1 **Biological and genetic characterization of Physostegia chlorotic mottle virus in Europe based**
2 **on host range, location, and time**

3

4 Coline Temple[1], Arnaud G. Blouin[1,9], Kris De Jonghe[2], Yoika Foucart[2], Marleen Botermans[3],
5 Marcel Westenberg[3], Ruben Schoen[3], Pascal Gentit[4], Michèle Visage[4], Eric Verdin[5], Catherine
6 Wipf-Scheibel[5], Heiko Ziebell[6], Yahya Z. A. Gaafar[6], Amjad Zia[6], Xiao-Hua Yan[6], Katja R.
7 Richert-Pöggeler[6], Roswitha Ulrich[7], Mark Paul S. Rivarez[8], Denis Kutnjak[8], Ana Vučurović[8],
8 Sébastien Massart[1*]

9

10 *(1) Plant Pathology Laboratory, TERRA-Gembloux Agro-Bio Tech, University of Liège*
11 *(ULIEGE), Gembloux, 5030, Belgium*
12 *(2) Plant Sciences Unit, Research Institute for Agriculture, Fisheries and Food (ILVO), Merelbeke,*
13 *9820, Belgium*

14 *(3) National Reference Centre of Plant Health, National Plant Protection Organization of the*
15 *Netherlands. P.O. Box 9102, 6700 HC Wageningen, The Netherlands*

16 *(4) Laboratoire de santé des végétaux, Agence nationale de sécurité sanitaire de l'alimentation, de*
17 *l'environnement et du travail (ANSES), Angers, 49100, France*

18 *(5) Unité de Pathologie Végétale, Institut national de recherche pour l'agriculture, l'alimentation*
19 *et l'environnement (INRAE), Avignon, 84000, France*

20 *(6) Institute for Epidemiology and Pathogen Diagnostics, Julius Kühn Institute (JKI) - Federal*
21 *Research Centre for Cultivated Plants, Braunschweig, 38104, Germany*

22 *(7) Regierungspräsidium Gießen, Wetzlar, 35578, Germany*

23 *(8) Department of Biotechnology and Systems Biology, National Institute of Biology (NIB),*
24 *Ljubljana, 1000, Slovenia*

25 *(9) Plant Protection Department, Agroscope, 1260, Nyon, Switzerland*

26 *Corresponding author: S. Massart: E-mail: sebastien.massart@uliege.be

27

28 **Abstract**

29 Application of high-throughput sequencing (HTS) technologies enabled the first
30 identification of Physostegia chlorotic mottle virus (PhCMoV) in 2018 in Austria. Subsequently,
31 PhCMoV was detected in Germany and Serbia on tomatoes showing severe fruit mottling and
32 ripening anomalies. We report here how pre-publication data-sharing resulted in an international
33 collaboration across eight laboratories in five countries enabling an in-depth characterization of
34 PhCMoV. The independent studies converged toward its recent identification in eight additional
35 European countries and confirmed its presence in samples collected 20 ago (2002). The natural
36 plant host range was expanded from two species to nine species across seven families, and we
37 confirmed the association of PhCMoV presence with severe fruit symptoms on economically
38 important crops such as tomato, eggplant, and cucumber. Mechanical inoculations of selected
39 isolates in greenhouse established the causality of the symptoms on a new indexing host range. In
40 addition, phylogenetic analysis showed a low genomic variation across the 29 near-complete
41 genomes sequences available. Furthermore, a strong selection pressure within a specific ecosystem
42 was suggested by nearly identical sequences recovered from different host plants through time.
43 Overall, this study describes the European distribution of PhCMoV on multiple plant hosts,

44 including economically important crops which the virus can cause severe fruit symptoms for. This
45 work demonstrates how to efficiently improve knowledge on an emergent pathogen by sharing
46 HTS data, and provides a solid knowledge foundation for further studies on plant rhabdoviruses.

47

48 **Keywords:** Emergent viruses, PhCMoV, datasharing, European distribution, high through put
49 sequencing, mechanical inoculation, biological characterization

50

51 **Funding:**
52 European Union´s Horizon 2020 Research and Innovation program under the Marie Sklodowska-
53 Curie, Grant Agreement no. 813542
54 Federal public service, public health, Belgium, Grant Agreement no. RT 18/3 SEVIPLANT
55 Euphresco project 'Phytosanitary risks of newly introduced crops (PRONC), Grant Agreement
56 no. 2018-A-293
57 European Union's Horizon 2020 research and innovation programme, Grant agreement No.
58 871029

59

60 **Introduction**

61 High throughput sequencing (HTS) technologies have drastically increased the pace of new
62 virus discoveries (Adams *et al.,* 2018). Following genome identification, biological
63 characterization is essential to evaluate the scientific, commercial, and regulatory impact of plant
64 pathogens (Massart *et al.,* 2017). Biological characterization of a new virus requires comprehensive
65 knowledge on host range, vector, transmission, symptomatology, and general understanding of the

epidemiology (Massart *et al.,* 2017). It requires studying of the virus to be done under controlled conditions, e.g., through mechanical inoculation or grafting (bioassays) (Roenhorst *et al.,* 2013). This is a long and complex process that does not follow the current pace of virus discoveries by HTS (Hou *et al.,* 2021). In this context, HTS data sharing across laboratories before publication can speed up the characterization of emerging viruses in plants, avoid duplication of effort and accelerate a more accurate pest risk analysis (Hammond et al., 2020). For example, it could describe the natural host range and symptoms associated with a new pathogen more extensively and identify crops that may have been impacted, or crops that could serve as reservoir. Merging HTS data from different sources (regions, countries) and data collected at different times (including historical samples) provides a better view of the spatial and temporal status and distribution of viruses, while improving knowledge on epidemiology from phylogenetic analyses. Additionally, historical samples and/or nucleic acids can be used to obtain valuable information on the viral origin, and gathering data from different sources about the conditions of discovery (host range, symptoms, etc.) can help to identify a possible route of invasion (Jones *et al.,* 2021).

Proving a causal relationship between a virus and a disease is one of the first steps in evaluating the risk associated with a new disease agent. However, complying with Koch's postulates is a time-consuming process that requires extensive bioassays (Fraile *et al.,* 2016; Adams *et al.,* 2018). To accelerate this characterization, Fox *et al.,* (2020) proposed a new approach based on epidemiological studies and statistical analysis that provide valuable insights into causal relationships. In that context, bringing together HTS data and bioassay results from various research laboratories offers a possibility to optimize the study of causal associations between a disease and a potential viral or virus-like agent.

88 Physostegia chlorotic mottle virus (PhCMoV) was first identified on *Physostegia*
89 *virginiana* collected from Austria by Illumina HTS in 2014 (Menzel *et al.,* 2018). Subsequently,
90 PhCMoV was detected in Germany and Serbia on tomatoes showing severe fruit marbling and
91 ripening anomalies (Gaafar *et al.,* 2018; Vučurović *et al.,* 2021). PhCMoV has a negative-sense,
92 single-stranded RNA (−ssRNA) genome of 13,321 nucleotides and belongs to the genus
93 *Alphanucleorhabdovirus* of the family *Rhabdoviridae* (Kuhn *et al.,* 2020). Plant rhabdoviruses are
94 believed to originate from insect viruses (Whitfield *et al.,* 2018; Dolja *et al.,* 2020); they are insect-
95 vector-transmitted in a persistent and propagative manner (Jackson *et al.,* 2005). Seed or pollen
96 transmission of plant rhabdoviruses has never been described (Jackson *et al.,* 2005).

97 Phylogenetic analyses of alphanucleorhabdoviruses revealed a close relationship between
98 PhCMoV and eggplant mottled dwarf virus (EMDV), potato yellow dwarf virus (PYDV),
99 constricta yellow dwarf virus (CYDV), and joa yellow blotch-associated virus (JYBaV) (Dietzgen
100 *et al*., 2021; Bejerman *et al.*, 2021). Those five alphanucleorhabdoviruses share the same genome
101 organization, which contains seven canonical open reading frames (ORFs) encoding (from 3' to
102 5') nucleoprotein (N), unknown function protein (X), phosphoprotein (P), putative movement
103 protein (Y), matrix protein (M), glycoprotein (G) and large polymerase protein (L) (Dietzgen *et*
104 *al*., 2021). These viruses infect dicotyledonous plants, and three of them (EMDV, PYDV and
105 CYDV) are transmitted by leafhoppers. Vectors are still to be identified for the two most recently
106 discovered viruses (JYBaV and PhCMoV). As genetically close plant rhabdoviruses are
107 transmitted by a particular type of vector (Dietzgen *et al.,* 2021), PhCMoV and JYBaV are quite
108 likely transmitted by a leafhopper, like how their close relatives alphanucleorhabdoviruses are.

109 Recent discoveries of PhCMoV in several European countries on various host plants –
110 associated with severe symptoms in some cases - suggest that it is an emerging virus potentially

111 harmful to economically important crops. Therefore, efficient and rapid characterization is required
112 to establish proper risk assessment and manage the disease. In that context, eight European
113 laboratories worked together to improve knowledge on PhCMoV biology, epidemiology, and
114 genetic diversity.

115

116 **Material and methods**

117 The PhCMoV isolates that are reported here were independently detected and studied in
118 different laboratories. PhCMoV was detected and identified from different plants during virus
119 surveillance programs and plant pathogen diagnostic processes. For the detection, HTS and
120 conventional sequencing (PCR and sanger sequencing) approaches were conducted. To confirm
121 the presence of the virus after HTS detection, RT-PCR or mechanical transmission tests were
122 performed. Ribo-depleted total RNA, double-stranded RNA (dsRNA) and Virion-Associated
123 Nucleic Acids (VANA) were used as extraction and virus enrichment strategies prior to HTS on
124 Illumina or Oxford Nanopore Technologies MinION platforms.

125 Host plant species, geographical location, date of collection, symptoms and sequencing
126 method for each sample are indicated in Table 1. All the sequences were deposited in the GenBank
127 database and the corresponding accession numbers are indicated in Table 1. The number of reads
128 generated and horizontal coverage for each sample is indicated in the Supplementary Table 1.
129 PhCMoV was detected from samples collected as part of surveys in Germany, Belgium, France,
130 the Netherlands, and Slovenia and from symptomatic plants of different origins (the Netherlands,
131 Russia, and Romania) submitted to the national reference laboratory in the Netherlands for
132 diagnostics. The context of sample discovery is descripted for each sample in the following section,

133 but the different sequencing methods and bioinformatic analyses are detailed in the Supplementary
134 method 1.

135 i/ Samples origin and analysis by HTS

136

137 **Samples Be_SL1, Be_SM1 and Be_GP1**

138 During a survey on *Solanaceae* in 2019 in Belgium, one plant of *Solanum lycopersicum* (Be_SL1)
139 was collected in a tomato production tunnel where multiple plants were showing deformed,
140 mottled, and discolored fruits (Supplementary Figure 1). During this survey, the leaves of five
141 plants of *Solanum melongena* (Be_SM1) showing strong vein clearing were collected in another
142 tunnel and pooled together. The virus enrichment method VANA and the library preparation was
143 performed on these two samples prior to HTS (Supplementary method 1) revealing the presence of
144 PhCMoV.

145 A year later, multiple eggplant and tomato plants exhibited similar symptoms to those
146 that were observed in 2019 within the same site we collected. Additionally, while inspecting
147 *Capsicum annuum* grown in one of the tunnels, a plant of *Galinsonga parviflora* (Be_GP1)
148 showing vein clearing was collected (Fig. 1h). RNA was extracted following the method described
149 by Oñate-Sánchez *et al.,* (2008) and the detection of PhCMoV in these samples was confirmed by
150 RT-PCR using the primers published by Gaafar *et al.,* (2018). The sample related to a new host
151 (Be_GP1) was sequenced by Illumina after total RNA extraction, DNase treatment and
152 ribodepletion (Supplementary method 1).

153

154 **Sample Be_SA1, Be_IB1, Be_IB2 and Be_PM1**

155 In the framework of a study on the phytosanitary risk of viruses in newly introduced crops in
156 Belgium (PRONC, FPS project), eight samples of *Stacchys affinis* (crosne) and 91 samples of
157 *Ipomoea batatas* (sweet potato) from imported vegetatively propagated starting material and seeds
158 were collected in 2019 and 2020 in different production sites, including two community-supported
159 agriculture (CSA) farms. The samples were taken randomly and not specifically based on the
160 presence of symptoms. In a follow up survey, asymptomatic plants of several common weeds,
161 including *Persicaria maculosa* (lady's thumb), *Chenopodium album* (lamb's quarters), *Solanum*
162 *nigrum* (black nightshade), grasses (e.g., *Digitaria sanguinalis* (hairy crabgrass), *Echinochloa*
163 *crus-galli* (cockspur grass)) and some other crops (*Physalis philadelphica* (tomatillo) and *Sechium*
164 *edule* (chayote)), growing around the crosne plants were sampled. The samples were sequenced by
165 Illumina after total RNA extraction, DNase treatment and ribodepletion (Supplementary method
166 1).
167

168 **Sample Ge_CS1**

169 During a survey in July 2020, nine cucumber samples (*Cucumis sativus* L.) showing mosaic leaf
170 curling, chlorotic spots and yellowing symptoms were collected in an organic farm in Hesse State,
171 Germany where the previously published PhCMoV isolates KY706238, MK948541 and
172 KY859866 had been discovered (Gaafar *et al.,* 2018). Using immunosorbent electron microscopy
173 (ISEM), cucumber mosaic virus was identified in five samples, while in one sample (Ge_CS1),
174 bacilliform particles were observed suggesting the presence of a rhabdovirus. To identify the virus,
175 double stranded RNA (dsRNA) extraction followed by MinION sequencing were performed
176 (Supplementary method 1).
177

178 **Sample SL_SL1**

179 In Slovenia, a survey of viruses in tomatoes and surrounding weeds was conducted in summer
180 2019. Thirty-five plant samples were collected within greenhouses at one farming site (10 tomato
181 plants with symptoms resembling viral infection (which include, but not limited to, leaf curling,
182 mosaic and yellowing leaves), 10 tomato plants without any visible disease symptoms and 15
183 samples from 12 wild species growing as weeds). The samples were sequenced by Illumina after
184 total RNA extraction, DNase treatment and ribodepletion (Supplementary method 1).

185

186 **Samples Nd_SL1, Ru_SL1, Nd_H1, Nd_H2, Ro_SL1 and Nd_CS1**

187 From 2017-2019 symptomatic plant samples from the Netherlands, Russia and Romania were
188 submitted to the NPPO of the Netherlands for diagnostic purposes. The samples were sequenced
189 by Illumina after total RNA extraction, DNase treatment and ribodepletion (Supplementary method
190 1).

191

192 **Sample Fr_SL1, Fr_SL2, Fr_SM2, Fr_SM3, Fr_SM4 and Fr_SM1**

193 A survey conducted on cucurbits viruses in the south of France (Provence-Alpes Côte d'Azur) in
194 summer 2008 revealed one cucumber sample with mosaic and yellowing leave symptoms
195 (sample: ‘C08-119’). ELISA performed with antisera produced for detecting the cucurbit-
196 infecting viruses EMDV, zucchini yellow mosaic virus, watermelon mosaic virus, cucurbit aphid-
197 borne yellows virus, cucumber mosaic virus, melon necrotic spot virus, moroccan watermelon
198 mosaic virus, papaya ringspot virus and algerian watermelon mosaic virus only revealed the
199 presence of EMDV (pers Eric Verdin).

200 In 2018, eggplant samples collected in Nouvelle-Aquitaine (Lot-et Garonne
201 department) with vein clearing and deformed leave symptoms were simultaneously analysed in
202 two french research institutes (ANSES and INRAE) by RT-PCR with primers published by
203 Alfaro-Fernández *et al.,* (2011). Sanger sequencing was performed on amplicons of eggplant
204 samples as well as cucumber samples collected in 2008. BLASTn homology search revealed the
205 presence of PhCMoV for these two samples (Fr_SM4, 'C08-119').
206 From 2002 to 2018 in Southeastern France, several eggplant and tomato plants showing
207 dwarfing, bumpy and marbling fruits and leaves, as well as deformations and vein clearing, were
208 collected. Dip preparations were prepared from young symptomatic tomato or eggplant leaves,
209 negatively stained with 1% phosphotungstic acid (PTA) and observed by transmission electron
210 microscopy revealed the presence of characteristic bullets-shaped particles suggesting the presence
211 of a rhabdovirus. Total RNA was extracted using the RNeasy Plant Mini kit® (Qiagen) according
212 to the manufacturer's instructions and tested by RT-PCR with a set of primers designed for the
213 detection of EMDV (Alfaro-Fernández *et al.,* 2011). The PCR products showed 78-81% nucleotide
214 sequence identity with EMDV, but since the PhCMoV sequence was not available at the time of
215 detection (2002, 2011, 2013, 2014), the virus in the samples were categorized as "unknown
216 nucleorhabdovirus" and set aside. Recently, these sequences were blasted to the updated NCBI
217 database and the infection with PhCMoV were confirmed (96% to 98 of nucleotide sequence
218 identity). Thereafter, the samples have been sequenced by HTS, Fr_SL1, Fr_SL2, Fr_SM2,
219 Fr_SM3 and Fr_SM4 following the same methods described for Be_GP1 and Fr_SM1 following
220 the same method described for Nd_SL1 (Supplementary method 1). Since 'C08-119' is the only
221 sample that was not fully sequenced, the sequence of the amplicon generated with the primers of

222 Alfaro-Fernández *et al.,* (2011) and obtained by Sanger sequencing is available in the
223 Supplementary method 2 and on NCBI under the accession 'RYS_C08-119-A2021'.

224 ii/ Bioassays

225

226 Since mechanical transmission assays were performed in two distinct laboratories, JKI and NPPO-
227 NL, the methods differ.

228 **Sample isolate: KY882264 (JKI)**

229 PhCMoV-infected *Nicotiana benthamiana* fresh leaves containing MW848528 isolate were used
230 to inoculate *Chenopodiastrum murale*, *Chenopodium quinoa, Datura metel*, *D. stramonium,*
231 *Hyoscyamus niger*, *Medicago sativa*, *N. benthamiana*, *N. occidentalis* 'hesperis', *N. occidentalis*
232 'P1', *N. tabacum* 'samsun', *Petroselinum crispum, Petunia* sp., *Physalis floridana*, *Solanum*
233 *lycopersicum* 'harzfeuer', *S. lycopersicum* 'linda'. Four plants per species were inoculated. The
234 method used for the inoculation was described before by Gaafar *et al.,* (2019). Briefly, symptomatic
235 leaves were homogenized in Norit inoculation buffer (50mM phosphate buffer, pH 7, containing
236 1mM ethylenediaminetetraacetic acid (Na-EDTA), 20mM sodium diethyldithiocarbamic acid (Na-
237 DIECA), 5mM thioglycolic acid, 0.75% activated charcoal and 30 mg Celite). Using a glass
238 spatula, the homogenate was gently rubbed onto the leaves which were then rinsed with water. The
239 inoculated plants were kept under greenhouse conditions (at 22 °C; photoperiod of 16 h light
240 [natural daylight with additional growth light Phillips IP65, 400 W] and 8 h dark). Symptoms were
241 observed four weeks post inoculation and the presence of PhCMoV was confirmed by RT-PCR
242 with the primers of Gaafar *et al.,* (2018).

243 **Sample isolate: Ru_SL1, Nd_SL1, Ro_SL1, Nd_CS1, Nd_H2 (NPPO-NL)**

244 In the Netherlands, different PhCMoV isolates were tested on selected herbaceous indicators
245 including *C. quinoa, D. stramonium, N. benthamiana, N. glutinosa, N. occidentalis P1, N.*
246 *tabacum* 'WB', *Physalis floridana, S. lycopericum*. Not all the plants were tested for all isolates,
247 but the combinations are presented in Table 2. Three plants per species were inoculated. The
248 method used for the inoculation protocol is described by Verhoeven & Roenhorst (2000). Briefly,
249 1g of infected frozen leaf material (*N. benthamiana* for Ru_SL1 and Nd_SL1 and original host
250 for Ro_SL1, Nd_CS1 and Nd_H2) was ground in 10 mL inoculation buffer [0.02 M phosphate
251 buffer pH 7.4, 2% (wlv) polyvinylpyrrolidone [(PVP; MW 10000)]. Plants were inoculated at a
252 young stage (3-6 leaves) by gently rubbing the inoculum onto carborundum-dusted leaves. After
253 inoculation, plants were rinsed with water and placed in a glasshouse at 18-25°C with
254 supplementary illumination for a day length of at least 14 h. Each isolate was inoculated to at
255 least two plants per plant species and inspected visually for symptoms during the following seven
256 weeks. The virus infection was confirmed by ELISA in all the inoculated plants (pers Marleen
257 Botermans and Ruben Schoen).

258 iii/ Phylogenetic analyses

259

260 For the phylogenetic analyses, all the PhCMoV known sequences to date were used. This includes
261 PhCMoV sequences published by Menzel *et al.,* (2018); Gaafar *et al.,* (2018); Gaafar *et al.,* 2021;
262 Vučurović *et al.,* (2021) and the 21 new PhCMoV sequences generated in this study.

263 Prior to genome analysis, PhCMoV genomes were all trimmed to start at the sequence
264 "CATGAGACT" (position 40 on genome KX636164) and end after "TGCACCTA" (position
265 13275 on genome KX636164). Phylogenetic analysis was carried out using the MEGA-X software
266 (v10.1.8) (Kumar *et al.,* 2018). Sequence alignments were performed on near-complete genome

267 using MUSCLE and the best DNA model was applied to the maximum-likelihood analysis
268 (GTR+G+I model). Support for the branching patterns in the phylogenetic trees was determined
269 by analyzing 1000 bootstrap replicates. For graphical representation, SIMPLOT software (version
270 3.5.1) was used to compare similarity of the genomic sequences of selected PhCMoV isolates to
271 the reference query KX636164 (Window: 200bp, Step: 20bp, Gapstrim: On, Hamming). To
272 improve the graphical representation, the analysis was limited to 16 PhCMoV isolates including
273 the most divergent ones (Nd_SL1 and Nd_H2). KX636164 genome has been chosen as a reference
274 because it is the first discovered PhCMoV isolate and longest genome (Menzel *et al.,* 2018).

275 Finally, to compare the genetic similarity between the different isolates for different
276 genomic regions, the sequence of the N, X, P, Y, M, G and L ORF were extracted using
277 Geneious software for all the isolates indicated in Table 1. Pairwise nucleotide and amino acid
278 sequences identities were calculated for all isolates based on MUSCLE alignment (Muscle
279 3.8.425 by Robert C. Edgar).

280

281 **Results**

282 **1. Natural host range and symptoms**

283 In addition to the detection of PhCMoV in new host species belonging to the *Lamiaceae* (*Stachys*
284 *affinis*) and *Solanaceae* (*Solanum melongena*) families, this study expands the natural host range
285 of PhCMoV to seven new plant families: *Cucurbitaceae* (*Cucumis sativus*), *Ranunculaceae*
286 (*Helleborus* sp.), *Convolvulaceae* (*Ipomoea batatas*), *Polygonaceae (Persicaria maculosa)* and
287 *Asteraceae* (*Galinsoga parviflora*) (Table 1). These detections enabled the description of PhCMoV
288 related symptoms on several hosts (Fig. 1, Table 1). Only samples with single infection by

289 PhCMoV are shown in Fig. 1 (eggplant*:* Be_SM1, Fr_SM1, Fr_SM2, Fr_SM3, Fr_SM4, cucurbits:
290 Nd_CS1, Ge_CS1; *Helleborus*: Nd_H1, Nd_H2; *G. parviflora*: Be_GP1; tomato: Nd_SL1,
291 Ru_SL1, Ro_SL1, Be_SL1, Be_PM1).

292 As described previously by Gaafar *et al.,* (2018) and Vučurović *et al*., (2021) infected
293 tomato fruit were unevenly ripened and mottled (Ru_SL1, Be_SL1, Ro_SL1) (Fig. 1a). In this
294 study, some of the tomato infected fruit were also deformed (Supplementary Figure 1). All
295 PhCMoV infected tomato plants that bore mature fruit at the time of collection showed
296 symptomatic fruit regardless of their growing conditions. The symptoms observed on tomato leaves
297 were more variable: no symptom was observed on the leaves of Be_SL1 and Ro_SL1, mottled
298 leaves were observed on Ru_SL1, and vein clearing and deformed leaves were observed on
299 Nd_SL1 (Figure 1b).

300 Like infected tomato, PhCMoV-infected eggplants showed deformed, unevenly ripened
301 and mottled fruit (Fr_SM2, Fr_SM3, Fr_SM4) (Fig. 1c). Fr_SM1 showed deformed fruit. On the
302 leaves, Be_SM1 and Fr_SM2 showed vein clearing (Fig. 1d), and Fr_SM3 showed, yellowing.
303 Fr_SM4 and Fr_SM1 exhibited vein clearing and deformed leaves. Fr_SM2 showed dwarfism.
304 Sample Be_SM1 grouped five eggplants, all of which showed vein clearing in new leaves. No
305 mixed infection occurred in this bulk sample, which strongly suggests that PhCMoV was the causal
306 agent of the symptoms observed on all the plants. No fruit was present at the time of sampling.

307 Infected cucumber fruit were pointed, deformed, and showed vertical chlorotic stripes
308 (Nd_CS1) (Fig. 1e). The leaves exhibited interveinal chlorosis and sunken veins (Supplementary
309 Figure S1), leaf curling, chlorotic spots, and yellowing symptoms (Fig. 1f).

310 Finally, *G. parviflora* (Be_GP1) and *Helleborus* sp. (Nd_H1, Nd_H2) leaves showed vein
311 clearing (Fig. 1 g, h). No symptom was observed on *Stachys affinis* or *Persicaria maculosa* at the
312 time of collection.

313 **2. Experimental host range and symptoms**

314 We conducted independent experiments to investigate the indexing host range of PhCMoV. The
315 results of Menzel *et al*., (2018) (isolate KX636164) and Gafaar *et al*., (2018) (isolate KY859866)
316 were grouped with our own present results to have a more complete overview (Table 2).

317 At JKI, PhCMoV (MW848528) was mechanically transmitted to *D. stramonium, D. metel*
318 and *N. benthamiana* and induced yellowing and vein clearing four weeks after inoculation.
319 Inoculation of the other 13 plant species tested failed (Table 2). This result differs from previous
320 published reports, where *C. quinoa* and *P. floribunda* were successfully inoculated whereas
321 inoculation of *D. stramonium* and *D. metel* failed.

322 In The Netherlands, five PhCMoV isolates where single infection occurred (Nd_SL1,
323 Nd_CS1, Nd_H2, Ru_SL1, Ro_SL1) were mechanically transmitted to different indicator plants
324 (*D. stramonium, N. benthamiana, N. occidentalis P1, N. tabacum 'WB', P. floribunda, S.*
325 *lycopersicum*). An overview is presented in Table 2.

326 In all experiments, *N. benthamiana* displayed systemic symptoms four to seven weeks post
327 inoculation (Table 2) and Nd_SL1, Nd_CS1, Nd_H2, Ru_SL1 induced systemic symptoms in *N.*
328 *occidentalis P1* four to seven weeks post inoculation.

329 **3. Extended distribution across Europe since 2002**

330 This study provides an overview of the wide European geographical distribution of PhCMoV: its
331 presence is confirmed in six additional countries besides Germany, Austria, and Serbia where the
332 virus was previously reported (Menzel *et al.,* 2018; Gaafar *et al.,* 2018; Vučurović *et al.,* 2021):
333 Russia, Romania, Slovenia, The Netherlands, Belgium, and France (Table 1).

334 Although most of the detections are recent, re-analysis of historic *S. melongena* samples
335 (Fr_SM1) showed that PhCMoV was present in France as early as 2002. A cucumber sample
336 collected in France in 2008 and originally diagnosed as EMDV by ELISA using in-house antiserum
337 was re-analysed and diagnosed as PhCMoV by RT-PCR. This shows that some EMDV antiserums
338 used by ELISA can cross-react with PhCMoV and lead to incorrect diagnosis.

339 **4. Phylogenetic analysis of the genomes**

340 In total, 21 new near-complete PhCMoV sequences were generated during this study, and their
341 evolutionary relationships were investigated alongside all PhCMoV, EMDV and PYDV complete
342 genomes available from the GenBank database on a maximum-likelihood (ML) tree
343 (Supplementary Figure 2). Supported by bootstrap values of 1000, the analysis did not show any
344 clustering according to host plant, country of origin or year of collection (Fig. 2). However, isolates
345 collected from the same site (same farm) A, B, N or T grouped together regardless of the collection
346 date or host plant (Fig. 2). This was particularly obvious for some of the samples from Germany,
347 namely Ge_CS1, KY706238, KY859866, MK978541, and MW848528. They were collected at the
348 same site (Hesse state) and grouped together despite their collection date (from 2003 to 2020) and
349 host plants (cucumber, tomato). Be_SL1, Be_GP1 and Be_SM1 were also collected on the same
350 farm (Gembloux, Belgium) one year apart on three distinct host plants, but have almost identical
351 genome sequences (100% nt id; Supplementary Figure 3). Similarly, Fr_SM2 and Fr_SM3 were

352 collected at the same location and clustered together (Fig. 2). Interestingly, Be_SA1 and Be_PM1
353 sampled from the same farm also clustered together, along with Nd_CS1 which was isolated from
354 a different country and host family (Fig. 2).Overall, all the sequences from samples collected on a
355 same site clustered together, but the clusters did not all represent a geographical point.

356 To better understand the evolutionary relationships among PhCMoV isolates, nucleotides
357 and amino acid identities were calculated from the alignment of nearly complete genome sequences
358 and for each ORF (Fig. 3b). Relatively low genetic variability was observed for the near-complete
359 genomic sequences (>93% nt id) in 28 isolates out of 29 (Fig. 3b). Nd_SL1 isolate was the most
360 divergent isolate with 81-82% of nucleotide sequence identity (nts id) compared to the other 28
361 genomes (Fig. 3b). However, when the amino acid sequence identities (aa id) of the different
362 isolates were compared, the variability of Nd_SL1 ranged among the average pairwise identities
363 of the other isolates for most ORFs (N, P, Y, M, G) (Fig. 3b). Using Simplot to observe the
364 sequence similarity along the genome, a clear drop was visible in the intergenic regions located in-
365 between the coding regions (Fig. 3a). Overall, for all isolates except Nd_SL1, the ORF encoding
366 protein L was the most conserved gene, with a percentage of aa id > 99%. It was followed by the
367 ORF encoding protein G (aa id > 97%), and by those encoding proteins M, Y and P (aa id > 96%),
368 N (aa id > 95%) and X (aa id > 88 %).

369

370 **Discussion**

371 By collaborating and sharing data before submitting the results for publication, eight European
372 research groups investigated Physostegia chlorotic mottle virus in detail and characterized its
373 genome and biology.

374 This study demonstrates the ability of PhCMoV to naturally infect seven host plants (annual
375 and perennial ones) in addition to the two previously known hosts across seven families including,
376 economically important crops (*S. lycopersicum, S. melonga, C. sativus)*, newly introduced crops in
377 Europe *(I. batatas, S. affinis),* wild plants *(G. paviflora, P. maculosa)* and ornamentals (*Helleborus*
378 sp). Similar observations have been made for other alphanucleorhabdoviruses, *e.g.*, EMDV with
379 more than 25 hosts recorded on CABI (2021) (https://www.cabi.org/), including crops and
380 perennial plants such as *Hibiscus* sp., *Hydrangea macrophylla*, *Agapanthus* or *Pittosporum* sp.
381 (CABI, 2021). This suggests that the host range of PhCMoV is likely to be wider than described
382 here, and additional perennial hosts might help the virus overwinter.

383 Our results outline PhCMoV symptomatology on a large range of plants collected in
384 fields, gardens, and greenhouses. Overall, the presence of the virus was associated with virus-like
385 symptoms on leaves (vein clearing, chlorosis, mottling...) and severe symptoms on fruit
386 (deformation, marbling, uneven ripening). Only two samples (*S. affinis and P. maculosa*) did not
387 exhibit any symptom, suggesting that asymptomatic plants might host the virus. We did not
388 describe the symptomalogy of PhCMoV on sweet potato because of co-infection. Considering
389 only the samples single infected with PhCMoV, the symptoms were often variable across plants
390 from the same species. These variations may be due to several biases. First, they could be due to
391 human perception since different people recorded the symptoms. Secondly, the plants
392 corresponded to different cultivars and were grown under heterogeneous conditions. In addition,
393 symptom expression may be different depending on the plant growth stage at the time of
394 infection. Nevertheless, the presence of the virus was always associated with obvious vein
395 clearing on the leaves of *G. paviflora*, eggplant and *Helleborus*. This symptom was also
396 described for EMDV on honeysuckle and eggplant (Martelli *et al*., 1987).

397 The severe symptoms observed on tomato fruit (marbling, mottling, uneven ripening)
398 confirmed previous reports (Table 1, Gaafar *et al.,* 2018, Vučurović *et al.,* 2021). Even though
399 remarkable, these symptoms were not specific to PhCMoV: similar observations were made in
400 the case of other viral infections (EMDV, (Blancard, 2009) pepino mosaic virus (Hanssen *et al.,*
401 2009), tomato brown rugose fruit virus (EPPO Bulletin, 2020)) and in the case of nutrient
402 disorder mostly referred as "blotchy ripening" (Adams *et al*., 1995). The symptoms observed on
403 tomato leaves were highly variable (mottling and vein clearing) and sometimes absent. Therefore,
404 tomato leaves do not represent a good indicator of PhCMoV presence.

405 Vein clearing was observed on the leaves of four out of five eggplant samples. Vein
406 clearing is not specific for PhCMoV as it is also representative of the presence of EMDV and
407 alfalfa mosaic virus (Martelli *et al.,* 1986; Sofy *et al.,* 2021) but it is generally associated with
408 viral presence on eggplant and can differentiate viral presence from that of other pathogens,
409 abiotic stress, or nutritional disorders. Interestingly, this symptom can be used to monitor the
410 spread of the virus in a parcel infected by PhCMoV. Finally, the number of samples per species
411 sampled on the other host plants was too low to be associated with a specific symptom.

412 To confirm the presence of PhCMoV and to study its mechanical transmission, infected
413 leaves collected in various sites were mechanically inoculated on different indicator hosts. In total,
414 four out of eighteen indicator plant species were successfully inoculated and showed systemic
415 symptoms (Table 2; *D. metel, D. stramonium, N. benthamiana, N. occidentalis* P1). In the previous
416 studies, *C. quinoa, N. occidentalis '37B', N. clevelandii, N. tabacum 'WB', Physalis floridana* were
417 also mechanically inoculated (Table 2; Menzel *et al.,* 2018; Gaafar *et al.,* 2018). This host range is
418 similar to the one of EMDV which includes: *N. clevandii, N. glutinosa, N. rustica, N. tabacum, P.*
419 *hybrida,* and *P. floridana* (Mavrič *et al.,* 2006; Katis *et al.,* 2011). No systemic symptom of EMDV

420 infection has ever been reported on *C. quinoa* and *D. stramonium*. Despite the overall high
421 sequence identity of the PhCMoV isolates analysed in this study, the results were variable across
422 laboratories. Some plants were successfully inoculated in some laboratories but not in others (for
423 example: *N. occidentalis P1, D. stramonium*) and the range of observed symptoms on a same host
424 plant species was variable. Inoculation success and symptom expression depend on environmental
425 conditions (Hull, 2014) and inoculum sources. In addition, at NPPO-NL, some symptoms were
426 recorded four to seven weeks post-inoculation on *N. occidentalis P1* and *N. benthamiana* which is
427 longer than the recommended period of three weeks (Roenhorst *et al.,* 2013). Indexing is very
428 important to maintain and study viruses in controlled conditions, to separate them in case of
429 multiple infection and to find the best host for virus purification. It would also be interesting to
430 inoculate several plant species in the same experimental conditions to compare the impact of
431 divergent isolates on symptom expression. Overall, all the studies converged toward *N.*
432 *benthamiana* being the best experimental PhCMoV host. Our study also showed that inoculated
433 plants suspected to host PhCMoV should be kept in a greenhouse for symptom observations for at
434 least seven weeks.

435 With the generation of 29 sequences of near-complete genome, PhCMoV is now the plant
436 rhabdovirus with the highest number of near-complete genomes available. These genomes
437 provided data for studying the virus genetics in relation to host range, geographical location, and
438 time. Despite genetic variability ranging between 82% and 100% of nt sequence identity (for the
439 near-complete genome), the 29 samples did not cluster according to country or host plant.
440 In addition, there was 100% identity between isolate KY706238 collected on tomato in 2003 and
441 isolate Ge_CS1 collected on cucumber from the same site in 2020. This genome conservation
442 over time was observed in four distinct sites across Europe (yellow boxes in Fig. 2). It suggests

443 that the genome of PhCMoV does not evolve rapidly once established in a suitable ecosystem.
444 This highlights the impact of the geographical dimension on the genetic evolution of PhCMoV
445 and is in line with observations on other plant rhabdoviruses (EMDV, RSMV) whose
446 phylogenetic clusters correlate with geographical localization, but not necessarily with the host
447 plant or the sampling date (Tang *et al.,* 2014; Yang *et al.,* 2018; Pappi *et al.,* 2015). Since plant
448 rhabdoviruses are transmitted from plant to plant by insects in a persistent and propagative
449 manner and no other way of natural transmission is known, insect vectors are likely to be the
450 cause of the strong selective pressure on the genetic diversity of plant rhabdoviruses (Power,
451 2000).

452 For the 29 isolates analysed in this study and collected from eight countries and eight host
453 plant species, the genetic diversity was very low (less than 3% at the nt level for the near-
454 complete genome). This low genetic diversity has been observed in other plant rhabdoviruses.
455 For example, Yang *et al.,* (2018) showed that the genome of 13 isolates of rice stripe mosaic
456 virus (RSMV) collected in various geographical regions in China showed 99.4% of nucleotide
457 sequence identity. In another study, Samarfard *et al.,* (2018) showed a 99% aa sequence identity
458 of protein N across 13 alfalfa dwarf virus (ADV) isolates from different regions in Argentina. In
459 our study, between 92 and 99 % of nt sequence identity was observed among the 29 available
460 PhCMoV genomes with only one outlier, Nd_SL1, with 81-82% of nt sequence identity with the
461 other 28 isolates (Fig. 3). However, this isolate was not an outlier at the protein level; for
462 instance, it had more than 96% aa identity with all the PhCMoV isolates for protein N, while the
463 nt sequence identity ranged between 85 and 87% for the corresponding gene. Similar
464 observations have been reported for the cytorhabdovirus lettuce necrotic yellows virus (LNYV):

465 the ORF encoding protein N of two subgroups were approximatively 80% identical at the nt level
466 and 96% identical at the aa level (Higgins *et al.,* 2016).

467 Overall, this study brings together some key elements on the genetic diversity of
468 PhCMoV and its potential drivers. It shows the importance of accumulating genomic sequences
469 from diverse isolates to draw robust conclusions. Viral genomes from samples of different origins
470 (new location, new host, or different collection date) support a better understanding of the genetic
471 diversity and evolution of this virus, but the presence of an exception (i.e. isolate Nd_SL1)
472 suggests that the genetic diversity of PhCMoV remains partly uncovered and that the results need
473 to be interpreted carefully. Considering the severity of the symptoms observed on economically
474 important crops, it is unclear why the virus remained unnoticed for at least the past two decades.
475 The lack of appropriate diagnostic tests might be one of the reasons for this delay, since cross-
476 reactions occurred with one of the EMDV antibodies in 2008. This suggests that additional
477 infections may have been misdiagnosed. In addition, samples collected in 2002 (Fr_SM1), 2008,
478 2011 (Fr_SL1), 2013 (Fr_SM2/3) and 2014 (Fr_SL2) were set aside for identification because the
479 PCR products showed 78% nt identity with EMDV and the PhCMoV sequence was not available
480 at the time. Our research highlights the strength of HTS in plant virus detection, and the wider
481 application of these technologies for virus detection might explain the sudden simultaneous
482 identifications throughout Europe. Another complementary hypothesis of the recent detections
483 might be that the virus was present in the environment, but went unnoticed because it did not
484 cause a problem (low incidence), and a recent change in the environment led to its emergence.
485 Whether the virus is more prevalent nowadays or whether it was overlooked in the past remains
486 unknown. However, the current situation requires rapid characterization and a common response
487 from European countries because simultaneous PhCMoV detections in several European

488 countries over a wide host range including economically important foodstuffs suggests that the
489 virus could be an emerging pathogen. In that context, pre-publication data sharing, and
490 collaboration have been valuable to improve knowledge about this virus and would be beneficial
491 in the future to efficiently evaluate the risk associated with any emerging disease and implement
492 management strategies.

493 One of the next priorities will be to identify the insect vector and its life cycle. EMDV,
494 PYDV and CYDV are the closest relatives of PhCMoV with a known vector, and those vectors
495 all belong to *Cicadellidae*, which makes leafhoppers prime candidates for transmitting PhCMoV
496 (Dietzgen *et al.,* 2021). Furthermore, according to the transmission tests carried out by Babaie *et*
497 *al.,* (2003) EMDV was transmitted by one specific leafhopper (*Agallia vorobjevi*) and not by the
498 other 13 leafhopper species present in and around EMDV-infected fields. This suggests specific
499 virus-insect transmission. A second priority line of research will be to determine in which hosts
500 the virus is present in winter. This ability of plant rhabdoviruseses to infect different host plants
501 across families is an important factor to be considered for controlling the disease because a large
502 diversity of plants can serve as a reservoir during the no-crop season. A third axis will be to
503 assess the impact of the virus in terms of yield and economical loss on different cultivars and
504 when the plants are inoculated at different developmental stages.
505 Finally, understanding the epidemiology of the virus and the reasons for its multiple recent
506 detections in Europe are key elements to be investigated in order to evaluate if it can present a
507 threat for vegetable production and how to prevent potential outbreaks.
508

509 **Acknowledgments:**

510 The authors would like to thank Pier de Koning (NPPO) for his help to analyse the data and
511 Katarina Bačnik and Olivera Maksimović Carvalho Ferreira (NIB) for assisting in sample
512 collection. We also acknowledge Laurent Minet (Hortiforum asbl / Centre Technique Horticole de
513 Gembloux) for spotting the first symptoms of the virus in Belgium, monitoring the site thoroughly
514 and providing expert advice and Frederic Dresen (ULIEGE) for his strong technical support,
515 advices and insightful guidance throughout the study. We are grateful for the support of the growers
516 and farmers who allowed us to access their properties and collect samples over multiple years.

517

518 **Literature cited**

519 Adams, P. and Ho, L.C. 1995. Uptake and distribution of nutrients in relation to tomato fruit
520 quality. Acta Hortic. 412, 374-387
521 Adams, I. P., Fox, A., Boonham, N., Massart, S., and De Jonghe., K. 2018. The impact of high
522 throughput sequencing on plant health diagnostics. Eur J Plant Pathol. 152:909–919.
523 Alfaro-Fernández, A., Córdoba-Sellés, C., Tornos, T., Cebrián, M. C., and Font, M. I. 2011. First
524 Report of Eggplant mottled dwarf virus in Pittosporum tobira in Spain. Plant Disease. 95:75–75.
525 Babaie, G., and Izadpanah, K. 2003. Vector Transmission of Eggplant Mottled Dwarf Virus in
526 Iran. Journal of Phytopathology. 151:679–682.
527 Bejerman, N., Dietzgen, R. G., and Debat, H. 2021. Illuminating the Plant Rhabdovirus
528 Landscape through Metatranscriptomics Data. Viruses. 13:1304.
529 Blancard. 2009. Les maladies de la tomate. Quae, pages 346-366.
530 Dietzgen, R. G., Bejerman, N. E., Goodin, M. M., Higgins, C. M., Huot, O. B., Kondo, H., et al.
531 2020. Diversity and epidemiology of plant rhabdoviruses. Virus Research. :197942.

532 Dietzgen, R. G., Bejerman, N. E., Mei, Y., Jee, C. L. J., Chabi-Jesus, C., Freitas-Astúa, J., et al.
533 2021. Joá yellow blotch-associated virus, a new alphanucleorhabdovirus from a wild solanaceous
534 plant in Brazil. Arch Virol. 166: 1615–1622.
535 Dolja, V. V., Krupovic, M., and Koonin, E. V. 2020. Deep Roots and Splendid Boughs of the
536 Global Plant Virome. Annual Review of Phytopathology. 58:23–53.
537 Fox, A. 2020. Reconsidering causal association in plant virology. Plant Pathology. 69:956–961.
538 Fraile, A., and García-Arenal, F. 2016. Environment and evolution modulate plant virus
539 pathogenesis. Curr Opin Virol. 17:50–56.
540 Gaafar, Y., Richert-Pöggeler, K., Maaß, C., Vetten, H.-J., and Ziebell, H. 2019. Characterisation
541 of a novel nucleorhabdovirus infecting alfalfa (Medicago sativa). Virology Journal. 16.
542 Gaafar, Y. Z. A., Abdelgalil, M. a. M., Knierim, D., Richert-Pöggeler, K. R., Menzel, W.,
543 Winter, S., et al. 2018. First Report of physostegia chlorotic mottle virus on Tomato (Solanum
544 lycopersicum) in Germany. Plant Disease. 102:255–255.
545 Gaafar, Y. Z. A., Westenberg, M., Botermans, M., László, K., De Jonghe, K., Foucart, Y., et al.
546 2021. Interlaboratory Comparison Study on Ribodepleted Total RNA High-Throughput
547 Sequencing for Plant Virus Diagnostics and Bioinformatic Competence. Pathogens. 10:1174.
548 Hammond, J., Adams, I. P., Fowkes, A. R., McGreig, S., Botermans, M., Oorspronk, J. J. A. van,
549 et al. Sequence analysis of 43-year old samples of Plantago lanceolata show that Plantain virus X
550 is synonymous with Actinidia virus X and is widely distributed. Plant Pathology. 70: 249– 258.
551 Hanssen, I. M., Paeleman, A., Vandewoestijne, E., Bergen, L. V., Bragard, C., Lievens, B., et al.
552 2009. Pepino mosaic virus isolates and differential symptomatology in tomato. Plant Pathology.
553 58:450–460.

554 Higgins, C. M., Chang, W.-L., Khan, S., Tang, J., Elliott, C., and Dietzgen, R. G. 2016. Diversity
555 and evolutionary history of lettuce necrotic yellows virus in Australia and New Zealand. Arch
556 Virol. 161:269–277.
557 Hou, W., Li, S., and Massart, S. 2020. Is There a "Biological Desert" With the Discovery of New
558 Plant Viruses? A Retrospective Analysis for New Fruit Tree Viruses. Front. Microbiol. 11: 2953
559 Hull, 2014, Plant virology, Elsevier, Fith edition, Chapt, Page 171
560 Jackson, A. O., Dietzgen R. G., Goodin M. M., Bragg J.N., Min Deng, M. 2005. Biology of Plant
561 Rhabdoviruses. Annual Review of Phytopathology. 43:1, 623-660 Pathology 152 (4): 909-19.
562 Jones, R. A. C., Boonham, N., Adams, I. P., and Fox, A. 2021. Historical virus isolate
563 collections: An invaluable resource connecting plant virology's pre-sequencing and post-
564 sequencing eras. Plant Pathology. 70:235–248.
565 Katis, N., Chatzivassiliou, E., Clay, C. M., Maliogka, V., Pappi, P., Efthimiou, K., et al. 2011.
566 Development of an IC-RT-PCR assay for the detection of Eggplant mottled dwarf virus and
567 partial characterization of isolates from various hosts in Greece. Journal of plant pathology.
568 93:253–362.
569 Kuhn, J. H., Adkins, S., Alioto, D., Alkhovsky, S. V., Amarasinghe, G. K., Anthony, S. J., et al.
570 2020. 2020 taxonomic update for phylum Negarnaviricota (Riboviria: Orthornavirae), including
571 the large orders Bunyavirales and Mononegavirales. Arch Virol. 165:3023–3072.
572 Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. 2018. MEGA X: Molecular
573 Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol. 35:1547–1549.
574 Martelli, G. P., and Cherif, C. 1987. Eggplant Mottled Dwarf Virus Associated with Vein
575 Yellowing of Honeysuckle. Journal of Phytopathology. 119:32–41.

576 Martelli, G. P., and Hamadi, A. 1986. Occurrence of eggplant mottled dwarf virus in Algeria.
577 Plant Pathology. 35:595–597.
578 Massart, S., Candresse, T., Gil, J., Lacomme, C., Predajna, L., Ravnikar, M., et al. 2017. A
579 Framework for the Evaluation of Biosecurity, Commercial, Regulatory, and Scientific Impacts of
580 Plant Viruses and Viroids Identified by NGS Technologies. Front. Microbiol. 8: 45.
581 Mavrič, I., Tušek Žnidarič, M., Virščček Marn, M., Dolničar, P., Mehle, N., Lesemann, D.-E., et
582 al. 2006. First report of Eggplant mottled dwarf virus in potato and tomato in Slovenia. Plant
583 Pathology. 55:566–566.
584 Menzel, W., Richert-Pöggeler, K., Winter, S., and Knierim, D. 2018. Characterization of a
585 nucleorhabdovirus from Physostegia. Acta Horticulturae. 1193:29–38.
586 Oñate-Sánchez, L., and Vicente-Carbajosa, J. 2008. DNA-free RNA isolation protocols for
587 Arabidopsis thaliana, including seeds and siliques. BMC Research Notes. 1:93.
588 Pappi, P. G., Dovas, C. I., Efthimiou, K. E., Maliogka, V. I., and Katis, N. I. 2013. A novel
589 strategy for the determination of a rhabdovirus genome and its application to sequencing of
590 Eggplant mottled dwarf virus. Virus Genes. 47:105–113.
591 Power, A. G. 2000. Insect transmission of plant viruses: a constraint on virus variability. Current
592 Opinion in Plant Biology. 3:336–340.
593 Roenhorst, J. W., Botermans, M., and Verhoeven, J. T. J. 2013. Quality control in bioassays used
594 in screening for plant viruses. EPPO Bulletin. 43:244–249.
595 Samarfard, S., Bejerman, N. E., and Dietzgen, R. G. 2018. Distribution and genetic variability of
596 alfalfa dwarf virus, a cytorhabdovirus associated with alfalfa dwarf disease in Argentina. Virus
597 Genes. 54:612–615.

598 Sofy, A. R., Sofy, M. R., Hmed, A. A., Dawoud, R. A., Refaey, E. E., Mohamed, H. I., et al.
599 2021. Molecular Characterization of the Alfalfa mosaic virus Infecting Solanum melongena in
600 Egypt and the Control of Its Deleterious Effects with Melatonin and Salicylic Acid. Plants.
601 10:459.
602 Tang, J., Elliott, C., Ward, L. I., and Iqram, A. 2014. Identification of Eggplant mottled dwarf
603 virus in PEQ Hibiscus syriacus plants imported from Australia. Australasian Plant Dis. Notes.
604 10:6.
605 Tomato brown rugose fruit virus. 2020. EPPO Bulletin. 50:529–534.
606 Verhoeven, J. T. J., and Roenhorst, J. W. 2000. Herbaceous test plants for the detection of
607 quarantine viruses of potato*. EPPO Bulletin. 30:463–467.
608 Vučurović, A., Kutnjak, D., Mehle, N., Stanković, I., Pecman, A., Bulajić, A., et al. 2021.
609 Detection of Four New Tomato Viruses in Serbia using Post-Hoc High-Throughput Sequencing
610 Analysis of Samples from a Large-Scale Field Survey. Plant Disease. 105:9, 2325-2332.
611 Whitfield, A. E., Huot, O. B., Martin, K. M., Kondo, H., and Dietzgen, R. G. 2018. Plant
612 rhabdoviruses—their origins and vector interactions. Current Opinion in Virology. 33:198–207.
613 Yang, X., Chen, B., Zhang, T., Li, Z., Xu, C., and Zhou, G. 2018. Geographic Distribution and
614 Genetic Diversity of Rice Stripe Mosaic Virus in Southern China. Front. Microbiol. 9: 3068.
615
616
617
618
619

620

621

622

623

624

625

626

627

628

629

630

1 **TABLES**

2 **<u>Table 1.</u> Sample references with collection year, localization (country and town if known), original host, symptoms, detection or**
3 **confirmation method, sequencing strategy and bioinformatics pipeline used. NCBI GenBank accession numbers for each**
4 **sequenced isolate and co-infection with other viruses are also presented.**

Isolate name	Collection date	Origin: country (region or city)	Site (farm)	Original host [laboratory host if sequenced]	Symptoms on fruits	Symptoms on leaves [laboratory host if sequenced]	Detection method (D)/ confirmation (C) (protocol used)	Sequencing strategy (protocol used)	Coinfection with other viruses : Bioinformatic (B) or PCR results (PCR)	Bioinformatic pipeline (assemblers/ analyses)	Reference	Genbank accession
Fr_SM1	2002	France (Provence-Alpes-Côte d'Azur)	Site B	*Solanum melongena*	deformed	vein clearing, deformation	D: RT-PCR + sequencing (Alfaro-Fernandez et *al.,* 2009) /	Total RNA (a)	B: no	CLC workbench / Geneious	This study	MW934551
KY706238	2003	Germany, (State of Hess)	Site N	*Solanum lycopersicum*	unknown	unknown	C: RT-PCR (Gaafar *et al.,* 2018)	Total RNA + ribodepletion (Gaafar *et al.*, 2018)	B: no	Geneious	Gaafar *et al.*, 2018	KY706238
Fr_SL1	2011	France (Corse)	Site C	*Solanum lycopersicum*	deformed, uneven ripening, mottled	dwarf, mottled	D: RT-PCR + sequencing (Alfaro-Fernandez et *al.,* 2009) /	Total RNA (d)	B: Potato virus Y	Spades / Geneious	This study	MZ574100
232-12	2012	Serbia (Rasina District)	Site Q	*Solanum lycopersicum*	mottled, uneven ripening	mottled	RT-PCR (Vučurović *et al.*, 2021)	small RNA sequencing (Vucurovic *et al.,* 2021)	B: no	CLC workbench / Geneious (Vucurovic *et al.*, 2021)	Vučurović *et al.*, 2021	MT269810
238-12	2012	Serbia (Rasina District)	Site R	*Solanum lycopersicum*	mottled, uneven ripening	mottled	RT-PCR (Vučurović *et al.*, 2021)	small RNA sequencing (Vucurovic *et al.,* 2021)	B: no	CLC workbench / Geneious (Vucurovic *et al.*, 2021)	Vučurović *et al.*, 2021	MT269811

323-12	2012	Serbia (Jablanica District)	Site S	*Solanum lycopersicum*	mottled	ns	RT-PCR (Vučurović *et al.*, 2021)	small RNA sequencing (Vucurovic *et al.,* 2021)	B: Southern tomato virus	CLC workbench / Geneious (Vucurovic *et al.*, 2021)	Vučurović *et al.*, 2021	MT269812
Fr_SM2	2013	France, (Maine et Loire)	Site D	*Solanum melongena*	deformed, uneven ripening, mottled	vein clearing, plant: dwarf	D: RT-PCR + sequencing (Alfaro-Fernandez et *al.,* 2009) /	Total RNA (d)	B : no	Spades / Geneious	This study	MZ574102
Fr_SM3	2013	France, (Maine et Loire	Site D	*Solanum melongena*	deformed, uneven ripening, mottled	yellowing	D: RT-PCR + sequencing (Alfaro-Fernandez et *al.,* 2009) /	Total RNA (d)	B: no	Spades / Geneious	This study	MZ574103
Fr_SL2	2014	France (Provence-Alpes-Côte d'Azur)	Site E?	*Solanum lycopersicum*	deformed, uneven ripening, mottled	severe necrosis and dotted tasks (apical leaves)	D: RT-PCR + sequencing (Alfaro-Fernandez et *al.,* 2009) /	Total RNA (d)	B: Pepino mosaic virus + Squash mosaic virus	Spades / Geneious	This study	MZ574101
KX636164	2014	Austria	Site O	*Physostegia virginiana*	na	deformed, chlorosis and mottled	RT-PCR (Gaafar *et al.,* 2018)	Total RNA + ribodepletion (Gaafar *et al.*, 2018)	B: no	Geneious	Menzel *et al., 2018*	KX636164
KY859866	2015	Germany, (State of Hess)	Site N	*Solanum lycopersicum [N. benthamiana]*	marbling and discoloration	ns	C: RT-PCR (Gaafar *et al.,* 2018)	Total RNA (Gaafar *et al.*, 2017)	B: no	Geneious	Gaafar *et al.*, 2018	KY859866
Nd_SL1	2017	Netherlands	Site F	*Solanum lycopersicum [N.benthamiana]*	na	deformed, vein clearing	D: same as seq strategy C: mechanical inoculation	Total RNA (a)	B: no	CLC workbench / Geneious	This study	OK646027
Ru_SL1	2017	Russia	Site G	*Solanum lycopersicum [N.benthamiana]*	uneven ripening, mottled	mottled	D: same as seq strategy C: mechanical inoculation	Total RNA (a)	B: no	CLC workbench / Geneious	This study	OK646028
MK978541	2017	Germany, (State of Hess)	Site N	*Solanum lycopersicum [N. benthamiana]*	marbling and discoloration	distortion and mild yellow spots	C: RT-PCR (Gaafar *et al.,* 2018)	dsRNA (Gaafar et al., 2020)	B: no	Geneious	Gaafar *et al.*, 2020	MK978541

MW848528	2017	Germany, (State of Hess but different site)	Site P	*Solanum lycopersicum [N. benthamiana]*	marbling and discoloration	mild yellow spots	D: ELISA using JKI-2051	Total RNA + ribodepletion (Gaafar et al., 2020)	B: no	Geneious	Gaafar *et al., 2021*	MW848528
Nd_CS1	2018	Netherlands (Zélande)	Site H	*Cucumis sativus*	pointed, deformed, vertical chlorotic stripes	interveinal chlorosis and sunken veins (rugosity)	D: same as seq strategy C: mechanical inoculation	Total RNA (a)	B: no	CLC workbench / Geneious	This study	OK646030
Nd_H1	2018	Netherlands, (Gelderland)	Site I	*Helleborus*	na	vein clearing, chlorotic patterns and rings.	same as seq strategy	Total RNA (a)	B: no	CLC workbench / Geneious	This study	OK646029
Nd_H2	2018	Netherlands (South Holland)	Site J	*Helleborus*	na	chlorosis next to veins and mosaic	D: same as seq strategy C: mechanical inoculation	Total RNA (a)	B: no	CLC workbench / Geneious	This study	OK646031
Fr_SM4	2018	France (Nouvelle Aquitaine)	Site K	*Solanum melongena*	deformed, uneven ripening, mottled	vein clearing, deformed	D: RT-PCR + sequencing (Alfaro-Fernandez et *al.,* 2009) /	Total RNA (d)	B: no	Spades / Geneious	This study	MZ574104
Be_SL1	2018	Belgium, (Gembloux)	Site A	*Solanum lycopersicum*	deformed, uneven ripening, mottled	vein clearing on apical leaves	C: RT-PCR (Gaafar *et al.,* 2018)	VANA (c)	B: no	Spades / Geneious	This study	MZ501244
Be_SM1	2019	Belgium, (Gembloux)	Site A	*Solanum melongena*	na	vein clearing	D: same as sequencing strategy / C: RT-PCR (Gaafar et al., 2018)	VANA (c)	B: no	Spades / Geneious	This study	MZ501245
Sl_SL1	2019	Slovenia	Site L	*Solanum lycopersicum*	deformed, uneven ripening, mottled	severe leaf curling and mottling P : dwarf	C: RT-PCR (Gaafar *et al.,* 2018)	Total RNA (f)	B and PCR: Tomato mosaic virus, Potato virus Y	CLC Genomics Workbench / SPAdes	This study	MW366749

Ro_SL1	2019	Romania	Site M	*Solanum lycopersicum [N.benthamiana]*	uneven ripening, mottled	na	D: same as seq strategy C: mechanical inoculation	Total RNA (a)	B: no	CLC workbench / Geneious	This study	OK646026
Be_IB1	2019	Belgium (Kruisem)	Site U	*Ipomoea batatas*	na	chlorosis, purple pattern	C: RT-PCR (own primers)	Total RNA (b)	B: Sweet potato feathery mottle virus	Own pipeline + VirusDetect + BWA/QUASR	This study	MZ389081
Be_SA1	2019	Belgium (Putte)	Site T	*Stachys affinis*	na	ns	C: RT-PCR (Gaafar *et al.,* 2018)	Total RNA (b)	B: no	Own pipeline + VirusDetect + BWA/QUASR	This study	MZ322957
Ge_CS1	2020	Germany, (State of Hess)	Site N	*Cucumis sativus*	na	mosaic, leaf curling, chlorotic spots and yellowing	C: RT-PCR (Gaafar *et al.,* 2018)	dsRNA e*	B: no	Minimap2 / Geneious	This study	MW081210
Be_GP1	2020	Belgium, (Gembloux)	Site A	*Galinsoga parviflora*	na	vein clearing	C: RT-PCR (Gaafar *et al.,* 2018)	Total RNA (d)	B: no	Spades / Geneious	This study	MZ574099
Be_PM1	2020	Belgium (Putte)	Site T	Persicaria maculosa	na	ns	C: RT-PCR (own primers)	Total RNA (b)	B: no	Own pipeline + VirusDetect + BWA/QUASR	This study	MZ389082
Be_IB2	2020	Belgium , import from Portugal	Site V	*Ipomoea batatas*	na	vein clearing , mosaic and stunting	C: RT-PCR (own primers)	Total RNA (b)	B: Sweet potato feathery mottle virus, Sweet potato chlorotic stunt virus, Potato virus Y	Own pipeline + VirusDetect + BWA/QUASR	This study	MW834321

Legend: a= protocol used by NVWA, b= protocol used by ILVO, c, d = protocol used by Uliege, e= protocol used by JKI, f= protocol used by NIB; All the samples were sequenced on Illumina plateform except for * = MinION ; na = non applicable (for example in the case that there is no fruit when the symptoms were recorded), ns = no symptoms observed

9 **<u>Table 2.</u> PhCMoV indexing host range study accross different laboratories (DSMZ, JKI and NVWA).**

	DSMZ- KX636164 (Menzel *et al.,* 2018)		JKI - KY859866 (Gaafar *et al.,* 2018) - HZ15-192		JKI - MW848528 (This study) - HZ16-558		NVWA - Ru_SL1 (This study)	NVWA - Nd_SL1 (This study)	NVWA - Ro_SL1 (This study)	NVWA - Nd_CS1 (This study)	NVWA - Nd_H2 (This study)
Inoculated test plant	**Symptoms**	**ELISA/ RT-PCR**	**Symptoms**	**ELISA/ RT-PCR**	**Symptoms**	**ELISA/ RT-PCR**	**Symptoms**	**Symptoms**	**Symptoms**	**Symptoms**	**Symptoms**
Chenopodium quinoa	-	-	y, m	+	-	-		-			
C. sativus	-	-									
Chenopodiastrum murale					-	-					
Datura stramonium	-	-			y	+	-	-			
D. metel	-	-			y	+					
Hyoscyamus niger					-	-					
Medicago sativa					-	-					
Nicotiana benthamiana	m	+	y, m	+	y, vc	+	m, r, g (5 wks p.i., 3/3)	m, r, g (5 wks p.i., 3/3)	m, r, g (4 wks p.i. (1/2))	vc, m, r, g (5 wks p.i. 2/3)	m, r, g (7 wks p.i., 3/3)
N. glutinosa	-	-			-	-		-			
N. occidentalis 'P1'	-	-	-	-	-	-	vc (4 wks p.i., 3/3)	vc, g (cl) (4 wks p.i., 3/3)	(0/2)	c (7 wks p.i., 1/3)	vc, g (7 wks p.i., 1/3)
N. tabacum samsunn					-	-					
N. tabacum 'WB'	vc	+					-	-			
N. clevelandii	m	+	y, m	+							
N. glutinosa '24A'	-	-			-	-					
N. hesperis	-	-			-	-					
N. occidentalis '37B'	vc	+	-	-							
Physalis floribunda							-	-			
Petroselinum crispum					-	-					
Petunia					-	-					
Physalis floridana	-	+			-	-					
Solanum lycopericum					-	-		-			

10

11 Legend: c = chlorosis, cl = chlorotic lesions, g = growth reduction, ic = interveinal chlorosis, m = mottle, nl = necrotic lesions, r = rugosity, vc =
12 vein clearing, y = yellowing, () = symptoms observed occasionally - = no symptoms, empty space = not tested, xx wks p.i. = number of weeks after
13 inoculation before the observation of the first systemic (?) symptom, x/x = number of plants showing symptoms/ number of inoculated plants

14 **FIGURE LEGENDS**

15 **Fig. 1. Pictures of natural PhCMoV infected plants** Symptoms of PhCMoV on infected *Solanum lycopersicum* fruits [Ro_SL1 (a)] and leaves
16 [Nd_SL1 (b)], *Solanum melongena* fruit [Fr_SM4 (c)] and leaves [Be_SM1(d)], *Cucumis sativus fruits* [Nd_CS1 (e)] and leaves [Ge_CS1
17 (f)], *Helleborus* leaves [Nd_H1 (g)] *Galinsoga parviflora* [Be_GP1 (h)]. No coinfections with other viruses occurred in these samples.

18 **Fig. 2. Phylogenetic tree inferring relationships of 29 PhCMoV isolates (among which 21 new genomes published in this study) based on**
19 **nucleotides alignment of near complete genomic sequences.** The phylogenetic tree was inferred by using the Maximum Likelihood method and
20 GTR+G+I model based on the full genome sequence MUSCLE alignment (nucleotides) of all the PhCMoV isolates known at this date. Each isolate
21 is labelled with its name and the information of the collection: country (flag), host, and year. Orange squares and letters highlight identical
22 collection sites (farm). The values on the branches show the percentage of support out of 1000 bootstrap replications, and the scale bar indicates
23 the number of nucleotides substitutions per site.

24 **Fig. 3. Differences and similarities between selected PhCMoV isolates in different ORF** a) graphic representation of nucleotide identities (%)
25 using SIMPLOT of 16 full genome sequences of PhCMoV (ref query = KX636164; Window: 200bp, Step: 20 bp, Gapstrim: On, Hamming across the
26 complete genome sequence and its genome organization. In red is the representation of the most divergent isolate Nd_SL1 b) Nucleotide and

27 amino acid sequence identities calculated for N, X, P, Y, M, G and L ORFs for all isolates studied. The identities (%) were calculated based on

28 MUSCLE alignment (Muscle 3.8.425 by Robert C. Edgar). The number of bp for the full genome sequence is indicated for KX636164.

Fig. 1. **Pictures of natural PhCMoV infected plants** Symptoms of PhCMoV on infected *Solanum lycopersicum* fruits [Ro_SL1 (a)] and leaves [Nd_SL1 (b)], *Solanum melongena* fruit [Fr_SM4 (c)] and leaves [Be_SM1(d)], *Cucumis sativus* fruits [Nd_CS1 (e)] and leaves [Ge_CS1 (f)], *Helleborus* leaves [Nd_H1 (g)] *Galinsoga parviflora* [Be_GP1 (h)]. No coinfections with other viruses occurred in these samples.

638x347mm (130 x 130 DPI)

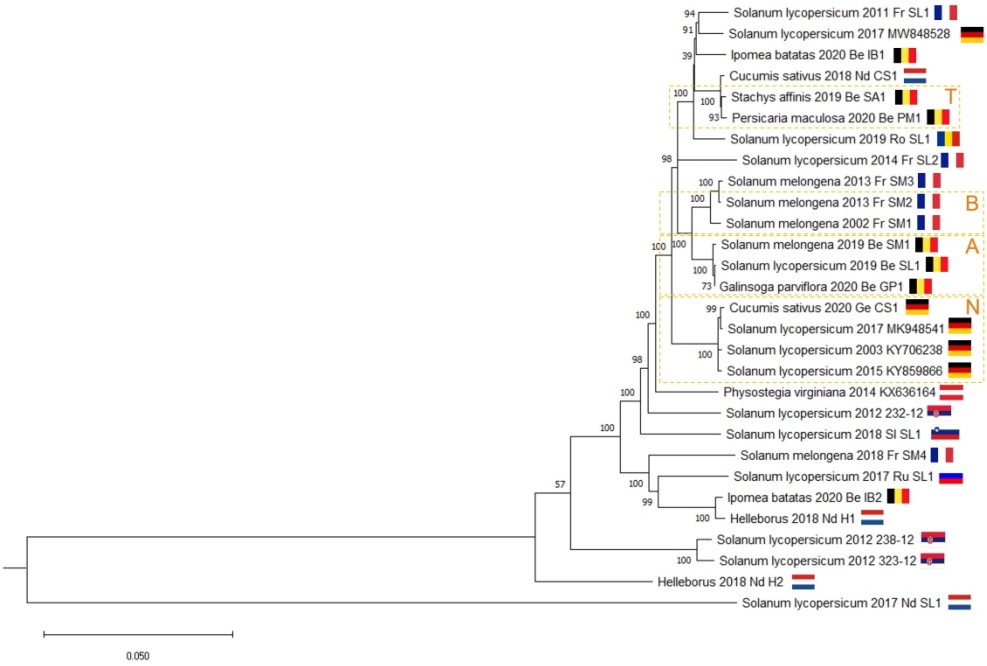
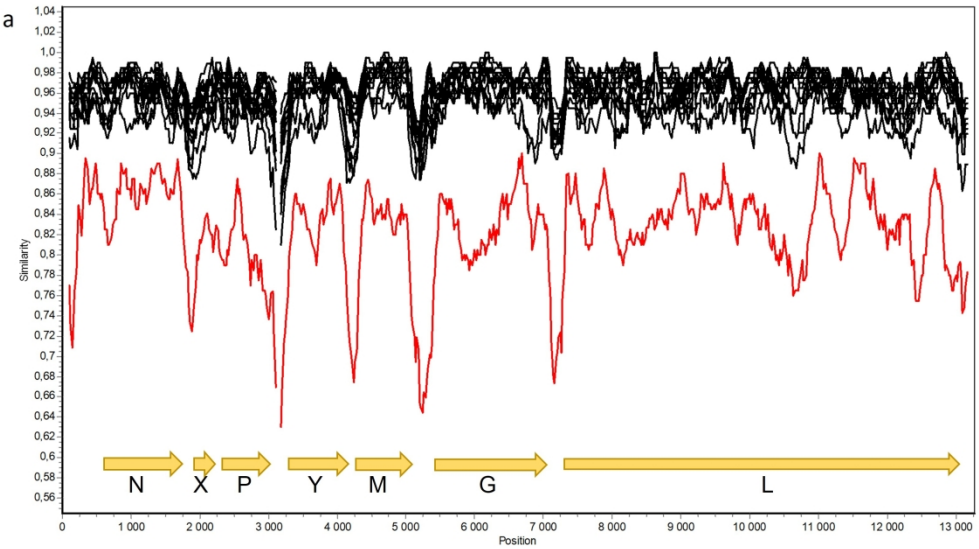


Fig. 2. Phylogenetic tree inferring relationships of 29 PhCMoV isolates (among which 21 new genomes published in this study) based on nucleotides alignment of near complete genomic sequences. The phylogenetic tree was inferred by using the Maximum Likelihood method and GTR+G+I model based on the full genome sequence MUSCLE alignment (nucleotides) of all the PhCMoV isolates known at this date. Each isolate is labelled with its name and the information of the collection: country (flag), host, and year. Orange squares and letters highlight identical collection sites (farm). The values on the branches show the percentage of support out of 1000 bootstrap replications, and the scale bar indicates the number of nucleotides substitutions per site.

929x654mm (96 x 96 DPI)

a



b

		Full genome	N	X	P	Y	M	G	L
PhCMoV isolates genome and ORF size range	nts	13320	1269-1482	306-366	885	864-1104	756-798	1857-1935	5841-5874
	aa	na	422-494	101-121	294-295	287-368	251-265	618-656	1938-1958
Range of sequence identity of Nd_SL1 to all known PhCMoV	nts ID (%)	82	85-87	76-81	79-81	83-85	84-85	82	83-84
	aa ID (%)	na	96-100	73-76	83-87	92-99	93-97	91-94	95-96
Range of sequence identities amongst all known PhCMoV isolates excluding Nd_SL1	nts ID (%)	>93	>94	>89	>93	>93	>95	>94	>94
	aa ID (%)	na	>95	>88	>95	>95	>96	>97	>98

Fig. 3. Differences and similarities between selected PhCMoV isolates in different ORF a) graphic representation of nucleotide identities (%) using SIMPLOT of 16 full genome sequences of PhCMoV (ref query = KX636164; Window: 200bp, Step: 20 bp, Gapstrim: On, Hamming across the complete genome sequence and its genome organization. In red is the representation of the most divergent isolate Nd_SL1 b) Nucleotide and amino acid sequence identities calculated for N, X, P, Y, M, G and L ORFs for all isolates studied. The identities (%) were calculated based on MUSCLE alignment (Muscle 3.8.425 by Robert C. Edgar). The number of bp for the full genome sequence is indicated for KX636164.

248x183mm (330 x 330 DPI)

Isolate name	Collection date	Origin: country (region or city)	Site (farm)	Original host [laboratory host if sequenced]	Symptoms on fruits	Symptoms on leaves [laboratory host if sequenced]
Fr_SM1	2002	France (Provence-Alpes-Côte d'Azur)	Site B	*Solanum melongena*	deformed	vein clearing, deformation
KY706238	2003	Germany, (State of Hess)	Site N	*Solanum lycopersicum*	unknown	unknown
Fr_SL1	2011	France (Corse)	Site C	*Solanum lycopersicum*	deformed, uneven ripening, mottled	dwarf, mottled
232-12	2012	Serbia (Rasina District)	Site Q	*Solanum lycopersicum*	mottled, uneven ripening	mottled
238-12	2012	Serbia (Rasina District)	Site R	*Solanum lycopersicum*	mottled, uneven ripening	mottled
323-12	2012	Serbia (Jablanica District)	Site S	*Solanum lycopersicum*	mottled	ns
Fr_SM2	2013	France, (Maine et Loire)	Site D	*Solanum melongena*	deformed, uneven ripening, mottled	vein clearing, plant: dwarf
Fr_SM3	2013	France, (Maine et Loire	Site D	*Solanum melongena*	deformed, uneven ripening, mottled	yellowing
Fr_SL2	2014	France (Provence-Alpes-Côte d'Azur)	Site E?	*Solanum lycopersicum*	deformed, uneven ripening, mottled	severe necrosis and dotted tasks (apical leaves)
KX636164	2014	Austria	Site O	*Physostegia virginiana*	na	deformed, chlorosis and mottled
KY859866	2015	Germany, (State of Hess)	Site N	*Solanum lycopersicum [N. benthamiana]*	marbling and discoloration	ns
Nd_SL1	2017	Netherlands	Site F	*Solanum lycopersicum [N.benthamiana]*	na	deformed, vein clearing
Ru_SL1	2017	Russia	Site G	*Solanum lycopersicum [N.benthamiana]*	uneven ripening, mottled	mottled
MK978541	2017	Germany, (State of Hess)	Site N	*Solanum lycopersicum [N. benthamiana]*	marbling and discoloration	distortion and mild yellow spots

MW848528	2017	Germany, (State of Hess but different site)	Site P	*Solanum lycopersicum [N. benthamiana]*	marbling and discoloration	mild yellow spots
Nd_CS1	2018	Netherlands (Zélande)	Site H	*Cucumis sativus*	pointed, deformed, vertical chlorotic stripes	interveinal chlorosis and sunken veins (rugosity)
Nd_H1	2018	Netherlands, (Gelderland)	Site I	*Helleborus*	na	vein clearing, chlorotic patterns and rings.
Nd_H2	2018	Netherlands (South Holland)	Site J	*Helleborus*	na	chlorosis next to veins and mosaic
Fr_SM4	2018	France (Nouvelle Aquitaine)	Site K	*Solanum melongena*	deformed, uneven ripening, mottled	vein clearing, deformed
Be_SL1	2018	Belgium, (Gembloux)	Site A	*Solanum lycopersicum*	deformed, uneven ripening, mottled	vein clearing on apical leaves
Be_SM1	2019	Belgium, (Gembloux)	Site A	*Solanum melongena*	na	vein clearing
Sl_SL1	2019	Slovenia	Site L	*Solanum lycopersicum*	deformed, uneven ripening, mottled	severe leaf curling and mottling P : dwarf
Ro_SL1	2019	Romania	Site M	*Solanum lycopersicum [N.benthamiana]*	uneven ripening, mottled	na
Be_IB1	2019	Belgium (Kruisem)	Site U	*Ipomoea batatas*	na	chlorosis, purple pattern
Be_SA1	2019	Belgium (Putte)	Site T	*Stachys affinis*	na	ns
Ge_CS1	2020	Germany, (State of Hess)	Site N	*Cucumis sativus*	na	mosaic, leaf curling, chlorotic spots and yellowing
Be_GP1	2020	Belgium, (Gembloux)	Site A	*Galinsoga parviflora*	na	vein clearing
Be_PM1	2020	Belgium (Putte)	Site T	Persicaria maculosa	na	ns
Be_IB2	2020	Belgium , import from Portugal	Site V	*Ipomoea batatas*	na	vein clearing , mosaic and stunting

Detection method (D)/ confirmation (C) (protocol used)	Sequencing strategy (protocol used)	Coinfection with other viruses : Bioinformatic (B) or PCR results (PCR)	Bioinformatic pipeline (assemblers/ analyses)	Reference	Genbank accession
D: RT-PCR + sequencing (Alfaro-Fernandez et *al.,* 2009) /	Total RNA (a)	B: no	CLC workbench / Geneious	This study	MW934551
C: RT-PCR (Gaafar *et al.,* 2018)	Total RNA + ribodepletion (Gaafar *et al*., 2018)	B: no	Geneious	Gaafar *et al.*, 2018	KY706238
D: RT-PCR + sequencing (Alfaro-Fernandez et *al.,* 2009) /	Total RNA (d)	B: Potato virus Y	Spades / Geneious	This study	MZ574100
RT-PCR (Vučurović *et al.*, 2021)	small RNA sequencing (Vucurovic *et al.,* 2021)	B: no	CLC workbench / Geneious (Vucurovic *et al.*, 2021)	Vučurović *et al.*, 2021	MT269810
RT-PCR (Vučurović *et al.*, 2021)	small RNA sequencing (Vucurovic *et al.,* 2021)	B: no	CLC workbench / Geneious (Vucurovic *et al.*, 2021)	Vučurović *et al.*, 2021	MT269811
RT-PCR (Vučurović *et al.*, 2021)	small RNA sequencing (Vucurovic *et al.,* 2021)	B: Southern tomato virus	CLC workbench / Geneious (Vucurovic *et al.*, 2021)	Vučurović *et al.*, 2021	MT269812
D: RT-PCR + sequencing (Alfaro-Fernandez et *al.,* 2009) /	Total RNA (d)	B : no	Spades / Geneious	This study	MZ574102
D: RT-PCR + sequencing (Alfaro-Fernandez et *al.,* 2009) /	Total RNA (d)	B: no	Spades / Geneious	This study	MZ574103
D: RT-PCR + sequencing (Alfaro-Fernandez et *al.,* 2009) /	Total RNA (d)	B: Pepino mosaic virus + Squash mosaic virus	Spades / Geneious	This study	MZ574101
RT-PCR (Gaafar *et al.,* 2018)	Total RNA + ribodepletion (Gaafar *et al*., 2018)	B: no	Geneious	Menzel *et al., 2018*	KX636164
C: RT-PCR (Gaafar *et al.,* 2018)	Total RNA (Gaafar *et al*., 2017)	B: no	Geneious	Gaafar *et al.*, 2018	KY859866
D: same as seq strategy C: mechanical inoculation	Total RNA (a)	B: no	CLC workbench / Geneious	This study	OK646027
D: same as seq strategy C: mechanical inoculation	Total RNA (a)	B: no	CLC workbench / Geneious	This study	OK646028
C: RT-PCR (Gaafar *et al.,* 2018)	dsRNA (Gaafar et al., 2020)	B: no	Geneious	Gaafar *et al.*, 2020	MK978541

D: ELISA using JKI-2051	Total RNA + ribodepletion (Gaafar et al., 2020)	B: no	Geneious	Gaafar *et al., 2021*	MW848528
D: same as seq strategy C: mechanical inoculation	Total RNA (a)	B: no	CLC workbench / Geneious	This study	OK646030
same as seq strategy	Total RNA (a)	B: no	CLC workbench / Geneious	This study	OK646029
D: same as seq strategy C: mechanical inoculation	Total RNA (a)	B: no	CLC workbench / Geneious	This study	OK646031
D: RT-PCR + sequencing (Alfaro-Fernandez et *al.,* 2009) /	Total RNA (d)	B: no	Spades / Geneious	This study	MZ574104
C: RT-PCR (Gaafar *et al.,* 2018)	VANA (c)	B: no	Spades / Geneious	This study	MZ501244
D: same as sequencing strategy / C: RT-PCR (Gaafar et al., 2018)	VANA (c)	B: no	Spades / Geneious	This study	MZ501245
C: RT-PCR (Gaafar *et al.,* 2018)	Total RNA (f)	B and PCR: Tomato mosaic virus, Potato virus Y	CLC Genomics Workbench / SPAdes	This study	MW366749
D: same as seq strategy C: mechanical inoculation	Total RNA (a)	B: no	CLC workbench / Geneious	This study	OK646026
C: RT-PCR (own primers)	Total RNA (b)	B: Sweet potato feathery mottle virus	Own pipeline + VirusDetect + BWA/QUASR	This study	MZ389081
C: RT-PCR (Gaafar *et al.,* 2018)	Total RNA (b)	B: no	Own pipeline + VirusDetect + BWA/QUASR	This study	MZ322957
C: RT-PCR (Gaafar *et al.,* 2018)	dsRNA e*	B: no	Minimap2 / Geneious	This study	MW081210
C: RT-PCR (Gaafar *et al.,* 2018)	Total RNA (d)	B: no	Spades / Geneious	This study	MZ574099
C: RT-PCR (own primers)	Total RNA (b)	B: no	Own pipeline + VirusDetect + BWA/QUASR	This study	MZ389082
C: RT-PCR (own primers)	Total RNA (b)	B: Sweet potato feathery mottle virus, Sweet potato chlorotic stunt virus, Potato virus Y	Own pipeline + VirusDetect + BWA/QUASR	This study	MW834321

	DSMZ- KX636164 (Menzel *et al.,* 2018)		JKI - KY859866 (Gaafar *et al.,* 2018) - HZ15-192		JKI - MW848528 (This study) - HZ16-558	
Inoculated test plant	**Symptoms**	**ELISA/ RT-PCR**	**Symptoms**	**ELISA/ RT-PCR**	**Symptoms**	**ELISA/ RT-PCR**
Chenopodium quinc	-	-	y, m	+	-	-
C. sativus	-	-				
Chenopodiastrum murale					-	-
Datura stramonium	-	-			y	+
D. metel	-	-			y	+
Hyoscyamus niger					-	-
Medicago sativa					-	-
Nicotiana bentham	m	+	y, m	+	y, vc	+
N. glutinosa	-	-			-	-
N. occidentalis 'P1'	-	-	-	-	-	-
N. tabacum samsunn					-	-
N. tabacum 'WB'	vc	+				
N. clevelandii	m	+	y, m	+		
N. glutinosa '24A'	-	-			-	-
N. hesperis	-	-			-	-
N. occidentalis '37B	vc	+	-	-		
Physalis floribunda						
Petroselinum crispum					-	-
Petunia					-	-
Physalis floridana	-	+			-	-
Solanum lycopericum					-	-

NVWA - Ru_SL1 (This study)	NVWA - Nd_SL1 (This study)	NVWA - Ro_SL1 (This study)	NVWA -Nd_CS1 (This study)	NVWA - Nd_H2 (This study)
Symptoms	**Symptoms**	**Symptoms**	**Symptoms**	**Symptoms**
	-			
-	-			
m, r, g (5 wks p.i., 3/3)	m, r, g (5 wks p.i., 3/3)	m, r, g (4 wks p.i. (1/2))	vc, m, r, g (5 wks p.i. 2/3)	m, r, g (7 wks p.i., 3/3)
	-			
vc (4 wks p.i., 3/3)	vc, g (cl) (4 wks p.i., 3/3)	(0/2)	c (7 wks p.i., 1/3)	vc, g (7 wks p.i., 1/3)
-	-			
-	-			
	-			

<u>Supplementary method 1</u>:

Samples Be_SL1, Be_SM1 and Be_GP1

The VANA protocol and library preparation used for the sample Be_SL1 and Be_SM1 followed the method described by Maclot et *al.,* (2021) after Palanga *et al.,* (2016) and Filloux *et al.,* (2015). In brief, 10 g of tissue were ground in 50 mL of Hanks' buffered salt solution (HBSS, composed of 0.137 M NaCl, 5.4 mM KCl, 0.25 mM Na2HPO4, 0.07 g glucose, 0.44 mM KH2PO4, 1.3 mM CaCl2, 1.0 mM MgSO4, 4.2 mM NaHCO3), using a tissue homogenizer. In a 50 ml falcon, the clarification was obtained from a centrifugation run of 10,000 g for 10 min at 4°C. Supernatant was then filtered through a 0.45 µm sterile syringe filter and 10,5 ml of supernatant were put into an ultracentrifuge tube (Beckman Ultra clear 14 mL tubes (#344085)). Then, a sucrose cushion, made of 1 ml of 30% sucrose in 0.2M potassium phosphate pH 7.0, was deposited at the bottom of the tube. Extract was then centrifuged at 40 000 rpm for 2 hours at 4°C using the 50Ti rotor (Beckman). The library preparation also followed the method of Maclot *et al*. (2021). Briefly, the pellet was suspended in 0,5 ml HBSS. From the resuspension, 200 µl was digested by 15 U bovine pancreas DNase I (Euromedex) and 1.9 U RNase A (Euromedex) suspension during 90 min at 37 °C. Total nucleic acids were 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 multiplex identifier (MID) linker was used in the PCR. Finally, an amplification step (PCR) was performed using HotStarTaq (Qiagen). Then, the Illumina library was prepared at GIGA Genomics (University of Liege, Belgium) using NEBNext Ultra II DNA library prep kit (New England BioLabs, US).

For Be_GP1, the RNA extract was DNase treated (DNase I, Amplification Grade, Thermofisher) prior to Total RNA sequencing. The sample was then prepared at GIGA using the TruSeq Stranded Total RNA with Ribo-Zero Plant kit.

All the three libraries were then sequenced on the Illumina NextSeq500 sequencer for the generation of 6.8 M; 0.2 M and 22 M of paired-end reads (2 x 150 base pair (bp)) for Be_SM1, Be_SL1 and Be_GP1 respectively. Resulting sequence reads were first demultiplexed for Be_SL1 and Be_SM1 according to the linker and trimmed from the adaptor, paired, and merged using the Geneious R11 software platform (https://www.geneious.com) prior to *de novo* assemble with SPAdes (Bankevich *et al.,* 2012). Contigs were compared using tBlastx against a database of viruses and viroids sequences downloaded from NCBI in November 2020 (RefSeq virus database). The complete PhCMoV genome was assembled for each sample by mapping reads with Geneious to a reference genome KX636164 (using parameters Medium-Low Sensitivity/ Fast). The consensus sequences were then extracted from the mapping (threshold: 50%). RT-PCR using the primers published by Gaafar *et al.,* (2018) was conducted on all the three sequence samples to confirm the virus presence.

Sample Be_SA1, Be_IB1, Be_IB2 and Be_PM1

Total RNA was extracted from these samples with the Spectrum™ Plant Total RNA Kit (Sigma) according to the manual instructions yet preceded by a prelysis step of 10 min at 56°C with 3% 2-mercaptoethanol and 2.5% PVP-40 added to the Spectrum kit lysis buffer. Total RNA was sent to Admera Health (New Jersey, USA) for RNA-depleted library preparation (NEBNext Ultra II with Ribo-Zero Plant) and Illumina sequencing of forty millions of paired-end reads (2 x 150bp) per sample on Illumina NextSeq sequencer. Resulting sequence reads were trimmed with Cutadapt (Martin, 2011) and merged with PEAR (Zhang *et al.,* 2014). Remaining ribosomal RNA sequences

were removed with SortMeRNA (Kopylova *et al.,* 2012) and a blast search on the non-rRNA reads was conducted against the VirusDetect database vrl_Plants_U239_U100 (Zheng *et al.,* 2017). Duplicate reads were removed using Picardtools, v.1.95 and clean reads were mapped against the reference genome KY706238 using BWA (v.0.7.8) The complete PhCMoV genomes were constructed with QUASR, v.6.0.8 for two samples of *I. batatas* (Be_IB1 and Be_IB2), one sample of *S. affinis* (Be_SA1) and one sample of *P. maculosa* (Be_PM1) that tested positive during the survey. Be_PM1 and Be_SA1 were collected at the same site while the other isolates were collected from distinct locations. RT-PCR using the primers published by Gaafar *et al.,* (2018) was conducted on Be_SA1 to confirm the virus presence. For Be_IB1 & IB2 and Be_PM1, the presence of PhCoMV was confirmed by RT-PCR with primers designed by ILVO (PhCMoV_1376 AGGCTCTCAAGAACAACCCG & PhCMoV_1800 TCATGGTGTGTTGGGTTTTT).

Sample Ge_CS1

To identify the virus, double stranded RNA (dsRNA) extraction followed by MinION sequencing were performed as described in (Gaafar *et al.,* 2019) and 1.15 million of reads were generated. Using BLASTn search, PhCMoV was the only identified virus in the sample. The new complete PhCMoV genome was assembled with Minimap2 (Li *et al*., 2018) by mapping the reads against the reference genome MK948541. Additional confirmation step was performed using RT-PCR with PhCMoV specific primers (Gaafar *et al.,* 2018).

In August 2020, rhabdovirus particles were observed by TEM from another cucumber sample (JKI-2086069) from Wiesbaden, Hesse. The presence of PhCMoV was confirmed by RT-PCR (Gaafar *et al.,* 2018).

Sample SL_SL1

Total RNA from individual plants was extracted using RNeasy Plant Mini kit ® (Qiagen) following the protocol described by Pecman *et al.,* (2017). Extracted RNAs were pooled by species or/and symptom expression (5 pooled samples in total) and after ribosomal RNA depletion and library preparation with Illumina TruSeq library preparation kit, the pools were sequenced using Illumina HiSeq2500 with the generation of 19 million of paired-end reads (2 x 150bp) (Macrogen, Korea). Bioinformatics analyses for detection of viruses were conducted using CLC Genomics Workbench (v. 20) and SPAdes (v. 3.14) as described before (Pecman *et al.,* 2017). PhCMoV was detected in a pool of six tomato plants with virus-like disease symptoms. Genome of the virus was assembled by mapping the reads to the reference genome of PhCMoV (NCBI GenBank accession number KY706238). RT-PCR confirmation of PhCMoV using the primers of Gaafar *et al.,* (2018) in individual samples revealed only one plant (Sl_SL1) out of six to be infected with the virus. RT-qPCR based on the protocol of Boben *et al.,* (2007) and Kogovšek *et al.,* (2008) of also revealed presence of tomato mosaic virus and potato virus Y in the same plant.

Samples Nd_SL1, Ru_SL1, Nd_H1, Nd_H2, Ro_SL1 and Nd_CS1

Total RNA was extracted from leaf samples Nd_SL1, Ru_SL1, Nd_H1, Nd_H2, Ro_SL1 and Nd_CS1, using the RNeasy Plant Mini Kit ® (Qiagen) as described in Botermans *et al*., (2013). Nearly complete genome sequences were obtained using HTS by a ribosomal RNA-depleted total RNA approach. DNase treated total RNA extract was sent to GenomeScan (Leiden, Netherlands) for generation of 2 Gb Illumina RNAseq paired-end reads (2 x 150bp) per sample. There, the RNA extract was ribosome depleted using the Ribo-zero rRNA removal plant leaf kit (Illumina). The Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs) was used to process the samples according to the protocol "NEBNext Ultra II Directional RNA Library Prep Kit for Illumina". Quality and yield sample preparation were measured with a Fragment Analyser (Agilent) prior to pooling for sequencing on a NovaSeq (Illumina). Sequence reads (PE150) were

analysed in CLC Genomics workbench v. 11.0.1 (Qiagen) and run in a custom workflow build for detection of de novo assembled viral contigs. First, a quality trim (quality limit = 0.05; ambiguous limit = 2) was performed, followed by a de novo assembly (map reads back to contigs = on; length fraction = 0.8; similarity fraction = 0.8; minimum contig length = 200) and consensus sequences extraction (low coverage threshold = 10; remove regions with low coverage = on; post-remove action = split). The de novo assembled contigs (>100 nt) were analysed using megaBLAST (maximum alignments per database sequence = 5; maximum E-value = 1e−6, minimum identity = 70%) and DIAMOND (Buchfink et al., 2015) with a local installation of the NCBI nr(/nt) databases. All contigs were analysed using the Pfam search option (database v 31). BLAST results were visualized in Krona (bit score threshold = 25) (Ondov et al., 2011). The same pipeline was repeated with 1% of all reads as *de novo* assembly of high coverage contigs can be problematic, resulting in fragmented assemblies. Viral sequences were analysed in Geneious R11 (Biomatters).

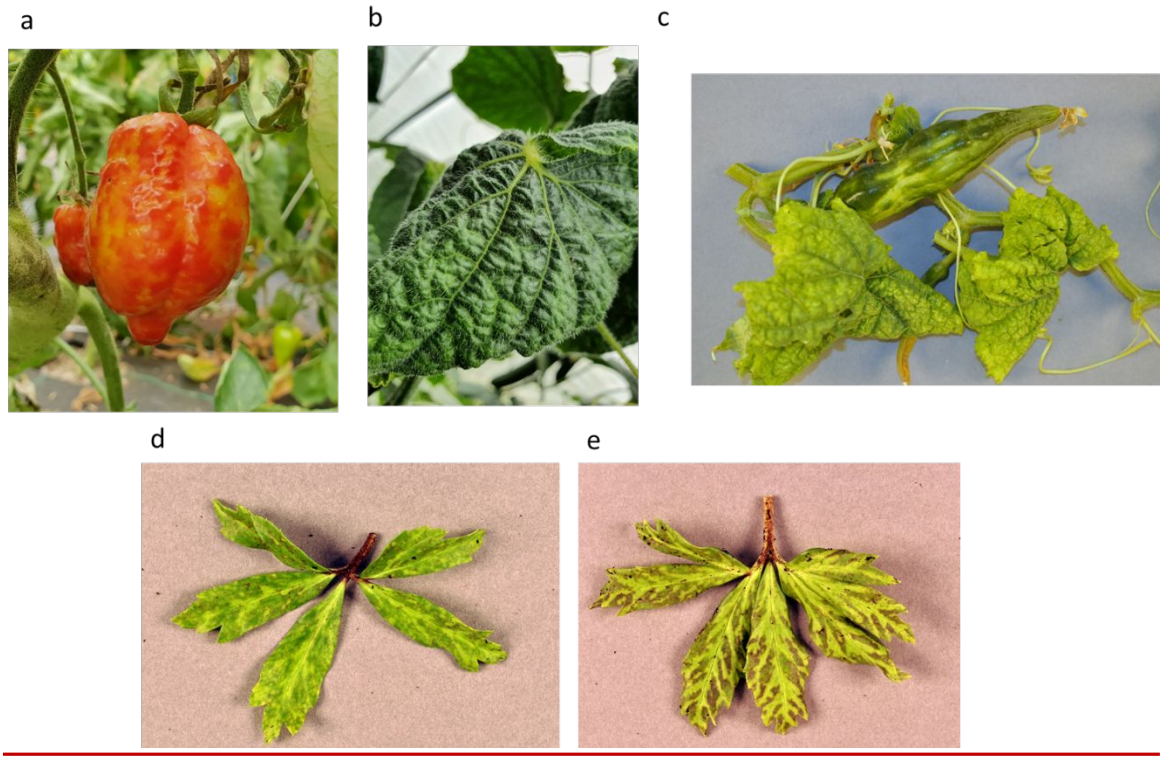
Supplementary method 2:

The sample 'C08-119' has been sequenced by sanger according to Alfaro-Fernández *et al.,* (2011).

>RYS_C08-119-A2021

CTACATCTTGGAATAtTATCCATACGGTAAAAGACAAGGCAATATCACCAAACAGGG
AGGAGCTGCATCACAACTTGATCACACGTGGCCATGTTTTCAATCAAGACATGAGAA
GGGTCATCCTGAAAACTATTATGACTCCATTAGGGCCCATGAGAGAGTTCCTGGAGA
AGGTAGGGGAGAGTGGTCTTGAGAATAAATATCTTATGATTGGGATCTACCCAAAAG
AGAGAGAATTGAAGGTCAAGCCCAGATTCTTCTCTTTGATGACCCATGAATTCAGAT
TGTATGTTACGTCTACTGAGGGGTTGTTGAACGATAAGGTTTTG

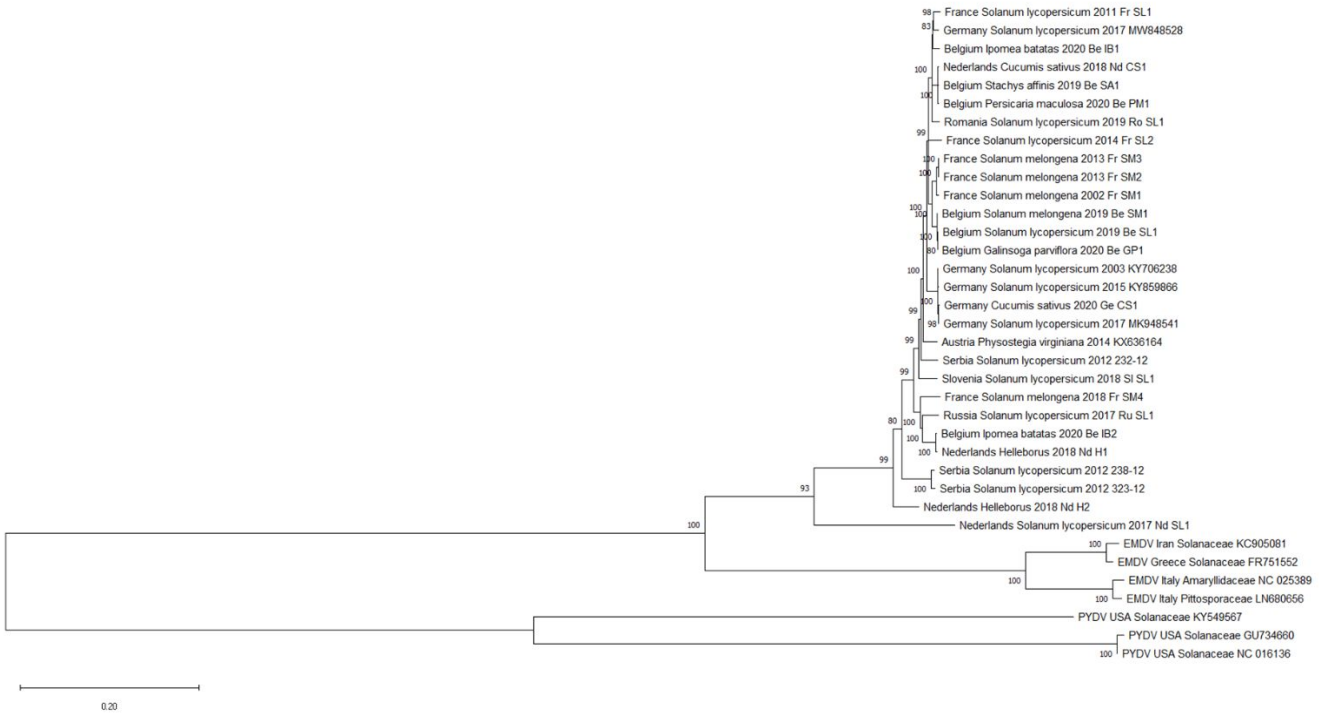
Supplementary Figure 1:



Pictures of natural PhCMoV infected plants

The pictures represent PhCMoV infected *solanum lycopersicum* fruits [Be_SL1 (a), *cucumis sativus* leaves [Ge_CS1 (b), Nd_CS1 (c)] and *helleborus* leaves [Nd_H2 (d,e)].

Supplementary Figure 2:



Phylogenetic tree inferring relationships between 29 physostegia chlorotic mottle virus isolates (among which 21 new genomes published in this study), 4 eggplant mottle dwarf virus isolates and 3 potato yellow dwarf virus isolates based on nucleotides alignment of near complete genomic sequences.

The evolutionary history was inferred by using the Maximum Likelihood method and GTR+G+I model based on the full genome sequence MUSCLE alignment (nucleotides) of all the PhCMoV, EMDV and PYDV isolates known at this date. Each isolate is labelled with its name, the information of the collection (country, host, and year) and with the NCBI accession number if the sequence has been found on the NCBI database. The values on the branches show the percentage of support out of 1000 bootstrap replications, and the scale bar indicates the number of nucleotides substitutions per site.

Supplementary Figure 3:

	Nd_SL1	Nd_H2	238-12	323-12	Sl_SL1	232-12	KX636164	Be_IB1	Fr_SL2	Ge_CS1	MK948541	KY706238	KY859866	Be_SL1	Be_SM1	Be_GP1	Fr_SM1	Fr_SM3	Fr_SM2	Ro_SL1	Be_PM1	Be_SA1	Nd_CS1	Fr_SL1	MW848528	Fr_SM4	Ru_SL1	Be_IB2	Nd_H1
Nd_SL1		82	82	82	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	82	82
Nd_H2	82		94	93	93	93	93	93	93	93	93	93	93	94	94	94	93	93	93	93	93	93	93	93	93	93	93	93	93
238-12	82	94		99	94	94	94	93	93	93	93	94	94	94	94	94	94	94	94	93	94	94	94	93	94	93	93	93	93
323-12	82	93	99		94	94	94	93	93	93	93	93	93	94	94	94	93	93	93	93	93	93	93	93	93	93	93	93	93
Sl_SL1	81	93	94	94		96	96	95	96	96	96	96	96	96	96	96	96	96	96	96	96	96	96	96	96	95	95	95	95
232-12	81	93	94	94	96		97	96	96	96	96	96	96	97	97	97	96	97	97	96	96	96	96	96	97	95	95	95	95
KX636164	81	93	94	94	96	97		96	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	95	95	95	95
Be_IB1	81	93	93	93	95	96	96		96	97	97	97	97	97	97	97	97	97	97	98	97	98	98	98	98	94	94	94	94
Fr_SL2	81	93	93	93	96	96	97	96		97	97	97	97	98	98	98	97	97	97	97	97	97	97	97	97	95	95	95	95
Ge_CS1	81	93	93	93	96	96	97	97	97		100	100	100	98	98	98	97	97	97	97	97	97	97	97	97	95	95	95	95
MK948541	81	93	93	93	96	96	97	97	97	100		100	100	98	98	98	97	98	98	97	97	97	97	97	97	95	95	95	95
KY706238	81	93	94	93	96	96	97	97	97	100	100		100	98	98	98	98	98	98	97	97	97	97	97	97	95	95	95	95
KY859866	81	93	94	93	96	96	97	97	97	100	100	100		98	98	98	97	97	97	97	97	97	97	97	97	95	95	95	95
Be_SL1	81	94	94	94	96	97	97	97	98	98	98	98	98		100	100	99	99	99	98	98	98	98	98	98	95	95	95	95
Be_SM1	81	94	94	94	96	97	97	97	98	98	98	98	98	100		100	99	99	99	98	98	98	98	98	98	95	95	95	95
Be_GP1	81	94	94	94	96	97	97	97	98	98	98	98	98	100	100		99	99	99	98	98	98	98	98	98	95	95	95	95
Fr_SM1	81	93	94	93	96	96	97	97	97	97	97	98	97	99	99	99		99	99	98	98	98	98	98	98	95	95	95	95
Fr_SM3	81	93	94	93	96	97	97	97	97	97	98	98	97	99	99	99	99		100	98	98	98	98	98	98	95	95	95	95
Fr_SM2	81	93	94	93	96	97	97	97	97	97	98	98	97	99	99	99	99	100		98	98	98	98	98	98	95	95	95	95
Ro_SL1	81	93	93	93	96	96	97	98	97	97	97	97	97	98	98	98	98	98	98		98	98	98	98	98	95	95	95	95
Be_PM1	81	93	94	93	96	96	97	97	97	97	97	97	97	98	98	98	98	98	98	98		100	100	98	98	95	95	95	95
Be_SA1	81	93	94	93	96	96	97	98	97	97	97	97	97	98	98	98	98	98	98	98	100		100	98	98	95	95	95	95
Nd_CS1	81	93	94	93	96	96	97	98	97	97	97	97	97	98	98	98	98	98	98	98	100	100		98	98	95	95	95	95
Fr_SL1	81	93	93	93	96	96	97	98	97	97	97	97	97	98	98	98	98	98	98	98	98	98	98		99	95	95	95	95
MW848528	81	93	94	93	96	97	97	98	97	97	97	97	97	98	98	98	98	98	98	98	98	98	98	99		95	95	95	95
Fr_SM4	81	93	93	93	95	95	95	94	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95		96	96	96
Ru_SL1	81	93	93	93	95	95	95	94	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95	96		97	97
Be_IB2	82	93	93	93	95	95	95	94	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95	96	97		100
Nd_H1	82	93	93	93	95	95	95	94	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95	96	97	100	

Percentage identities (nt) for full sequences from the 29 PhCMoV isolates

Supplementary Table 1:

Isolate name	Total number of reads (R1 + R2 non dedupliqués)	Number of reads mapping PhCMoV genomes	Coverage*	PhCMoV genome on which we mapped the read	Sequencer
Be_SM1	6,8 M	1,1 M	12 387	KX636164	Illumina Miseq
Fr_SM4	24 M	2,6 M	29 279	KX636164	Illumina Miseq
Fr_SL1	22 M	7,7	86 712	KX636164	Illumina Miseq
Fr_SM2	23 M	5,2 M	58 559	KX636164	Illumina Miseq
Fr_SM3	24 M	3,1 M	34 910	KX636164	Illumina Miseq
Fr_SL2	23 M	4 M	4 505	KX636164	Illumina Miseq
Be_SL1	0,2 M	0,1 M	1 126	KX636164	Illumina Miseq
Be_GP1	22 M	1,8 M	20 270	KX636164	Illumina Miseq
Nd_SL1	15 M	1,6 M	18 018	KX636164	Illumina Miseq
Ru_SL1	19 M	1,1 M	12 387	KX636164	Illumina Miseq
Nd_CS1	6,9 M	1,7 M	19 144	KX636164	Illumina Miseq
Nd_H1	7,5 M	1,8 M	20 270	KX636164	Illumina Miseq
Nd_H2	20 M	2,1 M	23 649	KX636164	Illumina Miseq
Ro_SL1	15 M	0,4 M	4 505	KX636164	Illumina Miseq

Fr_SM1	35 M	0,8 M	9 009	KX636164	Illumina Miseq
Be_IB1	24,7M	8385	100	KY706238	Illumina Miseq
Be_IB2	23,4M	813	10	KY706238	Illumina Miseq
Be_SA1	27,7M	31 544	397	KY706238	Illumina Miseq
Be_PM1	36,6M	34 374	447	KY706238	Illumina Miseq
Sl_SL1	15,2M	0,1M	1 126	KY706238	Illumina Miseq
Ge_CS1	1,1 M	9804	110	MK948541	Oxford Nanopore Technologies MinIon

Number of sequenced reads and horizontal coverage

*Coverage = (number of reads which map on the PhCMoV genome x lenght of one read) / the number of bases in the reference genome

References:

Alfaro-Fernández, A., Córdoba-Sellés, C., Tornos, T., Cebrián, M. C., and Font, M. I. 2011. First Report of Eggplant mottled dwarf virus in Pittosporum tobira in Spain. Plant Disease. 95:75–75.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. J Comput Biol. 19:455–477.

Boben, J., Kramberger, P., Petrovič, N., Cankar, K., Peterka, M., Štrancar, A., et al. 2007. Detection and quantification of Tomato mosaic virus in irrigation waters. Eur J Plant Pathol. 118:59–71.

Botermans, M., van de Vossenberg, B. T. L. H., Verhoeven, J. Th. J., Roenhorst, J. W., Hooftman, M., Dekter, R., et al. 2013. Development and validation of a real-time RT-PCR assay for generic detection of pospiviroids. Journal of Virological Methods. 187:43–50.

Buchfink, B., Xie, C. & Huson, D. 2015. Fast and sensitive protein alignment using DIAMOND. Nat Methods 12: 59–60.

Filloux, D., Dallot, S., Delaunay, A., Galzi, S., Jacquot, E., and Roumagnac, P. 2015. Metagenomics Approaches Based on Virion-Associated Nucleic Acids (VANA): An Innovative Tool for Assessing Without A Priori Viral Diversity of Plants. Methods Mol Biol. 1302:249–257.

Gaafar, Y. Z. A., Abdelgalil, M. a. M., Knierim, D., Richert-Pöggeler, K. R., Menzel, W., Winter, S., et al. 2018. First Report of physostegia chlorotic mottle virus on Tomato (Solanum lycopersicum) in Germany. Plant Disease. 102:255–255.

Gaafar, Y., Lüddecke, P., Heidler, C., Hartrick, J., Sieg-Müller, A., Hübert, C., et al. 2019. First report of Southern tomato virus in German tomatoes. New Disease Reports. 40:1–1.

Kogovšek, P., Gow, L., Pompe-Novak, M., Gruden, K., Foster, G. D., Boonham, N., et al. 2008. Single-step RT real-time PCR for sensitive detection and discrimination of Potato virus Y isolates. J Virol Methods. 149:1–11.

Kopylova, E., Noé, L., and Touzet, H. 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics. 28:3211–3217.

Li, H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 34:3094–3100.

Maclot, F. J., Debue, V., Blouin, A. G., Fontdevila Pareta, N., Tamisier, L., Filloux, D., et al. 2021. Identification, molecular and biological characterization of two novel secovirids in wild grass species in Belgium. Virus Research. 298:198397.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet Journal. 17:10–12.

Ondov, B. D., Bergman, N. H., and Phillippy, A. M. 2011. Interactive metagenomic visualization in a Web browser. BMC Bioinformatics. 12:385.

Palanga, E., Filloux, D., Martin, D. P., Fernandez, E., Gargani, D., Ferdinand, R., et al. 2016. Metagenomic-Based Screening and Molecular Characterization of Cowpea-Infecting Viruses in Burkina Faso. PLOS ONE. 11:e0165188.

Pecman, A., Kutnjak, D., Gutiérrez-Aguirre, I., Adams, I., Fox, A., Boonham, N., et al. 2017. Next Generation Sequencing for Detection and Discovery of Plant Viruses and Viroids: Comparison of Two Approaches. Front. Microbiol. 8: 1998.

Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. 2014. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics. 30:614–620.

Zheng, Y., Gao, S., Padmanabhan, C., Li, R., Galvez, M., Gutierrez, D., et al. 2017. VirusDetect: An automated pipeline for efficient virus discovery using deep sequencing of small RNAs. Virology. 500:130–138.