



OPEN

## Nearly complete genome sequences of the first two identified Colorado potato beetle viruses

Maria Antonets<sup>1,2,3✉</sup>, Sergei Bodnev<sup>1</sup>, Ulyana Rotskaya<sup>3</sup>, Elena Kosman<sup>3</sup>, Tatyana Tregubchak<sup>1</sup>, Tatyana Bauer<sup>1</sup>, Mamedyar Azaev<sup>1</sup>, Vadim Kryukov<sup>3</sup> & Denis Antonets<sup>2,4,5</sup>

The Colorado potato beetle is one of the most devastating potato pests in the world. However, its viral pathogens, which might have potential in pest control, have remained unexplored. With high-throughput sequencing of Colorado potato beetle samples derived from prepupal larvae which died from an unknown infection, we have identified two previously unknown RNA viruses and assembled their nearly complete genome sequences. The subsequent genetic and phylogenetic analysis demonstrated that the viruses, tentatively named *Leptinotarsa iflavirus 1* and *Leptinotarsa solinvi-like virus 1*, are the novel representatives of the *Iflaviridae* and *Solinviviridae* families, respectively. To the best of our knowledge, these are the first sequencing-confirmed insect viruses derived from Colorado potato beetle samples. We propose that *Leptinotarsa iflavirus 1* may be associated with a lethal disease in the Colorado potato beetle.

The Colorado potato beetle—*Leptinotarsa decemlineata* of Chrysomelidae family—is one of the most widespread and destructive Solanaceae plants pests in the world. Although the use of insecticides helps to decline the beetle's populations, the Colorado potato beetle is notorious for its ability to develop resistance to all chemicals registered to date<sup>1</sup>. That is why there is an urgent need for new effective methods to control Colorado potato beetle populations and also to develop new biocontrol strategies as the most environmentally friendly<sup>2</sup>. Currently, the existing methods of biological control of Colorado potato beetle populations use entomopathogenic fungi such as *Beauveria bassiana*, *Metarhizium* sp.<sup>3</sup> and bacteria *Bacillus thuringiensis*<sup>4</sup>, but not viruses. However, entomopathogenic virus-based bioagents that are species-specific and generally safe for non-target organisms are valuable tools for developing integrated pest management strategies. As for other insect pests, positive examples include the nucleopolydoviruses of the *Baculoviridae* family, which are currently used to control a number of Lepidoptera species<sup>4</sup>.

In a recent study we presented the results of metagenomic analysis of viral diversity in non-target Colorado potato beetle samples—namely, in public genomic and transcriptomic *Leptinotarsa decemlineata* NGS data obtained from the NCBI SRA database<sup>5</sup>. In these data, we have identified virus-associated genetic sequences belonging to more than 90 species and 32 families of viruses (excluding bacteriophages), a significant part of which were RNA viruses. It is worth noting that, targeted studies dedicated to the identification of Colorado potato beetle viruses have not been conducted to date.

When growing natural populations of the Colorado potato beetle under laboratory conditions<sup>6,7</sup>, we observed larvae death rate of 8–35% (upto 60%) when they completed feeding and entered into the prepupal stage (7–11 days after molting at instar IV). The dead larvae were characterized by specific phenotype with body turgor loss, body straightening and legs stretching (Fig. 1). 1–3 days post mortem the hemolymph acquired a dark tint characteristic to septicemia (Fig. 1C). To the best of our knowledge these particular disease specific symptoms were not described earlier in Colorado potato beetle. In the samples of dead Colorado potato beetle

<sup>1</sup>State Research Center of Virology and Biotechnology “Vector”, Rospotrebnadzor, 630559 Koltsovo, Russia. <sup>2</sup>Novel Software Systems LLC, Akademika Lavrentiev Ave. 6, 630090 Novosibirsk, Russia. <sup>3</sup>Institute of Systematics and Ecology of Animals SB RAS, Frunze Str. 11, 630091 Novosibirsk, Russia. <sup>4</sup>Novosibirsk State University, Pirogova Str. 2, 630090 Novosibirsk, Russia. <sup>5</sup>MSU Institute for Artificial Intelligence, Lomonosov Ave. 27, 119192 Moscow, Russia. ✉email: starchevskayamaria@mail.ru



**Figure 1.** Disease symptoms. (A)—the healthy prepupa; (B)—the prepupa 1 day after death; (C)—the prepupa 3 days after death.

prepupae we identified two novel RNA viruses, one of which was attributed to *Iflaviridae* family, and the second—to *Solinviviridae* family.

Members of the *Iflaviridae* family (order Picornavirales) have single-stranded (+)RNA genomes of 9–11 kb. Their genome contains a single open reading frame encoding a polyprotein that is proteolytically processed into functional viral proteins<sup>8</sup>. More than 550 genomic iflavirus sequences have been published to date to the NCBI GenBank database, of which 134 were published in 2021, 105 in 2022 and 42 in 2023. Iflaviruses infect arthropods with most of the described representatives infect insects. Iflaviruses target specific tissues of the insect host, often affecting crucial physiological processes. They can oppose the host immune responses by interfering with antiviral defense mechanisms and ensuring successful viral replication<sup>9</sup>. In some cases iflaviruses cause persistent infection, which makes them particularly interesting<sup>10</sup>.

Members of the *Solinviviridae* family (order Picornavirales) have single-stranded (+)RNA genomes of 10–11 kb. *Solinviviridae* viruses infect arthropods, with most described representatives infecting insects and crustaceans<sup>11,12</sup>. Some of these viruses are known to cause chronic infections and others, which are more virulent, cause systemic infections and acute mortality, however little is currently known about the effects of the most of these viruses on their hosts<sup>12</sup>. Only about 30 genome sequences of *Solinviviridae* viruses have been published to NCBI GenBank database to date.

## Results and discussion

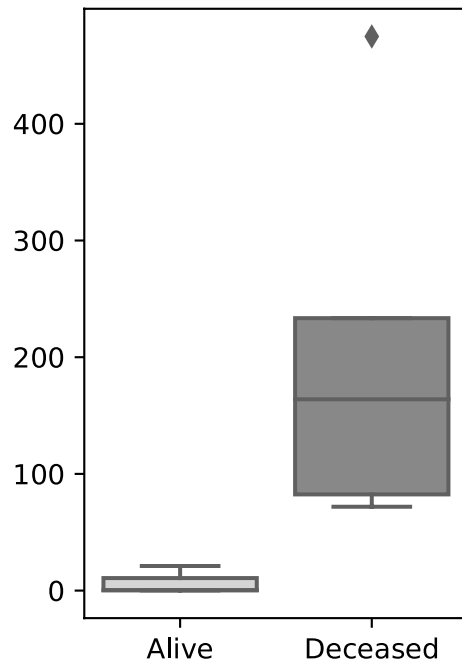
### Assessment of viral load with quantitative PCR

In our previous analysis of public Colorado potato beetle NGS data we were able to assemble several near full-length genomes of viruses related to *Iflaviridae* family<sup>5</sup> and thus we proposed that the observed Colorado potato beetle prepupal larvae death might be at least partially attributed to iflaviral infection. To test this possibility, we undertook a qPCR study to assess the quantities of viral genetic material in symptomatic dead prepupae in comparison to asymptomatic healthy ones. The viral RNA-dependent RNA polymerase (RdRp) gene was selected to assess the viral load and *L. decemlineata* ribosomal protein RP4 and RP18 genes were taken as reference ones, as it was proposed in<sup>13</sup>. The designed oligonucleotide primers are shown in Table 1. The qPCR study demonstrated the increased levels of viral RdRp in fresh dead symptomatic larvae (“deceased”) as compared to healthy asymptomatic ones (see Fig. 1A,B)—the mean relative RdRp expression in “deceased” group was about 32.9 times higher than in “alive” (Fig. 2). And the observed difference was found to be statistically significant ( $p = 0.0357$ ).

Thus, the significantly increased levels of viral RdRp were observed in tissues of freshly dead larvae with characteristic symptoms, which indicates active virus replication. Unfortunately, in current study it was impossible to attribute the larvae death to this particular virus infection. To prove that the virus causes lethal infection in Colorado potato beetle, it is necessary to purify the viral particles in order to determine their infectivity and lethal dose in further in vitro and in vivo studies or with producing infectious cDNA. And it is extremely important to determine the genome sequence of the target virus.

Gene name	NCBI GenBank accession number	Gene symbol	Primer sequence (5'–3')	Product size (bp)	PCR efficiency ( $\pm$ SD)	$T_m$ in qPCR, °C
<i>L. decemlineata</i> 60S ribosomal protein L4	XM_023165859	<i>Rp4</i>	Forward GAAACGAGCATT GCCCTTCC Reverse TCGCTGACACTG TAGGGTTG	119	2.08 $\pm$ 0.06	62
<i>L. decemlineata</i> 60S ribosomal protein L18	XM_023172940	<i>Rp18</i>	Forward TAGAATCCTCAA AGCAGGTGGC Reverse CTGGACCAAAGT GTT CACTGC	133	2.03 $\pm$ 0.04	62
Iflaviral RdRp		<i>RdRp</i>	Forward GAATATTCGGAT GACACTGTTGA Reverse TTAGAAACCTCA TGAGAATATGCAAG	82	1.85 $\pm$ 0.06	62

**Table 1.** The selected genes and oligonucleotide primer sequences used for qPCR. The primers for Rp4 and Rp18 were taken from paper Shi et al.<sup>13</sup> and modified by Rotskaya U.



**Figure 2.** Iflaviral RdRp relative expression in samples taken from alive and deceased Colorado potato beetle prepupae. The relative RdRp expression was determined as detailed in corresponding “Materials and methods” section. The RdRp expression values were normalized to reference genes: ribosomal proteins L4 and L18 (Rp4 and Rp18). The difference in viral loads was found to be statistically significant (Mann–Whitney two-tailed test,  $p=0.0357$ ).

### Sequencing and assembly of viral genome sequences

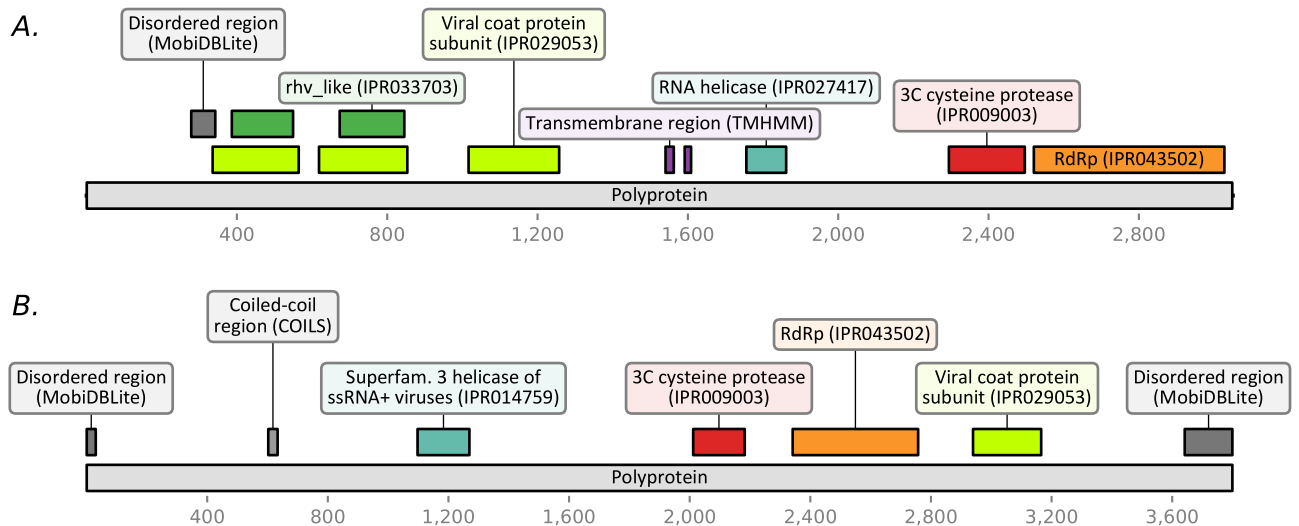
The sequencing yielded 473.5 × 2 thousand reads, with an average length of 124 nt, for the CPB3 sample and 480.5 × 2 thousand reads, with an average length of 127 nt, for the CPB6 sample. The de novo assembly yielded a nearly complete genome sequence of an iflavirus named *Leptinotarsa iflavirus 1* (in sample CPB3) with an average coverage depth of 834x (5.2% of total reads were back aligned to the assembled genomic sequence with at least 90% overlap and 95% identity) and a nearly complete genome sequence of solinvi-like virus named *Leptinotarsa solinvi-like virus 1* (in sample CPB6) with an average coverage depth of 1116x (8.7% of total reads were back aligned to the assembled genomic sequence with at least 90% overlap and 95% identity). The genetic sequences of *Leptinotarsa iflavirus 1* and *Leptinotarsa solinvi-like virus 1* were deposited to GenBank (with accession IDs OR613011 and OR613010, respectively).

Both viruses were found in both analyzed samples, but due to much lower coverage depth we were unable to assemble the contigs containing the complete CDS of *Leptinotarsa iflavirus 1* from sample CPB6, and contigs containing the complete CDS of *Leptinotarsa solinvi-like virus 1* from sample CPB3.

### Analysis of viral genetic sequences with InterProScan

The *Leptinotarsa iflavirus 1* genome sequence is 9887 nt long, excluding the polyA-tail, and contains a single ORF (positions 561–9707), 5'-UTR (560 nt) and 3'-UTR (180 nt). The ORF encodes a polyprotein of 3049 aa, which, according to the results of InterProScan, contains domains characteristic to iflaviruses and their mutual arrangement is also typical for the *Iflaviridae* family. These domains are the following: two rhv-like (rhinovirus-like) domains in two subunits of the Picornavirus/Calicivirus-like capsid protein (*IPR033703*, E-values of  $5.96 \times 10^{-13}$  and  $1.15 \times 10^{-21}$ , respectively), domain belonging to Calicivirus-like capsid proteins superfamily (*IPR029053*,  $2.5 \times 10^{-32}$ ), RNA helicase domain (*IPR000605*,  $1.4 \times 10^{-18}$ ), domain belonging to cysteine 3C peptidases superfamily (*IPR009003*,  $2.36 \times 10^{-16}$ ) and RNA-dependent RNA polymerase domain (*IPR043502*,  $4.65 \times 10^{-78}$ ) (Fig. 3A). The amino acid sequence of the *Leptinotarsa iflavirus 1* polyprotein has the highest homology with *Apis flavivirus 2* polyprotein (*UCR92484*); according to the BLASTp results their mutual identity is 37.72%.

The genome sequence of *Leptinotarsa solinvi-like virus 1* is 11,603 nt long, excluding the polyA tail, and contains a single ORF (coordinates 62–11458). This ORF encodes a polyprotein of 3800 aa, which, according to InterProScan, contains the following domains: RNA helicase domain (*IPR000605*, E-value:  $7.2 \times 10^{-20}$ ), a domain belonging to the superfamily of cysteine 3C peptidases (*IPR009003*,  $7.41 \times 10^{-6}$ ), RNA-dependent RNA polymerase domain (*IPR043502*,  $9.39 \times 10^{-55}$ ) and Calicivirus-like capsid protein domain (*IPR029053*,  $1.5 \times 10^{-18}$ ) (Fig. 3B). The amino acid sequence of *Leptinotarsa solinvi-like virus 1* polyprotein has the highest homology with the polyprotein of *Hangzhou Solinvi-like virus 1* (*UHR49784*), derived from samples of *Altica cyanea* leaf-feeding beetle (Chrysomelidae). According to the BLASTp results, their mutual sequence identity is 33.32%.



**Figure 3.** The results of viral polyproteins annotation with InterProScan. (A) The polyprotein of *Leptinotarsa iflavirus 1* (WNV56445) and (B) *Leptinotarsa solinvi-like virus 1* polyprotein (WNV56444).

### Phylogenetic analysis of viral RdRp proteins

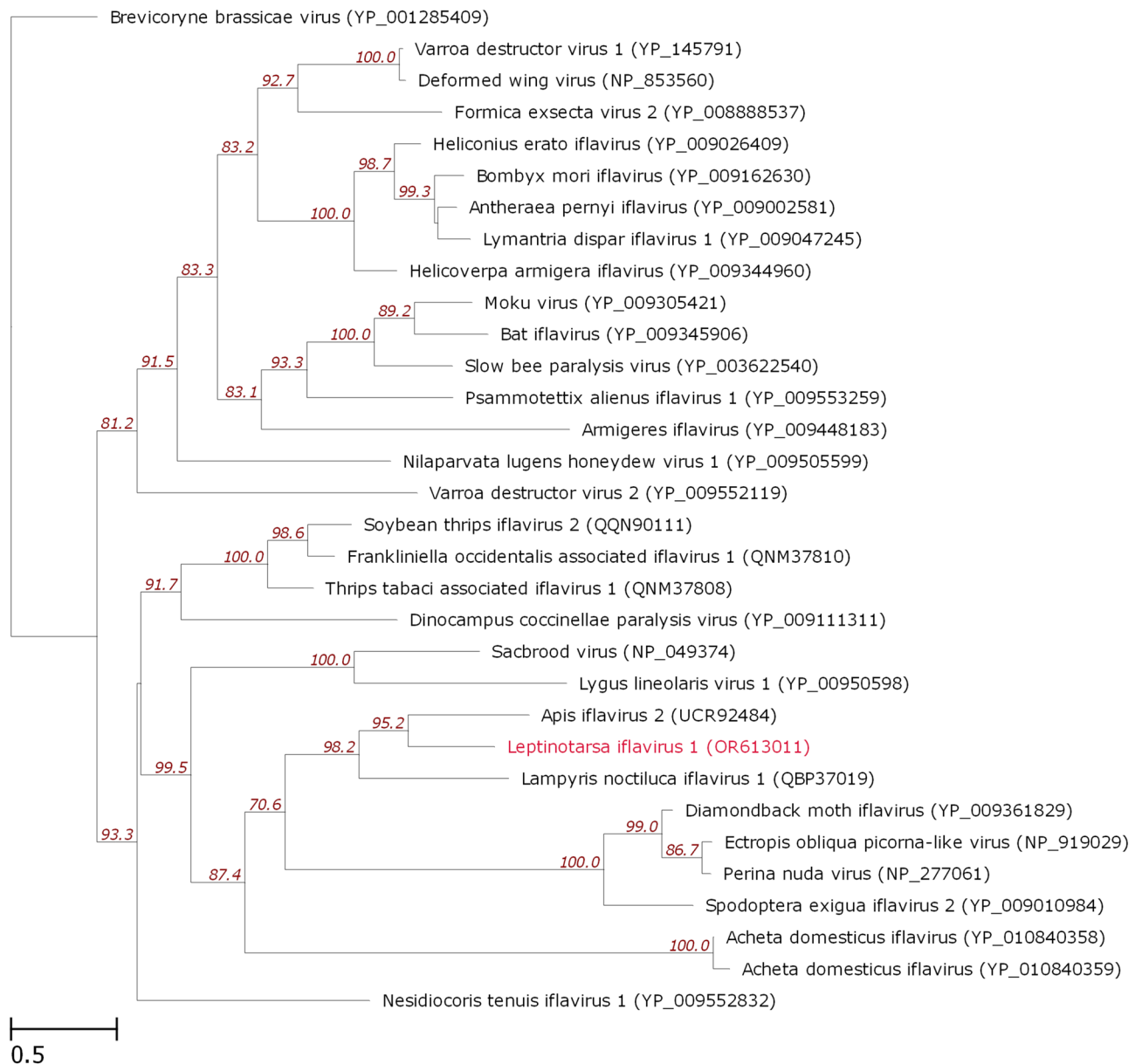
Since iflaviral sequences are very diverse, the phylogenetic analysis was performed using the most conserved polyprotein fragments containing RdRp. The multiple amino acid sequence alignment was built using the sequences with the highest homology to *Leptinotarsa iflavirus 1* RdRp (at least 30% identity) selected from NCBI nr database using BLASTp. The phylogenetic tree was constructed with IQ-Tree software using the maximum likelihood method (Fig. 4). The sequences closest to *Leptinotarsa iflavirus 1* RdRp were the RdRp of *Apis iflavirus 2* (UCR92484), derived from metagenomic samples of honey bees collected in China in the Henan province in 2017, and the RdRp of *Lampyrus noctiluca iflavirus 1* (QBP37019), obtained from metagenomic analysis of a firefly sample collected in Finland in 2017.

To construct the phylogenetic tree of *Leptinotarsa solinvi-like virus 1* RdRp, we used the multiple RdRp amino acid sequence alignment taken from ICTV (<https://ictv.global/sites/default/files/inline-images/Figure3-1.v2.aa.alignment.fst>, accessed on August 2023)<sup>11</sup>. The alignment was augmented with sequences closely related to *Leptinotarsa solinvi-like virus 1* RdRp extracted from NCBI nr database with BLASTp (with at least 30% identity). Figure 5 demonstrates a fragment of the resulting phylogram produced with IQ-Tree.

The *Solinviviridae* sequences extracted from ICTV were also found to include RdRp encoded by the putative viral genetic sequence *Leptinotarsa decemlineata* TSA (GenBank: *GEEF01170301*), obtained from transcriptomic assembly of CPB samples (BioProject ID: *PRJNA297027*). This RdRp has the highest homology with the *Leptinotarsa solinvi-like virus 1* RdRp among all the compared sequences; the nucleotide sequence identity was 97% (1212/1251). We have also performed pairwise comparison of the full-length polyproteins encoded by *Leptinotarsa solinvi-like virus 1* genome and *Leptinotarsa decemlineata* TSA sequence. The total length of the polyprotein sequence encoded by *Leptinotarsa decemlineata* TSA is 3819 aa. The identity of the amino acid sequences of the compared polyproteins is 96.86% (3699 out of 3819 aa), most amino acid substitutions were found to be conservative (the sequence homology according to BLOSUM62 equals to 98.56%). As compared to the *Leptinotarsa solinvi-like virus 1* polyprotein, a fragment of 19 aa is inserted. The comparison of nucleotide sequences demonstrated that the *Leptinotarsa decemlineata* TSA sequence lacks terminal fragments, including the polyA tail. The polyprotein encoding nucleotide sequence of *Leptinotarsa decemlineata* TSA contains 480 SNPs, including 379 transitions, 101 transversions, 360 synonymous and 120 missense substitutions (with 55 being conservative), and 1 deletion (57 nt). Thus, it appears that the *Leptinotarsa decemlineata* TSA sequence (*GEEF01170301*) corresponds to a near full-length genomic sequence of a related solinvi-like virus. In addition to *Leptinotarsa decemlineata* TSA, the other closest to *Leptinotarsa solinvi-like virus 1* RdRp were the two RdRp sequences of *Hangzhou solinvi-like virus 1* (NCBI Protein: *UHR49768* and *UHR49784*), derived from metagenomic samples of *Altica cyanea* leaf-feeding beetle (Chrysomelidae) collected in China in 2016.

Thus, for the first time the full-length genome sequences of viruses belonging to the order Picornovirales were obtained from natural samples of Colorado potato beetle. The phylogenetic analysis allowed to attribute these viruses as representatives of the iflaviruses and solinvi-like viruses—*Leptinotarsa iflavirus 1* and *Leptinotarsa solinvi-like virus 1*, respectively. The analysis of amino acid sequences of polyproteins encoded by the genomes of these viruses confirmed the presence of characteristic viral proteins and demonstrated the correspondence of their domain structure to that of other described iflaviruses and solinvi-like viruses, respectively.

However, we cannot be sure that the observed *L. decemlineata* larvae death was caused by these particular viruses, although the external symptoms (delayed septicemia) and qPCR data (near 30-fold increase of viral load as compared to healthy insects), as well as the detection of these particular viruses in tissues of dead symptomatic larvae with NGS, speak in support of this hypothesis. Noteworthy is the death of Colorado potato beetle at a certain period of their development (at transition to the prepupa stage), and we assume this might be associated



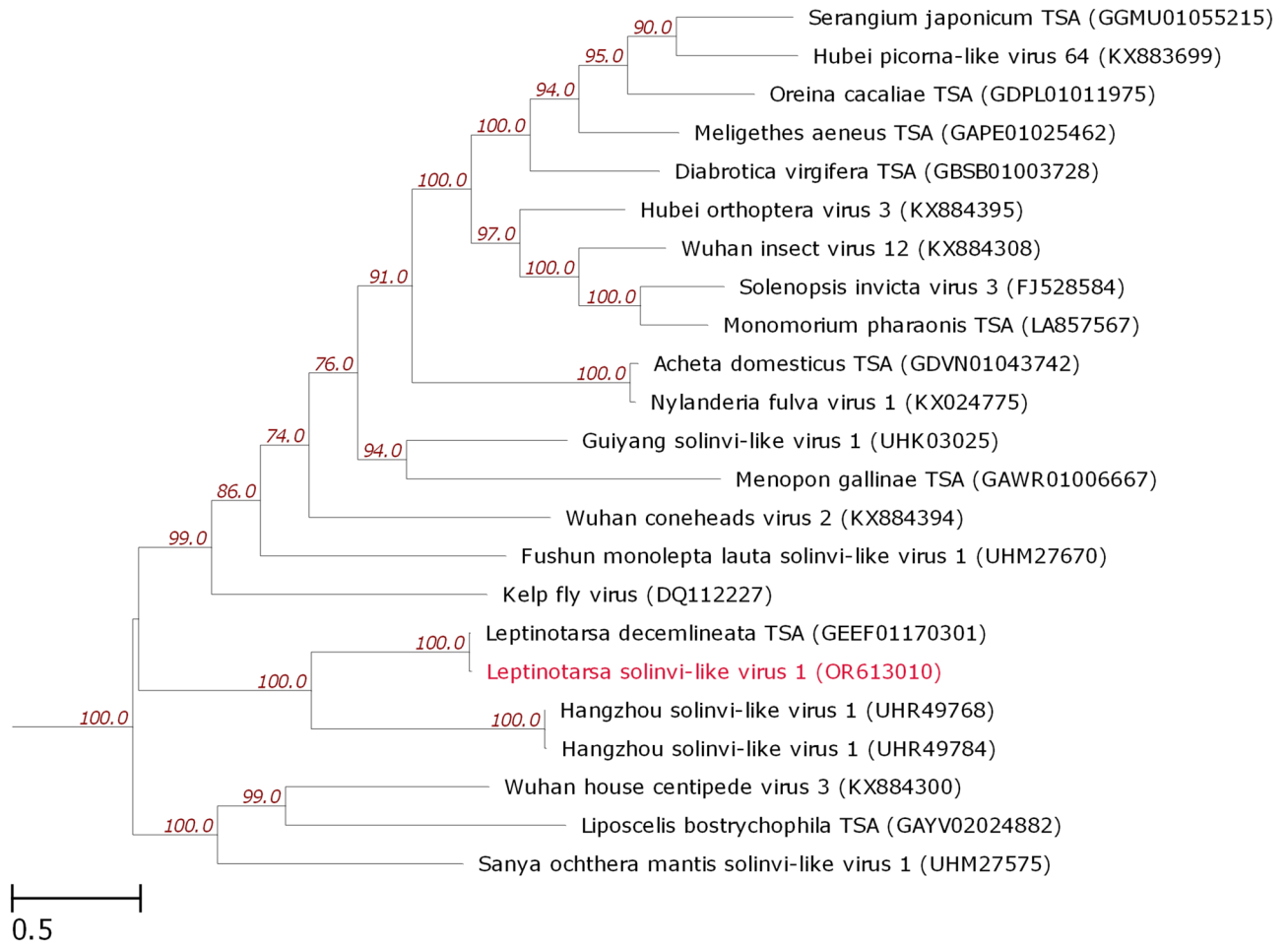
**Figure 4.** The phylogenetic tree of RdRp amino acid sequences of *Leptinotarsa iflavirus 1* and closely related iflaviruses and picorna-like viruses. The tree was constructed with IQ-Tree by the maximum likelihood method. The numbers next to the branching points denote the support indices ( $\geq 70\%$ ) determined by boot strap statistical analysis (1000 replications). The tree is drawn to scale, with branch lengths corresponding to evolutionary distances.

with increased insects' vulnerability to generalized viremia at the beginning of body restructuring period. Further studies will be aimed at isolating the viral particles and assaying their infectivity and pathogenicity in detail using both in vitro and in vivo models. We believe that our results may contribute to future development of novel Colorado potato beetle biological control methods.

## Materials and methods

### Biological samples collection

*Leptinotarsa decemlineata* III-IV instar larvae were collected from private potato fields free from any insecticide treatment (Karasuk, Novosibirsk region, Russian Federation; 53°43'N; 77°38'E). Due to these potato fields not being located in protected areas, there was no need for special permission to collect beetles. The landowners did not prevent access to the fields. Endangered or protected species were not used in this work. The insects were kept in laboratory conditions as described earlier<sup>6,7</sup>. The larvae were kept in ventilated plastic containers (300 ml; 10 larvae per container) at temperature of 24–25 °C and 16:8 h light/dark period daily fed with fresh *Solanum tuberosum* foliage. Insects that died in the prepupal stage with presumable viral infection symptoms (Fig. 1) were immediately frozen in liquid nitrogen and then stored at – 80 °C until RNA extraction.



**Figure 5.** The phylogenetic tree of RdRp amino acid sequences of *Leptinotarsa solinvi-like virus 1* and closely related solinvi-like viruses and solinvi-like viruses. The tree was constructed with IQ-Tree by the maximum likelihood method. The numbers next to the branching points denote the support indices ( $\geq 70\%$ ) determined by boot strap statistical analysis (1000 replications). The tree is drawn to scale, with branch lengths corresponding to evolutionary distances. The tree is a part of a bigger one (not shown).

### Samples preparation and qPCR

The whole larvae were frozen in liquid nitrogen and stored at  $-18\text{ }^{\circ}\text{C}$ . The samples were lyophilized at  $-53\text{ }^{\circ}\text{C}$  and 400 mTorr for 24 h. The lyophilized corpses were homogenized by micro pestles in liquid nitrogen and treated with Lira-reagent (BioLabMix Ltd., Novosibirsk, Russia). The subsequent total RNA extraction was performed according to Lira-reagent protocol. DNase treatment was carried out according to DNase I (RNase-free) kit (TransGen Biotech Co. Ltd., China) protocol. The reverse transcription of RNA to cDNA was performed with RevertAid<sup>TM</sup> M-MuLV Reverse Transcriptase (Fermentas, Vilnius, Lithuania) and 2.0 pMol of 9N primers.

qPCR was performed as described earlier<sup>14,15</sup>. The PCR reaction conditions were the following:  $95\text{ }^{\circ}\text{C}$  for 3 min, followed by 40 cycles of  $94\text{ }^{\circ}\text{C}$  for 15 s and  $64\text{ }^{\circ}\text{C}$  for 30 s, followed by the melting curve analysis ( $70\text{--}90\text{ }^{\circ}\text{C}$ ). The *L. decemlineata* ribosomal protein RP4 and RP18 genes were taken as reference, as it was proposed in<sup>13</sup>. The viral RNA-dependent RNA polymerase (RdRp) gene was selected to assess the viral load. Primers were designed with Primer-BLAST tool<sup>16</sup> and IDT OligoAnalyser 3.1<sup>17</sup>. The designed oligonucleotide primers were synthesized by Biosset Ltd. (Novosibirsk, Russia). The primer sequences are provided in Table 1.

### qPCR gene expression calculations and statistical analysis

Gene expression was calculated using the  $2^{-\Delta\Delta\text{Cq}}$  method with Bio-Rad CFX manager software (Bio-Rad, USA). The program Past 4.03 was used for analyses qPCR data. The statistical significance of observed differences in normalized viral RdRp expression values in dead larvae samples with presumable viral infection ( $N=5$ ) and in asymptomatic larvae samples ( $N=3$ ) was examined with two-sided Mann–Whitney test (the critical significance level was set to 0.05). The statistical analysis was performed with R statistical analysis environment (v.4.2.1)<sup>18</sup>.

### Samples preparation and sequencing

Individual dead CPB prepupae were homogenized in 1 ml of standard PBS manually in Eppendorf 1.5 ml micro-tubes using sterile Axigen disposable pestles (PES-15-B-SI, Axigen, USA). The homogenates were centrifuged

on Eppendorf Minispin centrifuge at 10000g for 10 min, the supernatants were collected into 15 ml tubes and diluted 10 times with sterile PBS. The resulting solutions were filtered through a Millex-HV filter pad with a pore diameter of 0.45 µm (Millex<sup>®</sup> PVDF syringe filter, Merck, USA) and the filter pads were additionally washed with 5 ml of sterile PBS. The resulting solutions were concentrated to a volume of 200 µl using Amicon<sup>®</sup> Ultra 4 50K centrifuge filters (Merk Millipore Ltd, Ireland) in Eppendorf 5804 centrifuge at 3000g with adapters for 15 ml tubes. The aliquots of 100 µl of the resulting solutions were supplied with 2 µl of MgCl<sub>2</sub> (100 mM), shaken and 10 µl of benzonase (Sigma, USA) with an activity of 250 units/µl was added. The suspension was incubated at 37 °C for 30 min and then EDTA was added to inactivate the benzonase.

Then total RNA was extracted from 100 µl of benzonase-treated sample with TriZol (ThermoFisher Scientific, USA) according to the recommended protocol. 40 µg of glycogen in the form of an aqueous solution with a concentration of 20 mg/ml was used as a co-precipitant. The precipitate was dissolved in 30 µl of deionized water, then the solution was used to obtain the double-stranded DNA fragments with SISPA (Sequence-Independent, Single-Primer Amplification) protocol, according to<sup>19</sup>.

The resulting dsDNA fragments were purified from unspent components and reaction products with AMPure beads (Beckman Coulter), and then used to prepare NGS libraries. Nucleic acid concentrations were measured with Qubit 3.0 using the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific). The NGS library preparation was carried out using the NEBNext Ultra II FS DNA Library Prep Kit for Illumina (NEB, USA), which performs fragmentation, end repair and dA-tailing, and adapter ligation with a single enzyme mix. The NEBNext Multiplex Oligos for Illumina (Index Primer Set 1) (NEB, USA) was used for multiplexing. The sequencing was performed on Illumina iSeq 100 platform (300 cycles).

### Bioinformatic analysis

The FastQ files obtained from sequencing were processed with fastp v0.20.0<sup>20</sup> to remove adapters, and short and low-quality sequences (with quality less than 20 and with length of no more than 30 nt). The preprocessed reads were assembled into contigs using MEGAHIT v1.2.9 assembler<sup>21</sup> with the meta-sensitive parameter enabled. The resulting contigs were analyzed using BLASTx v2.9.0+<sup>22</sup> against the NCBI nr database (<https://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/nr.gz>, accessed on August 2023). The quality of assembled viral contigs was checked by back-aligning the obtained reads with BWA MEM v0.7.17-r1188<sup>23</sup>. Target viral sequences were translated in 6 translation frames with EMBOSS transeq v6.6.0.0<sup>24</sup>. The open reading frames within the polyproteins were determined by homology with BLASTp v2.9.0+ (against NCBI nr database). The polyproteins domain annotation was carried out with InterProScan<sup>25</sup>. The genomic sequences annotations were visualized using the DnaFeatures-Viewer v3.1.2 Python3 package<sup>26</sup>. The phylogenetic analysis of *Leptinotarsa iflavirus 1* and *Leptinotarsa solinvi-like virus 1* RdRp proteins was performed using viral RdRp protein sequences with amino acid identity ≥ 30% selected with BLASTp from the NCBI nr database, and also the sequences taken from ICTV-provided alignment of *Soliniviridae* RdRp proteins (<https://ictv.global/sites/default/files/inline-images/Figure3-1.v2.aa.alignment.fst>, accessed on August 2023). The sequences were aligned with mafft v7.313<sup>27</sup> using the L-INS-i algorithm with 1000 iterations. The phylogenetic trees were constructed using the maximum likelihood method with IQ-Tree v2.0.3<sup>28</sup>. The following substitution models were used: LG + I + G4 for *Leptinotarsa iflavirus 1* and LG + F + R5 for *Leptinotarsa solinvi-like virus 1*. The selection of replacement models was carried out with ModelFinder algorithm<sup>29</sup>. Bootstrap analysis was performed with UFBoot algorithm<sup>30</sup> in 1000 replications. The phylogenetic trees were visualized with ete3 v3.1.2 Python3 package<sup>31</sup>.

### Data availability

The resulting FASTQ files were deposited to SRA with corresponding IDs: SRR25929780, SRR25929781, SRR25929782, SRR25929783, SRR25929784 (BioProject: PRJNA1013476). The data was made public on October 11, 2023.

Received: 29 October 2023; Accepted: 30 December 2023

Published online: 03 January 2024

### References

- Alyokhin, A., Baker, M., Mota-Sanchez, D., Dively, G. & Grafius, E. Colorado potato beetle resistance to insecticides. *Am. J. Potato Res.* **85**, 395–413. <https://doi.org/10.1007/s12230-008-9052-0> (2008).
- van Lenteren, J. C., Bolckmans, K., Köhl, J., Ravensberg, W. J. & Urbaneja, A. Biological control using invertebrates and microorganisms: Plenty of new opportunities. *BioControl* **63**, 39–59. <https://doi.org/10.1007/s10526-017-9801-4> (2018).
- Alyokhin, A. & Kryukov, V. Ecology of a potato field. In *Insect Pests Potato Glob. Perspect. Biol. Manag.* 451–462 (Elsevier, 2022). <https://doi.org/10.1016/B978-0-12-821237-0-00003-2>.
- Beas-Catena, A., Sánchez-Mirón, A., García-Camacho, F., Contreras-Gómez, A. & Molina-Grima, E. Baculovirus biopesticides: An overview. *J. Anim. Plant Sci.* **24**, 362–373 (2014).
- Starchevskaya, M. *et al.* The metagenomic analysis of viral diversity in Colorado potato beetle public NGS data. *Viruses* **15**, 1–21. <https://doi.org/10.3390/v15020395> (2023).
- Tomilova, O. G. *et al.* Changes in antifungal defence systems during the intermolt period in the Colorado potato beetle. *J. Insect Physiol.* **116**, 106–117. <https://doi.org/10.1016/j.jinsphys.2019.05.003> (2019).
- Kryukov, V. Y., Rotskaya, U. N., Yaroslavtseva, O. N., Noskov, Y. A. & Glupov, V. V. Expression of immunity- and stress-related genes during an intermolt period in the Colorado potato beetle. *Insects* **13**, 1168. <https://doi.org/10.3390/insects13121168> (2022).
- Le Gall, O. *et al.* Picornavirales, a proposed order of positive-sense single-stranded RNA viruses with a pseudo-T = 3 virion architecture. *Arch. Virol.* **153**, 715. <https://doi.org/10.1007/s00705-008-0041-x> (2008).
- Manual of Techniques in Invertebrate Pathology* (Elsevier, 2012) ISBN 9780123868992.
- De Miranda, J. R. *et al.* 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> (2010).

11. Brown, K. *et al.* ICTV virus taxonomy profile: Soliniviridae. *J. Gen. Virol.* **100**, 736–737. <https://doi.org/10.1099/jgv.0.001242> (2019).
12. Cruz-Flores, R., Andrade, T. P. D., Mai, H. N., Alenton, R. R. R. & Dhar, A. K. Identification of a novel solinivirus with nuclear localization associated with mass mortalities in cultured whiteleg shrimp (*Penaeus vannamei*). *Viruses* <https://doi.org/10.3390/v14102220> (2022).
13. Shi, X.-Q. *et al.* Validation of reference genes for expression analysis by quantitative real-time PCR in *Leptinotarsa decemlineata* (say). *BMC Res. Notes* **6**, 93. <https://doi.org/10.1186/1756-0500-6-93> (2013).
14. Kryukov, V. Y., Rotskaya, U. N., Yaroslavtseva, O. N., Noskov, Y. A. & Glupov, V. V. Expression of immunity- and stress-related genes during an intermolt period in the Colorado potato beetle. *Insects* <https://doi.org/10.3390/insects13121168> (2022).
15. Rotskaya, U. N., Kryukov, V. Y., Kosman, E., Tyurin, M. & Glupov, V. V. Identification of the ricin-B-lectin LdrBLk in the Colorado potato beetle and an analysis of its expression in response to fungal infections. *J. Fungi* <https://doi.org/10.3390/jof7050364> (2021).
16. Ye, J. *et al.* Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinform.* **13**, 134. <https://doi.org/10.1186/1471-2105-13-134> (2012).
17. Owczarzy, R. *et al.* IDT SciTools: A suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Res.* **36**, W163–W169. <https://doi.org/10.1093/nar/gkn198> (2008).
18. R Core Team R: A Language and Environment for Statistical Computing (2022).
19. Djikeng, A. *et al.* Viral genome sequencing by random priming methods. *BMC Genomics* **9**, 5. <https://doi.org/10.1186/1471-2164-9-5> (2008).
20. Chen, S., Zhou, Y., Chen, Y. & Gu, J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890. <https://doi.org/10.1093/bioinformatics/bty560> (2018).
21. Li, D., Liu, C. M., Luo, R., Sadakane, K. & Lam, T. W. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via Succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676. <https://doi.org/10.1093/bioinformatics/btv033> (2015).
22. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) (1990).
23. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324> (2009).
24. Rice, P., Longden, I. & Bleasby, A. EMBL: The European molecular biology open software suite. *Trends Genet.* **16**, 276–277. [https://doi.org/10.1016/S0168-9525\(00\)02024-2](https://doi.org/10.1016/S0168-9525(00)02024-2) (2000).
25. Jones, P. *et al.* InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240. <https://doi.org/10.1093/bioinformatics/btu031> (2014).
26. Zulkower, V. & Rosser, S. DNA features viewer: A sequence annotation formatting and plotting library for python. *Bioinformatics* **36**, 4350–4352. <https://doi.org/10.1093/bioinformatics/btaa213> (2020).
27. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780. <https://doi.org/10.1093/molbev/mst010> (2013).
28. Minh, B. Q. *et al.* IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534. <https://doi.org/10.1093/molbev/msaa015> (2020).
29. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermini, L. S. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589. <https://doi.org/10.1038/nmeth.4285> (2017).
30. Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522. <https://doi.org/10.1093/molbev/msx281> (2018).
31. Huerta-Cepas, J., Serra, F. & Bork, P. ETE 3: Reconstruction, analysis, and visualization of phylogenomic data. *Mol. Biol. Evol.* **33**, 1635–1638. <https://doi.org/10.1093/molbev/msw046> (2016).

## Acknowledgements

This work was supported by the Ministry of Science and Higher Education of the Russian Federation (agreement # 075-15-2019-1665). Insect collection, their laboratory maintenance, tissue sample preparation and qPCR were supported by the Russian Scientific Foundation (grant # 22-14-00309).

## Author contributions

Conceptualization, D.A. and M.A.; methodology, M.A., S.B., T.T., T.B., U.R., E.K., and D.A.; scripting and analytical pipelines, D.A. and M.A.; visualization, D.A., U.R. and V.K.; resources, and data curation, M.A., M.Az., V.K. and D.A.; biological samples collection, U.R. and V.K.; qPCR, U.R. and E.K.; DNA extraction and sequencing, T.T., T.B. and S.B.; original manuscript draft preparation, M.A., V.K. and D.A.; final manuscript review and editing, M.A., U.R., V.K. and D.A.; supervision, D.A.; funding acquisition, D.A. and V.K. All authors have read and agreed to the published version of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to M.A.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.





**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024