Degenerate PCR primers for potexvirus detection in seagrasses

1.1 Author names

Shen Jean Lim[1*] (https://orcid.org/0000-0003-4578-5318)
Karyna Rosario[1] (https://orcid.org/0000-0001-9847-4113)
Meredith E. Kernbach[1] (https://orcid.org/0000-0001-8161-057X)
Anthony J. Gross[1] (https://orcid.org/0000-0001-6392-5121)
Bradley T. Furman[2] (https://orcid.org/0000-0002-9284-8388)
Mya Breitbart[1] (https://orcid.org/0000-0003-3210-2899)

1.2 Affiliation(s)

[1] College of Marine Science, University of South Florida, St Petersburg, FL, USA
[2] Florida Fish and Wildlife Conservation Commission, Florida Fish and Wildlife Research Institute, Saint Petersburg, FL, USA

1.3 Corresponding author and email address

Shen Jean Lim (jeanlim@usf.edu)

1.4 Keywords

Potexvirus, primer, PCR, seagrass, marine, virome, *Alphaflexivirdae*

1.5 Repositories:

All sequence data have been deposited into NCBI GenBank under the accession numbers OR827689-OR827705, OR854648, OR863396, and OR879052-OR879056.

2. Abstract

The genus *Potexvirus* includes many species known to infect terrestrial plants and, more recently, a novel species, *Turtlegrass virus X* (TVX) infecting the seagrass *Thalassia testudinum*. We designed a degenerate primer pair, Potex7F and Potex7-RC, to amplify a 584 nt RNA-dependent RNA polymerase (RdRp) fragment from potexviruses infecting seagrasses. This primer pair was modified from Potex5/Potex2-RC primers published in 2002 that target terrestrial plant potexviruses, by including nucleotide bases found in TVX and related potexvirus sequences identified in seagrass viromes from Tampa Bay, Florida, USA. Using this newly developed primer set, we performed a reverse transcriptase polymerase chain reaction (RT-PCR) survey to screen for potexviruses in 63 opportunistically collected, apparently healthy seagrass samples. The survey examined the host species *T. testudinum*, *Halodule wrightii*, *Halophila stipulacea*, *Syringodium filiforme*, *Ruppia maritima*, and *Zostera marina*. PCR products were successfully amplified and sequenced from *T. testudinum* samples collected around Florida, USA, indicating prevailing potexvirus infection in the region.

3. Impact statement

Potexviruses are widespread in terrestrial plants; however, the recent discovery of a potexvirus in the seagrass *Thalassia testudinum* extends their host range to marine flowering

plants. Here we present degenerate potexvirus PCR primers, Potex7F and Potex7-RC, to enable the identification of potexvirus infections in seagrass habitats. Discovery of potexviruses in *T. testudinum* highlights the utility of this new primer set in uncovering the diversity, host range, and geographic range of potexviruses infecting seagrasses.

4. Data summary

All sequence data are available in NCBI GenBank under the accession numbers OR827689-OR827705, OR854648, OR863396, and OR879052-OR879056. The authors confirm all supporting data, code and protocols have been provided within the article or through supplementary data files.

5. Introduction

Potexviruses are positive-sense single-stranded RNA (ssRNA) viruses from the family *Alphaflexiviridae* that infect many agronomically important flowering plants (angiosperms), including potato, tobacco, and tomato (1). The type member of this genus is *Potato virus X* (PVX), from which the genus name *Potexvirus* is derived (2). Potexviruses have a broad geographic range that reflects the distribution of their hosts and the prevalence of global commerce (2, 3). Although each potexvirus species has a limited natural host range, some species can infect a wide range of host plants in experimental settings (1). Potexviruses are transmitted mechanically through contaminated agricultural equipment or plant-to-plant contact (1-3). Transmission via seed and aphid vectors and non-specific transmission through chewing insects has been reported but is less common (1-3). Most potexviruses cause persistent, asymptomatic infections or mild mosaic disease, while some cause dwarf, necrotic, or ringspot symptoms in their natural hosts (1, 2). Notably, PVX can cause severe symptoms when its plant host is co-infected with potyviruses (4-6).

In addition to infecting terrestrial plants, potexviruses are also capable of infecting seagrasses, the only marine angiosperms. A novel potexvirus species, *Turtlegrass virus X* (TVX), was discovered in apparently healthy *Thalassia testudinum* seagrass samples collected from Terra Ceia Aquatic Preserve, Tampa Bay, Florida, USA (7). The identification of TVX indicates that an important aspect of potexvirus diversity has been overlooked by focusing only on terrestrial plants. Further examination of this viral group in understudied aquatic habitats necessitates the development of methods to effectively identify a broad range of potexviruses. Degenerate primers used in reverse transcriptase polymerase chain reaction (RT-PCR) assays, such as Potex5/Potex-2RC (8), have proven useful in detecting diverse potexviruses. However, these existing potexvirus primers were designed based on species infecting terrestrial plants and could fail to amplify novel species with divergent sequences, such as TVX (7).

This study aimed to develop a broad RT-PCR assay that expands the detection of potexviruses to include those infecting seagrasses. Seagrass potexvirus sequences identified using viromics were aligned with published sequences from the TVX genome (7), PVX genome (9), and the degenerate Potex5/Potex-2RC primers (8) to evaluate sequence identity and primer coverage. We modified the existing degenerate potexvirus primer set to include

nucleotide bases found in the seagrass potexvirus sequences and used the new degenerate primers, named Potex 7F/Potex7-RC, to explore potexvirus diversity in seagrass samples.

6. Methods

Sample collection for viromes. The Tampa Bay region contains over 60 seagrass transects (10) that are monitored annually, with seagrass and geographic data available from the Tampa Bay Estuary Program (https://tbep.org/seagrass-assessment/). Seagrass samples from the species *T. testudinum*, *H. wrightii*, *S. filiforme*, and *R. maritima* were collected for viromics in January and February 2022 from eight Tampa Bay seagrass sites (S1T5, S1T16, S2T2, S3T8, S3T13, S4T5, S4T10, and BCB11; **Figure 1**, **Table 1**, and **Table S1**). From each site, ten seagrass leaves from multiple shoots were collected from each species, stored in bags, and transferred to the laboratory on ice. In the lab, seagrass leaves were rinsed with MilliQ water, cleaned with sterile razor blades to remove visible epiphytes, then stored at -80°C until virus particle purification.

Virome preparation. Virus-like particles (VLPs) were purified from ~150 mg pooled leaf tissue from each species at each site using previously described methods (7). Briefly, each sample was homogenized in a Zymo Research (ZR) BashingBead Lysis Tube (Irvine, CA, USA) containing 2 mm ceramic beads and 800 µL Suspension Medium (100 mM NaCl, 50 mM Tris-HCl (pH=7.5), and 8 mM $MgSO_4$) for 1.5 minutes at maximum speed (5 m/s) using a Fisherbrand™ Bead Mill 4 Homogenizer (Fisher Scientific, Waltham, MA, USA). Tissue homogenates were centrifuged at 6,000 *g* for 6 minutes at room temperature and syringe-filtered through 0.45-µm Sterivex filters (Millipore, Burlington, MA, USA). Chloroform was added to a 480 µL aliquot of each filtrate to a final concentration of 20%. The mixture was vortexed vigorously, incubated at room temperature for 10 minutes, and centrifuged at 16,100 *g* for 30 seconds. The aqueous layer (~170 µL) was treated with a nuclease mixture of 21 U of TURBO™ DNase (Invitrogen™, Waltham, MA, USA), 4.5 U of Baseline-ZERO™ DNase (Epicentre, Paris, France), 112.5 U Benzonase® endonuclease (EMD Millipore, Burlington, MA, USA), and 400 U Ambion™ RNase I in 1X Turbo™ DNase Buffer (Invitrogen™) for 1.5 hours at 37°C, followed by nuclease inactivation with 20 mM EDTA (pH=8.0).

From each purified viral fraction, RNA was extracted using Qiagen's RNeasy Mini kit (Valencia, CA, USA) with on-column DNase digestion. Extracted nucleic acids from VLPs and two extraction blanks were used for cDNA synthesis using the SuperScript™ IV First-Strand Synthesis System (Invitrogen™) with random hexamers provided by the manufacturer. cDNA products were pooled based on transect and species (**Table 1**) for second strand synthesis and sequencing. Second strand synthesis was performed on a total of 8 cDNA pools using New England Biolabs' DNA Polymerase I, Large (Klenow) Fragment (Ipswich, MA, USA). Double-stranded cDNA samples were purified with ZR's DNA Clean & Concentrator-25 kit. Purified cDNA was fragmented to 300 bp using the Covaris® M220 Focused-ultrasonicator (Woburn, MA, USA) at the University of South Florida's Genomics Program Sequencing core (Tampa, FL, USA). Fragmented cDNA samples were used for NGS library preparation with the xGen™ ssDNA & Low-Input DNA Library Preparation Kit (Integrated DNA Technologies, Coralville, IA, USA) following the manufacturer's protocol for DNA concentrations >1 ng/µL, except for the extraction blank pool for which the protocol for DNA concentrations <1 ng/µL

125 was used. All libraries (**Table 1**) were sequenced on the Illumina HiSeq platform (2 x 150 bp)
126 at Azenta Life Sciences (Burlington, MA, USA).
127
128 ***Virome sequence analysis.*** Sequencing reads were processed, assembled, and annotated
129 using methods described in Rosario *et al.* (2022) (11). Reads were trimmed using
130 Trimmomatic v0.36.0 (12) using the default parameters with the exception of a read head
131 crop of 10 bases. Read quality post-trimming was assessed using FastQC v0.11.5
132 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Trimmed reads from each
133 cDNA library were assembled using the pipeline recommended for PCR-amplified viral
134 metagenomes (13) which included read deduplication using BBtool's clumpify.sh (dedupe
135 subs=0 passes=2; https://sourceforge.net/projects/bbmap) and assembly using single-cell
136 SPAdes (14). From each assembly, contigs >1,000 bp were retained using the filter by
137 sequence length tool on the Galaxy server (https://usegalaxy.org) (15) and compared with
138 viral protein sequences in Release 93 of the National Center for Biotechnology Information
139 (NCBI)'s Reference Sequence (RefSeq) database (16) using BLAST+ v2.7.1 (17) with an e-value
140 threshold of $10x10^{-10}$. BLAST results were imported into MEGAN6 Community Edition (18) to
141 detect contigs similar to potexviruses. Pairwise average nucleotide identities (ANI) between
142 identified seagrass potexvirus sequences and the TVX genome (NCBI accession number:
143 NC_040644) (7) were computed using fastANI v0.1.3 (19), which only reports alignment(s)
144 with >80% ANI, and pairwise average amino acid identities (AAI) were calculated using
145 CompareM v0.1.2 (https://github.com/donovan-h-parks/CompareM).
146
147 ***Primer design*.** RNA-dependent RNA polymerase (RdRp) sequences from the potexvirus
148 contigs annotated from the Tampa Bay seagrass viromes were aligned using Multiple
149 Alignment using Fast Fourier Transform (MAFFT) v7.5.08 (5) with the L-INS-i iterative
150 refinement method incorporating local pairwise alignment information and the --
151 adjustdirectionaccurately option to generate reverse complement sequences if necessary.
152 Primer sequences of Potex5 and Potex-2RC were retrieved from Van der Vlugt and
153 Berendsen (2002) (8) and manually compared to the multiple sequence alignment to
154 evaluate primer coverage. The Potex5/Potex-2RC primers were modified to include
155 nucleotide bases found in the seagrass potexvirus sequences (**Figure 2**) and named Potex7F
156 (5'-CAYCARCARGCHAARGAYGW-3') and Potex7-RC (5'-AGCATVGCNSCRCYTG-3').
157
158 ***Primer validation and optimization*.** Potex7F/Potex7-RC were tested and optimized on cDNA
159 from *T. testudinum* samples known to contain TVX, based on the TVX-specific RT-PCR assay
160 described by Van Bogaert *et al.* 2019 (7). A gradient PCR reaction was performed to
161 determine the optimal annealing temperature of Potex7F/Potex7-RC for TVX detection. The
162 reaction contained 0.48 µM of each primer, 2 µL cDNA template, 1 µL GC enhancer, and 1X
163 AmpliTaq Gold™ 360 Master Mix (Applied Biosystems™, Waltham, MA, USA) in a 25 µL
164 reaction volume. Gradient PCR was performed under the following conditions: Initial
165 denaturation at 95°C for 10 minutes, 40 cycles of denaturation at 95°C for 30 seconds,
166 annealing at 44.3-51.4°C for 30 seconds, extension at 72°C for 1 minute, followed by
167 elongation at 72°C for 10 minutes and cooling at 11°C. The PCR product was visualized
168 following gel electrophoresis on a 1% (wt/vol) agarose gel stained with ethidium bromide.
169 All PCR reactions, except for the no template control, yielded visible bands. An annealing
170 temperature of 50.9°C produced a bright band with no nonspecific amplification; therefore,

171 this temperature was used for all RT-PCR assays. Potex7F/Potex7-RC were further validated
172 on cDNA samples synthesized from 1 µg RNA extracted from two terrestrial plants infected
173 with potexviruses. For these samples, gradient PCR was performed using lower annealing
174 temperatures of 40-50.9°C to generate sufficient PCR products for sequencing. All PCR
175 products were purified using the Zymoclean Gel DNA Recovery Kit, quantified using the
176 Qubit™ DNA high sensitivity (HS) assay (Invitrogen™), and Sanger sequenced bidirectionally
177 by TACGen (Richmond, CA, USA) or Eurofin Genomics (Louisville, KY, USA).
178
179 ***RT-PCR***. A RT-PCR assay using Potex 7F/7-RC was performed on seagrass leaves collected
180 from separate sampling efforts following the protocols detailed above. These included in-
181 house collections of 1) *T. testudinum*, *H. wrightii*, *S. filiforme*, and *R. maritima* from Tampa
182 Bay sites used for virome sequencing; 2) *T. testudinum* from the Dry Tortugas National Park,
183 Florida, USA collected between May 16th-20th, 2022; 3) *T. testudinum* from a systematic
184 seagrass survey at Terra Ceia Aquatic Preserve on August 1st, 2022; and 4) *T. testudinum*, *H.*
185 *wrightii*, and *S. filiforme* samples collected from Tampa Bay seagrass site S3T8 (Lassing Park)
186 on October 3rd, 2023 (**Figure 1** and **Table S1**). Seagrass samples were also received from
187 other research groups, including *T. testudinum* from Panacea (Florida, USA), *Zostera marina*
188 from York River (Virginia, USA), West Falmouth Harbor (Massachusetts, USA), Sitka (Alaska,
189 USA), New Zealand, and Kalmar (Sweden), and *Halophila stipulacea* from Jobos Bay National
190 Estuarine Research Reserve, Puerto Rico (**Table S1**).
191
192 From each species and site, 30-100 mg of leaves from multiple shoots were pooled for total
193 RNA extraction. Each pooled seagrass sample was homogenized in a ZR BashingBead Lysis
194 Tube containing 2 mm ceramic beads and 800 µL RNA lysis buffer for 5 minutes at maximum
195 speed using a Fisherbrand™ Bead Mill 4 Homogenizer. Tissue homogenates were
196 centrifuged at maximum speed (21,130 *g*) for 1 minute and RNA was extracted from the
197 total volume of the supernatant (~800 µL) using ZR's Quick-RNA™ Plant Miniprep kit. Each
198 RNA sample was quantified using the Qubit™ RNA high sensitivity (HS) assay. From each
199 sample, cDNA was synthesized from 8 µL RNA and 50 ng random hexamers using the
200 SuperScript™ IV First-Strand Synthesis System (Invitrogen™). PCR amplification was
201 performed on each cDNA sample at 50.9°C annealing temperature, followed by gel
202 electrophoresis, PCR product purification, DNA quantification, and Sanger sequencing, using
203 methods described above.
204
205 ***Amplicon sequence analysis*.** Amplicon sequences were compared with sequences from
206 NCBI's nucleotide (nt) and non-redundant protein (nr) collections (16) and the TVX genome
207 (7) using the megablast and/or blastx programs on NCBI's BLAST® server (20). Amplicon
208 sequence pairs were quality-trimmed and assembled into consensus sequences using the
209 default parameters of CAP3 (21) implemented in Unipro UGENE v48.1 (22). Forward and
210 reverse sequences from six PCR products could not be assembled due to non-overlapping
211 fragments. For each unassembled sequence pair, the forward or reverse sequence with
212 higher Q20 values was used for downstream analysis.
213
214 Sequences from 22 Potex7F/Potex7-RC PCR products were aligned with the two potexvirus
215 sequences from Tampa Bay seagrass viromes and reference genome sequences of TVX
216 (NC_040644) (7), *Bamboo mosaic virus* (NC_001642) (23), *Foxtail mosaic virus* (NC_001483)

(24), PVX (NC_011620) (9), and *Lolium latent virus* (NC_010434) (25) retrieved from the NCBI Virus database (26). Multiple sequence alignment was performed using the L-INS-I and --adjustdirectionaccurately options in MAFFT v7.5.08 (5). Pairwise sequence identities were calculated from this alignment using Clustal Omega v1.2.3 (27) with the --distmat-out, --percent-id, and --full options. The pairwise sequence identity matrix was clustered, based on Euclidean distances, using the hierarchical clustering with complete linkage method implemented in the heatmap.2 function of the gplots R package (https://github.com/talgalili/gplots).

For phylogenetic analysis, a 449 nt conserved region within the multiple sequence alignment was extracted using the default options of gblocks v0.91b (28). Molecular Evolutionary Genetics Analysis (MEGA) 11 (29) identified the best model for the conserved blocks to be the Hasegawa-Kishino-Yano model (30) with a discrete Gamma distribution and (HKY+G). Using this substitution model, a maximum likelihood tree with 1,000 bootstrap replicates was constructed from the conserved blocks using MEGA11 (29). The tree was condensed using MEGA11 (29) to eliminate interior branches with <50% bootstrap support. The final tree was visualized and annotated using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/Figtree/).

7. Results

We sequenced seven RNA viromes from four seagrass species, *T. testudinum*, *H. wrightii*, *S. filiforme*, and *R. maritima*, collected and pooled from eight seagrass sites (10) throughout Tampa Bay in January and February 2022 (**Table 1** and **Figure 1**). From these viromes, we detected a 3,039 nt potexvirus contig from pool 1 (containing *T. testudinum* samples pooled from upper and middle Tampa Bay; NCBI accession number: OR827689) and a 3,771 nt potexvirus contig from pool 7 (containing *R. maritima* samples pooled from throughout Tampa Bay; NCBI accession number: OR827690) that were most closely related to TVX. The partial potexvirus genomes shared <80% average nucleotide identity (ANI) and 80.7% average amino acid identity (AAI) with each other. The sequence from pool 1 shared 88.5% AAI with the TVX genome (7), while the sequence from pool 7 shared 84.9% AAI with the TVX genome.

Table 1. Details of viromes sequenced in this study.

Library	Seagrass species	Region of Tampa Bay	Sites
Pool 1	*Thalassia testudinum*	Upper and Middle	S1T5; S1T16; S3T13; S3T8
Pool 2	*Thalassia testudinum*	Lower	S4T5; S4T10; BCB11
Pool 3	*Syringodium filiforme*	Upper and Middle	S1T5; S1T16; S3T8
Pool 4	*Syringodium filiforme*	Lower	S4T5; S4T10; BCB11
Pool 5	*Halodule wrightii*	Upper and Middle	S1T5; S1T16; S2T2; S3T13; S3T8
Pool 6	*Halodule wrightii*	Lower	S4T5; S4T10; BCB11
Pool 7	*Ruppia maritima*	Upper, Middle, and Lower	S1T5; S1T16; S2T2; S3T8; S4T10
Pool 8	Blank cDNA	N/A	N/A

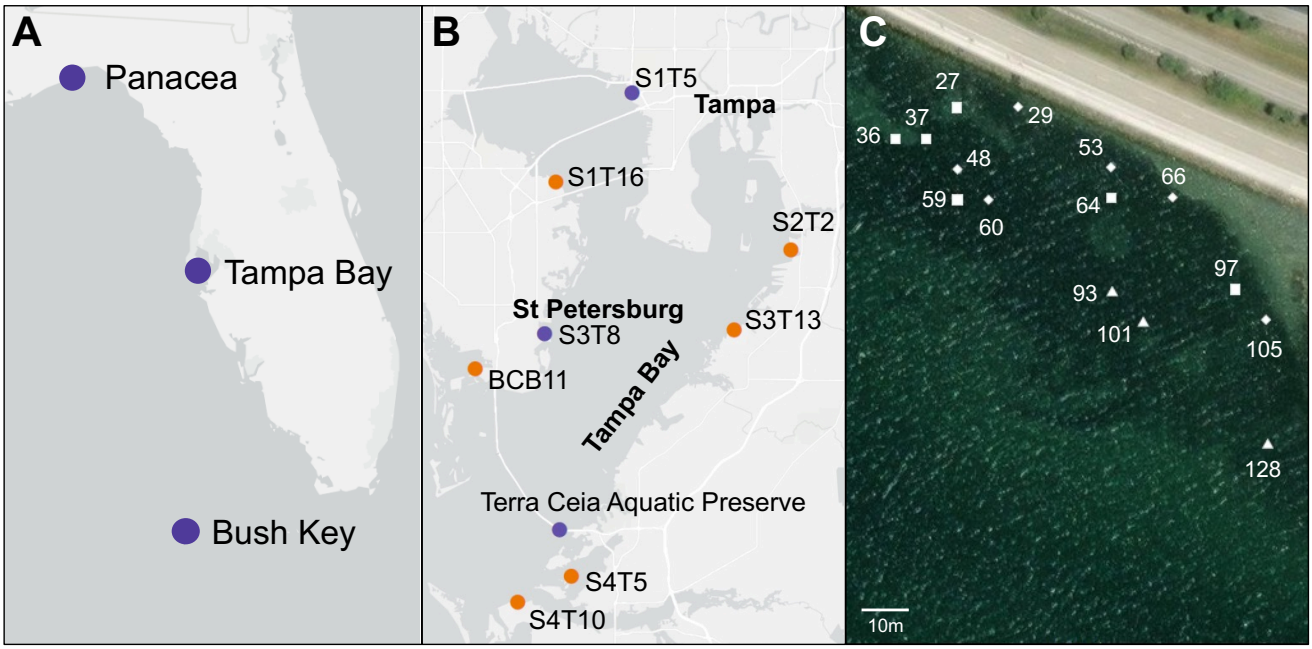

Figure 1. Map showing locations of seagrass sites in (A) Florida, (B) Tampa Bay, and (C) Terra Ceia Aquatic Preserve where samples were collected for viromics and/or RT-PCR. Purple markers in (A) and (B) denote sites where potexvirus products were amplified from *T. testudinum* with the Potex7F/Potex7-RC PCR primers, whereas samples from sites with orange markers tested negative. Marker shapes in (C) represent potexvirus subclades found in Terra Ceia Aquatic Preserve sampling sites and correspond to **Figure 3** and **Figure 4** (negative samples not shown). Metadata of seagrass samples is described in **Table S1**.

We used the Potex7F/Potex7-RC primer pair to test for potexviruses in 67 seagrass samples, including *T. testudinum*, *H. wrightii*, *H. stipulacea*, *S. filiforme*, *R. maritima*, and *Z. marina* from within and outside the USA (**Table S1**). PCR products were successfully amplified and sequenced from 20 *T. testudinum* samples from Florida sites, including Terra Ceia Aquatic Preserve, Tampa Bay seagrass sites S1T5 and S3T8 (Lassing Park), Panacea located in the Florida Panhandle, and Bush Key located in the Dry Tortugas National Park (**Figure 1**). Although a potexvirus contig was initially identified in the *R. maritima* virome (pool 7), PCR amplification on *R. maritima* samples collected from the same sites was not successful.

A Forward primers

PVX [NC_011620]	5'-	C	A	T	C	A	A	C	A	A	G	C	A	A	A	A	G	A	T	G	A	3'-
TVX [NC_040644]	5'-	C	A	C	C	A	G	C	A	G	G	C	A	A	A	G	G	A	C	G	A	3'-
Pool 1 [OR827689]	5'-	C	A	T	C	A	G	C	A	A	G	C	T	A	A	A	G	A	C	G	A	3'-
Pool 7 [OR827690]	5'-	C	A	C	C	A	G	C	A	G	G	C	A	A	A	G	G	A	C	G	T	3'-
Potex5 primer	5'-	C	A	Y	C	A	R	C	A	R	G	C	**M**	A	A	R	G	A	Y	G	**A**	3'-
Potex7F primer	5'-	C	A	Y	C	A	R	C	A	R	G	C	**H**	A	A	R	G	A	Y	G	**W**	3'-

B Reverse primers

PVX [NC_011620]	5'-	C	A	G	G	A	T	G	G	A	G	C	T	A	T	G	C	T	3'-
TVX [NC_040644]	5'-	C	A	G	G	A	C	G	G	A	G	C	G	A	T	G	C	T	3'-
Pool 1 [OR827689]	5'-	C	A	G	G	A	C	G	G	C	G	C	C	A	T	G	C	T	3'-
Pool 7 [OR827690]	5'-	C	A	G	G	A	T	G	G	C	G	C	T	A	T	G	C	T	3'-
Potex2 primer	5'-	C	A	R	G	A	Y	G	S	N	G	C	**Y**	A	T	G	C	T	3'-
Potex7 primer	5'-	C	A	R	G	A	Y	G	S	N	G	C	**B**	A	T	G	C	T	3'-
Potex2-RC primer	3'-	G	T	Y	C	T	R	C	S	N	C	G	**R**	T	A	C	G	A	5'-
Potex7-RC primer	3'-	G	T	Y	C	T	R	C	S	N	C	G	**V**	T	A	C	G	A	5'-

Figure 2. Alignment of PVX, TVX, and other seagrass potexvirus sequences with (A) forward (Potex5) and (B) reverse (Potex2-RC) primers designed by Van der Vlugt and Berendsen (2002) and in this study (Potex7F/Potex7-RC). Conserved positions in the alignment are highlighted in grey and modified nucleotide bases are highlighted in red.

Phylogenetic analysis placed the seagrass potexvirus RdRp sequences identified in viromes and Potex7F/Potex7-RC primer products from seagrasses in the same clade as the RdRp sequence fragment from the TVX genome (7) with 99% bootstrap confidence (**Figure 3**). Potexviruses that infect terrestrial plants formed sister clades to the TVX-containing clade. The TVX-containing clade comprised three subclades containing potexvirus sequences from Terra Ceia Aquatic Preserve (**Figure 3**). Average pairwise nucleotide sequence identities were 93±3% within each subclade and 79±3% between each subclade (**Figure 4**). Other potexvirus sequences obtained from upper Tampa Bay (site S1T5), Lassing Park (site S3T8), Panacea, Bush Key, and the *R. maritima* virome (pool 7) grouped within the TVX-containing clade but did not cluster with any other sequences (**Figure 3**).

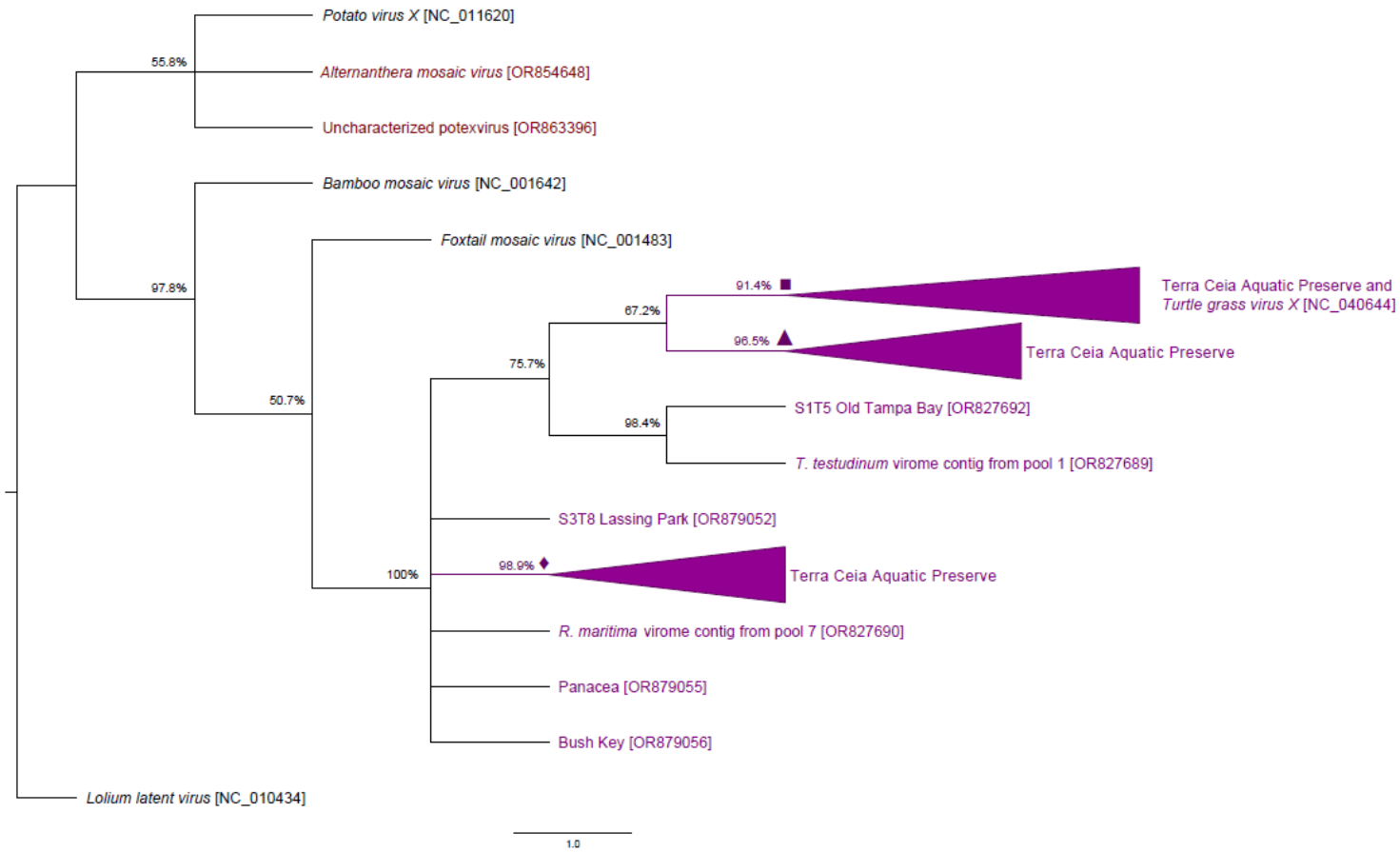

Figure 3. Condensed bootstrap consensus tree of potexvirus nucleotide sequences from seagrass viromes and PCR products (purple text), in relation to potexvirus sequences amplified from terrestrial plants (brown text) and potexvirus sequences from reference genomes (black text). The NCBI accession number for each sequence is indicated in square brackets. Accession numbers for sequences from Terra Ceia Aquatic Preserve samples are listed in **Figure 4**. The tree was constructed by MEGA11 from a 449 nt conserved region in the multiple sequence alignment. Branch lengths indicate the number of substitutions per site. Bootstrap values on tree nodes indicate the percentage of trees, based on 1,000 replicates, in which sequences within a node were clustered together. Shapes beside bootstrap values of Terra Ceia Aquatic Preserve subclades correspond to shapes in **Figure 1C**. *Lolium latent virus* was used as an outgroup.

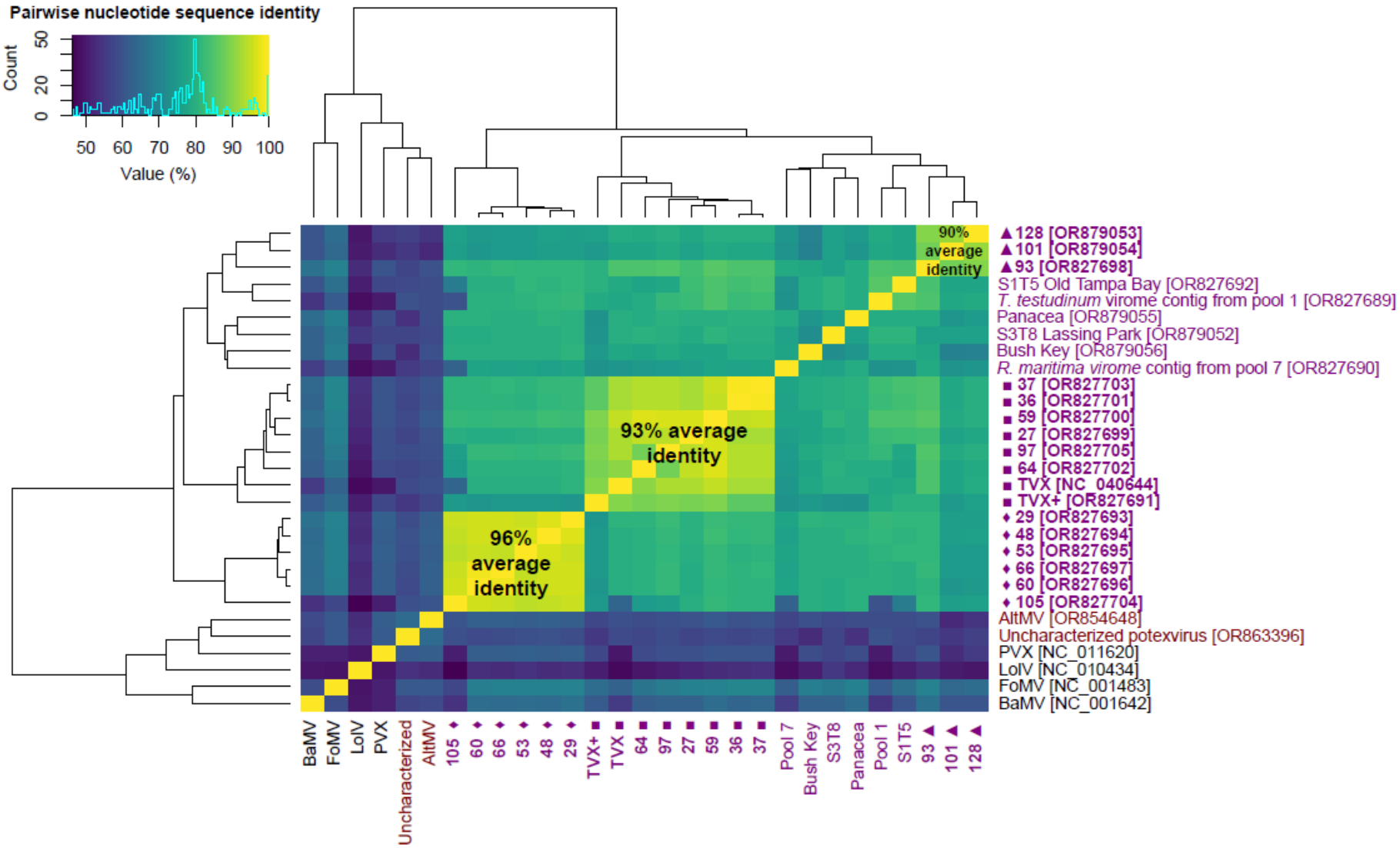
Pairwise nucleotide sequence identity
Count
0
20
50
50
60
70
80
90
100
Value (%)
90% average identity
93% average identity
96% average identity
▲128 [OR879053]
▲101 [OR879054]
▲93 [OR827698]
S1T5 Old Tampa Bay [OR827692]
T. testudinum virome contig from pool 1 [OR827689]
Panacea [OR879055]
S3T8 Lassing Park [OR879052]
Bush Key [OR879056]
R. maritima virome contig from pool 7 [OR827690]
■ 37 [OR827703]
■ 36 [OR827701]
■ 59 [OR827700]
■ 27 [OR827699]
■ 97 [OR827705]
■ 64 [OR827702]
■ TVX [NC_040644]
■ TVX+ [OR827691]
♦ 29 [OR827693]
♦ 48 [OR827694]
♦ 53 [OR827695]
♦ 66 [OR827697]
♦ 60 [OR827696]
♦ 105 [OR827704]
AltMV [OR854648]
Uncharacterized potexvirus [OR863396]
PVX [NC_011620]
LolV [NC_010434]
FoMV [NC_001483]
BaMV [NC_001642]
BaMV
FoMV
LolV
PVX
Uncharacterized
AltMV
105 ♦
60 ♦
66 ♦
53 ♦
48 ♦
29 ♦
TVX+ ■
TVX ■
64 ■
97 ■
27 ■
59 ■
36 ■
37 ■
Pool 7
Bush Key
S3T8
Panacea
Pool 1
S1T5
93 ▲
101 ▲
128 ▲

Figure 4. Histogram (upper left corner), heatmap, and dendrogram (hierarchical clustering) of pairwise nucleotide sequence identities between potexvirus sequences from seagrass viromes and PCR products (purple text), potexvirus sequences amplified from terrestrial plants (brown text), and potexvirus RdRp sequences from reference genomes (black text). The NCBI accession number for each sequence is indicated in square brackets. Terra Ceia Aquatic Preserve seagrass samples are highlighted in bold, with marker shapes corresponding to shapes in **Figure 1C**. Abbreviations: TVX, *Turtle grass virus X*; TVX+, amplicon sequence from *T. testudinum* samples known to contain TVX (positive controls); AltMV, *Alternanthera mosaic virus*; PVX, *Potato virus X*; LolV, *Lolium latent virus*; FoMV, *Foxtail mosaic virus*; BaMV, *Bamboo mosaic virus*.

8. Discussion

In this study, we developed a degenerate primer pair, Potex7F/Potex7-RC, to identify potexvirus infection in seagrasses. Potex7F/Potex7-RC was validated on TVX-infected seagrass samples and potexvirus-infected terrestrial plant samples. Gradient PCR showed higher optimal annealing temperatures for seagrass samples compared to terrestrial plant samples; therefore PCR optimization is crucial for the efficient detection of novel potexviruses.

Potex7F/Potex7-RC were used in a RT-PCR survey to test six seagrass species for potexviruses. Among the seagrass species examined, potexvirus infection was only detected in *T. testudinum*, the species from which TVX was identified. It should be noted that 31 of 54 samples (54%) screened were *T. testudinum,* so while potexviruses were not identified in the other seagrass species, future research may reveal a more widespread host range amongst marine angiosperms.

The RT-PCR survey confirmed that *T. testudinum* is a natural host of potexviruses and that potexvirus infection is prevalent in Florida seagrass beds, including northwest (Panacea) and southwest (Tampa Bay and Dry Tortugas National Park) Florida. Within a single seagrass meadow at Terra Ceia Aquatic Preserve, three distinct potexvirus sequence variants were observed. Among all sites sampled in this study, the most extensive potexvirus survey was conducted within the Terra Ceia Aquatic Preserve. Systematic RT-PCR surveys at other sites, coupled with viromics, will be useful in elucidating potexvirus diversity in seagrasses and in resolving potexvirus sequence variants into species or strains. The sampling scheme in this study was limited to *T. testudinum* populations along the western coast of Florida adjacent to the Gulf of Mexico. However, *T. testudinum* has a widespread geographic range across the Western Atlantic Ocean, with high genetic diversity that is generally partitioned into two clusters corresponding to the Gulf of Mexico or the Caribbean phylogeographic regions (31). Future potexvirus surveys covering the total distribution range of *T. testudinum*, the related species *T. hemprichii*, and other seagrass species have the potential to uncover relationships between potexvirus diversity, host diversity, and their phylogeographic distributions.

9. Author statements

9.1 Author contributions

SJL: Data curation, Formal analysis, Software, Visualization, Writing - original draft
KRC: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Writing – review & editing.
MEK: Investigation, Writing – review & editing
AJG: Investigation, Writing - review & editing
BTF: Investigation, Writing - review & editing
MB: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing - review & editing.

9.2 Conflicts of interest

The author(s) declare that there are no conflicts of interest.

9.3 Funding information

This project was supported by award OCE-2219547 from the National Science Foundation to MB, KRC, and BTF and a Proposal Enhancement Grant from the University of South Florida Internal Awards Program to MB. This research was made possible by support from the Florida Institute of Oceanography for collection of the Dry Tortugas samples.

9.4 Ethical approval

N/A

9.5 Consent for publication

N/A

9.6 Acknowledgements

Fieldwork was conducted under permit no. FKNMS-2022-026 issued by the Florida Keys National Marine Sanctuary, permit no. DRTO-2022-SCI-0004 issued through the U.S. Department of the Interior National Park Service and permit no. SAL-19-1172A-SR issued through the Florida Fish and Wildlife Conservation Commission. We acknowledge the contributions of Luis Alvarado-Marchena, Victoria Congdon, Keith Keel, Isabella Ritchie, Taylor Little, Jessye Kirkham, and Mike Wheeler in field work and sample collection at these sites. Isabella Ritchie provided cDNAs from *T. testudinum* samples from Terra Ceia Aquatic Preserve. Salvador P Lopez, Jr and Scott Adkins from the Agricultural Research Service, United States Horticultural Research Laboratory, U.S. Department of Agriculture (USDA), Fort Pierce, FL, USA provided cDNAs from potexvirus-infected plant samples. We thank Cypress Rudloe, Maya Groner, Ian Hewson, Karin Holmfeldt, Natalia Lopez Figueroa, Carolyn Malmstrom, and Jordan Rede for providing seagrass samples from other sites.

10. References

1. Ryu KH, Hong JS. Potexvirus. In: Mahy BWJ, Van Regenmortel MHV, editors. Encyclopedia of Virology (Third Edition). Oxford: Academic Press; 2008. p. 310-3.
2. Verchot J. *Potato virus X*: A global potato-infecting virus and type member of the *Potexvirus* genus. Molecular Plant Pathology. 2022;23(3):315-20.
3. AbouHaidar MG, Gellatly D. Potexviruses. In: Granoff A, Webster RG, editors. Encyclopedia of Virology (Second Edition). Oxford: Elsevier; 1999. p. 1364-8.
4. Rochow WF, Ross AF. Virus multiplication in plants doubly infected by potato viruses X and Y. Virology. 1955;1(1):10-27.
5. Pacheco R, Garcia-Marcos A, Barajas D, Martianez J, Tenllado F. PVX-potyvirus synergistic infections differentially alter microRNA accumulation in *Nicotiana benthamiana*. Virus Research. 2012;165(2):231-5.
6. Pollari M, Sipari N, Poque S, Himanen K, Makinen K. Effects of poty-potexvirus synergism on growth, photosynthesis and metabolite status of *Nicotiana benthamiana*. Viruses. 2022;15(1):121.
7. 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. Limnology and Oceanography Letters. 2019;4(1):1-8.
8. van der Vlugt RAA, Berendsen M. Development of a general potexvirus detection method. European Journal of Plant Pathology. 2002;108(4):367-71.

9. Huisman MJ, Linthorst HJ, Bol JF, Cornelissen JC. The complete nucleotide sequence of *Potato virus X* and its homologies at the amino acid level with various plus-stranded RNA viruses. Journal of General Virology. 1988;69(8):1789-98.
10. Sherwood ET, Greening HS, Johansson JOR, Kaufman K, Raulerson GE. Documenting seagrass recovery since the 1980's and reviewing the benefits. Southeastern Geographer. 2017;57(3):294-319.
11. Rosario K, Van Bogaert N, Lopez-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:e13875.
12. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20.
13. Roux S, Trubl G, Goudeau D, Nath N, Couradeau E, Ahlgren NA, et al. Optimizing *de novo* genome assembly from PCR-amplified metagenomes. PeerJ. 2019;7:e6902.
14. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. Journal of Computational Biology. 2012;19(5):455-77.
15. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Čech M, et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Research. 2018;46(W1):W537-W44.
16. Sayers EW, Beck J, Bolton EE, Bourexis D, Brister JR, Canese K, et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Research. 2021;49(D1):D10-D7.
17. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10:421.
18. Huson DH, Beier S, Flade I, Gorska A, El-Hadidi M, Mitra S, et al. MEGAN Community Edition - interactive exploration and analysis of large-scale microbiome sequencing data. PLoS Computational Biology. 2016;12(6):e1004957.
19. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nature Communications. 2018;9(1):5114.
20. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. Nucleic Acids Res. 2008;36(Web Server issue):W5-9.
21. Huang X, Madan A. CAP3: A DNA sequence assembly program. Genome Research. 1999;9(9):868-77.
22. Okonechnikov K, Golosova O, Fursov M, team U. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics. 2012;28(8):1166-7.
23. Lin NS, Lin BY, Lo NW, Hu CC, Chow TY, Hsu YH. Nucleotide sequence of the genomic RNA of bamboo mosaic potexvirus. Journal of General Virology. 1994;75(9):2513-8.
24. Bancroft JB, Rouleau M, Johnston R, Prins L, Mackie GA. The entire nucleotide sequence of *Foxtail mosaic virus* RNA. Journal of General Virology. 1991;72(9):2173-81.
25. Vaira AM, Maroon-Lango CJ, Hammond J. Molecular characterization of *Lolium latent virus*, proposed type member of a new genus in the family *Flexiviridae*. Archives of Virology. 2008;153(7):1263-70.
26. Brister JR, Ako-Adjei D, Bao Y, Blinkova O. NCBI viral genomes resource. Nucleic Acids Research. 2015;43(Database issue):D571-7.

27. Sievers F, Higgins DG. Clustal Omega for making accurate alignments of many protein sequences. Protein Science. 2018;27(1):135-45.
28. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Molecular Biology and Evolution. 2000;17(4):540-52.
29. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Molecular Biology and Evolution. 2021;38(7):3022-7.
30. Hasegawa M, Kishino H, Yano T. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution. 1985;22(2):160-74.
31. van Dijk K-j, Bricker E, van Tussenbroek BI, Waycott M. Range-wide population genetic structure of the Caribbean marine angiosperm *Thalassia testudinum*. Ecology and Evolution. 2018;8(18):9478-90.