WARWICK
THE UNIVERSITY OF WARWICK

Manuscript version: Author's Accepted Manuscript
The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:
http://wrap.warwick.ac.uk/143365

How to cite:
Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:
Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

warwick.ac.uk/lib-publications

Sequence analysis of 43-year old samples of *Plantago lanceolata* show that *Plantain virus X* is synonymous with *Actinidia virus X* and is widely distributed

Journal:	*Plant Pathology*
Manuscript ID	PP-20-245.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	04-Sep-2020

Complete List of Authors:	Hammond, John; USDA ARS Adams, Ian; Fera Science Ltd; University of Newcastle upon Tyne, Insitute of Agri-food Research and Innovation Fowkes, Aimee; Fera Science Ltd McGreig, Sam; Fera Science Ltd Botermans, Marleen; National Plant Protection Organisation van Oorspronk, Joanieke; National Plant Protection Organisation Westenberg, Marcel; National PLant Protection Organisation Verbeek, Martin; Wageningen University & Research Dullemans, Annette; Wageningen University & Research Stijger, Christina; Wageningen University & Research Blouin, Arnaud; University of Liege, TERRA - Plant Pathology Massart, Sebastien; University of Liege, TERRA - Plant Pathology De Jonghe, Kris; Instituut voor Landbouw-, Visserij- en Voedingsonderzoek, Plant Sciences Heyneman, Maaike; Instituut voor Landbouw-, Visserij- en Voedingsonderzoek, Plant Sciences Walsh, John; University of Warwick, Life Sciences Fox, Adrian; Fera Science Ltd

Topics:	diagnostics, taxonomy & phylogenetics
Organisms:	viruses & viroids
Other Keywords:	plantain virus X, actinidia virus X, historic isolates, sequencing

SCHOLARONE™
Manuscripts

1 **Sequence analysis of 43-year old samples of *Plantago lanceolata* show that**
2 ***Plantain virus X* is synonymous with *Actinidia virus X* and is widely distributed**

3 John Hammond[1], Ian Adams[2,8], Aimee R. Fowkes[2], Sam McGreig[2], Marleen
4 Botermans[3], Joanieke J.A. van Oorspronk[3], Marcel Westenberg[3], Martin Verbeek[4],
5 Annette M. Dullemans[4], Christina C.M.M. Stijger[4], Arnaud G. Blouin[5], Sebastien
6 Massart[5], Kris De Jonghe[6], Maaike Heyneman[6], John A. Walsh[7], Adrian Fox[2*]

7 1. USDA-ARS, Beltsville, Maryland, USA
8 2. Fera Science Ltd, York, YO41 1LZ, UK
9 3. National Plant Protection Organization, Geertjesweg 15, 6706 EA Wageningen,
10 The Netherlands
11 4. Wageningen University and Research, Droevendaalsesteeg 1, 6708 PB
12 Wageningen, The Netherlands
13 5. University of Liège, Gembloux Agro-BioTech, TERRA, Laboratory of
14 Phytopathology, Passage des déportés, 2, 5030 Gembloux, Belgium
15 6. Flanders Research Institute for Agriculture, Fisheries and Food (ILVO), Plant
16 Sciences, Burg. Van Gansberghelaan 96, 9820 Merelbeke, Belgium
17 7. School of Life Sciences, University of Warwick, Wellesbourne, Warwick CV35
18 9EF, UK
19 8. Institute for Agri-Food Research and Innovation, Newcastle University,
20 Newcastle, NE1 7RU, UK

21 *Corresponding author: adrian.fox@fera.co.uk

22 Keywords: plantain virus X, actinidia virus X, historic isolates, sequencing,

For Peer Review

23 **Abstract**

24 *Plantain virus X* was first recognized by the ICTV as a species in the genus *Potexvirus*
25 in 1982. However, because no sequence was available for plantain virus X (PlVX),
26 abolishing the species was proposed to the *Flexiviridae* working group of the ICTV in
27 2015. This initiated efforts to sequence the original isolates from *Plantago lanceolata*
28 samples. Here we report the full genome sequencing of two original isolates of PlVX,
29 which have demonstrated the virus to be synonymous to *Actinidia virus X* a species
30 previously reported from kiwifruit (*Actinidia* sp.) and blackcurrant (*Ribes nigrum*). PlVX
31 was previously noted to be widespread in the UK in *P. lanceolata*. This report
32 additionally presents novel data on the distribution and diversity of PlVX, collected at the
33 same site as the original UK isolates, and from three independent surveys, two in The
34 Netherlands and one in Belgium. This study also includes two new host records for
35 PlVX, *Browallia americana* and *Capsicum annuum* (sweet pepper), indicating the virus
36 is more widespread and infects a broader range of hosts than previously reported. This
37 stresses the importance of surveys of non-cultivated species to gain insight into viral
38 distribution and host range. This study also demonstrates the value of generating
39 sequence data for isolates retained in virus collections. Additionally, it demonstrates the
40 potential value in pre-publication data-sharing for giving context to virus detections such
41 as the four independent studies here which, when combined, give greater clarity to the
42 identity, diversity, distribution and host range of plantain virus X.

43

44

For Peer Review

45 **Introduction**

46 By 1982, *Plantago* species were known to support natural infection by at least 30
47 distinct viruses from diverse families, and to be experimental hosts to at least 13 more
48 (Hammond, 1982, Kostin & Volkov, 1976). More recently five additional viruses have
49 been reported from *Plantago lanceolata* in Finland, together with fragmentary
50 sequences of an apparent ophiovirus and an endornavirus (Susi et al., 2017, Susi et al.,
51 2019), although a caulimovirus may be the same as one reported by Hammond
52 (Hammond, 1981). *P. lanceolata* was first reported as a natural host of the potexvirus
53 plantain virus X (PlVX) in 1981 (Hammond, 1981, Hammond, 1982, Hammond & Hull,
54 1981). PlVX was detected in 51 of 130 plants during a survey of *P. lanceolata* collected
55 from eight of nine distinct regions of England (United Kingdom, UK) sampled, from
56 southwest Cornwall, east to Berkshire, Cambridgeshire, Essex and Norfolk, north to
57 Warwickshire, Cumbria and Northumberland. The virus, which induces no symptoms in
58 *P. lanceolata* was purified and physically and biologically characterized (Hammond &
59 Hull, 1981). It was first recognized by the ICTV as a species in the genus *Potexvirus* in
60 1982 (initially described as plantago virus X, corrected to *Plantain virus X* in 1991).
61 However, because no sequence was available, abolishing the species was proposed to
62 the Flexiviridae working group of the ICTV in 2015, prompting an effort to sequence the
63 original isolates from the samples collected in Cambridge, England, which had been
64 deposited in the collection of the late Alan Brunt, currently held at University of Warwick,
65 UK. A further reason to obtain sequence from PlVX was to evaluate the possibility that
66 PlVX might be synonymous with another potexvirus, *Plantago asiatica mosaic virus*,

For Peer Review

67 which was reported infecting *P. asiatica* in the Russian Far East (Kostin & Volkov, 1976)
68 at about the same time as PlVX was being studied in the UK.

69 In 2007 in New Zealand, virus research on symptomatic kiwifruit (*Actinidia chinensis*)
70 identified the presence of a novel potexvirus, recognized by the ICTV as *Actinidia virus*
71 *X* (Pearson et al., 2011, Blouin et al., 2013). Although actinidia virus X (AVX) was
72 inoculated to herbaceous indicators on three occasions, attempts to inoculate back to
73 *Actinidia* showed a rapid decrease of virus titre on systemic leaf to become
74 undetectable after two months suggesting that *Actinidia* may not be a suitable host. The
75 virus was later reported from a blackcurrant (*Ribes nigrum*) intercepted on entry to
76 Canada (James & Phelan, 2016), however the geographic source of this plant was not
77 reported.

78 This study reports the full genomes of two original isolates of PlVX stored in a virus
79 collection for 43 years and sequenced using high-throughput sequencing (HTS). These
80 sequence data unexpectedly demonstrated *Plantain virus X* to be synonymous to
81 *Actinidia virus X*. In addition, contemporary isolates of PlVX were collected from the
82 same sample sites in Cambridge, UK, as originally sampled. Isolates of PlVX from three
83 independent studies in mainland Europe are also presented. One study in The
84 Netherlands was investigating the potential presence and distribution of PlVX in *P.*
85 *lanceolata*. A second study in The Netherlands was investigating the potential risk of
86 viral inoculum present in wildflower banks where the virus was also revealed infecting *P.*
87 *lanceolata*. A study in Belgium was an ~~HTS~~ metagenomic based survey of the viral
88 status of the family *Solanaceae,* which revealed the presence of PlVX in *Browallia*
89 *americana*. An additional isolate sequence was identified from a reexamination of a viral

For Peer Review

90 sequence from a *Capsicum annuum* sample imported from Ethiopia into the
91 Netherlands in 2011. These studies were brought together through pre-publication data-
92 sharing amongst the European plant health virology community. The sequences of all
93 these isolates are presented here enhancing the knowledge of the distribution, diversity
94 and identity of PlVX.

95 **Materials and methods**

96 The virus isolates reported here are gathered from a range of sources. Each of the
97 collaborating laboratories detected virus isolates independently of each other and these
98 originated from historic isolate collections, prior laboratory interceptions, field survey
99 studies and targeted sampling of *P. lanceolata* in the UK, The Netherlands and Belgium.
100 Each laboratory conducted a combination of HTS and conventional sequencing,
101 however, specific methods differed between each laboratory. ~~High throughput~~
102 ~~sequencing was carried out by either~~ A range of metagenomic-based approaches were
103 used, ~~a~~including ribosomal RNA depleted total RNA or Virion-Associated Nucleic Acids
104 (VANA)~~ approach~~followed by HTS (Illumina). Conventional PCR was carried out by
105 either broad spectrum RT-PCR for potexviruses (Van der Vlugt & Berendsen, 2002) or
106 specific RT-PCR (primers from this study). The methods used relating to each sample
107 have been highlighted in table 1.~~ and detailed in Supplementary information S.1-S.7~~.

108

109 *Historic isolates of PlVX*

110 Sample references 76/16 and 77/48: A subset of the original samples collected from
111 Cambridge, England, had been stored dried over Calcium chloride and were deposited

For Peer Review

112 in the collection of the late Dr Alan Brunt. This collection was subsequently deposited
113 with The University of Warwick. Two isolates labelled PlVX 76/16 and 77/48 were
114 subsampled, and the subsamples were sent to Fera Science Ltd, York, UK, for high
115 throughput sequencing by a ribosomal RNA-depleted total RNA approach. (see
116 Supplementary information S.1). RNA was extracted from the two sub-samples using an
117 RNeasy plant mini kit (Qiagen, UK). Indexed TruSeq complete plant libraries, including
118 RNA depletion step, were produced from the RNA and sequenced using a MiSeq V2
119 600 cycle kit (Illumina). The resulting data was then analyzed as described in Fox et al.
120 (2019). Phylogenetic trees were produced using the Maximum likelihood algorithm and
121 500 bootstraps in MEGA 7 (Kumar et al., 2016). Pairwise identities were calculated
122 using the same software.

123

124

125 *Collection of contemporary samples of* P. lanceolata *from the United Kingdom*
126 Sample references FR3, FR4, FR6, FR9, WC1, WC3 and WC5: Contemporary samples
127 of *P. lanceolata* were collected in September 2018 from, and near, the same sites in
128 Cambridge, England, as originally sampled in 1976-78. Fourteen plants from two sites
129 within less than 1 km of each other near Cambridge were tested at Fera Science Ltd,
130 York UK. Nine plants came from the Fulbourn Road (FR prefix) and a further five from
131 Worts Causeway (WC prefix). Samples were tested as both bulked sub-samples and
132 individual samples using conventional RT-PCR (Van der Vlugt & Berendsen, 2002) and
133 sequencing of PCR product (see Supplementary information S.2). Initially plants were
134 sub-sampled and tested as three bulked samples: A (WC 1-5), B (FR 1-5), C (FR 6-9),

plantpath@bspp.org.uk

For Peer Review

135 and after all three tested positive by RT-PCR (Van der Vlugt & Berendsen, 2002) for the
136 presence of potexviruses the samples were extracted and tested individually. Leaves
137 were sampled randomly from around each plant to a total subsample of approximately
138 0.5 g (actual sample weights ranged from 0.25 g-1.47 g), these were diluted 1 in 10 in
139 extraction buffer (Adams et al., 2013) by weight, and subsequently ground in
140 homogenization bags (BIOREBA, Switzerland). RNA extraction was done by magnetic
141 bead extraction using Invimag Virus DNA/RNA mini-kit (Invitek GmbH). The samples
142 were tested using general potexvirus PCR primers (Van der Vlugt & Berendsen, 2002).
143 RT-PCR was carried out using Thermo Scientific Verso 1-Step RT-PCR Hot-Start Kit
144 (ThermoFisher) following the manufacturer's instructions, but using 25 µM of primers
145 and an annealing temperature of 52 °C (Van der Vlugt & Berendsen, 2002). PCR
146 products were analysed by electrophoresis using 1% agarose gel stained with ethidium
147 bromide; an amplicon of approximately 600 bp was obtained and sequenced by
148 Eurofins (Germany) to enable sequence identification.

149

150

151 *Collection and testing of contemporary samples of P. lanceolata from the Netherlands*

152 Sample references 614590 and 39014434: In the Netherlands, an investigation was
153 conducted to try and identify the presence of PlVX. Leaf samples were collected in four
154 different regions (Wageningen, Haaksbergen, Klarenbeek and Buren). Twenty leaf
155 samples were collected from each region and bulked in one sample. These were
156 submitted to the National Reference Laboratory of the Netherlands (NPPO-NL),

For Peer Review

157 Wageningen, Netherlands. Two samples tested positive using conventional RT-PCR
158 (Van der Vlugt & Berendsen, 2002) and sequencing of PCR product. (see
159 Supplementary information S.3). RNA was extracted from about 1 g frozen leaf tissue
160 according to (Botermans et al., 2013) iConventional RT-PCR was carried out for the
161 presence of potexviruses (Van der Vlugt & Berendsen, 2002). RT-PCR reactions were
162 carried out using SuperScript One-Step RT-PCR with Platinum Taq DNA Polymerase
163 (Invitrogen). Two samples gave positive results (references 614590 and 39014434),
164 PCR products were bi-directionally sequenced (Van De Vossenberg & Van der Straten,
165 2014) to enable (partial) identification by sequence analysis.
166
167
168 Sample reference AVX-2018-001: A separate project was conducted by Wageningen
169 University and Research (WUR), The Netherlands, to examine the potential of wild-
170 flower strips to act as a source of virus inoculum for crops. Samples of *P. lanceolata*
171 were taken from a 3-year old flower strip, located on a blueberry farm in the Overijssel
172 region of The Netherlands. Plant samples were ground in a phosphate inoculation buffer
173 and mechanically inoculated onto a set of four test plants, i.e. *Chenopodium quinoa*,
174 *Nicotiana benthamiana*, *N. glutinosa*, and *N. occidentalis* 'P1'. These inoculated
175 (indicator) plants developed heavy necrosis in the *N. occidentalis* plants and chlorotic or
176 necrotic local lesions in the *C. quinoa* plants. Electron microscopy on samples of the
177 necrotic *N. occidentalis* plants clearly indicated the presence of filamentous particles of
178 approximately 500 nm in length, indicating infections with a member of the genus
179 *Potexvirus* (family *Alphaflexiviridae*). A sample from symptomatic (necrotic) *N.*

For Peer Review

180 *occidentalis* 'P1' was tested using high throughput sequencing by a ribosomal RNA-
181 depleted total RNA approach (see Supplementary information S.4).. A sample from
182 symptomatic (necrotic) *N. occidentalis* 'P1' was extracted using the RNeasy plant mini
183 kit (Qiagen) according the manufacturer's guidelines, and tested by conventional RT-
184 PCR using the generic potexvirus primers Potex-5 (fw) and Potex-2RC according to
185 Van der Vlugt and Berendsen (2002) using the Access RT-PCR system (Promega) and
186 visualisation of the amplicon on a 1% Agarose gel stained with GelRed (Biotium). The
187 obtained amplicon was directly sequenced by Sanger sequencing using the same
188 primers at Macrogen Europe (Amsterdam, The Netherlands). Additionally, total RNA
189 was DNase treated and ribosome depleted using the Ribo-zero rRNA removal plant leaf
190 kit (Illumina) and two dual unique indexed libraries produced using the TruSeq stranded
191 total RNA library prep kit (Illumina) as per the manufacturer's instruction. The resulting
192 library was pooled with other indexed libraries, diluted to 10pM, mixed with 5% PhiX
193 library (Illumina) and sequenced on an Illumina MiSeq using a 500 cycle V2 kit at
194 Wageningen University and Research. Reads were split per sample by corresponding
195 molecular identifiers (MIDs) using CASAVA 1.8 software (Illumina) with no mismatch in
196 the MID region allowed. Data analysis was performed using CLC Genomics Workbench
197 12.0.2 (Qiagen). After quality trimming (settings: quality limit 0.05%; short reads <100 nt
198 and broken pairs were discarded), reads were used for *de novo* assembly. Contigs with
199 a minimum length of 500 nt were subsequently analyzed, using BLASTn and BLASTx.

200 The 3'UTR of the RNA segment was determined by sequencing RT-PCR fragments
201 generated using segment-specific forward primers corresponding to sequences located
202 at the 3' end of ORF5 in combination with an oligo dT primer. The 5'UTR was

For Peer Review

203 determined using a Roche 5′/3′ RACE Kit according to the manufacturer's protocol. In
204 this case, the reverse primers were based on the 5' region of ORF1. Sanger sequencing
205 was performed by Macrogen Europe (Amsterdam, The Netherlands).

206

207

208 Capsicum annuum *sample ex. Ethiopia, imported into the Netherlands*

209 Sample reference 5422861: Another study was a re-examination of a viral sequence
210 obtained previously: In 2011, symptoms were observed on a *Capsicum annuum* sample
211 imported from Ethiopia into the Netherlands. The leaves showed mild mottle symptoms.
212 A leaf sample of *C. annuum* was ground in a phosphate inoculation buffer and
213 mechanically inoculated onto a set of eight test plants, i.e. *Ca. annuum* 'Westlandse
214 Grote Zoete', *C. quinoa*, *Datura stramonium*, *N. benthamiana*, *N. glutinosa*, and *N.
215 occidentalis* 'P1', *N. tabacum* 'White Burley' and *Solanum lycopersicum* 'Money-maker'.
216 These inoculated (indicator) plants showed small necrotic local lesions and systemic
217 mild mottle in the *Ca. annuum* plants, small necrotic local and systemic lesions and
218 systemic growth reduction in the *N. occidentalis* 'P1' plants, chlorotic or necrotic local
219 lesions and systemic vein clearing in the *C. quinoa* plants, systemic rugosity on the
220 lower leaf of one *N. benthamiana*. On *N. glutinosa*, *N. tabacum* 'White Burley', *S.
221 lycopersicum* and *D. stramonium* no symptoms were observed. Electron microscopy on
222 a sample of an *N. benthamiana* plant clearly indicated the presence of filamentous
223 particles of approximately 575 nm in length, indicating infections with a member of the
224 genus *Potexvirus* (family *Alphaflexiviridae*). Subsequently, RNA was extracted from

For Peer Review

225 about 1 g leaf tissue of a symptomatic *N. benthamiana* and processed as described for
226 the *P. lanceolata* from the Netherlands ~~(Supplementary information S.3)~~, in order to
227 obtain sequence data (not presented). Additionally, whole genome sequence was
228 obtained using high throughput sequencing by a ribosomal RNA-depleted total RNA
229 approach ~~(see Supplementary information S.5)~~. Total RNA from frozen C. annuum
230 (5422861) leaf sample was DNase treated and sent to GenomeScan (Leiden, the
231 Netherlands) for generation of 2 Gb Illumina RNAseq 150PE (paired-end) data per
232 sample. The RNA extract was ribosome depleted using the Ribo-zero rRNA removal
233 plant leaf kit (Illumina). The Ultra II Directional RNA Library Prep Kit for Illumina (New
234 England Biolabs, MA, USA) was used to process the samples according to the protocol
235 "NEBNext Ultra II Directional RNA Library Prep Kit for Illumina". Quality and yield after
236 sample preparation were measured with a Fragment Analyzer (Agilent, CA, USA) prior
237 to pooling for sequencing on an Illumina NovaSeq (Illumina, CA, USA).

238 RNAseq data were analyzed in CLC Genomics workbench v11.0.1 (Qiagen, Germany)
239 and run in a custom workflow build for detection of de novo assembled viral contigs.
240 First, a quality trim (quality limit = 0.05; ambiguous limit = 2) was performed, followed by
241 a de novo assembly (map reads back to contigs = on; length fraction = 0.8; similarity
242 fraction = 0.8, minimum contig length = 200) and consensus sequences extraction (low
243 coverage threshold= 10; remove regions with low coverage = on; post-remove action =
244 split). The de novo assembled contigs (>100 nt) were analyzed using BLASTn
245 (maximum alignments per database sequence = 5; maximum E-value = 1e-6, minimum
246 identity = 70%) and DIAMOND (Buchfink et al., 2015) with a local installation of the
247 NCBI nr(/nt) databases. Blast results were visualized in Krona (bit score threshold = 25)

For Peer Review

248 (Ondov et al., 2011). The same pipeline was repeated with 1% of all reads as de novo
249 assembly of high coverage contigs can be problematic resulting in fragmented
250 assemblies. Viral sequences were analysed in Geneious R11 (Biomatters, New
251 Zealand).

252

253

254 *Survey of solanaceous species in Belgium*

255 Sample reference 'GBVC_AVX_01': In Belgium, during a survey of viruses in
256 Solanaceae, 21 species belonging to twelve genera (Supplementary information Table
257 1) were collected from the Meise Botanic Garden (province of Flemish Brabant) by
258 ILVO. Samples were tested in bulk using a VANA approach. ~~(see Supplementary~~
259 ~~information S.6)~~ The samples were pooled together (200 mg per plant) as a bulked
260 sample and virions were purified as Virion-Associated Nucleic Acids (VANA) after
261 Palanga et al. (2016). In brief, the pooled samples were ground in 10 volumes of Hank's
262 buffered salt solution (HBSS; 0.137 M NaCl, 5.4 mM KCl, 0.25 mM Na2HPO4, 1 g/L
263 glucose, 0.44 mM KH2PO4, 1.3 mM CaCl2,1.0 mM MgSO4, and 4.2 mM NaHCO3).
264 The extract was clarified by centrifugation (8,000 g for 10 minutes), and filtration (0.45
265 µm). The virions were collected by ultracentrifugation, using 10,5 mL of the filtrate on 1
266 mL 30% sucrose cushion (2 hours ~150,000 g Beckman 50 Ti rotor). The pellet was
267 suspended in 1 mL of HBSS. From the resuspension, 200 µL were digested by 15 U of
268 bovine pancreas DNase I (Euromedex) and 1.9 U of RNase A suspension (90 minutes
269 incubation at 37°C) to degrade non-encapsidated nucleic acid. Total nucleic acids were

For Peer Review

extracted with PureLink Viral RNA/DNA kit (Invitrogen) and reverse transcribed (for the RNA) with Superscript III (Life technologies) into DNA. The second strand of cDNA was synthetized with the use of large Klenow fragment polymerase (Promega). Individual barcodes (Tagged dodecamers) were added to each pool in the RT and Klenow steps, and the corresponding MID was used in the PCR. Finally, an amplification step (PCR) was performed using HotStarTaq (Qiagen). After the library preparation with the TruSeq mRNA stranded kit, the samples were then sequenced on the Illumina NextSeq500 at GIGA Genomics (University of Liege, Belgium). The resulting sequence reads were trimmed from the adaptor, paired and merged using the Geneious R11 software platform (https://www.geneious.com) and *de novo* assembled with SPAdes (Bankevich et al., 2012).

~~and i~~Individual plant infections were confirmed using specific primers designed for the detection of the target nucleic acid and sequence analysis of the PCR product.~~ (see Supplementary information S.7)~~ Primers AcVX 4,544 F (GCACGCCAGTATCATGCTCCAGA) and AcVX 4,782 R (TGCTGGTGCCTTCTTGTCCTGTC) were designed in the polymerase region from the sequence obtained by HTS, in order to identify the infected plants. The 21 samples were screened by RT-PCR using RNA extracted from the frozen samples with RNeasy plant mini kit (Qiagen). Reverse transcription was performed with the enzyme Tetro (Bioline) with random hexamers and PCR was carried out with Mango Taq (Bioline) for 40 cycles of 94°C for 15 seconds, 60°C for 20 seconds and 72°C for 20 seconds, preceded by an initial incubation of 1 minute at 94°C and followed by a final extension at

For Peer Review

292 72°C for 3 minutes. The anticipated amplicon size is 239 bp. Sanger sequencing was
293 performed by Macrogen Europe (Amsterdam, The Netherlands).
294
295
296 **Results**

297 *Analysis of HTS results from historic samples*

298 Genome sequencing of PlVX 76/16 and 77/48 yielded 474,188 and 427,300 paired end
299 reads of 300 nt respectively. Assembly of these reads yielded contigs of 6805 nt (PlVX
300 76/16, NCBI GenBank accession no MN334616) produced from 1042 reads (~46x
301 coverage) and 6850 nt (PlVX 77/48, access no. MN334615) produced from 2734
302 reads (~120x coverage). These contigs had high identity to the reference sequence of
303 actinidia virus X (access no. NC_028649.1).
304

305 *Results from contemporary P. lanceolata samples using RT-PCR and Sanger*
306 *sequencing*

307 Initially, the samples from the United Kingdom were screened in three bulks. All three
308 bulks tested positive in RT-PCR and produced consensus sequences (access no.
309 MN557306- MN557312, data not presented). To determine how many plants were
310 infected, the plants were then extracted and tested individually. In total six plants tested
311 positive, four plants from along Fulbourn Road (FR3, FR4, FR6 and FR9) and two from
312 along Worts Causeway (WC1 and WC5) (see table 1).

For Peer Review

313 Two out of four Dutch samples from Klarenbeek (6144590) and Buren (39014434),
314 province Gelderland, tested positive in RT-PCR and produced consensus sequences
315 (access no. MN432890 and MN432891; see table 1).

316

317 *Results on The Netherlands isolate AVX-2018-001*

318 The Netherlands isolate from the wild-flower strip in the region Overijssel gave a
319 positive reaction in the generic potexvirus RT-PCR. BLASTn analysis on the sequence
320 derived from direct Sanger sequencing on the obtained amplicon indicated 90% identity
321 to isolate L5 of AVX (access no. KC568202) and 85% to isolate RV3124 (access no.
322 NC_028649). The isolate was then designated AVX-2018-01. ~~HTS genome~~ Genome
323 sequencing of AVX-2018-001 yielded ~~8,775,250~~ 740,070 nt paired end reads. Assembly
324 of these reads yielded in a contig of 6868 nt with high identity to AVX (access no.
325 NC_028649.1). This contig was mapped by 1,~~605~~ 616 reads (~50x coverage). The
326 UTRs were confirmed by 5' and 3' RACE, which resulted in a full-length sequence of
327 6882 nt (AVX-2018-001, NCBI access no. MT123349).

For Peer Review

328

329 *Results on* Capsicum annuum *sample*

330 *Capsicum* leaf sample (5422861) tested positive in RT-PCR producing amplicons of the
331 expected size. In 2011 the species name *Actinidia virus X* was not yet described, and
332 no sequences were available in NCBI GenBank, the resulting consensus sequences
333 appeared to belong to an undescribed species in the genus *Potexvirus* and could not be
334 further identified. After reexamination of the sequence in 2019, however, it showed

335 highest identity with AVX (access no. KC568202). The subsequent Illumina sequencing
336 of sample 5422861 yielded 5,422,861 ~~16,546,406~~ paired end reads of 150 nt. Assembly
337 of these reads yielded a contig of 6882 nt - produced from 204,280 reads (~4432x
338 coverage ~~5422861~~, access no. MN756626) with high identity to the reference sequence
339 of AVX (access no. NC_028649.1) and 100% nt identity to the fragment which was
340 obtained in 2011. Additionally, a contig of 15,666 nt with highest identity to sequences
341 of bell pepper endornavirus was obtained.

342

343 *Results from survey of solanaceous species in Belgium*

344 In total, 720,220 unique reads were obtained from the pooled solanaceous samples.
345 Large contigs were obtained from the *de novo* assembly. One 7,004 nt contig matched
346 AVX when compared to the NCBI GenBank virus RefSeq database by tBLASTx. This
347 contig was mapped by 97,173 reads (~2000x coverage) using Geneious R11
348 (www.geneious.com)

349 From the PCR confirmation, only one of the 21 samples, originating from a *B.*
350 *americana* plant was found infected by AVX. The *B. americana* plant was sampled in
351 October and the plant was already deteriorating due to autumn conditions. There were
352 no specific symptoms that indicated virus presence. A 239 bp amplicon was Sanger
353 sequenced and the sequence was submitted to GenBank (access no. MT150906). The
354 sequence was identical to that obtained from the original sample pool by HTS. The full
355 genome showed 89% nt identity to the corresponding fragment of the full genome of the

For Peer Review

356 kiwifruit isolate (access no. KC568202) and 83% identity to the *Ribes* isolate (access
357 no. KR872420.1).
358
359 *Sequence analysis of near whole genomes of historic and contemporary isolates of*
360 *plantain virus X*
361 Figure 1 shows a phylogenetic tree produced from the genomes of PlVX, AVX and
362 related potexviruses. The AVX-like contig produced from PlVX 77/48 has open reading
363 frames for the five coding regions expected for a potexvirus and also has an open
364 reading frame related to the ORF6 reported for the *Ribes* isolate of AVX (James &
365 Phelan, 2016). The contig from PlVX 76/16 contains the same ORFs but lacks the last
366 few amino acids from the ORF5 which was not completely sequenced. Comparison
367 between the two nucleotide sequences (PlVX 77/48 and PlVX 76/16) shows they have
368 99.7% identity. All the PlVX genomes from this study and the *Actinidia* AVX genome
369 have more than 91% identity to each other, except that the *Ribes* AVX isolate (access
370 no. KR872420.1) is an outlier with 83-84% identity to the other PlVX / AVX isolates. The
371 ICTV species demarcation for the genus *Potexvirus* suggests coat protein or
372 polymerase gene sequences with less than about 72% constitute different species. The
373 coat protein genes of PlVX / AVX have more than 84% identity and the polymerase
374 genes more than 82% identity. The PlVX and AVX isolates form a clade with the *Ribes*
375 AVX isolate (James & Phelan, 2016) (access no. KR872420.1) as the sister to the rest
376 of the PlVX and AVX isolates.
377

For Peer Review

378 *Sequence analysis of partial nucleotide sequence of a potexvirus from contemporary*
379 *and historic samples*

380 The sequences produced from the contemporary United Kingdom and Netherlands *P.*
381 *lanceolata* samples encode part of the polymerase gene. A comparison of these
382 sequences with all of the PlVX and AVX full genomes shows that all sequences have
383 greater than 85% nucleotide identity. Figure 2 shows a phylogenetic tree produced from
384 the partial polymerases of PlVX, AVX and related potexviruses. The diversity of the
385 contemporary isolates from Cambridge almost encompassed all of the Dutch, Belgian,
386 original PlVX isolates, and the *Actinidia* isolate of AVX, with only the *Ribes* isolate of
387 AVX falling outside the group of FR and WC isolates, but clearly closely related (Figure
388 2). Interestingly the two historic isolates (76/16 and 77/48) were more closely related to
389 each other despite greater geographic separation than the more recent isolates.

390

391 **Discussion**

392 This report presents the first sequence of the complete coding region of plantain virus X
393 from 43 year old preserved samples of the virus collected when the virus was first
394 described (Hammond, 1981, Hammond, 1982, Hammond & Hull, 1981). Additionally,
395 samples of *P. lanceolata* which were sampled over 40 years apart and from different
396 geographic regions were shown to be infected with the same virus. On the basis of the
397 data presented it is likely that PlVX is more widespread and naturally occurring in *P.*
398 *lanceolata*. Whereas PlVX was previously reported only from eight regions of England, it
399 has now been identified in *P. lanceolata* from two regions in the Netherlands and in two

plantpath@bspp.org.uk

For Peer Review

400 new (Solanaceous) hosts namely *Browallia americana* from Belgium and *Capsicum*
401 *annuum* imported from Ethiopia into the Netherlands. The biological impact of the virus
402 on *B. americana* is unclear, since no clear symptoms were present at the time of
403 sampling, as is also the case in *P. lanceolata*. Although the leaves of the *C. annuum*
404 plant showed mild mottle symptoms, the impact of the PlVX isolate is unclear since it
405 was in mixed infection with bell pepper endornavirus, although this virus is typically
406 regarded as symptomless (Otulak-Kozieł et al., 2020). It is probable that the virus has
407 not been previously reported to be more widespread because it does not seem to evoke
408 clear symptoms. It is unusual to test plants for asymptomatic infections and surveys of
409 non-cultivated species are rare (Roossinck & Garcia-Arenal, 2015).

410 An additional motivation for seeking fresh cultures of PlVX to sequence was the
411 possibility that PlVX might prove to be synonymous with *Plantago asiatica mosaic virus*
412 (Genus *Potexvirus*). Plantago asiatica mosaic virus (PlAMV) was initially isolated from
413 *P. asiatica* in the Russian Far East (Kostin & Volkov, 1976). PlAMV was subsequently
414 reported to infect naturalized lilies (*Lilium maximowiczii*; syn. of *L. leichtlinii* var.
415 *maximowiczii*) in Japan in 2006 (Ozeki et al., 2006). PlAMV was then discovered
416 infecting commercial Asiatic and Oriental lily hybrids in the Netherlands (EPPO, 2011),
417 however, the sequence of the Dutch isolate and many similar 'European-type' isolates
418 detected in commercially-grown lilies from various other countries around the globe fall
419 into a clade distinct from those of the original Japanese lily isolates, various other
420 isolates from wild plants in Japan, isolates from *P. asiatica* from Russia and Korea, and
421 'Nandina mosaic' isolates from the USA and Japan (Komatsu et al., 2017, Hammond &
422 Reinsel, 2018). The origin of the 'European-type' infection in lilies is therefore unknown,

For Peer Review

423 but as *P. lanceolata* is a common weed of agricultural fields across Europe, there was a
424 possibility that it might have been the source of PlAMV infection in lilies in the
425 Netherlands. PlAMV has not (to our knowledge) been reported to naturally infect *P.*
426 *lanceolata*, but does infect this host experimentally (Hammond, 2018, Kock et al., 2011).
427 In the UK, PlAMV has not been previously reported except for interceptions of infected
428 imported lilies (Harju et al., 2018), and as such is still subject to plant health regulation.
429 Should PlVX have proven to be synonymous with PlAMV, and present in the wider
430 environment in Europe, then regulatory controls on PlAMV would be no longer
431 necessary. Importantly, from a plant health perspective the sequence data presented in
432 this report confirm that PlVX is not synonymous with plantago asiatica mosaic virus, so
433 in areas where PlAMV has not been reported to occur plant health monitoring and
434 action against this virus may still be justified.

435 Naturally, PlVX might also have proven to be synonymous with another potexvirus such
436 as one of the two other potexviruses reported naturally infecting *Plantago* species.
437 Plantago severe mottle virus was reported from *P. major* in Canada (Rowhani &
438 Peterson, 1980), and initially recognized as a species of the genus *Potexvirus*, but due
439 to lack of an extant culture and absence of any sequence data, the species was
440 abolished in 2015 (Adams & Kreuze, 2015). Another potexvirus was isolated from *P.*
441 *major*, *Taraxacum officinale*, and *Callistephus chinensis* in Argentina and provisionally
442 named Argentine plantago virus (APlaV); APlaV was found to be serologically related to
443 papaya mosaic, plantago severe mottle, and boussingaltia mosaic viruses (all four of
444 which are now considered strains of papaya mosaic virus), but not to PlVX (Gracia et
445 al., 1983).

For Peer Review

446 Unexpectedly, the data presented here indicate that the closest match in the NCBI
447 GenBank database is the potexvirus actinidia virus X (AVX) with a high degree of
448 sequence homology in both the coat protein (93% nucleotide; 99% amino acid identity)
449 and the polymerase (89% nucleotide; 98% amino acid identity), in addition to 83-91%
450 genome nucleotide identity. As the ICTV criteria for species discrimination in the genus
451 *Potexvirus* are less than ~72% nucleotide identity or ~80% amino acid identity between
452 their respective CP or polymerase genes (ICTV 9th Report), the PlVX isolates detected
453 in each sample and AVX belong to the same species. Actinidia virus X was first
454 detected in 2005 (KC568202) following virus isolation from *Actinidia chinesis* cv.
455 Hort16A (Kiwifruit) in New Zealand (Pearson et al., 2011, Blouin et al., 2013). The virus
456 was subsequently reported from *Ribes nigrum* (blackcurrant) being tested during import
457 to Canada for cultivation (~~Origin~~ origin of material is not reported) (James & Phelan,
458 2016). There are no further reports of the occurrence or distribution of this virus. As a
459 (then-considered) non-European virus reported to infect soft-fruits such as *Ribes*, AVX
460 could be considered to be a candidate for plant health regulatory status within the
461 European Union, however, the data presented here confirm that the virus is already
462 present and broadly distributed in the wider environment. Within this study, the *P.*
463 *lanceolata* sample from a Netherlands Blueberry farm in the region Overijssel was found
464 in an area planted as a flower strip following a preceding cropping with blackcurrants
465 and it was initially hypothesized that the infection present in the *Plantago* was likely to
466 originate from the blackcurrants. However, the concurrent finding of PlVX infecting *P.*
467 *lanceolata* samples from elsewhere in The Netherlands without this cropping history,
468 and from contemporary samples from the UK with no obvious correlation to the

For Peer Review

469 presence of *Ribes* spp., this is now considered likely to be coincidental rather than
470 causal.

471 These data demonstrate that *Plantain virus X* is synonymous with *Actinidia virus X*, and
472 in light of the original dates of publication of the two viruses relative to each other,
473 *Plantain virus X* should be considered to have precedence of nomenclature over
474 *Actinidia virus X*. This is further supported by the findings that *A. chinensis* is not a
475 preferred host (Blouin et al., 2013) and when *A. chinensis* was mechanically infected
476 with the virus the titre fell rapidly and could not be detected after 2 months (Pearson et
477 al., 2011).

478 Given the presence of the same virus in naturally occurring infections in multiple
479 cultivated and wild hosts, and from three continents, it is likely that the virus is more
480 widespread than is currently reported and further work should be considered to
481 investigate the distribution of this virus. We propose the sequence of the 1977 isolate
482 77/48 as the reference sequence for PlVX (access no. MN334615), as that of isolate
483 76/16 is missing a portion of ORF4, encoding the TGB3 protein.

484 Contradicting statements have been made regarding diversity of virus populations in
485 wild (or uncultivated) ecosystems relative to agricultural ecosystems. Roossinck et al.
486 (2015) noted that plant virus diversity in wild ecosystems is far greater than in
487 agricultural situations, and that most viruses affecting plants in communities of wild
488 plants do not induce obvious symptoms in such hosts (as is also the case for PlVX in *P.*
489 *lanceolata*). However, viral genetic diversity and virulence have also been reported to
490 be greater in isolates from cultivated hosts than from wild hosts, as a result of selection
491 pressures imposed on viruses under managed agricultural systems (Pagán, 2018).

For Peer Review

492 Interestingly, the partial RdRp sequences of contemporary PlVX isolates from within 1
493 km of each other near Cambridge show significant diversity within a clade that includes
494 the two over 40 year-old PlVX isolates, the *Actinidia* AVX isolate from New Zealand and
495 the Dutch samples, with only the *Ribes* isolate identified as AVX showing as slightly
496 distinct to the group of contemporary PlVX isolates (Figure 2). That the level of diversity
497 between current isolates from a single host species collected within less than 1km of
498 each other is greater than that observed between three Dutch isolates from two regions,
499 and other isolates separated in time, space, and two additional host species suggests
500 the apparent plasticity of PlVX, and its potential ability to adapt to other taxonomically
501 diverse hosts, nevertheless, more isolates need to be characterized to support this
502 hypothesis.

503 Although PlVX was originally considered as a virus of little economic consequence,
504 apparently restricted to a single host in which no significant symptoms were observed
505 (Hammond, 1981, Hammond & Hull, 1981), it is now apparent that PlVX (as AVX) can
506 naturally infect economically significant horticultural hosts and symptomatic infections
507 have been observed in *Actinidia* (Blouin et al., 2013)*, Ribes* (James & Phelan, 2016)
508 and *Capsicum* (this study). This is comparable to the emergence and current economic
509 importance of PlAMV, about 30 years after the first report in *P. asiatica* in the Russian
510 Far East (Kostin & Volkov, 1976), and at about the same time in the USA under the
511 name nandina mosaic virus (Moreno et al., 1976). PlAMV is currently known to naturally
512 infect hosts from nine taxonomically diverse plant families (Hammond, 2018). The
513 experimental host range of PlVX includes species in five taxonomically diverse plant
514 families in addition to the Plantaginaceae (Hammond & Hull, 1981), with natural hosts in

For Peer Review

515 the Actinidiaceae (*Actinidia*), Grossulariaceae (*Ribes*) and Solanaceae (*Browallia,*
516 *Capsicum*) further extending the host taxonomic diversity, suggesting a probable ability
517 to infect additional crop plants. Given the emergence of the two potexviruses PlVX and
518 PlAMV from different species of *Plantago* over approximately the same timeframe, the
519 essentially worldwide distribution of *Plantago* species, and current knowledge of more
520 than 30 viruses naturally infecting *Plantago* spp. (Kostin & Volkov, 1976, Hammond,
521 1982, Susi et al., 2017, Susi et al., 2019), additional viruses can probably be expected
522 to be revealed with potential consequences for taxonomically diverse economic crops
523 under conditions of environmental and agricultural flux. Further surveys of the viruses in
524 *Plantago* and other uncultivated species are therefore warranted.

525 The contemporary samples collected from the Netherlands and Belgium were identified
526 as being the same virus as being investigated in the UK through one-to-one data
527 sharing between research groups. With the increasing adoption of HTS for both frontline
528 diagnostics and baselining studies, this type of pre-publication data sharing becomes
529 crucial for addressing the biosecurity implications of using the technology, such as
530 mitigating the risks of unnecessary regulatory action (MacDiarmid et al., 2013). The
531 concurrent detections of the same virus by separate research groups, in a range of
532 samples, using multiple tools and sequencing strategies, gives a depth of internal
533 validation to the findings that could not be achieved from groups working in isolation. To
534 formalize this type of pre-publication data sharing, an initiative is now being taken
535 forward through the Euphresco plant health research coordination network
536 (www.Euphresco.net). The value of linking of historic isolates to give context to recent
537 virus detections is evident from this report. A current project, Virus Curate, under the

For Peer Review

538 Euphresco network (www.euphresco.net) aims to close the knowledge gap between
539 viruses published on the basis of biological and serological data, and the generation of
540 sequence data from isolates of these viruses held in virus collections. As demonstrated
541 in this report, these data will become invaluable as HTS is increasingly applied for
542 baseline and landscape scale virome studies to help give taxonomic context, host range
543 and distribution data to support risk assessment of viruses revealed through such
544 studies.

545 **Acknowledgements**

546 The historic UK PlVX isolates sequenced in this study were sourced from the virus
547 isolate collection of the late Dr Alan Brunt, held at the University of Warwick. The
548 sequencing of historic isolates was funded by Defra through the Euphresco topic "Virus
549 curate". Work on contemporary samples was funded through the Defra-Fera Long-term
550 services agreement. The sequencing of the AVX-2018-001 isolate was funded by the
551 Dutch Ministry of Economic Affairs in the Topsector Program "Horticulture and Starting
552 Materials" under the theme "Plant Health" (Optimal Diagnostics through the use of
553 innovative detection methods, project number: 1605-029). The Belgian work was funded
554 by the FPS Health, Food Chain Safety and Environment under the project "RT 18/3
555 SEVIPLANT".

556

557 **Data Availability**

558 All sequence data used within this article are available through NCBI GenBank under
559 the listed accession numbers

For Peer Review

560

561 **References**

562 Adams IP, Abidrabo P, Miano DW*, et al.*, 2013. High throughput real-time RT-PCR
563 assays for specific detection of cassava brown streak disease causal viruses, and their
564 application to testing of planting material. *Plant Pathology* **62**, 233-42.
565 Adams M, Kreuze J, 2015. Taxonomic Proposal 2015.014aP: Abolish 3 species from
566 the genus Potexvirus, family *Alphaflexiviridae*. In.
567 Bankevich A, Nurk S, Antipov D*, et al.*, 2012. SPAdes: a new genome assembly
568 algorithm and its applications to single-cell sequencing. *Journal of computational*
569 *biology* **19**, 455-77.
570 Blouin A, Pearson M, Chavan R*, et al.*, 2013. Viruses of kiwifruit (Actinidia species).
571 *Journal of Plant Pathology*, 221-35.
572 Botermans M, Van De Vossenberg B, Verhoeven JTJ*, et al.*, 2013. Development and
573 validation of a real-time RT-PCR assay for generic detection of pospiviroids. *Journal of*
574 *virological methods* **187**, 43-50.
575 Buchfink B, Xie C, Huson DH, 2015. Fast and sensitive protein alignment using
576 DIAMOND. *Nature methods* **12**, 59.
577 Eppo, 2011. New pest records in EPPO member countries: Plantago asiatica mosaic
578 virus (Potexvirus, PlAMV) found on Lilium spp. in the Netherlands. *EPPO Reporting*
579 *Service* **2011/082**.
580 Fox A, Fowkes A, Skelton A*, et al.*, 2019. Using High Throughput Sequencing in support
581 of a plant health outbreak reveals novel viruses in Ullucus tuberosus (Basellaceae).
582 *Plant Pathology* **68**, 576-87.

For Peer Review

583 Gracia O, Koenig R, Lesemann D, 1983. Properties and classification of a potexvirus
584 isolated from three plant species in Argentina. *Phytopathology* **73**, 1488-92.
585 Hammond J, 1981. Viruses occurring in Plantago species in England. *Plant Pathology*
586 **30**, 237-43.
587 Hammond J, 1982. Plantago as a host of economically important viruses. In. *Advances*
588 *in virus research.* Elsevier, 103-40. (27.)
589 Hammond J, 2018. Datasheet: Plantago asiatica mosaic virus. . *CABI Invasive Species*
590 *Compendium*.
591 Hammond J, Hull R, 1981. Plantain virus X: A new potexvirus from Plantago lanceolata.
592 *Journal of general virology* **54**, 75-90.
593 Hammond J, Reinsel M. Sequence variability between Plantago asiatica mosaic virus
594 isolates. *Proceedings of the XIV International Symposium on Virus Diseases of*
595 *Ornamental Plants 1193*, *2018*, 1-8.
596 Harju V, Forde S, Tozer H, *et al.*, 2018. Plantago asiatica mosaic virus detected in
597 Lilium in the UK. *New Disease Reports* **38**, 25-.
598 James D, Phelan J, 2016. Complete genome sequence of a strain of Actinidia virus X
599 detected in Ribes nigrum cv. Baldwin showing unusual symptoms. *Archives of virology*
600 **161**, 507-11.
601 Kock MD, Lemmers M, Van Schadewijk T, 2011. Identificatie en inventarisatie van een
602 nieuw Potexvirus in lelie. In. *Onderzoeksrapport PPO-WUR.* (32.)
603 Komatsu K, Yamashita K, Sugawara K, *et al.*, 2017. Complete genome sequences of
604 two highly divergent Japanese isolates of Plantago asiatica mosaic virus. *Archives of*
605 *virology* **162**, 581-4.

For Peer Review

606 Kostin V, Volkov Y, 1976. Some properties of the virus affecting Plantago asiatica [in
607 Russian]. *Virusnye Bolezni Rastenij Dalnego Vostoka* **25**, 205–10.
608 Kumar S, Stecher G, Tamura K, 2016. MEGA7: molecular evolutionary genetics
609 analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**, 1870-4.
610 Macdiarmid R, Rodoni B, Melcher U, Ochoa-Corona F, Roossinck M, 2013. Biosecurity
611 implications of new technology and discovery in plant virus research. *PLoS pathogens*
612 **9**, e1003337.
613 Moreno P, Attathom S, Weathers L, 1976. Identification, transmission and partial
614 purification of a potexvirus causing a disease of Nandina plants in California.
615 *Proceedings of the American Phytopathological Society* **3**, 319.
616 Ondov BD, Bergman NH, Phillippy AM, 2011. Interactive metagenomic visualization in a
617 Web browser. *BMC bioinformatics* **12**, 385.
618 Otulak-Kozieł K, Kozieł E, Escalante C, Valverde RA, 2020. Ultrastructural Analysis of
619 Cells From Bell Pepper (Capsicum annuum) Infected With Bell Pepper Endornavirus.
620 *Frontiers in Plant Science* **11**.
621 Ozeki J, Takahashi S, Komatsu K*, et al.*, 2006. A single amino acid in the RNA-
622 dependent RNA polymerase of Plantago asiatica mosaic virus contributes to systemic
623 necrosis. *Archives of virology* **151**, 2067-75.
624 Pagán I, 2018. The diversity, evolution and epidemiology of plant viruses: A
625 phylogenetic view. *Infection, Genetics and Evolution* **65**, 187-99.
626 Palanga E, Filloux D, Martin DP*, et al.*, 2016. Metagenomic-Based Screening and
627 Molecular Characterization of Cowpea-Infecting Viruses in Burkina Faso. *PLOS ONE*
628 **11**, e0165188.

For Peer Review

629 Pearson M, Cohen D, Chavan R, Blouin A, 2011. Actinidia is a natural host to a wide
630 range of plant viruses. *Acta horticulturae* **913**, 467-72.
631 Roossinck MJ, Garcia-Arenal F, 2015. Ecosystem simplification, biodiversity loss and
632 plant virus emergence. *Curr Opin Virol* **10**, 56-62.
633 Roossinck MJ, Martin DP, Roumagnac P, 2015. Plant Virus Metagenomics: Advances
634 in Virus Discovery. *Phytopathology* **105**, 716-27.
635 Rowhani A, Peterson J, 1980. Characterization of a flexuous rod-shaped virus from
636 Plantago. *Canadian Journal of Plant Pathology* **2**, 12-8.
637 Susi H, Filloux D, Frilander MJ, Roumagnac P, Laine AL, 2019. Diverse and variable
638 virus communities in wild plant populations revealed by metagenomic tools. *PeerJ* **7**,
639 e6140.
640 Susi H, Laine AL, Filloux D*, et al.*, 2017. Genome sequences of a capulavirus infecting
641 Plantago lanceolata in the Aland archipelago of Finland. *Arch Virol* **162**, 2041-5.
642 Van De Vossenberg B, Van Der Straten M, 2014. Development and validation of real-
643 time PCR tests for the identification of four Spodoptera species: Spodoptera eridania,
644 Spodoptera frugiperda, Spodoptera littoralis, and Spodoptera litura (Lepidoptera:
645 Noctuidae). *Journal of economic entomology* **107**, 1643-54.
646 Van Der Vlugt RA, Berendsen M, 2002. Development of a general potexvirus detection
647 method. *European Journal of Plant Pathology* **108**, 367-71.

648

For Peer Review

Table 1. Sample references with origin host, source, and sequencing method (Supplementary information reference to the specific sequencing method) used for each sample. NCBI GenBank accession numbers for each sequenced isolate are also presented.

Sample reference	**Host**	**Source**	**Diagnostic/Sequencing method (Supplementary files reference)**	**NCBI GenBank Accession Number**
76/16	*Plantago lanceolata*	Plant virus collection, University of Warwick, UK	HTS ~~(S.1)~~	MN334616
77/48	*P. lanceolata*	Plant virus collection, University of Warwick, UK	HTS ~~(S.1)~~	MN334615
FR3	*P. lanceolata*	Cambridgeshire. UK	PCR Product ~~(S.2)~~	MN557306
FR4	*P. lanceolata*	Cambridgeshire. UK	PCR Product ~~(S.2)~~	MN557307
FR6	*P. lanceolata*	Cambridgeshire. UK	PCR Product ~~(S.2)~~	MN557309
FR9	*P. lanceolata*	Cambridgeshire. UK	PCR Product ~~(S.2)~~	MN557308
WC1	*P. lanceolata*	Cambridgeshire. UK	PCR Product ~~(S.2)~~	MN557312
WC3	*P. lanceolata*	Cambridgeshire. UK	PCR Product ~~(S.2)~~	MN557310
WC5	*P. lanceolata*	Cambridgeshire. UK	PCR Product ~~(S.2)~~	MN557311
6144590	*P. lanceolata*	Klarenbeek, Gelderland, The Netherlands	PCR Product ~~(S.3)~~	MN432890
39014434	*P. lanceolata*	Buren, Gelderland The Netherlands	PCR Product ~~(S.3)~~	MN432891
AVX-2018-001	*N. occidentalis 'P1'*, ex. *P. lanceolata*	Overijssel, The Netherlands	HTS ~~(S.4)~~	MT123349
5422861	*Capsicum annuum*	The Netherlands	HTS ~~(S.5)~~	MN756626
GBVC_AVX_01	*Browallia americana*	Flemish Brabant, Belgium	HTS ~~(S.6)~~	MN923516
GBVC_AVX_01	*B. americana*	Flemish Brabant, Belgium	PCR product ~~(S.7)~~	MT150906

For Peer Review

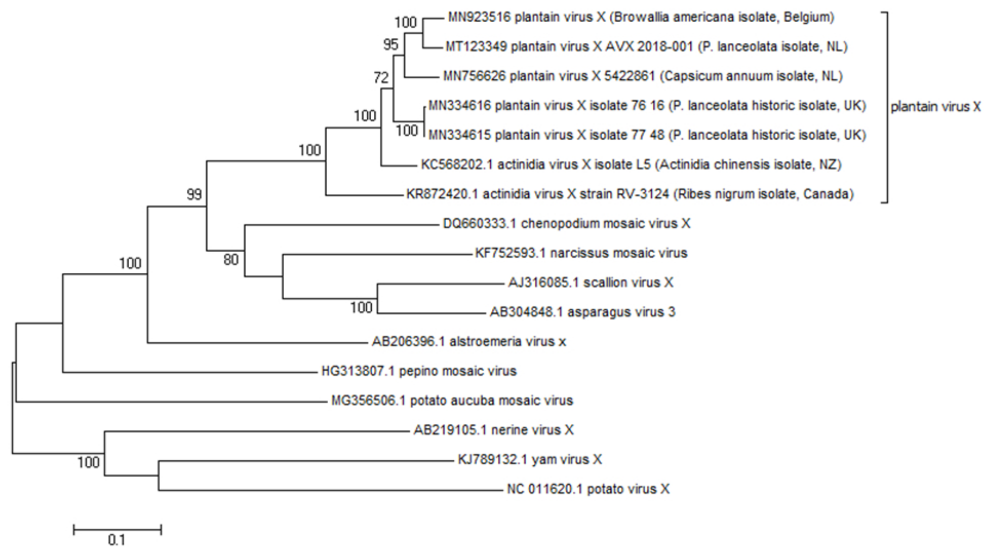

Figure 1. Maximum likelihood phylogenetic tree (500 bootstraps) produced from genomes of PlVX, AVX and related Potexvirus genomes. PlVX isolates include host species and countries of isolation (UK: United Kingdom, NZ: New Zealand, NL: The Netherlands)

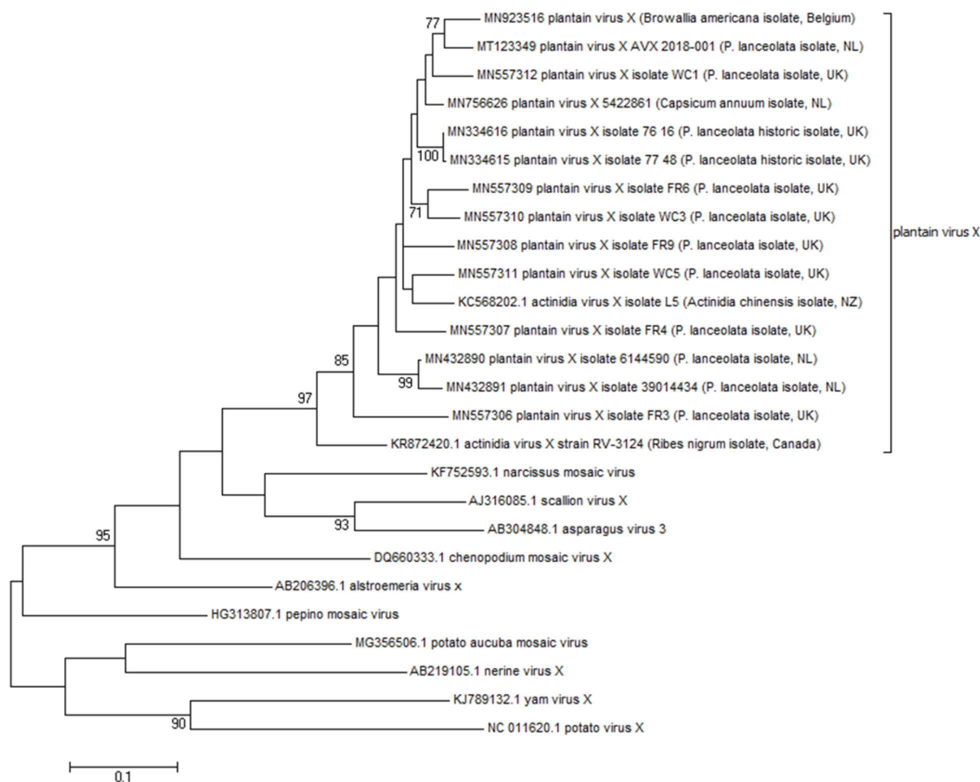

Figure 2. Maximum likelihood phylogenetic tree (500 bootstraps) produced from partial polymerase nucleotide sequences of PlVX, AVX and related potexviruses. PlVX isolates include host species and countries of isolation (UK: United Kingdom, NZ: New Zealand, NL: The Netherlands)