Environmental Microbiome

RESEARCH

Open Access

Wild and globally traded ornamental aquatic plants harbor diverse plant viruses, including notable crop pathogens

Lana Vogrinec[1,2], Katarina Bačnik[1], Martina Bačič[3], Nataša Mehle[1,4], Aimee R. Fowkes[5], Živa Lengar[1], Val Harju[5], Ian P. Adams[5], Adrian Fox[5] and Denis Kutnjak[1*]

Abstract

Background Aquatic plants play key roles in ecosystems, serving as primary producers and providing habitat for other aquatic life. While many are ecologically important, some invasive species, often introduced through the ornamental plant trade, pose threats to various aquatic ecosystems. Although viral infections have been documented to some extent in aquatic crops, the viral diversity in wild and ornamental aquatic plants remains largely unexplored. Investigating the viral communities of aquatic plants is important, as their direct contact with water allows for the potential long-distance transmission of stable viruses released from infected individuals. Invasive aquatic plants exacerbate this issue by introducing novel microbes, including viruses, to new regions, increasing the potential threat to native plant populations.

Results Here, we investigated the viral communities of diverse aquatic plants by mining publicly available transcriptome data of 79 wild aquatic species and sequencing the RNA from 14 plant species (some of them of different ornamental varieties), sourced from hobby aquascaping stores. Plant viruses from various families were detected in taxonomically diverse aquatic plants, ranging from algae to angiosperms. Alongside sequences of known crop pathogens, such as turnip yellows virus, cucumber mosaic virus, and lettuce chlorosis virus, we identified contigs of putative novel viral species belonging to several plant-infecting viral families. Most notably, we discovered sequences of known and novel begomoviruses, which may be causing observed ornamental phenotypes in two different aquatic plants. Further, we identified a novel potyvirus that appears to be globally present in multiple ornamental plants from the genus *Sagittaria*. We detected it in three plants sourced from online stores in Slovenia, as well as six plants intercepted during the import process into the UK.

Conclusions Our findings expand on the so far limited knowledge of aquatic plant viruses, revealing known and putative novel plant viral species across diverse aquatic plant taxa. The detection of crop viruses, including regulated pathogens, in ornamental aquatic plants highlights the risks associated with their unregulated global trade. Further research into viruses of aquatic plants may provide insights into their role in ecosystems as well as their potential impact on agriculture.

*Correspondence:
Denis Kutnjak
denis.kutnjak@nib.si

Full list of author information is available at the end of the article

Keywords Aquatic plants, Macrophytes, Viromes, High-throughput sequencing, Data mining

Background

Aquatic plants or macrophytes are multicellular photosynthetic organisms inhabiting freshwater and marine environments, including streams, slow rivers, wetlands, and seasonally flooded areas [1]. They encompass a wide range of taxonomically diverse groups, including macroalgae of phyla *Chlorophyta*, *Rhodophyta*, *Xanthophyta*, and *Phaeophyta*, bryophytes (*Bryophyta*), ferns (*Pteridophtya*), and seed plants (*Spermatophyta*) [2]. Essential to ecosystems, aquatic plants provide habitat for other aquatic organisms, impact hydrological regimes, and serve as primary producers [1]. In ecology, they are often used as biological indicators, offering insights into water quality and habitat conditions [3–5]. Some species (e.g., rice) are cultivated for consumption, thus also holding considerable economic value for society. Conversely, several aquatic plants, such as the water hyacinth (*Eichhornia crassipes*) [6, 7], rank among the most aggressive invasive species worldwide, posing a significant threat to freshwater ecosystems [2]. One of the primary sources for the introduction of invasive aquatic plants is the global trade of ornamental plants used as decorations in aquariums and ponds [8]. Several studies have shown that these plants are easily obtained through online stores and informal marketplaces where their sale remains unregulated [9, 10]. Imported plants can carry microbes, including viruses, which may survive long distances and establish themselves in new environments together with their host. This could result in the emergence of novel pathogens that threaten native populations of wild and cultivated aquatic species [11].

Although aquatic plants dominate numerous water-related ecosystems [12] and, as such, represent an abundant group of potential viral hosts, our knowledge of the viruses that infect them remains limited. Addressing this gap is important, as viruses harbored by these plants can enter nearby water bodies through roots or surface runoff during rainfall [13, 14], allowing them to travel over longer distances. Extensive research has shown that plant viruses are present in diverse types of water bodies [13], with some extremely stable viruses persisting in these environments over extended periods as infective virions [14, 15]. Moreover, experimental evidence suggests that healthy plants can be infected through their roots when exposed to infected water [14, 16] highlighting the role of water as a medium for virus transmission.

To provide a framework for our study, we reviewed the existing literature on viral infections in aquatic plants. Excluding the well-documented viral pathogens of rice (*Oryza sativa*) [17], we found evidence of 31 viruses associated with 18 aquatic plant species (Supplementary Table 1), emphasizing the limited understanding of viral diversity in this group of organisms. Notably, most of the studies have focused on viruses in aquatic crops, especially rice, which is known to be susceptible to over 25 viruses that significantly threaten its production [17]. Examples that strongly impact rice yields worldwide include rice black-streaked dwarf virus (*Fijivirus*) [18], southern rice black-streaked dwarf virus (*Fijivirus*) [19], rice yellow mottle virus (*Sobemovirus*) [20], and rice stripe necrosis virus (*Benyvirus*) [21]. In water chestnut (*Eleocharis dulcis*), a consumable aquatic vegetable, high-throughput sequencing (HTS) revealed the presence of an ilarvirus (water chestnut virus A) [22] and a soymovirus (water chestnut soymovirus 1) [23], while symptoms of chlorosis, mosaicism, and stunting have been linked to an infection with cucumber mosaic virus (CMV, *Cucumovirus*) [24]. Similarly, a symptomatic infection with CMV has been confirmed in lotus (*Nelumbo nucifera*) [25], which is also susceptible to dasheen mosaic virus (*Potyvirus*) [26] and apple stem grooving virus (*Capillovirus*) [27], both causing mild to severe symptoms. Additionally, sequences of a potyvirus (sweet potato latent virus) and several badnaviruses (lotus badnavirus 1–3) were detected in lotus plants with HTS [28].

While viral infections in aquatic crops have been documented to some extent, viruses of wild aquatic plants remain poorly understood. In 1975, a putative tobamovirus, termed Chara australis virus, was identified in samples of a large charophyte alga *Chara australis* [29, 30]. Subsequent analysis with HTS revealed unusual viral architecture, with different proteins showing the highest similarity to three different, distantly related viral taxa [31]. In England, leaf curling in common cordgrass (*Sporobolus anglicus*) was linked to an infection with a potyvirus, spartina mottle virus [32], while common reed plants (*Phragmites australis*) in Turkey with symptoms of mosaicism, chlorosis, and streaking were found to be infected with three crop pathogens: maize dwarf virus (*Potyvirus*), sugarcane mosaic virus (*Potyvirus*) and barley yellow dwarf virus-PAV (*Luteovirus*) [33].

The limited understanding of viral infections in wild aquatic plants extends to marine species. Seagrasses stand as a notable exception, having received more attention due to their ecological importance. Several seagrass viruses were identified through HTS and data mining, spanning various viral families such as *Alphaflexiviridae* [34, 35], *Amalgaviridae* [36], *Betaflexiviridae* [35, 37], *Bromoviridae* [38], *Endornaviridae* [38], *Rhabdoviridae* [39] and *Virgaviridae* [38]. Besides seagrasses, several kelp species from the orders *Laminariales* and *Ectocarpales* [40–42] were shown to harbor viruses from the

family *Phycodnaviridae*, which otherwise mainly infect unicellular eukaryotic algae.

To date, only one extensive metagenomic study specifically focusing on aquatic plant viromes has been conducted. Sampling of aquatic plants from four freshwater springs in Florida revealed several putative plant viral sequences, with the majority detected in a single species, *Vallisneria americana*. Two known crop pathogens, oat blue dwarf virus (*Marafivirus*) and potato virus Y (*Potyvirus*), were identified, while the majority of the remaining contigs belonged to novel members of the *Potyviridae* and *Partitiviridae* families [43]. Additionally, data mining efforts have contributed to viral discovery in aquatic plants to some extent. A key example is the comprehensive study performed by Mifsud et al., which utilized the 1000 Plant Transcriptome Initiative data to uncover novel viruses across diverse plant species, including some aquatic plants [44].

In this study, we aimed to broaden the knowledge of aquatic plant viruses and evaluate risks of possible viral spread connected to the global trade/transfer of some ornamental aquatic species. We used two approaches: (1) we mined the data from the 1000 Plant Transcriptome Initiative (1KP) to explore the viromes of diverse aquatic plants, including algae, non-vascular, and seed plants, and (2) we employed high-throughput RNA sequencing to detect viral sequences in 14 species of decorative aquatic plants commonly used in aquascaping.

Methods

Mining of aquatic plant transcriptomes for viral sequences

Collection of 1KP transcriptome data for selected aquatic plants

We first searched for viruses associated with aquatic plants by mining publicly available RNA sequencing data from the 1000 Plant Transcriptomes Initiative (1KP), which generated transcriptomes from 1153 plant species spanning various levels of diversity [42]. Due to their unavailability at the time of this study, 29 transcriptomes were not considered. From this dataset, we selected 79 aquatic plant species for analysis, ranging from macroalgae to angiosperms (Supplementary Table 2). The relevant data for chosen samples, including unique 1KP IDs, were retrieved from the sample list [42]. The assembled contigs generated by the 1KP were downloaded from the data repository (accessed 2022-12-07) [45]. The raw read data were acquired from the National Center for Biotechnology Information Sequence Read Archive (NCBI SRA) database using the wget command and data accession links (accessed 2022-12-08).

Identification, filtering, and classification of viral sequences

To identify putative viral sequences, the assembled contigs were first compared to the NCBI non-redundant (nr) database (accessed 2022-07-16) using DIAMOND (v. 0.9.34) [46] with the BLASTX alignment option (E-value = 10, word size = 3, matrix = BLOSUM62, gap costs = existence: 11 extension: 1). The last common ancestor (LCA) based taxonomic classification of DIAMOND results was performed using MEGAN6 [47] (v. 6.24.20-0, database megan-map-Feb2022). Putative viral contigs and their taxon paths were exported from MEGAN6 for further analysis. Raw reads were filtered for quality (limit = 0.05, maximum 2 ambiguous nucleotides allowed) and trimmed to remove the sequencing adapters using the CLC Genomics Workbench (CLC-GWB, v.23, QIAGEN, USA). Lastly, reads were mapped to assembled contigs with CLC-GWB v.23 (length fraction: 0.8, similarity fraction: 0.8) to determine the total mapped read count and average coverage, defined as the number of times each position in the contig was covered by mapped reads, averaged across the contig length.

Putative viral contigs shorter than 500 nucleotides (nt) were removed due to the less reliable taxonomical classification of such short sequences [48]. To ensure the use of the most up-to-date databases available at the time of analysis, putative viral contigs ≥ 500 nt were re-analyzed with a BLASTX similarity search against the NCBI nr database (accessed 2024-02-07) using BLAST + command line utility (v. 2.15.0, parameters: E-value = 10, word size = 3, matrix = BLOSUM62, gap costs = existence: 11 extension: 1). The BLASTX results were taxonomically classified with MEGAN7 (beta-release, v. 7.0.9) using the latest MEGAN-NCBI nr mapping database (megan-nr-r1) (Supplementary Table 3). Putative viral contigs identified through BLASTX and MEGAN7 were categorized into four groups according to the International Committee on Taxonomy of Viruses (ICTV) [49] host assignations: contigs that were classified in viral families with only plant hosts, in viral families with diverse hosts including plants, in viral families with non-plant hosts, and those that could not be classified to the family level. The families corresponding to each category are listed in Supplementary Table 4. The results were visualized using RStudio (v. 2024.09.1 + 394) and Corel Draw X7.

Contigs with similarity to members of viral families that infect only plants or those that have diverse hosts, including plants, were considered as putative plant viral contigs and were analyzed further. If multiple contigs produced identical first BLASTX hits and shared overlapping identical regions, these regions were aligned using MEGA11 [50] (v.11.0.13), and manual assembly was attempted to obtain longer sequences. For successfully re-assembled contigs (scaffolds), read mapping, BLASTX, and MEGAN7 analyses were repeated as described above. Final contig/scaffold sequences and the results of the BLASTX analysis are reported in Supplementary Table 5. Contigs/scaffolds likely representing complete or

near-complete viral genomes were identified by comparing their length to the expected genome sizes for the relevant viral genus or family. For these contigs, an additional search with the BLASTN against the NCBI core nucleotide database (accessed 2024-10-16) was performed (E-value = 0.05, word size = 11, gap costs = existence: 5 extension: 2). Pairwise identities between each contig and their closest BLASTN hit were calculated using SDT (v1.2) [51] to determine whether they represent known or putative novel viral species (Supplementary Table 6). The type of sequence used for SDT (nucleotide/protein, complete sequence/specific gene product), and the classification requirements were based on the species demarcation criteria for each genus or family, as reported by the ICTV [49].

Detection of viruses in ornamental aquatic plants using high-throughput sequencing

Collection and processing of aquatic plants, selected for sequencing

To explore the presence of viruses in globally traded aquatic plants, we acquired 17 different ornamental aquatic plants belonging to 14 species, either from online stores or from a home aquarium (Table 1). For three species, two variations per species were included: a 'wild-type' variation without any obvious ornamental patterns (*Hygrophila polysperma*, *Alternanthera reineckii* 'mini', *Echinodorus cordifolius*) and a variation that exhibited ornamental patterns resembling symptoms of viral infections (*Hygrophila polysperma* 'rosanervig', *Alternanthera reineckii* 'rosanervig', *Echinodorus cordifolius* 'marble queen'). For each sample, 500–1000 mg of fresh leaf material was collected. In case of evident algae growth, the surface of the leaves was gently cleaned with a clean paper towel beforehand.

RNA extraction, library construction, and sequencing

RNA extraction was performed from the collected leaf material for downstream analysis. Leaf tissue was homogenized using the FastPrep-24 lysis system (MP Biomedicals, USA), and RNA was isolated from each plant sample using the RNeasy Plant Mini Kit (QIAGEN, USA) according to the manufacturer's instructions. Quality and quantity of the RNA were checked using Nanodrop ONE spectrophotometer (Thermo Fischer Scientific, USA) and QUBIT3 fluorometer (Invitrogen, USA), using the High Sensitivity RNA kit (Thermo Fischer Scientific, USA).

Extracted RNA was sequenced, either in the form of pooled samples (AQ1-AQ4) that contained up to four individual plant RNA extracts or in the form of individual RNA extracts (AQ5-AQ7). The exact composition of samples and the origin of plants are listed in Table 1. Sample pool AQ1 consisted of plants from a home aquarium that contained various other organisms, such as snails, fish, and flatworms. Sample pools AQ2-AQ4 and samples AQ5-AQ7 contained plants that were bought from aquaristic stores, where they were either grown in soil or sterile tissue cultures. The total RNA extracted

Table 1 List of sequenced aquatic plants. Variation, botanical family, and inclusion in a corresponding sequencing sample are given, along with information about the source of plants and observed phenotypes. Sample pools AQ1 – AQ4 contained up to four plant species, while the remaining samples contained only a single species. Plants marked with an * were grown in tissue culture under sterile conditions before purchase

Species	Variation	Family	Corresponding sample / sample pool designation	Source	Ornamental patterns / viral-like symptoms
Anubias barteri		*Araceae*	AQ1	home aquarium	none
Lemna minor		*Araceae*	AQ1	home aquarium	none
Cryptocoryne sp.		*Araceae*	AQ1	home aquarium	none
Lindernia rotundifolia		*Linderniaceae*	AQ2	online store	vein yellowing
Echinodorus cordifolius	marble queen	*Alismataceae*	AQ2	online store	a marbling pattern of darker and lighter green patches
Hygrophila polysperma	rosanervig	*Acanthaceae*	AQ2	online store	vein yellowing
Alternanthera reineckii	rosanervig	*Amaranthaceae*	AQ2	online store	vein yellowing
*Alternanthera reineckii**	mini	*Amaranthaceae*	AQ3	online store	none
Hygrophila polysperma		*Acanthaceae*	AQ3	online store	none
Echinodorus cordifolius		*Alismataceae*	AQ3	online store	none
Bacopa monnieri		*Plantaginaceae*	AQ4	online store	none
*Eleocharis acicularis**		*Cyperaceae*	AQ4	online store	none
*Lobelia cardinalis**	mini	*Campanulaceae*	AQ4	online store	none
*Myriophyllum sp.**	Guyana	*Haloragaceae*	AQ5	online store	none
*Sagittaria subulata**		*Alismataceae*	AQ6	online store	none
Ludwigia repens		*Onagraceae*	AQ7	online store	none

from *Phaseolus vulgaris* co-infected with Phaseolus vulgaris alphaendornaviruses 1–3 was used as an alien control [52].

The sample pools and samples were sent to Novogene (China) for library preparation and high-throughput sequencing. There, total RNA was first depleted of ribosomal RNA, followed by sequencing library preparation using TruSeq RNA Library Prep Kit (Illumina, USA). Libraries were shotgun sequenced using the Novaseq X Plus platform (Illumina, USA).

Contig assembly, classification, and characterization of putative viral sequences

To identify viral sequences in the samples, reads were assembled into contigs and analyzed. Following adapter trimming and quality filtering with CLC-GWB v.23 (QIAGEN, USA), the assembly of reads was performed with SPAdes [53] (v. 3.15.5) using the rnaviralspades.py mode. Using the same pipeline and criteria as described in the data mining section of Methods, putative viral contigs longer than 500 nt were identified (Supplementary Table 8) and grouped into four categories based on family-level classification. Relative abundance of reads for each viral family was calculated for each sample based on the number of reads mapping to contigs longer than 500 nt (Supplementary Table 8). This value was calculated by dividing the sum of reads mapped to such contigs for each viral family by the total number of reads in the sample. Using the approach outlined above, contigs from plant-associated viral families were examined in detail. Contigs with identical BLASTX hits were aligned in the overlapping regions with MEGA11 [50] (v.11.0.13), and manually reassembled, when overlapping to obtain longer scaffolds. The final list of putative plant virus contigs/scaffolds is reported in Supplementary Table 9.

To further characterize putative plant virus contigs, open reading frames (ORF) were predicted using SnapGene (v. 6.1.2.) with the following conditions: minimum length = 75 amino acids; START codon required, except at DNA ends. To determine whether the contigs contain complete coding sequences (CDS), the predicted ORF structure was compared to the genome organization of the members of the families into which they were classified. For contigs with complete CDS, an additional BLASTN search and pairwise analysis with SDT (v1.2) [51] were performed as described in the data mining section of Methods, to determine whether they represent known or putative novel viral species (Supplementary Table 6) [49]. The polyprotein amino acid sequence of Sagittaria virus A (SgVA) was aligned with that of potato virus A using MEGA11 [50] (v.11.0.13) to predict polyprotein cleavage sites and the associated mature proteins, using UniProt [54] entry Q85197 as reference.

Screening of imported Sagittaria plants for viruses

In addition to ornamental aquatic plants acquired online, 19 plants from the genus *Sagittaria* were intercepted upon import by UK Plant Health and Seed Inspectors at Manchester airport and sent to Fera Science Ltd for testing (Supplementary Table 11). The samples were screened by ELISA for various viruses and tested positive for potyvirus. For ELISA testing, samples were tested in duplicate wells on Nunc Maxisorp microtitre plates (Merck, Germany), following the manufacturer's instructions. As no uninfected *Saggittaria* material was available, petunia leaves were used as the negative control. ELISA reactions were assessed using a spectrophotometer microplate reader (Thermo Fischer Scientific, USA) at $\lambda = 405$ nm (A_{405}). Samples were considered positive when the mean absorbance values were greater than three times the A_{405} of the negative control. To identify the virus, nucleic acids were extracted using Invimag Virus RNA/DNA Mini-kit (Invitek GmbH, Germany) and tested by RT-PCR for potyviruses using primers developed by Van der Vlugt et al. (1999) [55]. (Supplementary Table 10). The resulting PCR products of positive samples were sent for Sanger sequencing (MWB GmbH, Germany). Two of the positive samples were also tested by HTS using a TruSeq kit on the Illumina MiSeq as described in Fowkes et al. (2021) [56].

DNA sequencing of plants with detected begomoviral sequences

To obtain complete genomic sequences of begomoviruses detected in the sample pool AQ2, DNA sequencing was performed on individual plants from this pool. Following the homogenization with the FastPrep-24 lysis system (MP Biomedicals, USA), DNA was extracted from frozen plant tissue with QuickPick™ SML Plant DNA kit (BioNobile, Finland) and KingFisher mL Purification System (Thermo Fischer Scientific, USA) [57]. DNA extracts of individual plants (*L. rotundifolia*, *E. cordifolius* 'marble queen', *H. polysperma* 'rosanervig', *A. reineckii* 'rosanervig') were used for library preparation and sequencing, performed by Novogene (China) using the NovaSeq X Plus platform (Illumina, USA), yielding 6–55 million 150 bp long paired-end reads per sample. Quality filtering of raw reads and adapter trimming were performed using CLC-GWB v.23 (QIAGEN, USA). The reads were assembled into contigs with SPAdes [53] using default settings (v. 3.15.5) and further manual corrections as described in the data mining section of the Methods.

Detection of plant viruses in individual ornamental aquatic plant samples

To identify potential host plants, PCR assays were developed for putative plant virus contigs with complete CDS using the Primer3 web tool [58] (v. 4.1.0). PCRs

were performed on extracted nucleic acids of individual plants from the sample pools in which such viruses were detected. Additionally, two more *Sagittaria* plants were acquired (*S. terres*, *S. subulata 'pusilla'*) from a different online store and tested for SgVA using the designed PCR assay for this virus. PCR cycling conditions and primer sequences for each assay are listed in Supplementary Table 10. For RNA viruses, RT-PCRs were performed on extracted RNA with the OneStep RT-PCR kit (QIAGEN, USA) on the ProFlex PCR system (Thermo Fisher Scientific, USA) following the manufacturer's instructions. Positive controls (PKI-1, PKI-2) consisted of DNA extracted from *Chenopodium quinoa* leaves infected with Schlumbergera virus X amplified with virus-specific primers. For begomoviruses, PCRs were performed on extracted DNA using the Platinum *Taq* DNA polymerase (Thermo Fisher Scientific, USA) and the GeneAmp PCR System 9700 (Thermo Fisher Scientific, USA) according to the manufacturer's instructions. A PCR assay using SPG1/SPG2 universal begomoviral primers [59] and the DNA extracted from tomato leaf curl New Delhi begomovirus was used as a positive control (PKI-3). Three additional PCR assays for begomoviruses recommended in EPPO PM7/152 were also performed. The procedures were adapted from Wyatt and Brown (1996), Li et al. (2004), and Saison and Gentit (2015) [59–61]. Amplicons were visualized using agarose electrophoresis in a 1% agarose gel stained with ethidium bromide.

Rolling circle amplification and restriction digestion analysis

To confirm that begomoviruses, their satellites and defective molecules are present in plants in episomal circular form, a rolling circle amplification (RCA) was performed on DNA extracted from plants associated with begomoviral sequences using the TempliPhi Amplification kit (Cytiva, USA) as per the manufacturer's instructions. The pUC19 vector provided in the kit was used as a positive control. DNA concentration after the RCA was measured using the Nanodrop ONE spectrophotometer (Thermo Fischer Scientific, USA). Restriction digestion of RCA products was performed to verify the presence of the expected begomoviral sequences. Enzymes with a single restriction site per viral or satellite sequence were selected to produce linearized products corresponding to the full genome size (EcoRI for *Lindernia rotundifolia* extracts and HindIII for *Alternantherea reineckii 'rosanervig'* extracts). Restriction reactions were prepared in a total volume of 50 µl, consisting of 5 µl of enzyme-specific buffer (EcoRI: 10 x NEBuffer EcoRI for EcoRI, HindIII: 10 x rCutSmart), 1 µl of enzyme, 2 µl of amplified DNA, and 42 µl of nuclease-free water. Digestion was performed at 37 °C for 15 min, followed by heat inactivation for 20 min at 65 °C (EcoRI) or 80 °C (HindIII). Digested fragments were visualized on Agilent TapeStation Agilent Tape Station (Agilent Technologies, USA), using the D5000 ScreenTape Assay according to the manufacturer's instructions.

Transmission electron microscopy (TEM)

Visualization of SgVA virions was performed with TEM on leaf tissue from plants infected with SgVA. Leaves showing disease symptoms were selected, and approximately 3 mm pieces were excised and transferred to tubes with 300 µl of 0.1 M phosphate buffer (40 mM $NaH_2PO4 \times 2H2O$, 50 mM Na_2HPO_4, 2% PVP-10.000, pH 6.8). The tissue was homogenized using a plastic rod and incubated at room temperature for 5 min. To prepare the grid, 20 µl of a sample was applied to a freshly glow-discharged copper grid (400 mesh, formvar-carbon coated) for 5 min, then washed, and stained with one droplet of 1% (w/v) water solution of uranyl acetate. Two grids were prepared and observed with a transmission electron microscope TALOS L120 (Thermo Fisher Scientific, USA), operating at 100 kV. Analysis was performed on 5–10 grid squares which were analyzed at various locations, with micrographs taken randomly with the Ceta 16 M camera (Thermo Fisher Scientific, USA).

Phylogenetic analysis of selected putative novel viruses

Phylogenetic analysis was performed to determine the phylogenetic relationships of two newly discovered viruses with their closely related species.

For SgVA, the ICTV Virus Metadata Resource Table was used (VMR, v.21-221122_MSL37) to identify species from the *Potyvirus* genus (exemplar isolates) with complete coding sequences available in the NCBI Virus RefSeq database. The list was further expanded by performing BLASTX as described above and adding the hits that were missing from the VMR. The polyprotein amino acid sequences of the four SgVA isolates were aligned with those of the viruses in the compiled list, using ryegrass mosaic virus (*Rymovirus*) as an outgroup. The alignment was performed in MEGA11 [50] (v.11.0.13), using the MUSCLE plug-in, followed by trimming with TrimAl [62] (v.1.4, method automated1). The trimmed alignment is available in Supplementary File 1. Using the alignment, a maximum likelihood tree was constructed using IQ-TREE [63] (v.2) with the automatically determined best-fit substitution model (LG + F + I + G4) and the ultrafast bootstrap approximation with 1000 replicates [64]. The tree was visualized with iTol (v.6) [65]. Using SDT [51] (v1.2), the pairwise identity for polyprotein amino acid sequences was calculated for the four SgVA isolates (NIB, 1KP, FERA-1, FERA-2) and the most closely related potyviruses from the tobacco etch virus (TEV) clade.

For Lindernia rotundifolia yellow vein virus (LRYVV), phylogenetic analysis was performed as described in Brown et al., 2015 [66]. A BLASTN search of the

complete DNA-A sequence of LRYVV against the NCBI nr nucleotide database (accessed 2024-07-24) was first performed to select the closely related viruses for phylogenetic analysis. The first 100 hits were extracted and, in case of multiple hits for the same virus (e.g., different isolates), the one with the highest percentage identity to LRYVV was selected. The alignment was performed in MEGA11 [50] (v.11.0.13) with the MUSCLE plug-in using the DNA-A sequence of LRYVV and those of its closest BLASTN hits. The complete alignment is available in Supplementary File 2. This alignment was then used to construct the phylogenetic network using SplitsTree [67] CE (v. 6.3.32) with the P distance [68] and Neighbor-Net [69] methods with default parameters. The bootstrap splits method with default options was applied to calculate bootstrap supports [70]. Additionally, pairwise identities were calculated for the complete DNA-A nucleotide sequences of LRYVV and its closest BLASTN hits using SDT [51] (v1.2).

Results

Transcriptome mining reveals the presence of viral contigs across taxonomically diverse aquatic plants

We searched for viral sequences in the transcriptomes of 79 aquatic plants spanning a broad taxonomic range, from algae to angiosperms (Supplementary Table 2). A total of 746 contigs were classified as viral with MEGAN, out of which 280 were longer than 500 nucleotides and were analyzed further (Supplementary Table 3, Fig. 1). Viral contigs were present in 76% of the analyzed transcriptomes (Fig. 1a), encompassing all groups of aquatic plants except clubmosses, for which no viral sequences were detected.

Contigs that MEGAN classified into families with non-plant hosts were identified in most groups and were particularly prevalent in algae (Fig. 1b). Similarly, contigs that could not be assigned to a viral family were also detected across all groups, with several classified only at the domain (Viruses) or realm (*Riboviria*) level (Fig. 1b, Supplementary Table 3).

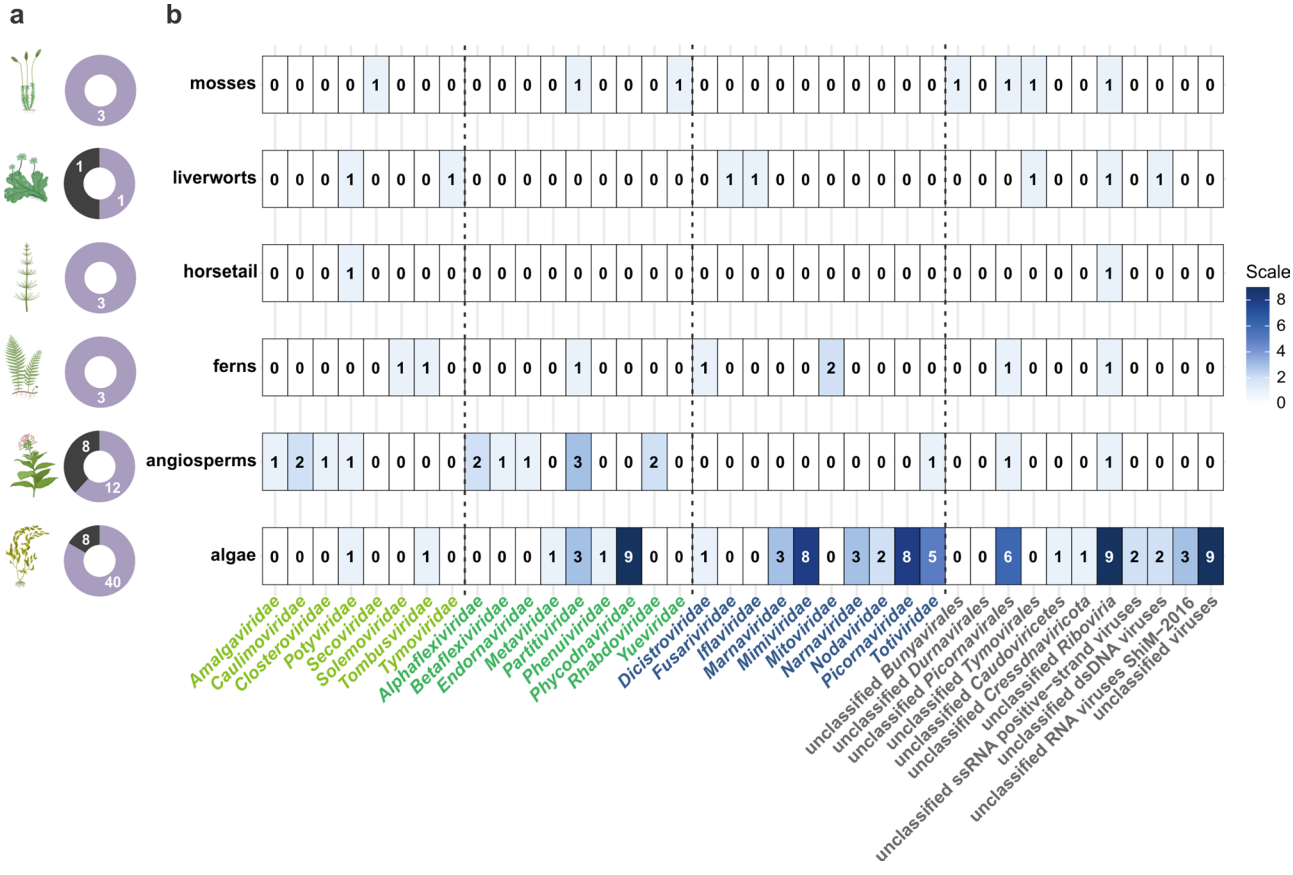

Fig. 1 Detection of viral contigs in aquatic plant transcriptomes from the 1000 Plant Transcriptome Initiative (1KP). a) Number of species with (purple portion) and without (black portion) detected viral contigs (longer than 500 nt) per plant group. The sum of both sections reflects the total number of species analyzed per plant group. b) A graphical representation of plant species with detected viral contigs (longer than 500 nt) across viral families for each plant group. Numbers in squares and the heatmap color code represent the number of analyzed species per plant group in which viral contigs (> 500 nt) were detected. The colors of taxons on the y-axis represent four categories, into which viral contigs (> 500 nt) were sorted based on family-level taxonomic classification with MEGAN: those from families with exclusively plant hosts (light green), those from families with diverse hosts, including plants (dark green), those from families with non-plant hosts (blue) and those not classified to family level (grey). The four categories are also distinguished by dotted lines

Transcriptomes from all groups of analyzed aquatic plants contained viral contigs, which were classified into families with exclusively plant hosts using MEGAN (Fig. 1b). They were most prevalent in angiosperms, where they were detected in 25% of analyzed species. Similarly, contigs from families associated with diverse hosts, including plants, were also most frequently found in angiosperms (45% of analyzed species), followed by algae (29% of analyzed species) (Fig. 1a, b). A substantial fraction of algae transcriptomes contained contigs with the highest identities to members of the *Phycodnaviridae* family, including several that were classified into the genus *Phaeovirus*. The protein sequence identities to the closest BLASTX hits for *Phycodnaviridae* contigs were relatively low (13–66%), suggesting that these represent novel viral species (Supplementary Table 3).

Despite examining only a few transcriptomes of ferns, liverworts, and mosses, we detected putative plant virus contigs in almost all of them (Fig. 1a). In heart-leaved spearmoss *Calliergon cordifolium*, several contigs were taxonomically placed with MEGAN as 'unclassified *Bunyavirales*' (recently promoted to class *Bunyaviricetes* [71]) (Supplementary Table 3). Five of those were highly similar to known viruses (Heart-leaved spear-moss bunyaviruses 1–5), previously discovered by Mifsud et al. in the same dataset [44]. The remaining sequences ($N = 25$) exhibited the highest similarity to several viruses from the *Bunyavirales* order, which were previously identified in the transcriptomes of a liverwort (*Odontoschisma prostratum)* and two mosses (*Aulacomnium heterostichum*, *Syntrichia princeps*) [44]. In BLASTX analysis, the amino acid identities of these contigs to their closest hits ranged from 31 to 84% (Supplementary Table 3). The closest hits include brown screw moss bunyavirus, Aulacomnium heterostichum bunyaviruses 1, 2, and 4, and liverwort-associated bunyaviruses 2–4.

Aquatic plant transcriptomes contain sequences of known plant viruses as well as several putative novel plant viruses

In line with the objectives of this study, we primarily focused our investigation on viruses with the potential to infect plants. We detected either partial or complete coding sequences of known and possible novel plant viruses in the transcriptomes of various aquatic plants. Key examples of viral sequences that were classified at least to the family level are detailed below. A complete overview of putative plant virus contigs, along with the results of read mapping and the BLASTX analysis, is reported in Supplementary Table 5.

In the transcriptomes of non-flowering plants, we uncovered an array of viral sequences, providing insights into potential plant viruses associated with this group. Several plant virus contigs were detected in the transcriptome of the liverwort *Scapania nemorosa*. Among them, one (contig ID 2153071_2144868_2164019, Supplementary Table 5) exhibited high similarity to common milkweed potyvirus (GenBank accession OX380443, genus *Potyvirus*, family *Potyviridae*) with 99.9% identity in the aligned region of the polyprotein amino acid sequence. Common milkweed potyvirus was first identified through data mining in the angiosperm *Asclepias syriaca*, utilizing the same 1KP dataset as examined here [44]. In the transcriptome of the moss *Scuoleria aquatica*, we detected a sequence (contig ID 2010568, Supplementary Table 5) that shared 61% amino acid identity in the aligned region of the RNA-dependent RNA polymerase (RdRp) gene to the goldenrod fern yue-like virus (GenBank accession OX380389, family *Yueviridae*), which was originally detected in the 1KP transcriptome of the fern *Pityrogramma trifoliata* [44]. In the dataset from the fern *Pilularia globulifera*, we identified a contig corresponding to the known crop pathogen, turnip yellows virus (GenBank accession JQ862472, genus *Polerovirus*, family *Solemoviridae*) [72]. The results of the SDT analysis confirmed its classification within the same viral species, *Polerovirus TUYV* (Supplementary Table 6).

The greatest richness of viral sequences with putative plant hosts was observed in angiosperms (Table 2). In the transcriptome of the seagrass *Posidonia australis*, we detected the complete coding sequence of a putative novel member of the *Betaflexiviridae* family (contig ID 2013547) (Supplementary Table 6). In BLASTX analysis, the contig had 46% amino acid identity in the aligned region to the RdRp gene sequence of banana mild mosaic virus (Table 2). In the transcriptome of seaside arrowgrass *Triglochin maritima*, we identified several contigs with the highest identities to cryptic viruses of the *Partitiviridae* family (Table 2). Two of those represented complete RNA1 (contig ID 2002317) and RNA2 (contig ID 2002316) coding sequences of putative novel viruses (Supplementary Table 6), while one partial contig (contig ID 2084669_208467) had 99% amino acid identity in the aligned region to the RdRp gene of beet cryptic virus 1 (genus *Alphapartitivirus*, family *Partitiviridae*) based on the BLASTX analysis. We discovered the complete genome of the orchid pathogen, Cymbidium mosaic virus [73] (contig ID 2004053_2003800, genus *Potexvirus*, family *Alphaflexiviridae*) (Table 2, Supplementary Table 6), in the dataset from *Cyperus papyrus*. Furthermore, a partial sequence was detected in *Myriophyllum aquaticum* (contig ID 2022808), which exhibited 100% amino acid identity in the aligned region of the minor coat protein to lettuce chlorosis virus, a known crop pathogen (genus *Crinivirus*, family *Closteroviridae*) (Table 2). The transcriptome of *Sagittaria latifolia* revealed several putative plant virus contigs, including the complete genome of Schlumbergera virus X (contig ID 2008980, genus *Potexvirus*, family *Alphaflexiviridae*) (Table 2, Supplementary

Table 2 Putative plant virus contigs/scaffolds detected in angiosperm transcriptomes and the results of the BLASTX analysis. Scaffolds that were manually re-assembled from contigs to obtain longer sequences are marked with combined IDs (e.g. 2028716_2143514), denoting the original sequences. [a] denotes viruses lacking the ICTV taxonomic classification

plant species	contig ID	contig length	closest BLASTX hit	accession number of the closest BLASTX hit	% identity to the closest BLASTX hit	MEGAN LCA-based taxonomic classification
Bacopa caroliniana	2177945	505	carnation etched ring virus	NP_612577	79.52	*Caulimovirus*
Bacopa caroliniana	2039416	569	carnation etched ring virus	ABX80504	75.00	*Caulimovirus*
Bacopa caroliniana	2017771	1559	carnation etched ring virus	CAH68829	74.86	*Caulimovirus*
Bacopa caroliniana	2039415	1401	cauliflower mosaic virus	BAO53218	76.04	*Caulimovirus tessellobrassicae*
Bacopa caroliniana	2019346	1671	soybean Putnam virus	YP_006607893	37.38	*Caulimovirus*
Ceratophyllum demersum	2005427	957	grapevine Roditis leaf discoloration-associated virus	AMN88380	49.47	*Badnavirus*
Cyperus papyrus	2004053_2003800	6238	Cymbidium mosaic virus	QQY02623	98.87	*Potexvirus cymbidii*
Hydrocotyle umbellata	2028716_2143514	13,181	Hydrocotyle umbellata endornavirus	QPO25406	98.81	Hydrocotyle umbellata endornavirus[a]
Myriophyllum aquaticum	2013022	1593	Cucumis virus 1	DAZ90697	41.67	unclassified *Rhabdoviridae*
Myriophyllum aquaticum	2022808	2655	lettuce chlorosis virus	QHB15134	100.00	*Crinivirus lactucachlorosi*
Posidonia australis	2013547	7627	banmivirus BanMMV	WAB21307	46.29	*Banmivirus musae*
Posidonia australis	2060969	1455	Dactylorhiza cryptic virus 2	QED42874	50.13	Dactylorhiza cryptic virus 2[a]
Sagittaria latifolia	2006917_2002862_2002861	7093	Bacopa monnieri virus 2	YP_010798893	45.93	*Rhabdoviridae*
Sagittaria latifolia	2006916	1403	Cuscuta virus 1	DAF42410	42.34	*Rhabdoviridae*
Sagittaria latifolia	2008219 (GenBank: PV019449)	9616	potato virus A	QPC97649	48.10	*Potyvirus atuberosi*
Sagittaria latifolia	2008980	6573	Schlumbergera virus X	AJF19167	97.61	*Potexvirus ecschlumbergerae*
Triglochin maritima	2085810	1503	Amaranthus tuberculatus amalgavirus 2	DAZ90991	61.69	*Amalgavirus*
Triglochin maritima	2084669_2084671	1388	beet cryptic virus 1	YP_002308574	99.13	*Alphapartitivirus betae*
Triglochin maritima	2002317	1996	beet cryptic virus 1	YP_002308574	82.96	*Partitiviridae*
Triglochin maritima	2002316	1722	carrot cryptic virus	YP_009508045	62.08	*Partitiviridae*
Triglochin maritima	2015565	1328	Fragaria chiloensis cryptic virus	QZN83647	81.57	unclassified *Partitiviridae*
Triglochin maritima	2018869	1213	Fragaria chiloensis cryptic virus	YP_001274393	63.96	unclassified *Partitiviridae*
Triglochin maritima	2085303	846	Rhodiola cryptic virus 1	QED42887	67.16	Rhodiola cryptic virus 1[a]
Triglochin maritima	2085775	1377	Rhodiola cryptic virus 1	QED42887	48.20	Rhodiola cryptic virus 1[a]
Triglochin maritima	2018450	1426	Rose cryptic virus 1	YP_001686788	57.57	unclassified *Partitiviridae*
Triglochin maritima	2007048	575	Rose partitivirus	ANQ45204	60.34	*Partitiviridae*
Utricularia sp.	2135711	641	Heterobasidion partitivirus 3	YP_009508058	67.65	*Partitiviridae*

Table 6). In the same plant, we detected two partial sequences resembling members of the *Rhabdoviridae* family (contig IDs 2006917_2002862_2002861, 2006916) (Table 2). Finally, we also discovered a sequence of a putative novel potyvirus, designated Sagittaria virus A (isolate 1KP, contig ID 2008219, GenBank accession PV019449) (Table 2, Supplementary Table 6) in *S. latifolia*. The sequence of Sagittaria virus A had 48% amino acid identity in the aligned region of the RdRp gene to its closest BLASTX hit, potato virus A (genus *Potyvirus*, family *Potyviridae*).

Ornamental aquatic plants used in aquascaping contain various viral nucleic acids, including those from known and novel plant viruses

Next, we explored the viromes of ornamental aquatic plants through high-throughput sequencing, focusing on globally traded species used for decorative purposes in aquariums and ponds. We sequenced 7 samples containing 17 different plants commonly used in aquascaping. A summary of sequencing data, including raw and trimmed read counts, total and viral contig numbers, and percentages of viral reads per sample, is listed in Supplementary Table 7. Sequencing resulted in 18,052,640–55,249,032 150 bp long paired-end reads per library. Mapping of reads to viral contigs revealed, on average 1.99% viral reads per sample, ranging from 0.0001 to 11.57%. Apart from the alphaendornaviruses known to be present in the sample, no other viral contigs were assembled from the alien control sequencing data, indicating an absence of cross-contamination. Out of the 596 putative viral contigs initially detected across samples, 192 remained after filtering (> 500 nt) and additional BLASTX analysis. We categorized these contigs into four groups based on MEGAN family-level classification and predicted host ranges as described in Methods (Fig. 2a, b, Supplementary Table 8).

Similar to the outcomes of data mining, contigs that were classified into families with predicted non-plant hosts were identified in most samples (Fig. 2a, b). Additionally, contigs that could not be classified at the family level were identified in four of the analyzed samples

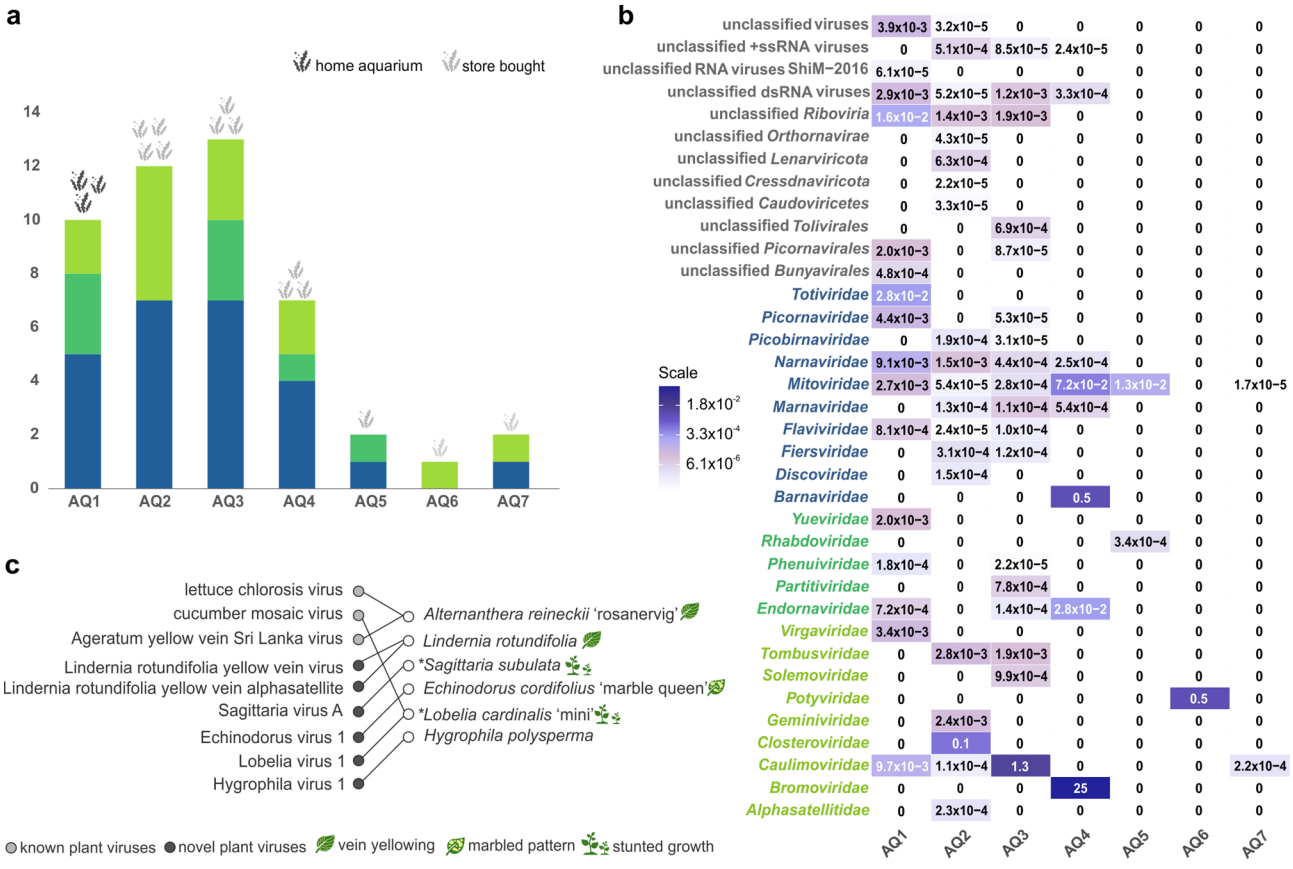

Fig. 2 Composition and distribution of viruses detected in ornamental aquatic plants with high-throughput sequencing. a) Viral family richness across analyzed samples. Stacked bar plots represent the number of viral families detected per sample based on the classification of contigs (> 500 nt) with MEGAN, including only the contigs that were classified at least to the family level by MEGAN analysis. Colors correspond to the three host-based categories inferred by host associations of corresponding families (families with members associated with exclusively plant hosts - light green, families with members associated with diverse hosts, including plants - dark green and families with members associated with non-plant hosts - blue). The origin and number of plant individuals for each sample or sample pool are noted above the graph with grey plant icons and explained in the legend above the plot. b) Relative abundance of reads mapped to viral contigs (> 500 nt) across samples and viral families. Each square represents the percentage of reads mapped per viral family, with color intensity reflecting logarithmically transformed values. Viral family names on the left are color-coded according to the three host-based categories described in a), with grey representing taxonomic bins for contigs that could not be classified to family level with MEGAN. c) Overview of detected viral contigs representing complete coding sequences from known and putative novel plant viruses. Associations of detected viruses with predicted plant hosts, as determined with PCR tests for each of the listed viruses on individual plant nucleic acid extracts, are shown as line connections

(Supplementary Table 3). In particular, pooled sample AQ1 contained a large fraction of reads that mapped to viral contigs, which were assigned only at the *Riboviria* realm level ("unclassified *Riboviria*" category) (Fig. 2b).

Putative plant virus contigs, identified by their classification into families with exclusively plant hosts using MEGAN, were detected across all samples. A detailed overview, along with the results of the BLASTX analysis and read mapping, is provided in Supplementary Table 9. Viral contigs that were classified into the *Caulimoviridae* family were detected in three samples, with the highest abundance of mapped reads observed in AQ3. High read abundance was also observed for contigs classified into the *Bromoviridae* family in pooled sample AQ4 and the *Closteroviridae* family in pooled sample AQ2 (Fig. 2b).

Among the putative plant virus contigs listed above, several were found to contain the complete coding sequences (CDS) (Table 3; Fig. 2c). BLASTN analysis revealed that several contigs detected in datasets of pooled samples AQ2 and AQ4 had high nucleotide sequence identities to two known crop pathogens with segmented genomes, lettuce chlorosis virus (isolate NIB-1, GenBank accessions PV019450- PV019451) and cucumber mosaic virus (isolate NIB-1, GenBank accessions PV019452- PV019453) (Table 3). Their classification within the respective viral species was further supported by SDT analysis (Supplementary Table 6). With PCR testing, we confirmed the presence of lettuce chlorosis virus in RNA extracts of *Alternanthera reineckii* 'rosanervig' (Fig. 2c, Supplementary Fig. 2.1) and cucumber mosaic virus in those of *Lobelia cardinalis* 'mini' (Fig. 2c, Supplementary Fig. 2.2).

We also detected several contigs with complete CDS, which were determined to represent putative novel viral species. In pooled sample AQ2, we identified a putative novel virus of the *Tombusviridae* family (Supplementary Table 6), exhibiting the highest genome-wide identity to Colocasia umbra-like virus (Table 3, Supplementary Fig. 1.1). PCR testing on individual RNA extracts from pooled plants revealed the presence of this virus in *Echinodorus cordifolius* 'marble queen' (Fig. 2c, Supplementary Fig. 2.1), prompting the name Echinodorus virus 1 (GenBank accession PV019455). In the same dataset (pooled sample AQ2), we identified several contigs that were classified into the genus *Begomovirus*, a discovery that is further discussed in the next section.

In pooled sample AQ3, we discovered a putative novel caulimovirus (Supplementary Table 6), and using PCR analysis, we detected its presence in the RNA extracts of *Hygrophila polysperma*, prompting the name Hygrophila virus 1 (GenBank accession PV019457) (Table 3; Fig. 2c, Supplementary Fig. 1.2, 2.1). Similarly, in pooled sample

Table 3 Detected plant virus contigs/scaffolds with complete coding sequences and the results of the BLASTN analysis. Scaffolds that were manually reassembled from contigs are marked with joined IDs (e.g., AQ2_12346_58280), denoting the original sequences that were combined. [a] indicates contigs of known plant viruses

contig ID	virus name	virus family	contig length	closest BLASTN hit	accession number of the closest BLASTN hit	% identity to the closest BLASTN hit	E-value
AQ2_124_73867	lettuce chlorosis virus [a]	*Closteroviridae*	8653	Lettuce chlorosis virus (RNA1)	NC_012909	99	0.0
AQ2_91	lettuce chlorosis virus [a]	*Closteroviridae*	8603	Lettuce chlorosis virus isolate BR (RNA 2)	MN216392	98	0.0
AQ4_745	cucumber mosaic virus [a]	*Bromoviridae*	3321	Cucumber mosaic virus isolate TAA_CMV (RNA1)	MZ666931	93	0.0
AQ4_991	cucumber mosaic virus [a]	*Bromoviridae*	3097	Cucumber mosaic virus isolate RP9 (RNA 2)	KC527693	94	0.0
AQ4_3200	cucumber mosaic virus [a]	*Bromoviridae*	2220	Cucumber mosaic virus isolate PV-0506 (RNA3)	KX525736	96	0.0
AQ2_6106	Ageratum yellow vein Sri Lanka [a] virus	*Begomoviridae*	2752	Ageratum yellow vein Sri Lanka virus (DNA-A)	NC_002981	95	0.0
AQ2_12346_58280	Lindernia rotundifolia yellow vein virus	*Begomoviridae*	2741	Lindernia anagallis yellow vein virus	KC172827	89	0.0
AQ2_26591	Lindernia rotundifolia yellow vein alphasatellite	*Alphasatellitidae*	1404	Tobacco curly shoot virus associated DNA 1	FM212564	90	0.0
AQ6_189	Sagittaria virus A	*Potyviridae*	9653	Malva vein clearing virus strain MVCV-SX	MN116683	69	2e-101
AQ2_3721	Echinodorus virus 1	*Tombusviridae*	2953	Colocasia umbra-like virus isolate PNG K	ON086742	67	1e-61
AQ4_2	Lobelia virus 1	*Endornaviridae*	16,663	Helianthus annuus alphaendornavirus isolate BJ	NC_040799	67	9e-84
AQ3_270	Hygrophila virus 1	*Caulimoviridae*	7825	Figwort mosaic virus	NC_003554	71	0–0

AQ4, we detected a novel virus from the family *Endornaviridae* (Supplementary Table 6) and, with PCR, confirmed its presence in the RNA extracts of *Lobelia cardinalis* 'mini', thus designating it Lobelia virus 1 (GenBank accession PV019456) (Table 3; Fig. 2c, Supplementary Fig. 1.3, 2.1). In sample AQ6, which consisted solely of *Sagittaria subulata*, we detected the sequence of Sagittaria virus A (isolate NIB-1, GenBank accession PV019448) – a novel potyvirus we had initially discovered in the transcriptome of *Sagittaria latifolia* (isolate 1KP, GenBank accession PV019449) (Table 3; Figs. 2c and 5, Supplementary Table 6, Supplementary Fig. 2.1). The detailed characterization of this virus is provided in a later part of this manuscript.

Identification of known and putative novel begomoviruses in two different ornamental aquatic plants

The discovery of begomoviral sequences in the HTS dataset from the pooled sample AQ2 led us to conduct a more detailed investigation, due to their inclusion on the European Union quarantine list: Annex II A of the commission implementing regulation (EU) 2019/2072 [74]. From the RNA sequencing data, we assembled two begomoviral contigs: one corresponding to the known monopartite begomovirus Ageratum yellow vein Sri Lanka virus (AYVSLV, isolate NIB-1, GenBank accession PV019458) and another representing a putative novel begomovirus.

With newly designed virus-specific PCR assays, we established the presence of AYVSLV in the DNA extract of *Alternanthera reineckii* 'rosanervig' (Figs. 2c and 3, Supplementary Fig. 2.3), while the putative novel begomovirus was detected in the DNA extracted from *Lindernia rotundifolia*, prompting the name Lindernia rotundifolia yellow vein virus (LRYVV, isolate NIB-1, GenBank accession PV019459) (Figs. 2c and 4, Supplementary Fig. 2.3). No product for either virus was detected using the PCR tests from the EPPO standard on diagnostics (PM7/152).

From the DNA sequencing data of *A. reineckii* 'rosanervig', we assembled the complete genomic sequence of AYVSLV (Fig. 3a, Supplementary Fig. 1.4). The sequence is 2752 nucleotides long and is 94.6% identical to the DNA-A AYVSLV reference sequence (GenBank accession NC_002981), assigning it to the same species according to the species demarcation criteria for this genus (Supplementary Table 6) [75]. Additionally, another sequence that showed similarity to known begomoviruses was discovered in the DNA sequencing dataset of the same plant. It is approximately half the size of the AYVSLV (1352 nt) and exhibits characteristics typical for defective DNAs [76–78], prompting the tentative name Ageratum yellow vein Sri Lanka virus D-DNA (AYVSLV D-DNA). The AYVSLV D-DNA sequence (Fig. 3a, Supplementary Fig. 1.5) is largely derived from AYVLSV, featuring an additional 301 nt region of unknown origin with low similarity to various eukaryotic sequences. It contains two short ORFs: ORF1 is homologous to the 5' end of AYVLSV AV2, while ORF2 is partly homologous to the 3' end of the same protein and partly to the 5' end of the replication (Rep) protein. ORF1 also encompasses the intergenic region with the conserved nonanucleotide sequence (TA ATATT/AC) containing the origin of replication. Notably, two additional repeats of the intergenic region follow the ORF1, starting directly after the nick site.

The putative host plant, *A. reineckii* 'rosanervig', displayed vein yellowing, resembling symptoms caused by a viral infection (Fig. 3b). Another variety of *A. reineckii* ('mini') lacking these patterns was sequenced in a separate pooled sample, AQ3, where no begomoviral sequences were detected. The episomal nature of AYVSLV and its defective DNA was confirmed by restriction digestion analysis of RCA products (Supplementary Fig. 2.4).

From the DNA sequencing data of *L. rotundifolia*, we assembled the complete DNA-A sequence of Lindernia rotundifolia yellow vein virus (Fig. 4a, Supplementary Fig. 1.6). The genome is 2741 nucleotides long and, based on SDT analysis (Supplementary Table 6, Supplementary Fig. 1.8), shares 88.9% genome-wide nucleotide identity with Lindernia angallis yellow vein virus (LAYVV, GenBank accession AY795900), thus representing a novel species in the *Begomovirus* genus. The genome organization is characteristic of monopartite begomoviruses, containing 7 predicted ORFs, including AV2, which is typical for the Old World begomoviruses [75]. It also contains the intergenic region with the conserved stem-loop forming nonanucleotide (TAATATT/AC) (Fig. 4a). The phylogenetic network analysis (Fig. 4b) shows LRYVV and LAYVV forming a distinct cluster. In addition to LRYVV, a complete genomic sequence of a novel alphasatellite, which we named Lindernia rotundifolia yellow vein alphasatellite (LRYVV alphasatellite, isolate NIB-1, GenBank accession PV019460), was also detected in the dataset from the same plant (Table 3; Fig. 2c, Supplementary Fig. 1.7). According to the SDT analysis (Supplementary Table 6), the LRYVV alphasatellite sequence shares 87.4% identity with the genomic nucleotide sequence of tobacco curly shoot alphasatellite, qualifying it as a novel species within the family *Alphasatellitidae* [79].

Similar to *A. reineckii*, *L. rotundifolia* also exhibited the pattern of vein yellowing (Fig. 4c), which we speculate could be a symptom of infection with LRYVV, possibly also in combination with the Lindernia rotundifolia yellow vein alphasatellite. Both LRYVV and its alphasatellite were also shown to be present in the plant in an episomal form (Supplementary Fig. 2.4).

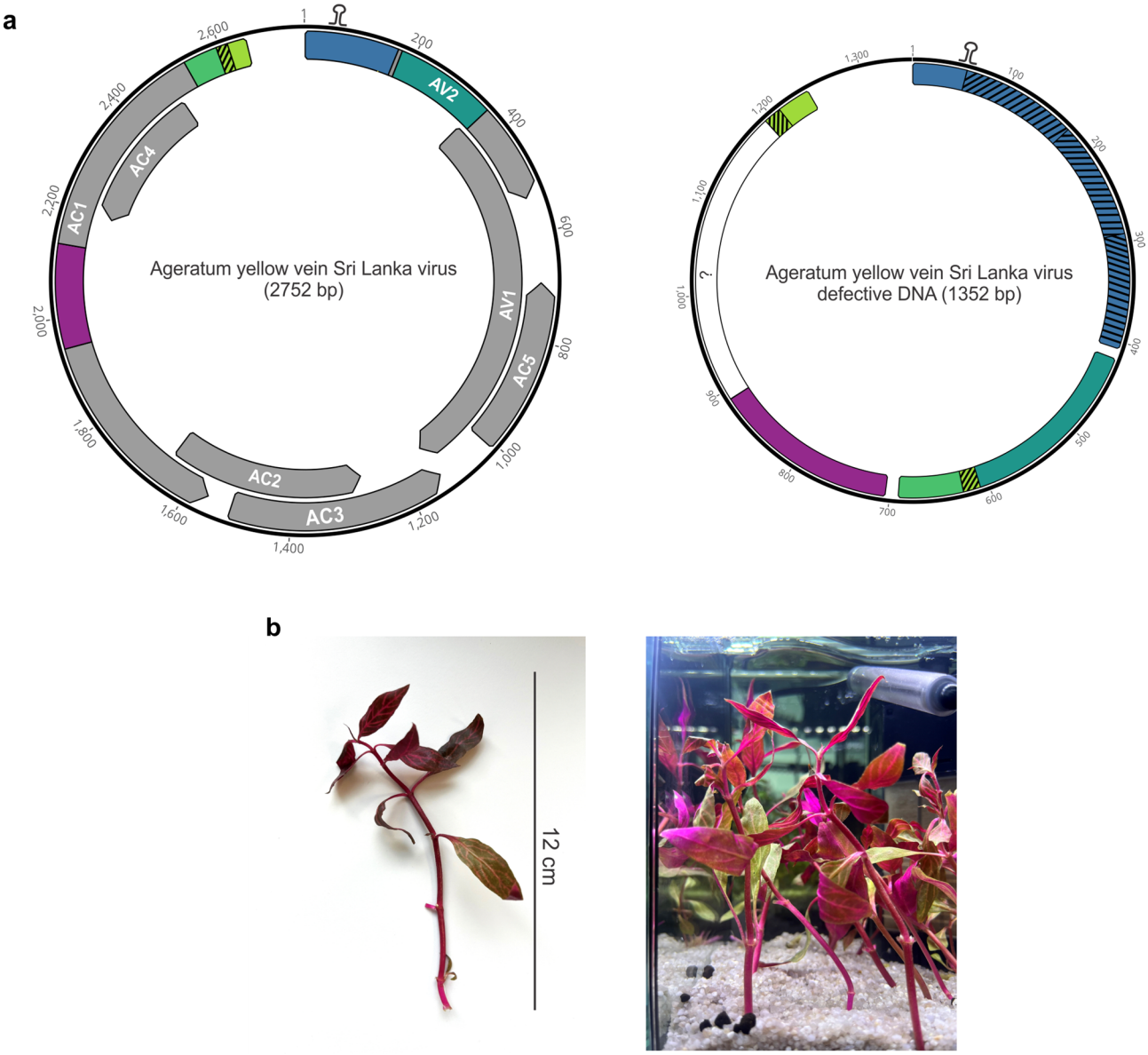

Fig. 3 Characteristics of Ageratum yellow vein Sri Lanka virus and its defective DNA. a) Genome organization of Ageratum yellow vein Sri Lanka virus (AYVSLV, left) and Ageratum yellow vein Sri Lanka virus defective DNA (AYVSLV D-DNA, right) discovered in *Alternanthera reineckii* 'rosanervig'. AV2 - pre-coat protein, AV1 - coat protein (CP), AC5 - C5 protein, AC3 - replication enhancer protein (REn), AC2 - transcriptional activator protein (TrAP), AC1 - replication-associated protein (Rep), AC4 - C4 protein. The colored blocks indicate homologous regions between the two sequences, while the striped pattern denotes regions where parts of the sequence are repeated. In AYVSLV D-DNA, the question mark (?) denotes the region of unknown origin. The conserved nonanucleotide (TAATATT/AC) is indicated with a stem-loop. b) A sample of *A. reineckii* 'rosanervig' positive for AYVSLV and AYVSLV D-DNA, exhibiting vein yellowing (left) and the same plant cultivated in the aquarium (right)

Discovery of a novel potyvirus by data mining and HTS analysis of purchased plants in several species from the genus Sagittaria

As noted earlier, we discovered a sequence of a putative novel potyvirus, Sagittaria virus A (SgVA, isolate 1KP, GenBank accession PV019449), by mining the transcriptome of the aquatic plant, *Sagittaria latifolia*. We later detected a near-identical sequence in the sequencing library of *Sagittaria subulata*, an aquatic plant that is cultivated and sold for ornamental use. It shared 98% identity across the polyprotein amino acid sequence to SgVA isolate 1KP and was thus designated SgVA isolate NIB-1 (GenBank accession PV019448) (Fig. 5a, d). These findings prompted us to conduct a more detailed investigation into the occurrence and diversity of this virus.

S. subulata plants that contained SgVA isolate NIB-1 were cultivated in an aquarium, where they initially showed stunted growth, followed by yellowing and

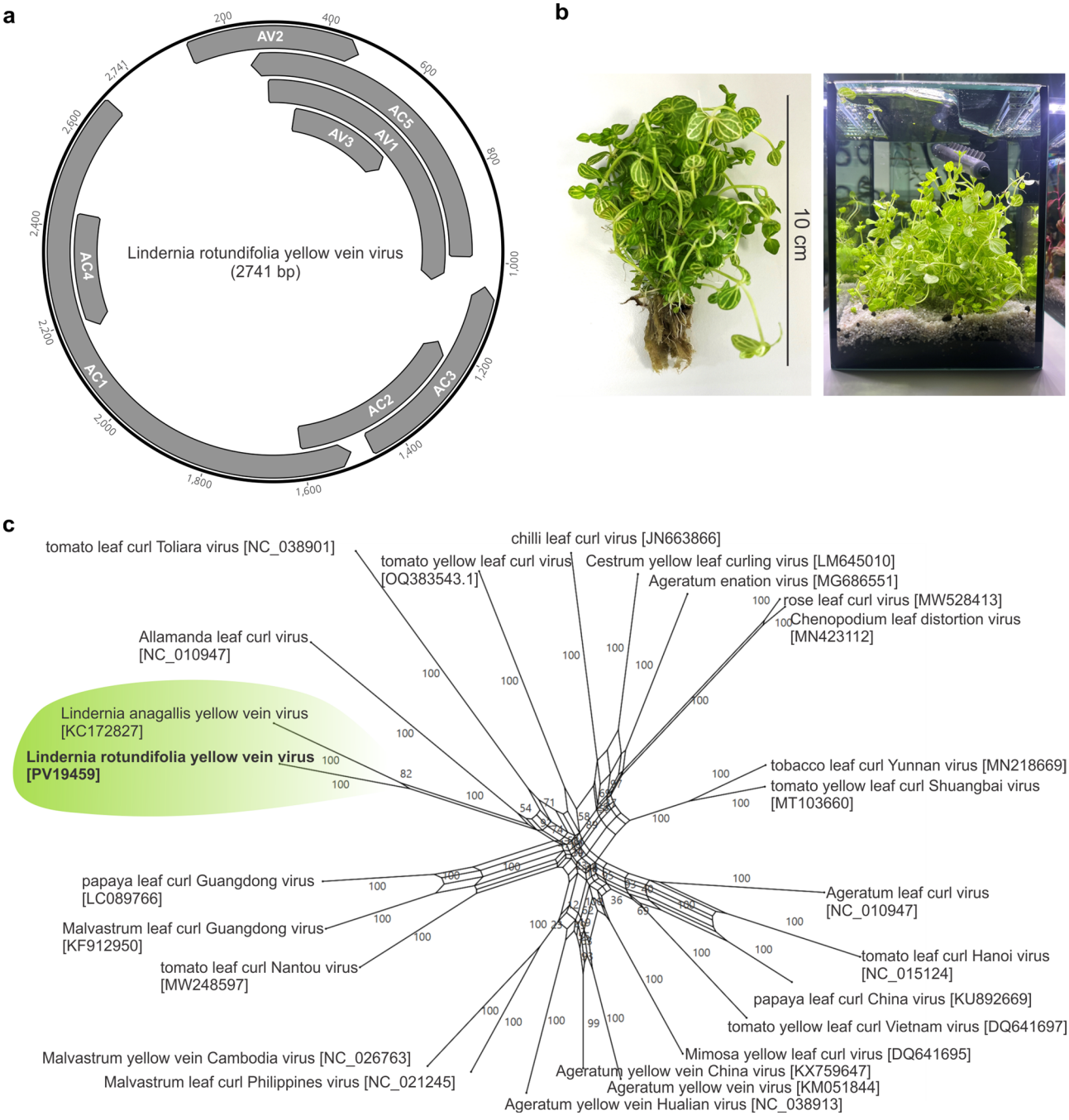

Fig. 4 Genomic characterization and phylogenetic network analysis of Lindernia rotundifolia yellow vein virus. a) Genome organization of the Lindernia rotundifolia yellow vein virus, discovered in *Lindernia rotundifolia*. AV1 - coat protein (CP), AV2 - pre-coat protein, AC5 - C5 protein, AV3 - V3 protein, AC3 - replication enhancer protein (REn), AC2 - transcriptional activator protein (TrAP), AC1 - replication-associated protein (Rep), AC4 - C4 protein. b) A specimen of *Lindernia rotundifolia*, positive for Lindernia rotundifolia yellow vein virus (left), and the same plant grown in an aquarium (right). c) Phylogenetic network inferred for Lindernia rotundifolia yellow vein virus and a set of closely related begomoviral sequences using the Neighbor-Net method in SplitsTree6. The numbers on the edges represent bootstrap support values for each split. GenBank accession numbers for each sequence are listed in brackets next to the virus name

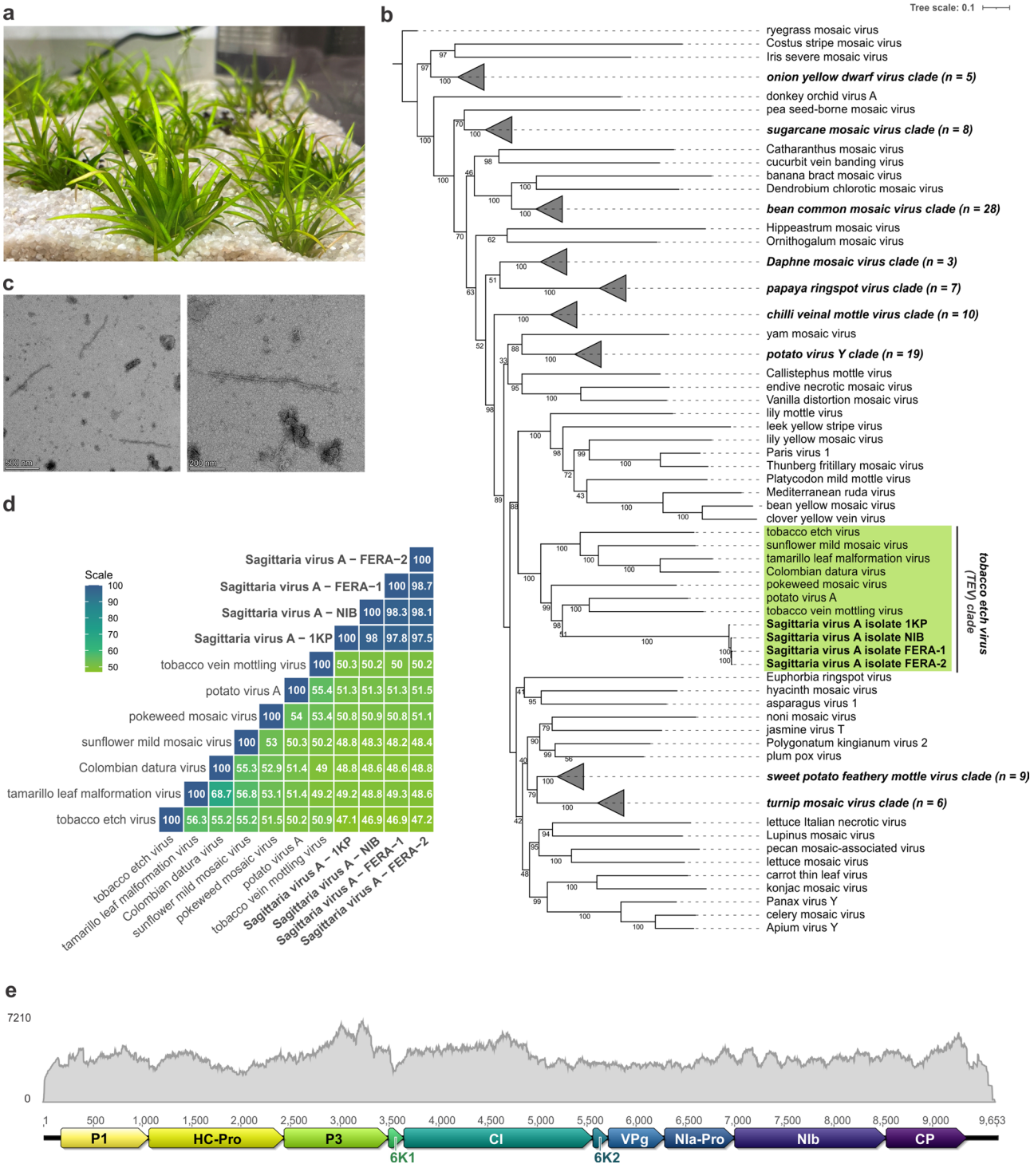

Fig. 5 (See legend on next page.)

(See figure on previous page.)

Fig. 5 Characterization and phylogenetic analysis of Sagittaria virus A. a) Cultivated *Sagittaria subulata* plants, infected with Sagittaria virus A – isolate NIB. b) A maximum likelihood phylogenetic tree derived from the alignment of the polyprotein sequences from four Sagittaria virus A isolates (NIB, FERA-1, FERA-2, 1KP) and selected members of the *Potyvirus* genus with complete coding sequences available in the NCBI RefSeq database. The tree was constructed using the LG+F+I+G4 model. Some clades are collapsed for clarity and are denoted after representative members according to the ICTV tree for the family Potyviridae. The number of collapsed sequences for each clade is shown in brackets. The tobacco etch virus (TEV) clade, which includes the novel Sagittaria virus A, is shown in green. The numbers next to the branches represent statistical support as determined by the ultrafast bootstrap analysis. The tree was rooted using an outgroup (ryegrass mosaic virus, genus *Rymovirus*). c) TEM micrograph of *S. subulata* plant extracts containing Sagittaria virus A. d) Heatmap of pairwise identities (%) of the amino acid polyprotein sequences from the four Sagittaria virus A isolates (bolded) and their closest relatives from the TEV clade, calculated with SDT. e) Genome organization of the Sagittaria virus A with predicted mature proteins and cleavage sites. P1 - protein 1, HC-Pro - helper component-protease, P3 - protein 3, 6K1 - 6 kDa peptide 1, CI - cylindrical inclusion protein, 6K2 - 6 kDa peptide 2, VPg - viral protein genome-linked, Nla-Pro - nuclear inclusion-a protease, NIb - nuclear inclusion-b protein, CP - capsid protein. The grey coverage plot above the genome graph illustrates the read mapping depth across the genome

thinning of the leaves, which appeared several months after planting (Fig. 5a). The infected plant tissue preparation was visualized under an electron microscope and revealed flexible filamentous structures typical of potyviruses (Fig. 5c).

SgVA isolate NIB shares 50.3% amino acid identity in the polyprotein sequence with its closest relative, potato virus A (Fig. 5d, Supplementary Table 6), supporting its designation as a novel species in the *Potyvirus* genus [80]. The genomic sequence is 9.6 kbp long and contains the entire polyprotein sequence, including the predicted cleavage sites that produce mature proteins (Fig. 5e).

Following this discovery, we explored the possible presence of SgVA among traded ornamental plants from the genus *Sagittaria*. Six samples that had been imported into the UK from Singapore and the Netherlands tested positive for a potyvirus in ELISA screening (Supplementary Table 11). Sanger sequencing of the RT-PCR products confirmed the presence of SgVA in all cases. Two of those samples, originating from Singapore, underwent additional HTS. This resulted in the reconstruction of two more SgVA genomic sequences (isolate FERA-1, GenBank accession PQ653482, and FERA-2, GenBank accession PQ653481). The two FERA isolates shared a 98% identity in the amino acid sequence of the polyprotein with each other, as well as with SgVA isolate NIB and SgVA isolate 1KP, confirming they are of the same species (Fig. 5d). The phylogenetic analysis revealed that all isolates of SgVA clustered together, with the plant isolates representing one group and 'isolate 1KP' representing a separate group within the cluster. They were placed into the tobacco etch virus (TEV) clade [75], clustering with tobacco vein mottling virus and potato virus A (Fig. 5b). Additionally, we tested two more samples acquired through different online stores in Slovenia with RT-PCR, and both proved to be positive for infection with SgVA, pointing to its potential distribution via global ornamental plant trade (Supplementary Table 11).

Discussion

Characterization of the underexplored viromes of non-cultivated and ornamental aquatic plants that we performed in this study revealed a high level of plant virus diversity across a range of aquatic plant species, spanning from simple multicellular algae to seed plants. While the majority of discovered viral sequences likely correspond to novel viral species, we also detected several sequences of known crop viruses, including some regulated pathogens.

Within the evolutionary basal aquatic plant groups, we examined the transcriptomes of 48 diverse macroscopic algae. Among the diverse viral sequences identified in these datasets, perhaps the most interesting finding is the discovery of 14 viral contigs resembling members of the *Phaeovirus* genus (*Phycodnaviridae*) in 9 algae species (Supplementary Table 5). While members of the *Phycodnaviridae* family primarily infect unicellular algae [81], sequences of viruses from the genus *Phaeovirus* have been detected in filamentous brown macroalgae from the orders *Ectocarpales* and *Laminariales* [40, 41]. Our findings expand the putative host range of phaeoviruses to rhodophytes (*Pyropia yezoensis* and *Grateloupia filicina*), whose viral communities have not been extensively studied [82]. To date, only several viral sequences resembling members of the order *Picornavirales* [83] and the family *Totiviridae* [84, 85] were detected in these hosts. Similar to algae, only a few studies [86–88] have explored the viral communities associated with other non-vascular plants such as liverworts and mosses. In one previous study, a comprehensive data mining analysis of the 1KP datasets uncovered 16 bunya-like sequences in libraries derived from five moss and liverwort species, forming a new clade within the *Bunyavirales* order [44]. Here, we detected several contigs resembling novel members of the *Bunyavirales* order in heart-leaved spearmoss, including five that had already been documented by Mifsud et al. in the same dataset (heart-leaved spearmoss bunyavirus 1–5) [44], expanding on the previously limited understanding of viral diversity in mosses.

While investigating the viromes of non-vascular aquatic plants provided important insights, we directed much of our attention to seed plants, specifically angiosperms, given their broad distribution and relevance to human food production.

The majority of viral sequences detected in angiosperms in this study originate from double-stranded viruses of the family *Caulimoviridae*. Several partial contigs resembling members of the genera *Caulimovirus*, *Cavemovirus*, and *Badnavirus* were detected in two aquatic plant transcriptomes (Supplementary Table 5), as well as in four pooled aquarium plant samples (Supplementary Table 9, Fig. 2b), with badnaviruses being the most prevalent. Viruses from the *Caulimoviridae* cause some economically important diseases of tropical and subtropical crops [89, 90], while some can integrate into the host genome, resulting in the formation of endogenous viral elements (EVEs) [91, 92]. Thus, some sequences identified here may also represent EVEs, which warrants further investigation.

We also discovered several partial contigs related to members of the *Rhabdoviridae* family in the transcriptomes of *Sagittaria latifolia* and *Myriophyllum aquaticum* (Table 2), as well as in the sequencing library of *Myriophyllum sp.* 'Guyana' (Supplementary Table 9). This discovery reflects the broad host range of the members of the *Rhabdoviridae* family, which are known to infect various monocot and dicot plants [93]. Viruses from this family have previously been identified in the transcriptomes of several decorative aquatic plants, including water hyssop (*Bacopa monnieri*) [94], water lily (*Nymphaea alba*) [95], and lace plant (*Aponogeton madagascariensis*) [96], as well as in the seagrass *Zostera* marina [39] (Supplementary Table 1).

Furthermore, we identified several sequences with the highest similarity to members of the *Partitiviridae* family in the transcriptomes of seaside arrowgrass, *Triglochin maritima*, and the seagrass *Posidonia australis* (Table 2). Partitiviruses are characterized by their persistent lifestyle, asymptomatic infections, and their inability to transmit horizontally [97, 98]. A similar trend was observed in the only other comprehensive study of wild aquatic plant viromes performed to date, where cryptic viruses from the family *Partitiviridae* were detected in seven aquatic plant species [43]. Out of these, five belong to the same order (*Alismatales*), as *Triglochin maritima* and *Posidonia australis.* The impact of cryptic viruses on plant hosts is largely unknown [98].

While the majority of viral sequences detected in this study represent putative novel viral species, as seen with other HTS-based studies [99–101], we also identified several known crop viruses in different species. We detected lettuce chlorosis virus, known to affect crops such as lettuce and sugar beet [102], as well as beans [103] and tomato [104], in the transcriptome of *M. aquaticum* and in the sequencing library of the ornamental aquatic plant *A. reineckii* 'rosanervig'. In another decorative plant, *Lobelia cardinalis* 'mini', we discovered the complete genome of cucumber mosaic virus with high-throughput sequencing (Table 3). This virus is known to infect an extensive range of plant species, including crops and wild plants [105]. This observation supports prior findings that cucumber mosaic virus can infect several species from the genus *Lobelia*, including ornamental hybrids [106]. Other notable examples detected across transcriptomes of various species include turnip yellows virus, Schlumbergera virus X, and Cymbidium mosaic virus (Supplementary Table 5), each known to affect a variety of economically important plants [73, 107–110]. Our findings indicate that aquatic plants could play a role as reservoirs for viruses infecting crops, especially given their ecological role and broad distribution. We hypothesize that aquatic plants may link aquatic and terrestrial environments by acquiring viral particles from the water, which can then be spread to nearby vegetation by vectors feeding on their emergent parts.

The potential of aquatic plants to harbor possibly pathogenic viruses, including such that infect crops, raises additional concerns about the unregulated trade of ornamental aquatic plants. The commercial trade of these plants is known to be a major source of invasive species introduction, threatening various freshwater and marine ecosystems [10, 11]. Although the sale of some exotic species is banned in many countries, studies have shown that they are often easily attainable through online retailers or informal marketplaces [9, 10, 111]. Such trade not only introduces invasive species but also facilitates the spread of their microbes, which may spill over to the native hosts. Native populations can be less resistant, leading to their decline and increasing the fitness of invasive species [112, 113].

We detected numerous viruses in the sequencing libraries of ornamental aquatic plants. However, for illustration, the risks associated with their trade and possible introduction to new ecosystems are discussed in-depth for (novel) begomoviruses and a novel potyvirus discovered in this study. Begomoviruses are quarantine pathogens in the European Union [114], the UK, and Switzerland [115]. Both *A. reineckii* and *L. rotundifolia* harboring begomoviral sequences displayed ornamental patterns of pink or yellow veins, a desirable trait that can enhance their sale. The absence of these patterns in a different *A. reineckii* variety ('mini'), in which no such viruses were detected, suggests that the vein yellowing could be caused by a viral infection. This hypothesis is further supported by the episomal nature of detected begomoviruses, indicating an active replication in the host plant. The appealing visual traits associated with virus-infected aquatic plants may lead to their selection by the growers, potentially contributing to the faster spread of these viruses through global trade.

While the detection of begomoviral sequences is concerning, the adaptability of these plants to new

environments should be considered when evaluating the associated risks. Both plant species in which begomoviral sequences were detected are native to warmer regions [116, 117], which could, to some extent, limit their establishment in colder climates. However, the ability of many exotic aquatic plant species to adapt to various environments [118–122] suggests they may also thrive in temperate habitats. Alien aquatic plants have been shown to survive in colder climates in ecological niches, such as thermal streams [123], with water lettuce (*Pistia stratiotes*) being a key example [122]. If virus-infected alien aquatic plants were to be established in such ecosystems, the potential pathways for virus transmission to crops should be considered. Begomoviruses are transmitted by the silverleaf whitefly (*Bemisia tabaci)*, which is already present in several European countries [124] and is predicted to expand further as global temperatures rise [125], increasing the risk for virus spillovers to native vegetation. For feeding and virus transmission to occur, the leaves of infected plants must be accessible to the vector. Submerged aquatic plants may emerge from the water in response to environmental stress [126] or due to water level fluctuations [127, 128], providing an opportunity for vector feeding and possible transmission of viruses to surrounding vegetation.

We also detected a novel potyvirus, Sagittaria virus A, in several ornamental aquatic plants from the genus *Sagittaria* obtained in two European countries, Slovenia and the UK. The UK samples were sourced from Singapore and the Netherlands, while the Slovenian samples were of unknown origin. Several hypothesis, or their combination, may explain these findings. Firstly, aquatic plants in nurseries and stores are often cultivated in shared spaces, where water may facilitate the spread of viruses between them [14]. Secondly, as implied above for begomovirus-infected plants, these specimens might be specifically selected for their desirable traits, such as stunted growth. Finally, despite seemingly different sources, all plants may originate from the same cultivation facility or are later vegetatively propagated before being distributed to smaller growers and stores. We also detected the sequence of SgVA in the transcriptome of *S. latifolia*, for which no metadata about sample origin was available. *S. latifolia* originates from North America and is widely cultivated as an ornamental pond plant in Europe. The polyprotein sequences of isolates from ornamental plants were found to be slightly more similar to each other than to the sequence detected in *S. latifolia*. These findings could point to the common source of ornamental plant specimens, but due to the lack of information on the origin of *S. latifolia*, definitive conclusions cannot be drawn.

Our findings highlight the great diversity of viral sequences in aquatic plants, recovered by data mining or analysis of newly generated RNA sequencing data; however, certain limitations must be considered. Both approaches represent the metagenomics-based discovery of viral sequences, which is a very powerful tool for the detection of viruses, but cannot provide information about, e.g., viral replication in the associated plant samples. Thus, putative host plant associations are necessarily done based on the relatedness of discovered viral sequences to sequences of other plant-infecting viruses. This is sometimes difficult, especially for highly divergent viruses that cannot be classified into known viral families, and some of such sequences were also detected in this study. Additionally, mining datasets not specifically designed for viral discovery can obscure the identification of plant viral sequences due to the specific preparation of samples that could limit viral detection. Often, metadata linked with sample origin and preparation is limited, which further hinders the interpretation of data. Reliable interpretation of such findings often requires additional biological characterization (e.g., infectivity assays), which can be particularly challenging in aquatic plants due to their requirement for specialized systems for cultivation and the poorly understood transmission mechanisms for the viruses that infect them. While our study focused on high-throughput sequencing approaches to characterize the viromes of selected aquatic plants, future work could benefit from an in-depth biological characterization of some of the interesting findings following a proposed framework [129]. This could involve local and global epidemiological surveys, transmission trials and further molecular characterizations, such as cloning of full-length viral genomes—particularly for circular DNA viruses such as begomoviruses—followed by Sanger sequencing or additional transmission experiments.

Conclusions

- The analysis of publicly available and newly generated HTS data has expanded the understanding of viruses associated with aquatic plants. Both known and novel viruses from two of the largest plant viral families, *Potyviridae* and *Geminiviridae*, were identified, along with viruses from several other economically important families, including major crop pathogens. The detected viral sequences represent a variety of genome types, reflecting the diverse viral populations in aquatic plants.
- Viral sequences were found across all investigated types of aquatic plants, encompassing a range of organizational complexities – from evolutionarily basal non-vascular algae to economically significant angiosperms. Data mining revealed numerous viral contigs in the transcriptomes of non-vascular plants, adding to the previously limited understanding of their viral diversity.

- The discovery of potentially pathogenic viruses in ornamental aquatic plants, including regulated species, underscores the risks associated with global trade in ornamental aquatic flora. To fully assess these risks, further research is needed to explore the biological properties of the identified viruses, particularly their potential for waterborne transmission.
- Viral infections may induce ornamental phenotypic changes that are selectively favored in globally traded aquatic plants, potentially aiding their spread. The establishment of such exotic plants in new ecosystems, along with subsequent virus spillover to native vegetation, would depend on several ecological factors - conditions that may become more favorable with the accelerating pace of global warming.

Abbreviations

AYVSLV	Ageratum yellow vein Sri Lanka virus
BLASTN	Basic Local Alignment Search Tool for Nucleotide Sequences
BLASTX	Basic Local Alignment Search Tool for Translated Nucleotide Sequences
CDS	Coding Sequence
CLC-GWB	CLC Genomics Workbench
D-DNA	Defective DNA
EVE	Endogenous Viral Elements
HTS	High-Throughput Sequencing
ICTV	International Committee on Taxonomy of Viruses
LAYVV	Lindernia anagallis yellow vein virus
LCA	Last Common Ancestor
LRYVV	Lindernia rotundifolia yellow vein virus
NCBI	National Center for Biotechnology Information
ORF	Open Reading Frame
RCA	Rolling Circle Amplification
RdRp	RNA-dependent RNA polymerase
SgVA	Sagittaria virus A
SRA	Sequence Read Archive
TEM	Transmission Electron Microscopy
TEV	Tobacco etch virus
VMR	Virus Metadata Resource
1KP	1000 Plant Transcriptomes Initiative

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40793-025-00783-6.

Supplementary Material 1: **Supplementary file 1** (fasta): Trimmed alignment of the potyviral sequences used for the construction of the phylogenetic tree for the placement of Sagittaria virus A in the ***Potyvirus*** genus.

Supplementary Material 2: **Supplementary file 2** (fasta): Trimmed alignment of the begomoviral sequences used for the construction of the phylogenetic network for the placement of Lindernia rotundifolia yellow vein virus in the ***Begomovirus*** genus.

Supplementary Material 3: **Supplementary Fig. 1** (pdf): Genome organizations of putative novel plant viruses with complete coding sequences, discovered in this study, along with the visualized read mapping depth. A heatmap of the pairwise identities for Lindernia rotundifolia yellow vein virus and its closest BLASTN hits is also shown.

Supplementary Material 4: **Supplementary Fig. 2** (pdf): Agarose gels showing PCR-based detection of viral contigs with complete CDS in individual plant extracts from pooled samples and the results of restriction digestion analysis of begomoviral genomes after rolling circle amplification.

Supplementary Material 5: **Supplementary table 1** (xlsx): Overview of viruses discovered in aquatic plants up to date. Only viral sequences with a GenBank accession are shown. Viruses of rice are not included, as they are reviewed elsewhere (https://pmc.ncbi.nlm.nih.gov/articles/PMC9609659/).

Supplementary Material 6: **Supplementary Table 2** (xlsx): List of analyzed macrophyte species and their metadata (1KP ID, SRA accession, plant family, major plant group, habitat).

Supplementary Material 7: **Supplementary Table 3** (xlsx): Summary of viral contigs over 500 nucleotides detected in macrophyte transcriptomes, along with the corresponding BLASTX analysis results.

Supplementary Material 8: **Supplementary Table 4** (xlsx): Criteria used for categorizing viral contigs based on ICTV predicted hosts.

Supplementary Material 9: **Supplementary Table 5** (xlsx): Putative plant virus contigs, identified in macrophyte transcriptomes, and the results of the BLASTX analysis. Scaffolds that were manually re-assembled from contigs are marked with joined IDs (e.g. 2004053_2003800), denoting the original sequences that were combined.

Supplementary Material 10: **Supplementary table 6** (xlsx): Summary of species demarcation results for contigs representing near-complete viral genomes. For each contig, the closest BLASTN hit (lowest e-value) and the corresponding SDT pairwise alignment results are shown. *Due to the first BLASTN hit corresponding to a partial sequence, the second hit was used for contig AQ4_2. Taxonomic assignments (columns 'genus' and 'family') are based on both MEGAN classification of the query contig and the taxonomy of its closest BLASTN hit.

Supplementary Material 11: **Supplementary Table 7** (xlsx): List of sequenced samples and summarized results of the bioinformatic analysis of the sequencing data. For each sample, the table provides the number of raw reads, total contigs assembled with SPAdes, viral contigs identified with MEGAN6 and the number of reads mapped to these viral contigs.

Supplementary Material 12: **Supplementary Table 8** (xlsx): Summary of viral contigs over 500 nucleotides detected with sequencing in aquatic plants, along with the corresponding BLASTX analysis results.

Supplementary Material 13: **Supplementary Table 9** (xlsx): Putative plant virus contigs, identified in ornamental aquatic plants, and the results of the BLASTX analysis. Scaffolds that were manually re-assembled from contigs are marked with joined IDs (AQ2_124_73867), denoting original sequences that were combined. Contigs of detected begomoviruses were assembled from DNA-sequencing data.

Supplementary Material 14: **Supplementary Table 10** (xlsx): List of PCR primers and cycling conditions used for determining putative plant hosts for selected viruses.

Supplementary Material 15: **Supplementary Table 11** (xlsx): List of Sagittaria plants, tested for Sagittaria virus A, along with sample metadata (sample type, testing site, origin, symptoms, PCR results).

Acknowledgements

The authors would like to thank Jakob Brodarič and Zarja Miović for their help with the nucleic acid extractions and Magda Tušek Žnidarič for sharing her expertise in electron microscopy.

Author contributions

LV, DK, KB and MB conceived and designed the study. LV, ARF and VH carried out the experiments. LV, AF, KB, IPA and DK contributed to the data analysis and interpretation of results. ŽL performed the electron microscopy experiments. NM designed and supervised the diagnostic testing for begomoviruses. LV wrote the original manuscript draft. All authors edited the manuscript and agreed to the final submitted version. DK, NM, AF and IAP acquired funding and supervised the project.

Funding

This research was funded by the Slovenian Research and Innovation Agency (ARIS) core (P4-0407, P4-0165, P1-0212) and project (J4-4553) financing, as well as by the Defra-Fera long-term services agreement.

Data availability
All raw sequencing data generated in this study were deposited in the NCBI SRA database with the BioProject IDs [PRJNA1213295] (https:/www.ncbi.nlm.nih.gov/bioproject/?term=(PRJNA1213295)%20AND%20bioproject_sra%5bfilter%5d%20NOT%20bioproject_gap%5bfilter%5d) and [PRJNA1185943](https:/www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA1185943) . Complete coding sequences of viruses discovered in this study are available in the NCBI GenBank database under the accession numbers PV019448–PV019460 and PQ653481–PQ653488.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]National Institute of Biology, Večna pot 121, Ljubljana 1000, Slovenia
[2]International Postgraduate School Jožef Stefan, Jamova cesta 39, Ljubljana 1000, Slovenia
[3]Department of Biology, Biotechnical Faculty, University of Ljubljana, Večna pot 111, Ljubljana 1000, Slovenia
[4]School for Viticulture and Enology, University of Nova Gorica, Rožna Dolina, Slovenia
[5]Fera Science Ltd, Sand Hutton, York YO41 1LZ, UK

Received: 14 April 2025 / Accepted: 2 September 2025
Published online: 22 October 2025

References

1. Lesiv MS, Polishchuk AI, Antonyak HL. Aquatic macrophytes: ecological features and functions. Biol Stud. 2020;14:79–94. https://doi.org/10.30970/SBI.1402.619.
2. Chambers PA, Lacoul P, Murphy KJ, Thomaz SM. Global diversity of aquatic macrophytes in freshwater. Hydrobiologia. 2008;595:9–26. https://doi.org/10.1007/S10750-007-9154-6/FIGURES/3.
3. Gebler D, Szoszkiewicz K, Pietruczuk K. Modeling of the river ecological status with macrophytes using artificial neural networks. Limnologica. 2017;65:46–54. https://doi.org/10.1016/J.LIMNO.2017.07.004.
4. Krtolica I, Cvijanović D, Obradović Đ, Novković M, Milošević D, Savić D, et al. Water quality and macrophytes in the Danube river: artificial neural network modelling. Ecol Indic. 2021;121:107076. https://doi.org/10.1016/J.ECOLIND.2020.107076.
5. Bytyqi P, Czikkely M, Shala-Abazi A, Fetoshi O, Ismaili M, Hyseni-Spahiu M, et al. Macrophytes as biological indicators of organic pollution in the lepenci river basin in Kosovo. J Freshw Ecol. 2020;35:105–21. https://doi.org/10.1080/02705060.2020.1745913.
6. Parolin P, Rudolph B, Bartel S, Bresch C, Poncet C. Worldwide invasion pathways of the South American Eichhornia crassipes. Acta Hortic. 2012;937:1133–40. https://doi.org/10.17660/ACTAHORTIC.2012.937.141.
7. Lowe S, Browne M, Boudjelas S, De Poortr M. 100 of the world's worst invasive alien species: A selection from the global invasive species database. The Invasive Species Specialist Group (ISSG); 2000.
8. Champion PD, Clayton JS, Hofstra DE. Nipping aquatic plant invasions in the bud: weed risk assessment and the trade. Hydrobiologia. 2010;656:167–72. https://doi.org/10.1007/S10750-010-0446-X/METRICS.
9. Peres CK, Lambrecht RW, Tavares DA, Chiba de Castro WA. Alien express: the threat of aquarium e-commerce introducing invasive aquatic plants in Brazil. Perspect Ecol Conserv. 2018;16:221–7. https://doi.org/10.1016/J.PECON.2018.10.001.
10. Olden JD, Whattam E, Wood SA. Online auction marketplaces as a global pathway for aquatic invasive species. Hydrobiologia. 2021;848:1967–79. https://doi.org/10.1007/S10750-020-04407-7/FIGURES/6.
11. Havel JE, Kovalenko KE, Thomaz SM, Amalfitano S, Kats LB. Aquatic invasive species: challenges for the future. Hydrobiologia. 2015;750:147–70. https://doi.org/10.1007/S10750-014-2166-0/FIGURES/1.
12. Thomaz SM, Cunha ER. da. The role of macrophytes in habitat structuring in aquatic ecosystems: methods of measurement, causes and consequences on animal assemblages' composition and biodiversity. Acta Limnologica Brasiliensia. 2010;22:218–36. https://doi.org/10.4322/ACTALB.02202011
13. Mehle N, Ravnikar M. Plant viruses in aqueous environment - survival, water mediated transmission and detection. Water Res. 2012;46:4902–17. https://doi.org/10.1016/J.WATRES.2012.07.027.
14. Mehle N, Gutiérrez-Aguirre I, Prezelj N, Delić D, Vidic U, Ravnikar M. Survival and transmission of potato virus Y, Pepino mosaic virus, and potato spindle tuber viroid in water. Appl Environ Microbiol. 2014;80:1455–62. https://doi.org/10.1128/AEM.03349-13/SUPPL_FILE/ZAM999105144SO1.PDF.
15. Bačnik K, Kutnjak D, Pecman A, Mehle N, Tušek Žnidarič M, Gutiérrez Aguirre I, et al. Viromics and infectivity analysis reveal the release of infective plant viruses from wastewater into the environment. Water Res. 2020;177. https://doi.org/10.1016/J.WATRES.2020.115628.
16. Mehle N, Bačnik K, Bajde I, Brodarič J, Fox A, Gutiérrez-Aguirre I, et al. Tomato brown rugose fruit virus in aqueous environments – survival and significance of water-mediated transmission. Front Plant Sci. 2023;14:1187920. https://doi.org/10.3389/FPLS.2023.1187920/BIBTEX.
17. Wang P, Liu J, Lyu Y, Huang Z, Zhang X, Sun B, et al. A review of Vector-Borne rice viruses. Viruses. 2022;14:2258. https://doi.org/10.3390/V14102258.
18. Wu N, Zhang L, Ren Y, Wang X. Rice black-streaked Dwarf virus: from multi-party interactions among plant–virus–vector to intermittent epidemics. Mol Plant Pathol. 2020;21:1007–19. https://doi.org/10.1111/MPP.12946.
19. Qiang Q, Wang Ang W, Jian J, Yang Ang Y, Guo Uo -H, -Hui G et al. Zhou Hou 1 1 Z,. The Complete Genome Sequence of Two Isolates of Southern rice black-streaked dwarf virus, a New Member of the Genus Fijivirus. Journal of Phytopathology. 2010;158:733–7. https://doi.org/10.1111/J.1439-0434.2010.01679.X
20. Omiat EG, Asante MD, Traoré VSE, Oppong A, Ifie BE, Ofosu KA, et al. Genetic diversity and epidemic histories of rice yellow mottle virus in Ghana. Virus Res. 2023;329:199106. https://doi.org/10.1016/J.VIRUSRES.2023.199106.
21. Bagayoko I, Celli MG, Romay G, Poulicard N, Pinel-Galzi A, Julian C, et al. Genetic diversity of rice Stripe necrosis virus and new insights into evolution of the genus benyvirus. Viruses. 2021;13. https://doi.org/10.3390/V13050737.
22. Lv L, Wu X, Weng J, Lai Y, Han K, Lu Y, et al. Complete genome sequence of a putative novel ilarvirus isolated from Eleocharis dulcis. Arch Virol. 2021;166:3477–81. https://doi.org/10.1007/S00705-021-05249-X.
23. Zhang F, Yang Z, Hong N, Wang G, Wang A, Wang L. Identification and characterization of water chestnut soymovirus-1 (WCSV-1), a novel soymovirus in water chestnuts (Eleocharis dulcis). BMC Plant Biol. 2019;19:1–14. https://doi.org/10.1186/S12870-019-1761-7/FIGURES/9.
24. Liu J, Wang YF, Hong N, Wang GP, Wang LP. The first report of cucumber mosaic virus infecting water chestnut in China. Plant Dis. 2014;98:164. https://doi.org/10.1094/PDIS-06-13-0663-PDN.
25. Venkataravanappa V, Krishna Reddy M. Viral diseases of ornamental plants-II. Viral Diseases of Field and Horticultural Crops. 2024;:789–804. https://doi.org/10.1016/B978-0-323-90899-3.00098-7
26. Yu X, Sheng JJ, Zheng XW, Diao Y, Zheng XF, Xie KQ et al. First Report of Dasheen mosaic virus Infecting Lotus (Nelumbo nucifera) in China. https://doi.org/101094/PDIS-01-15-0106-PDN. 2015;99. https://doi.org/10.1094/PDIS-01-15-0106-PDN
27. He Z, Chen W, Chen C, Liu X, Li L. First Report of Apple Stem Grooving Virus Infecting Lotus (Nelumbo nucifera) in China. https://doi.org/101094/PDIS-09-18-1626-PDN. 2019;103. https://doi.org/10.1094/PDIS-09-18-1626-PDN
28. Wang H, Liu X, Gan H, Chen W, Wu P, Li L, et al. Genomic and biological characterization of a novel strain of sweet potato latent virus isolated from Lotus (Nelumbo nucifera Gaertn). J Plant Pathol. 2019;101:1077–84. https://doi.org/10.1007/S42161-019-00338-Y/METRICS.
29. Skotnicki A, Gibbs A, Wrigley NG. Further studies on Chara corallina virus. Virology. 1976;75:457–68. https://doi.org/10.1016/0042-6822(76)90043-X.
30. Gibbs A, Skotnicki AH, Gardiner JE, Walker ES, Hollings M. A tobamovirus of a green Alga. Virology. 1975;64:571–4. https://doi.org/10.1016/0042-6822(75)90136-1.
31. Gibbs AJ, Torronen M, Mackenzie AM, Wood JT, Armstrong JS, Kondo H, et al. The enigmatic genome of Chara australis virus. J Gen Virol. 2011;92:2679–90. https://doi.org/10.1099/VIR.0.033852-0.

32. JONES P. Leaf mottling of spartina species caused by a newly recognised virus, spartina mottle virus. Ann Appl Biol. 1980;94:77–81. https://doi.org/10.1111/J.1744-7348.1980.TB03898.X.
33. Ilbağı H. Common Reed (Phragmites communis) is a natural host of important cereal viruses in the Trakya region of Turkey. Phytoparasitica. 2006;34:441–8. https://doi.org/10.1007/BF02981198/METRICS.
34. Van Bogaert N, Rosario K, Furman BT, Hall MO, Greco AM, Breitbart M. Discovery of a novel potexvirus in the seagrass thalassia testudinum from Tampa bay, Florida. Limnol Oceanogr Lett. 2019;4:1–8. https://doi.org/10.1002/lol2.10098.
35. Bejerman N, Debat H. Exploring the tymovirales landscape through metatranscriptomics data. Arch Virol. 2022;167:1785–803. https://doi.org/10.1007/s00705-022-05493-9.
36. Park D, Goh CJ, Kim H, Hahn Y. Identification of two novel amalgaviruses in the common eelgrass Zostera Marina and in Silico analysis of the Amalgavirus + 1 programmed ribosomal frameshifting sites. Plant Pathol J. 2018;34:150–6. https://doi.org/10.5423/PPJ.NT.11.2017.0243.
37. Goh CJ, Park D, Lee JS, Davey PA, Pernice M, Ralph PJ, et al. Zostera virus T - a novel virus of the genus tepovirus identified in the eelgrass, Zostera muelleri. Acta Virol. 2019;63:366–72. https://doi.org/10.4149/AV_2019_409.
38. Rede JE, Breitbart M, Lundquist C, Nagasaki K, Hewson I. Diverse RNA viruses discovered in multiple seagrass species. PLoS ONE. 2024;19. https://doi.org/10.1371/journal.pone.0302314.
39. Shin C, Choi D, Hahn Y. Identification of the genome sequence of Zostera associated varicosavirus 1, a novel negative-sense RNA virus, in the common eelgrass (Zostera marina) transcriptome. Acta Virol. 2021;65:373–80. https://doi.org/10.4149/av_2021_404.
40. Ruiz Martínez E, Mckeown DA, Schroeder DC, Thuestad G, Sjøtun K, Sandaa RA, et al. Phaeoviruses present in cultured and natural Kelp species, Saccharina latissima and laminaria hyperborea (Phaeophyceae, laminariales), in Norway. Viruses. 2023;15. https://doi.org/10.3390/v15122331.
41. McKeown DA, Schroeder JL, Stevens K, Peters AF, Sáez CA, Park J, et al. Phaeoviral infections are present in macrocystis, Ecklonia and Undaria (Laminariales) and are influenced by wave exposure in ectocarpales. Viruses. 2018;10. https://doi.org/10.3390/v10080410.
42. Leebens-Mack JH, Barker MS, Carpenter EJ, Deyholos MK, Gitzendanner MA, Graham SW, et al. One thousand plant transcriptomes and the phylogenomics of green plants. Nat 2019. 2019;574:7780. https://doi.org/10.1038/s41586-019-1693-2.
43. Rosario K, Van Bogaert N, López-Figueroa NB, Paliogiannis H, Kerr M, Breitbart M. Freshwater macrophytes harbor viruses representing all five major phyla of the RNA viral Kingdom orthornavirae. PeerJ. 2022;10. https://doi.org/10.7717/peerj.13875.
44. Mifsud JCO, Gallagher RV, Holmes EC, Geoghegan JL. Transcriptome mining expands knowledge of RNA viruses across the plant Kingdom. J Virol. 2022;96. https://doi.org/10.1128/JVI.00260-22.
45. Data packages for One Thousand Plant transcriptomes and phylogenomics of green plants. CyVerse Data Commons. 2019. https://datacommons.cyverse.org/browse/iplant/home/shared/commons_repo/curated/oneKP_capstone_2019. Accessed 17 Jan 2025.
46. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015;12:59–60. https://doi.org/10.1038/NMETH.3176.
47. Huson DH, Beier S, Flade I, Górska A, El-Hadidi M, Mitra S, et al. PLoS Comput Biol. 2016;12:e1004957. https://doi.org/10.1371/JOURNAL.PCBI.1004957. MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data.
48. Moubset O, François S, Maclot F, Palanga E, Julian C, Claude L, et al. Virion-Associated nucleic Acid-Based metagenomics: A decade of advances in molecular characterization of plant viruses. Phytopathology. 2022;112:2253–72. https://doi.org/10.1094/PHYTO-03-22-0096-RVW/ASSET/IMAGES/LARGE/PHYTO-03-22-0096-RVWT4-1669974912649.JPEG.
49. ICTV Report Chapters | ICTV. https://ictv.global/report. Accessed 8 Jan 2025.
50. Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol. 2021;38:3022–7. https://doi.org/10.1093/MOLBEV/MSAB120.
51. Muhire BM, Varsani A, Martin DP. SDT: A virus classification tool based on pairwise sequence alignment and identity calculation. PLoS ONE. 2014;9:e108277. https://doi.org/10.1371/JOURNAL.PONE.0108277.
52. PM 7/151 (1) Considerations for the use of high throughput sequencing in plant health diagnostics1. EPPO Bulletin. 2022;52:619–42. https://doi.org/10.1111/epp.12884
53. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using spades de Novo assembler. Curr Protoc Bioinf. 2020;70:e102. https://doi.org/10.1002/CPBI.102.
54. Consortium TU, Bateman A, Martin M-J, Orchard S, Magrane M, Adesina A, et al. UniProt: the universal protein knowledgebase in 2025. Nucleic Acids Res. 2025;53:D609–17. https://doi.org/10.1093/NAR/GKAE1010.
55. Van Der Vlugt RAA, Steffens P, Cuperus C, Barg E, Lesemann DE, Bos L, et al. Further evidence that Shallot yellow Stripe virus (SYSV) is a distinct potyvirus and reidentification of Welsh onion yellow Stripe virus as a SYSV strain. Phytopathology. 1999;89:148–55. https://doi.org/10.1094/PHYTO.1999.89.2.148.
56. Fowkes AR, McGreig S, Pufal H, Duffy S, Howard B, Adams IP, et al. Integrating high throughput sequencing into survey design reveals turnip yellows virus and soybean Dwarf virus in pea (Pisum Sativum) in the united Kingdom. Viruses. 2021;13. https://doi.org/10.3390/V13122530.
57. Mehle N, Nikolić P, Rupar M, Boben J, Ravnikar M, Dermastia M. Automated DNA extraction for large numbers of plant samples. Methods Mol Biol. 2013;938:139–45. https://doi.org/10.1007/978-1-62703-089-2_12.
58. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3–new capabilities and interfaces. Nucleic Acids Res. 2012;40. https://doi.org/10.1093/NAR/GKS596.
59. Li R, Salih S, Hurtt S. Detection of geminiviruses in Sweetpotato by polymerase chain reaction. Plant Dis. 2004;88:1347–51. https://doi.org/10.1094/PDIS.2004.88.12.1347.
60. Wyatt SD, Brown JK. Detection of subgroup III geminivirus isolates in leaf extracts by degenerate primers and polymerase chain reaction. Phytopathology. 1996;86:1288–93. https://doi.org/10.1094/PHYTO-86-1288.
61. Saison A, Gentit P. Development of a polyvalent detection method for Begomoviruses presenting a threat to the European tomato industry. TESTA EPPO Conference on diagnostics for plant pests.
62. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. TrimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:1972–3. https://doi.org/10.1093/BIOINFORMATICS/BTP348.
63. Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 2016;44:W232–5. https://doi.org/10.1093/NAR/GKW256.
64. Minh BQ, Nguyen MAT, Von Haeseler A. Ultrafast approximation for phylogenetic bootstrap. Mol Biol Evol. 2013;30:1188–95. https://doi.org/10.1093/MOLBEV/MST024.
65. Letunic I, Bork P. Interactive tree of life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res. 2024;52:W78–82. https://doi.org/10.1093/NAR/GKAE268.
66. Brown JK, Zerbini FM, Navas-Castillo J, Moriones E, Ramos-Sobrinho R, Silva JCF, et al. Revision of begomovirus taxonomy based on pairwise sequence comparisons. Arch Virol. 2015;160:1593–619. https://doi.org/10.1007/S00705-015-2398-Y.
67. Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol. 2006;23:254–67. https://doi.org/10.1093/MOLBEV/MSJ030.
68. Hamming RW. Error detecting and error correcting codes. Bell Syst Tech J. 1950;29:147–60. https://doi.org/10.1002/J.1538-7305.1950.TB00463.X.
69. Bryant D, Moulton V. Neighbor-Net: an agglomerative method for the construction of phylogenetic networks. Mol Biol Evol. 2004;21:255–65. https://doi.org/10.1093/MOLBEV/MSH018.
70. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evol (N Y). 1985;39:783. https://doi.org/10.2307/2408678.
71. Kuhn JH, Brown K, Adkins S, Torre JC, de la, Digiaro M, Ergünay K, et al. Promotion of order bunyavirales to class bunyaviricetes to accommodate a rapidly increasing number of related Polyploviricotine viruses. J Virol. 2024;98. https://doi.org/10.1128/JVI.01069-24.
72. Congdon BS, Baulch JR, Filardo FF, Nancarrow N. Turnip yellows virus variants differ in host range, transmissibility, and virulence. Arch Virol. 2023;168. https://doi.org/10.1007/S00705-023-05851-1.
73. Yusop MSM, Mohamed-Hussein ZA, Ramzi AB, Bunawan H. Cymbidium mosaic virus infecting orchids: what, how, and what next?? Iran J Biotechnol. 2022;20:e3020. https://doi.org/10.30498/IJB.2021.278382.3020.
74. Implementing regulation –. 2021/2069 - EN - EUR-Lex. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32021R2069. Accessed 17 Jan 2025.
75. The ICTV Report on Virus Classification and Taxon Nomenclature (10th report). Genus: begomovirus. ICTV. 2017. https://ictv.global/report/chapter/geminiviridae/geminiviridae/begomovirus. Accessed 17 Jan 2025.
76. Al Rwahnih M, Alabi OJ, Westrick NM, Golino D, Rowhani A. Description of a novel monopartite geminivirus and its defective subviral genome in

grapevine. Phytopathology. 2017;107:240–51. https://doi.org/10.1094/PHYTO-07-16-0282-R/ASSET/IMAGES/LARGE/PHYTO-07-16-0282-R_T5-1483426329394.JPEG.
77. Stanley J, Saunders K, Pinner MS, Wong SM. Novel defective interfering DNAs associated with ageratum yellow vein geminivirus infection of ageratum conyzoides. Virology. 1997;239:87–96. https://doi.org/10.1006/VIRO.1997.8856.
78. Huang C, Xie Y, Zhao L, Ren H, Li Z. A naturally occurring defective DNA satellite associated with a monopartite begomovirus: evidence for recombination between alphasatellite and betasatellite. Viruses. 2013;5:2116. https://doi.org/10.3390/V5092116.
79. Varsani A, Martin DP, Randles JW, Vetten HJ, Thomas JE, Fiallo-Olivé E, et al. Taxonomy update for the family alphasatellitidae: new subfamily, genera, and species. Arch Virol. 2021;166:3503–11. https://doi.org/10.1007/S00705-021-05232-6/FIGURES/2.
80. The ICTV Report on Virus Classification and Taxon Nomenclature. (10th report). Potyviridae. https://ictv.global/report_9th/RNApos/Potyviridae. Accessed 17 Jan 2025.
81. Wilson WH, Van Etten JL, Allen MJ. The phycodnaviridae: the story of how tiny giants rule the world. Curr Top Microbiol Immunol. 2009;328:1. https://doi.org/10.1007/978-3-540-68618-7_1.
82. Benites LF, Alves-Lima C. Viruses of rhodophyta: lack of cultures and genomic resources pose a threat to the growing red algal aquaculture industry. Appl Phycol. 2022;3:285–99. https://doi.org/10.1080/26388081.2022.2059701.
83. Lachnit T, Thomas T, Steinberg P. Expanding our Understanding of the seaweed holobiont: RNA viruses of the red Alga Delisea pulchra. Front Microbiol. 2016;6(JAN):167691. https://doi.org/10.3389/FMICB.2015.01489/BIBTEX.
84. Rousvoal S, Bouyer B, López-Cristoffanini C, Boyen C, Collén J. Mutant swarms of a totivirus-like entities are present in the red macroalga chondrus Crispus and have been partially transferred to the nuclear genome. J Phycol. 2016;52:493–504. https://doi.org/10.1111/JPY.12427.
85. Chiba Y, Tomaru Y, Shimabukuro H, Kimura K, Hirai M, Takaki Y, et al. Viral RNA genomes identified from marine macroalgae and a diatom. Microbes Environ. 2020;35:ME20016. https://doi.org/10.1264/JSME2.ME20016.
86. Debat H, Garcia ML, Bejerman N. Expanding the repertoire of the Plant-Infecting ophioviruses through metatranscriptomics data. Viruses. 2023;15:840. https://doi.org/10.3390/V15040840/S1.
87. Polischuk V, Budzanivska I, Shevchenko T, Oliynik S. Evidence for plant viruses in the region of Argentina islands, Antarctica. FEMS Microbiol Ecol. 2007;59:409–17. https://doi.org/10.1111/J.1574-6941.2006.00242.X.
88. Ros-Moner E, Jiménez-Góngora T, Villar-Martín L, Vogrinec L, González-Miguel VM, Kutnjak D, et al. Conservation of molecular responses upon viral infection in the non-vascular plant Marchantia polymorpha. Nat Commun. 2024;15:8326. https://doi.org/10.1038/S41467-024-52610-0.
89. Bhat AI, Hohn T, Selvarajan R, Badnaviruses. The current global scenario. Viruses. 2016;8:177. https://doi.org/10.3390/V8060177.
90. Borah BK, Sharma S, Kant R, Johnson AMA, Saigopal DVR, Dasgupta I. Bacilliform DNA-containing plant viruses in the tropics: commonalities within a genetically diverse group. Mol Plant Pathol. 2013;14:759. https://doi.org/10.1111/MPP.12046.
91. Vassilieff H, Geering ADW, Choisne N, Teycheney PY, Maumus F, Endogenous Caulimovirids. Fossils, Zombies, and Living in Plant Genomes. Biomolecules. 2023;13(1069):2023;13:1069. https://doi.org/10.3390/BIOM13071069
92. Diop SI, Geering ADW, Alfama-Depauw F, Loaec M, Teycheney PY, Maumus F. Tracheophyte genomes keep track of the deep evolution of the caulimoviridae. Sci Rep. 2018;8:572. https://doi.org/10.1038/S41598-017-16399-X.
93. Ammar ED, Tsai CW, Whitfield AE, Redinbaugh MG, Hogenhout SA. Cellular and molecular aspects of rhabdovirus interactions with insect and plant hosts. Annu Rev Entomol. 2009;54 Volume 54, 2009:447–68. https://doi.org/10.1146/ANNUREV.ENTO.54.110807.090454/1
94. Sidharthan VK, Baranwal VK. Mining of the water hyssop (Bacopa monnieri) transcriptome reveals genome sequences of two putative novel rhabdoviruses and a solendovirus. Arch Virol. 2021;166:1985–90. https://doi.org/10.1007/S00705-021-05061-7.
95. Bejerman N, Dietzgen RG, Debat H. Illuminating the plant rhabdovirus landscape through metatranscriptomics data. Viruses. 2021;13:1304. https://doi.org/10.3390/V13071304/S1.
96. Bejerman N, Dietzgen RG, Debat H. Unlocking the hidden genetic diversity of varicosaviruses, the neglected plant rhabdoviruses. Pathogens. 2022;11:1127. https://doi.org/10.3390/PATHOGENS11101127/S1.
97. Boccardo G, Lisa V, Luisoni E, Milne RG. Cryptic plant viruses. Adv Virus Res. 1987;32 C:171–214. https://doi.org/10.1016/S0065-3527(08)60477-7.
98. Roossinck MJ. Lifestyles of plant viruses. Philosophical Trans Royal Soc B: Biol Sci. 2010;365:1899. https://doi.org/10.1098/RSTB.2010.0057.
99. Dominguez-Huerta G, Zayed AA, Wainaina JM, Guo J, Tian F, Pratama AA et al. Diversity and ecological footprint of Global Ocean RNA viruses. Science (1979). 2022;376:1202–8. https://doi.org/10.1126/SCIENCE.ABN6358/SUPPL_FILE/SCIENCE.ABN6358_DATA_S1_AND_S2.ZIP
100. Edgar RC, Taylor J, Lin V, Altman T, Barbera P, Meleshko D, et al. Petabase-scale sequence alignment catalyses viral discovery. Nat 2022. 2022;602:7895. https://doi.org/10.1038/s41586-021-04332-2.
101. Neri U, Wolf YI, Roux S, Camargo AP, Lee B, Kazlauskas D, et al. Expansion of the global RNA Virome reveals diverse clades of bacteriophages. Cell. 2022;185:4023–e403718. https://doi.org/10.1016/J.CELL.2022.08.023.
102. Duffus JE, Liu HY, Wisler GC, Li R. Lettuce chlorosis virus - A new whitefly-transmitted closterovirus. Eur J Plant Pathol. 1996;102:591–6. https://doi.org/10.1007/BF01877027.
103. Ruiz ML, Simón A, García MC, Janssen D. First Report of Lettuce chlorosis virus Infecting Bean in Spain. https://doi.org/101094/PDIS-09-13-0939-PDN. 2014;98:857. https://doi.org/10.1094/PDIS-09-13-0939-PDN
104. Zhang SB, Zhang DY, Liu Y, Luo XW, Liu MY, Du J et al. First report of Lettuce chlorosis virus infecting tomato in China. https://doi.org/101094/PDIS-09-16-1315-PDN. 2017;101:846. https://doi.org/10.1094/PDIS-09-16-1315-PDN
105. Jacquemond M. Cucumber mosaic virus. Adv Virus Res. 2012;84:439–504. https://doi.org/10.1016/B978-0-12-394314-9.00013-0.
106. Cardin L, Delecolle B, Moury B. Occurrence of Mosaic Caused by Cucumber mosaic virus in Lobelia Hybrids in France and Italy. https://doi.org/101094/PD-89-0340B. 2007;89:340–340. https://doi.org/10.1094/PD-89-0340B
107. Stevens M, Mcgrann G, Clark B. Turnip yellows virus (syn beet Western yellows virus): an emerging threat to European oilseed rape production? 2008.
108. Slavíková L, Ibrahim E, Alquicer G, Tomašechová J, Šoltys K, Glasa M, et al. Weed hosts represent an important reservoir of turnip yellows virus and a possible source of virus introduction into oilseed rape crop. Viruses. 2022;14. https://doi.org/10.3390/v14112511.
109. Sanches MM, Lamas NS, Reis MBA, Arieta-Sosa JG, Romano E, Melo FL, et al. Genome assembly of Schlumbergera virus X infecting prickly Pear (Opuntia cochenillifera) in Brazil. Genome Announc. 2015;3:133–48. https://doi.org/10.1128/GENOMEA.00133-15.
110. Janssen D, García C, Ruiz L. First report of Schlumbergera virus X in Dragon fruit (Hylocereus spp.) in Spain. https://doi.org/101094/PDIS-09-21-1879-PDN. 2022;106:2004. https://doi.org/10.1094/PDIS-09-21-1879-PDN
111. June-Wells M, Vossbrinck CR, Gibbons J, Bugbee G. The aquarium trade: A potential risk for nonnative plant introductions in connecticut, USA. Lake Reserv Manag. 2012;28:200–5. https://doi.org/10.1080/07438141.2012.693575.
112. Rúa MA, Pollina EC, Power AG, Mitchell CE. The role of viruses in biological invasions: friend or foe? Curr Opin Virol. 2011;1:68–72. https://doi.org/10.1016/J.COVIRO.2011.05.018.
113. Najberek K, Olszańska A, Tokarska-Guzik B, Mazurska K, Dajdok Z, Solarz W. Invasive alien species as reservoirs for pathogens. Ecol Indic. 2022;139:108879. https://doi.org/10.1016/J.ECOLIND.2022.108879.
114. EUR-Lex –. 02019R2072-20220714 - EN - EUR-Lex. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02019R2072-20220714%26qid=1667573350824#tocId27. Accessed 27 Feb 2025.
115. Begomovirus (1BEGOG)[Categorization]| EPPO Global Database. https://gd.eppo.int/taxon/1BEGOG/categorization. Accessed 27 Feb 2025.
116. Alternanthera reineckii Briq. Plants of the World Online. Kew Science. https://powo.science.kew.org/taxon/urn:lsid:ipni.org:names:10461-2. Accessed 17 Jan 2025.
117. Lindernia rotundifolia (L.). Alston. Plants of the World Online. Kew Science. https://powo.science.kew.org/taxon/urn:lsid:ipni.org:names:1061345-2. Accessed 17 Jan 2025.
118. Shi M, Fu J. Alligator weed alternanthera Philoxeroides (Mart.) griseb. Biological invasions and its management in China. 2017;:163–73. https://doi.org/10.1007/978-981-10-3427-5_12
119. Pan XY, Geng YP, Sosa A, Zhang WJ, Li B, Chen JK. Invasive alternanthera philoxeroides: biology, ecology and management. Acta Phytotaxonomica Sinica. 2007;45:884–900. https://doi.org/10.1360/APS06134.
120. Rimac A, Stanković I, Alegro A, Gottstein S, Koletić N, Vuković N, et al. The Brazilian Elodea (Egeria densa Planch.) invasion reaches Southeast Europe. Bioinvasions Rec. 2018;7:381–9. https://doi.org/10.3391/BIR.2018.7.4.05.
121. Lacet JB, Moura A, do N, Dantas ÊW. Invasion of Egeria densa Planch. in semiarid reservoirs. Revista Brasileira de Botanica. 2019;42:491–7. https://doi.org/10.1007/S40415-019-00548-9

122. Šajna N, Haler M, Škornik S, Kaligarič M. Survival and expansion of pistia stratiotes L. in a thermal stream in Slovenia. Aquat Bot. 2007;87:75–9. https://doi.org/10.1016/J.AQUABOT.2007.01.012.
123. Šajna N, Urek T, Kušar P, Šipek M. The importance of thermally abnormal waters for Bioinvasions—A case study of pistia stratiotes. Divers (Basel). 2023;15:421. https://doi.org/10.3390/D15030421/S1.
124. Bemisia tabaci (BEMITA). EPPO Global Database. https://gd.eppo.int/taxon/BEMITA. Accessed 17 Jan 2025.
125. Milenovic M, Eickermann M, Junk J, Rapisarda C. Life history parameters of bemisia tabaci MED (Hemiptera: Aleyrodidae) in the present and future climate of central europe, predicted by physically realistic Climatic chamber simulation. Environ Entomol. 2023;52:502–9. https://doi.org/10.1093/EE/NVAD023.
126. Han S, Xing Z, Jiang H, Li W, Huang W. Biological adaptive mechanisms displayed by a freshwater plant to live in aquatic and terrestrial environments. Environ Exp Bot. 2021;191:104623. https://doi.org/10.1016/J.ENVEXPBOT.2021.104623.
127. Wang P, Zhang Q, Xu YS, Yu FH. Effects of water level fluctuation on the growth of submerged macrophyte communities. Flora: morphology, distribution, functional ecology of plants. 2016;223:83–9. https://doi.org/10.1016/J.FLORA.2016.05.005
128. Macek P, Rejmánková E, Houdková K. The effect of long-term submergence on functional properties of Eleocharis cellulosa Torr. Aquat Bot. 2006;84:251–8. https://doi.org/10.1016/J.AQUABOT.2005.11.003.
129. Fontdevila Pareta N, Khalili M, Maachi A, Rivarez MPS, Rollin J, Salavert F, et al. Managing the deluge of newly discovered plant viruses and viroids: an optimized scientific and regulatory framework for their characterization and risk analysis. Front Microbiol. 2023;14:1181562. https://doi.org/10.3389/FMICB.2023.1181562/XML.

Publisher's note