This corroborates earlier findings as the fatbody was previously reported as one of the sites for the virus multiplication in ApIV (Geng et al., 2017). Vertical transmission is one of the most commonly observed phenomena in many viruses infecting insects such as honey bees (Chen et al., 2006). Iflaviruses have previously shown to be transmitted vertically in *A. pernyi* silkworms (Geng et al., 2017). In the current study, we were also able to show vertical mode of transmission in the field samples infected with Iflavirus.

Co-infection is another phenomenon observed for Iflaviruses, where commonly both infecting pathogens establish a successful infection in the host. In this study, a total of nine adult silk moth samples out of 32 were found to be co-infected with Iflavirus and *Nosema* (a well-known pathogen which causes devastating pébrine disease in silkworms; Velide et al., 2013). The samples tested were infected adult silk moths without any disease symptoms, indicating that AmIV also causes covert infections in *A. mylitta* larvae and allows transstadial transmission of the virus. In a parallel study, there is evidence that co-infection of *Nosema* along with invertebrate iridescent virus (IIV) was one of the reasons for the colony collapse in honey bees (Bromenshenk et al., 2010). Controlled studies of co-infection of both AmIV and *Nosema* are necessary to understand the synergistic (if any) effect of the co-infecting pathogens. Co-infections of iflaviruses with other viruses have been noted earlier. For example, Iflavirus and baculovirus co-infection was beneficial for the survival of Iflavirus in *Spodoptera exigua* by increasing the physical stability and infectivity of the infecting Iflavirus (Jakubowska et al., 2016). In another study it has been found that in *S. exigua*, the susceptibility to NPV infections increases with covert infections with Iflavirus (Carballo et al., 2020); however, we did not find any NPV in the AmIV-infected larvae (data not shown). In contrast, the infection of a novel Iflavirus was inhibited by an infection of Rice dwarf virus in the Green Rice Leafhopper (Jia et al., 2020).

This study gives details on various aspects required to understand the disease etiology of the novel Iflavirus associated with vomiting disease in wild *A. mylitta* silkworms. The study revealed the presence of the AmlV as covert infections in field samples, its genome organization, phylogenetic analysis, infectivity through inoculation, tissue tropism, vertical transmission and co-infection. Further insights into the infecting virus may provide more details on the host pathogen interactions, immune responses, and identification of novel proteins, which may lead to design of better control measures that could be implemented at field level to prevent the occurrence and spread of this disease.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Kangayam M. Ponnuel: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – original draft. **Joachim R. de Miranda:** Investigation, Software, Visualization, Writing – review & editing. **Olle Terenius:** Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – review & editing. **Wenli Li:** Data curation, Investigation, Methodology, Writing – review & editing. **Katsuhiko Ito:** Data curation, Methodology. **Diksha Khajje:** . **G Shamitha:** . **Anupama Jagadish:** Writing – original draft, Writing – review & editing. **Himanshu Dubey:** Data curation, Formal analysis, Investigation, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Rakesh K. Mishra:** Resources.

Acknowledgements

The authors would like to thank the Central Silk Board (CSB), Govt. of India for providing the infrastructure support to carry out research. The study was also supported by the Swedish Research Council (Grant No: 2017–05463).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.virusres.2022.198703.

References

Arunkumar, K., Tomar, A., Daimon, T., Shimada, T., Nagaraju, J., 2008. WildSilkbase: an EST database of wild silkworms. *BMC Genom.* 9, 338. <https://doi.org/10.1186/1471-2164-9-338>.

Bailey, T.L., Johnson, J., Grant, C.E., Noble, W.S., 2015. The MEME suite. *Nucleic Acids Res.* 43, W39. <https://doi.org/10.1093/nar/gkv416>.

Beaurepaire, A., Piot, N., Doublet, V., Antunez, K., Campbell, E., Chantawannakul, P., Chejanovsky, N., Gajda, A., Heermsan, M., Panziera, D., Smaghe, G., Yañez, O., de Miranda, J.R., Dalmon, A., 2020. Diversity and global distribution of viruses of the Western honey bee, *Apis mellifera*. *Insects* 11, 239. <https://doi.org/10.3390/insects11040239>.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.

Bromenshenk, J.J., Henderson, C.B., Wick, C.H., Stanford, M.F., Zulich, A.W., Jabbour, R.E., Deshpande, S.V., McCubbin, P.E., Seccomb, R.A., Welch, P.M., Williams, T., Firth, D.R., Skowronski, E., Lehmann, M.M., Bilimoria, S.L., Gress, J., Wanner, K.W., Cramer, R.A., 2010. Iridovirus and microsporidian linked to honey bee colony decline. *PLoS One* 5. <https://doi.org/10.1371/journal.pone.0013181>.

Carballo, A., Williams, T., Murillo, R., Caballero, P., 2020. Iflavirus covert infection increases susceptibility to nucleopolyhedrovirus disease in *Spodoptera exigua*. *Viruses* 12, E509. <https://doi.org/10.3390/v12050509>.

Chen, Y., Evans, J., Feldlaufer, M., 2006. Horizontal and vertical transmission of viruses in the honey bee, *Apis mellifera*. *J. Invertebr. Pathol.* 92, 152–159. <https://doi.org/10.1016/j.jip.2006.03.010>. Society of Invertebrate Pathology 2006 Special Issue.

Chen, Y.P., Becnel, J.J., Valles, S.M., 2012. Chapter 5 - RNA viruses infecting pest insects. In: Vega, F.E., Kaya, H.K. (Eds.), *Insect Pathology*, 2nd ed. Academic Press, San Diego, pp. 133–170. <https://doi.org/10.1016/B978-0-12-384984-7.00005-1>.

Cheng, Z., Yang, J., Xia, H., Qiu, Y., Wang, Z., Han, Y., Xia, X., Qin, C.-F., Hu, Y., Zhou, X., 2013. The nonstructural protein 2C of a Picorna-like virus displays nucleic acid helix destabilizing activity that can be functionally separated from its ATPase activity. *J. Virol.* 87, 5205–5218. <https://doi.org/10.1128/JVI.00245-13>.

de Miranda, J.R., Granberg, F., Onorati, P., Jansson, A., Berggren, Å., 2021. Virus prospecting in crickets – discovery and strain divergence of a novel Iflavirus in wild and cultivated *Acheta domestica*. *Viruses* 13, 364. <https://doi.org/10.3390/v13030364>.

de Miranda, J.R., Cordon, G., Budge, G., 2010b. The Acute bee paralysis virus-Kashmir bee virus-Israeli acute paralysis virus complex. *J. Invertebr. Pathol.* 103 (Suppl 1), S30–47. <https://doi.org/10.1016/j.jip.2009.06.014>.

de Miranda, J.R., Dainat, B., Locke, B., Cordon, G., Berthoud, H., Gauthier, L., Neumann, P., Budge, G.E., Ball, B.V., Stoltz, D.B., 2010a. Genetic characterization of slow bee paralysis virus of the honeybee (*Apis mellifera* L.). *J. Gen. Virol.* 91, 2524–2530. <https://doi.org/10.1099/vir.0.022434-0>.

de Miranda, J.R., Genersch, E., 2010. Deformed wing virus. *J. Invertebr. Pathol.* 103 (Suppl 1) <https://doi.org/10.1016/j.jip.2009.06.012>. S48–61.

Eberle, K.E., Wennmann, J.T., Kleespies, R.G., Jehle, J.A., 2012. Chapter II - Basic techniques in insect virology. In: Lacey, L.A. (Ed.), *Manual of Techniques in Invertebrate Pathology*, 2nd ed. Academic Press, San Diego, pp. 15–74. <https://doi.org/10.1016/B978-0-12-386899-2.00002-6>.

Esvaran, V.G., Mohanasundaram, A., Mahadeva, S., Gupta, T., Ponnuel, K.M., 2019. Development and comparison of real-time and conventional PCR tools targeting β -tubulin gene for detection of *Nosema* infection in silkworms. *J. Parasit. Dis.* 43, 31–38. <https://doi.org/10.1007/s12639-018-1053-4>.

Forslund, K., Pekkari, I., Sonhammer, E.L., 2011. Domain architecture conservation in orthologs. *BMC Bioinform.* 12, 326. <https://doi.org/10.1186/1471-2105-12-326>.

Fujiyuki, T., Takeuchi, H., Ono, M., Ohka, S., Sasaki, T., Nomoto, A., Kubo, T., 2004. Novel insect picorna-like virus identified in the brains of aggressive worker honeybees. *J. Virol.* 78, 1093–1100. <https://doi.org/10.1128/jvi.78.3.1093-1100.2004>.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R., Bairoch, A., 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 31, 3784–3788. <https://doi.org/10.1093/nar/gkg563>.

Geng, P., Li, W., de Miranda, J.R., Qian, Z., An, L., Terenius, O., 2017. Studies on the transmission and tissue distribution of *Antheraea pernyi* Iflavirus in the Chinese oak silkworm *Antheraea pernyi*. *Virology* 502, 171–175. <https://doi.org/10.1016/j.virol.2016.12.014>.

Geng, P., Li, W., Lin, L., Miranda, J.R.de, Emrich, S., An, L., Terenius, O., 2014. Genetic characterization of a novel Iflavirus associated with vomiting disease in the Chinese oak silkworm *Antheraea pernyi*. *PLoS One* 9, e92107. <https://doi.org/10.1371/journal.pone.0092107>.

Ghosh, R.C., Ball, B.V., Willcocks, M.M., Carter, M.J., 1999. The nucleotide sequence of sacbrood virus of the honey bee: an insect picorna-like virus. *J. Gen. Virol.* 80 (Pt 6), 1541–1549. <https://doi.org/10.1099/0022-1317-80-6-1541>.

Gorbalenya, A.E., Koonin, E.V., 1989. Viral proteins containing the purine NTP-binding sequence pattern. *Nucleic Acids Res.* 17, 8413–8440. <https://doi.org/10.1093/nar/17.21.8413>.

Gorbalenya, A.E., Donchenko, A.P., Blinov, V.M., Koonin, E.V., 1989. Cysteine proteases of positive strand RNA viruses and chymotrypsin-like serine proteases. A distinct protein superfamily with a common structural fold. *FEBS Lett.* 243, 103–114. [https://doi.org/10.1016/0014-5793\(89\)80109-7](https://doi.org/10.1016/0014-5793(89)80109-7).

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mueceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat. Biotechnol.* 29, 644–652. <https://doi.org/10.1038/nbt.1883>.

Isawa, H., Asano, S., Sahara, K., Iizuka, T., Bando, H., 1998. Analysis of genetic information of an insect picorna-like virus, infectious flacherie virus of silkworm: evidence for evolutionary relationships among insect, mammalian and plant picorna (-like) viruses. *Arch. Virol.* 143, 127–143. <https://doi.org/10.1007/s007050050273>.

Jakubowska, A.K., Murillo, R., Carballo, A., Williams, T., van Lent, J.W.M., Caballero, P., Herrero, S., 2016. Iflavirus increases its infectivity and physical stability in association with baculovirus. *PeerJ* 4. <https://doi.org/10.7717/peerj.1687>.

Jakubowska, A.K., D'Angiolo, M., González-Martínez, R.M., Millán-Leiva, A., Carballo, A., Murillo, R., Caballero, P., Herrero, S., 2014. Simultaneous occurrence of covert infections with small RNA viruses in the lepidopteran *Spodoptera exigua*. *J. Invertebr. Pathol.* 121, 56–63. <https://doi.org/10.1016/j.jip.2014.06.009>.

Jia, W., Wang, F., Li, J., Chang, X., Yang, Y., Yao, H., Bao, Y., Song, Q., Ye, G., 2020. A novel Iflavirus was discovered in Green rice leafhopper *Nephotettix cincticeps* and its proliferation was inhibited by infection of rice dwarf virus. *Front. Microbiol.* 11, 621141. <https://doi.org/10.3389/fmicb.2020.621141>.

Jiang, P., Liu, Y., Ma, H.C., Paul, A.V., Wimmer, E., 2014. Picornavirus morphogenesis. *Microbiol. Mol. Biol. Rev.* 78, 418–437. <https://doi.org/10.1128/MMBR.00012-14>.

Jones, D.T., Taylor, W.R., Thornton, J.M., 1992. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* 8, 275–282. <https://doi.org/10.1093/bioinformatics/8.3.275>.

Kamer, G., Argos, P., 1984. Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. *Nucleic Acids Res.* 12, 7269–7282. <https://doi.org/10.1093/nar/12.18.7269>.

- King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J., 2012. Family - iflavirus. *Virus Taxonomy*. Elsevier, San Diego, pp. 846–849. <https://doi.org/10.1016/B978-0-12-384684-6.00072-0>.
- Koonin, E.V., Dolja, V.V., 1993. Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. *Crit. Rev. Biochem. Mol. Biol.* 28, 375–430. <https://doi.org/10.3109/10409239309078440>.
- Kryazhimskiy, S., Plotkin, J.B., 2008. The population genetics of dN/dS. *PLoS Genet.* 4, e1000304. <https://doi.org/10.1371/journal.pgen.1000304>.
- Kumar, P., Baig, M., Sengupta, K., Govindaiah., 1990. *Handbook on Pest and Disease Control of Mulberry and Silkworm*. UN, New York.
- Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. <https://doi.org/10.1093/molbev/msy096>.
- Lanzi, G., de Miranda, J.R., Boniotti, M.B., Cameron, C.E., Lavazza, A., Capucci, L., Camazine, S.M., Rossi, C., 2006. Molecular and biological characterization of deformed wing virus of honeybees (*Apis mellifera* L.). *J. Virol.* 80, 4998–5009. <https://doi.org/10.1128/JVI.80.10.4998-5009.2006>.
- Leggett, R.M., Ramirez-Gonzalez, R.H., Clavijo, B.J., Waite, D., Davey, R.P., 2013. Sequencing quality assessment tools to enable data-driven informatics for high throughput genomics. *Front. Genet.* 4. <https://doi.org/10.3389/fgene.2013.00288>.
- Liljas, L., Tate, J., Lin, T., Christian, P., Johnson, J.E., 2002. Evolutionary and taxonomic implications of conserved structural motifs between picornaviruses and insect picorna-like viruses. *Arch. Virol.* 147, 59–84. <https://doi.org/10.1007/s705-002-8303-1>.
- Lu, J., Zhang, J., Wang, X., Jiang, H., Liu, C., Hu, Y., 2006. *In vitro* and *in vivo* identification of structural and sequence elements in the 5' untranslated region of *Ectropis obliqua* picorna-like virus required for internal initiation. *J. Gen. Virol.* 87, 3667–3677. <https://doi.org/10.1099/vir.0.82090-0>.
- Maciel-Vergara, G., Ros, V.I.D., 2017. Viruses of insects reared for food and feed. *J. Invertebr. Pathol.* 147, 60–75. <https://doi.org/10.1016/j.jip.2017.01.013>.
- Mahobia, G., Yadav, G., 2010. Seasonal variation in the diseases of tropical tasar silkworm (*Antheraea mylitta* D.) under Bastar plateau of Chhattisgarh and their control measures. *Uttar Pradesh J. Zool.* 30 (1), 7–10. <http://mbimph.com/index.php/UPJOZ/article/view/223> (Accessed: 21August2021).
- Murakami, R., Suetsugu, Y., Kobayashi, T., Nakashima, N., 2013. The genome sequence and transmission of an Iflavirus from the brown planthopper, *Nilaparvata lugens*. *Virus Res.* 176, 179–187. <https://doi.org/10.1016/j.virusres.2013.06.005>.
- Ongus, J.R., Peters, D., Bonmatin, J.-M., Bengsch, E., Vlak, J.M., van Oers, M.M., 2004. Complete sequence of a picorna-like virus of the genus Iflavirus replicating in the mite *Varroa destructor*. *J. Gen. Virol.* 85, 3747–3755. <https://doi.org/10.1099/vir.0.80470-0>.
- Procházková, M., Škubník, K., Füzik, T., Mukhamedova, L., Pridal, A., Plevka, P., 2020. Virion structures and genome delivery of honeybee viruses. *Curr. Opin. Virol.* 45, 17–24. <https://doi.org/10.1016/j.coviro.2020.06.007>.
- Qanungo, K.R., Kundu, S.C., Mullins, J.I., Ghosh, A.K., 2002. Molecular cloning and characterization of *Antheraea mylitta* cytoplasmic polyhedrosis virus genome segment 9. *J. Gen. Virol.* 83 (Pt 6), 1483–1491. <https://doi.org/10.1099/0022-1317-83-6-1483>. Jun.
- Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., Lopez, R., 2005. InterProScan: protein domains identifier. *Nucleic Acids Res.* 33, W116–W120. <https://doi.org/10.1093/nar/gki442>.
- Ryan, M.D., Flint, M., 1997. Virus-encoded proteinases of the picornavirus super-group. *J. Gen. Virol.* 78 (Pt 4), 699–723. <https://doi.org/10.1099/0022-1317-78-4-699>.
- Suyama, M., Torrents, D., Bork, P., 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* 34, W609–W612. <https://doi.org/10.1093/nar/gkl315>.
- Valles, S.M., Chen, Y., Firth, A.E., Guérin, D.M.A., Hashimoto, Y., Herrero, S., de Miranda, J.R., Ryabov, E., ICTV Report Consortium, 2017. ICTV virus taxonomy profile: Iflaviridae. *J. Gen. Virol.* 98, 527–528. <https://doi.org/10.1099/jgv.0.000757>.
- Van Oers, M.M., 2010. Genomics and biology of iflaviruses. *Insect Virol.* 231–250.
- Velide, L., Bhagavanulu, M.V.K., Rao, A.P., 2013. Study on impact of parasite (*Nosema* species) on characters of tropical tasar silkworm *Antheraea mylitta* drury. *J. Environ. Biol.* 34, 75–78.
- Venkataraman, S., Prasad, B.V.L.S., Selvarajan, R., 2018. RNA dependent RNA polymerases: insights from structure, function and evolution. *Viruses* 10, E76. <https://doi.org/10.3390/v10020076>.
- Yañez, O., Piot, N., Dalmon, A., de Miranda, J.R., Chantawannakul, P., Panziera, D., Amiri, E., Smagghe, G., Schroeder, D., Chejanovsky, N., 2020. Bee Viruses: Routes of Infection in Hymenoptera. *Front. Microbiol.* 11, 943. <https://doi.org/10.3389/fmicb.2020.00943>.
- Yuan, H., Xu, P., Yang, X., Graham, R.I., Wilson, K., Wu, K., 2017. Characterization of a novel member of genus Iflavirus in *Helicoverpa armigera*. *J. Invertebr. Pathol.* 144, 65–73. <https://doi.org/10.1016/j.jip.2017.01.011>.