Molecular and biological characterization of novel and known Sequiviridae members infecting lettuce

Laurence Svanella-Dumas, Iason Tsarmpopoulos, Armelle Marais, Chantal Faure, Sebastien Theil, Miroslav Glasa, Lukas Predajna, Jonathan Gaudin, Sixing Tian, Laetitia Porcher, et al.

Molecular and biological characterization of novel and known *Sequiviridae* members infecting lettuce

Laurence Svanella-Dumas[1], Iason Tsarmpopoulos[1], Armelle Marais[1], Chantal Faure[1], Sébastien Theil[1*], Miroslav Glasa[2], Lukas Predajna[2], Jonathan Gaudin[3], Sixing Tian[4], Laëtitia Porcher[4], Pascal Gentit[4], Milena Leite De Oliveira[5], Renate Krause-Sakate[6], and Thierry Candresse[1+]

[1]INRAE, Univ. Bordeaux, UMR Biologie du fruit et Pathologie, CS20032, 33882 Villenave d'Ornon Cedex, France

[2]Biomedical Research Center of the Slovak Academy of Sciences, Institute of Virology, Dúbravská cesta 9, 84505 Bratislava, Slovak Republic

[3]INRAE, Bordeaux Sciences Agro, UMR Santé et Agroécologie du Vignoble, CS20032, 33882 Villenave d'Ornon Cedex, France

[4]ANSES, Plant Health Laboratory, Unité de Bactériologie, Virologie et détection des OGM, 7 rue Jean Dixméras, 49044 Angers Cedex 01, France

[5]Laboratorio Central Multisuarios (LACEM), Faculdade de Ciências Agronômicas, School of Agriculture, Sao Paulo State University (UNESP), Botucatu, Brazil

[6]Department of Plant Protection, Faculdade de Ciências Agronômicas, School of Agriculture, Sao Paulo State University (UNESP), Botucatu, Brazil

*Present address: UMR545 Fromage, INRAE, VetAgro Sup, Université Clermont Auvergne, 20 Côte de Reyne, 15000 Aurillac, France

[+]Corresponding author: T. Candresse; thierry.candresse@inrae.fr

Running title: characterization of lettuce-infecting *Sequiviridae*

Abstract: 173 words
Text (Introduction + M&M + Results + Discussion + Acknowledgements): 6265 words
3 Tables + 3 Figures

The nucleotide sequences reported here have been deposited in Genbank under accession numbers MT348705-10, MT559283-84 and OQ446443-47

ABSTRACT

High-throughput sequencing of two lettuces showing virus-like symptoms in France provided evidence of infection by members of the *Secoviridae* family. One plant (JG1) had a complex mixed infection that involved, among others, a novel *Waikavirus* (lettuce waikavirus 1) and two isolates of a sequivirus related to lettuce mottle virus (LeMoV). The second lettuce plant (JG2) was singly infected by LeMoV. Complete genomic sequences were obtained for all four isolates and, in addition, near complete genome sequences were obtained for other LeMoV or LeMoV-related isolates (from French cultivated and wild lettuces, and from a Brazilian cultivated lettuce) and for two isolates of another *Asteraceae*-infecting sequivirus, dandelion yellow mosaic virus (DaYMV). Analysis of these genomic sequences allows to propose tentative genome organization for the various viruses and to clarify their phylogenetic relationships. Sequence and host range comparisons point to significant differences between the two sequivirus isolates identified in the JG1 plant and LeMoV isolates from France and Brazil, suggesting they belong to a novel species for which the name lettuce star mosaic virus is proposed.

Keywords: Lettuce, France, Brazil, lettuce mottle virus, High-throughput sequencing, etiology.

INTRODUCTION

Lettuce (*Lactuca sativa* L.) is the most important annual leafy vegetable crop. It is primarily consumed fresh and is considered an important dietary vegetable worldwide. It belongs to the Asteraceae, the largest family of flowering plants. Cultivated lettuce was first documented on the walls of Egyptian tombs 4500 years ago and was soon cultivated in most parts of Europe and, later, in other parts of the world (De Vries 1997). The main producer of lettuce is Asia, especially India and China which, according to FAO in 2019, ranks first in annual production, followed by North and Central America, and Europe. In the European Union, Spain, Italy, Germany and France are the major lettuce-producing countries, with 74% of the total European lettuce production at almost 1.75 million tons in 2017 (Le Fustec 2020). In France, lettuce is grown year around in open fields, under plastic houses or glasshouses, with a total acreage of 8,300 hectares corresponding to about 500 million plants annually (Le Fustec 2020).

A number of pests and pathogens affect lettuce crops, including viruses belonging to the genera *Alfamovirus*, *Crinivirus*, *Cucumovirus*, *Cytorhabdovirus*, *Fabavirus, Ophiovirus*, *Polerovirus*, *Potyvirus*, *Sequivirus*, *Tombuvirus*, *Orthotospovirus* and *Varicosavirus* (Ciuffo et al. 2016; Ephytia 2019; Jadao et al. 2007; Lebeda et al. 2014; Moreno et al. 2012; Pavan et al. 2008). Many of these viruses cause epidemic diseases and considerable yield losses. Thanks to developments in high throughput sequencing (HTS) combined with specific bioinformatic tools (Lefebvre et al. 2019; Villamor et al. 2019), numerous plant viruses have been recently characterized, allowing major advances in the study of the etiology of viral diseases in perennial or annual crops (Wu et al. 2015; Maliogka et al. 2018; Maree et al. 2018). However, in many cases, past the identification and genome characterization of novel viral agents, the challenge is to demonstrate their potential implication in diseases and to

provide information about their biology, allowing to evaluate the risks these novel agents may pose (Massart et al. 2017).

In recent years, several novel viruses have been characterized in lettuce crops (Ciuffo et al. 2016; Svanella-Dumas et al. 2018, Verbeek et al. 2014), while the main viruses infecting lettuce in Europe or in Central America have been respectively reviewed in Moreno et al. (2012) and Pavan et al. (2008). Lettuce mosaic virus (LMV, *Potyvirus*) is probably the most important virus of lettuce worldwide (Pavan et al. 2008). LMV causes typical mosaic symptoms that may occasionally turn necrotic. However, some other lettuce viruses can cause almost indistinguishable mosaic symptoms, such as lettuce mottle virus (LeMoV) and dandelion yellow mosaic virus (DaYMV), both of them from the genus *Sequivirus* in the family *Secoviridae* (Jadao et al. 2007). DaYMV was the first sequivirus reported in lettuce and in plants of dandelion (*Taraxacum officinale*) in different European countries (Kassanis 1944, 1947; Bos et al. 1983; Jadao et al. 2007). LeMoV, differing from DaYMV in host range, is a distinct sequivirus infecting lettuce in Brazil and in Chile (Krause-Sakate et al. 2005; Jadao et al. 2007). In Brazil, and in particular in Sao Paulo State, LeMoV was in the early 2000s the most frequent virus in lettuce crops (De Marchi et al. 2012, Krause-Sakate et al. 2007) but it has so far never been reported outside of South America. For both DaYMV and LeMoV, only very short, partial genome sequences are available (Jadao et al. 2007) and there is very little information available on their genetic variability.

We report here the biological and molecular characterization by HTS of several secoviruses from wild and cultivated lettuce in France and Brazil and from dandelion in Slovakia. Besides providing the first complete genomic sequence for LeMoV and near-complete genome sequences for DaYMV, the results presented here describe a novel *Sequivirus* and a novel *Waikavirus* detected in French cultivated lettuces.

MATERIALS AND METHODS

Plant samples and virus isolates

To identify the etiological agent(s) involved, three cultivated lettuces (*Lactuca sativa* L.) presenting systemic mosaic symptoms with some necrotic points were analyzed by HTS. The first one (cv. Shangore, JG1) was collected in 2014 near Sainte-Livrade (southwestern France) and the second one (cv. Guétary, JG2) came from an organic plot under a plastic greenhouse and was collected in 2015 near Angers (western France). The last (21-015C, variety unknown) was collected in southern France in 2022. The mosaic symptoms on the JG2 plant, shown in Fig 1A, had more pronounced necrosis than those on the JG1 plant, while besides mosaic, the 21-015C plant showed stunting and leaf deformation.

Several isolates of lettuce mottle virus (LeMoV) were analyzed. The P22 isolate was obtained from an asymptomatic wild lettuce plant (*L. serriola*) collected in spring 2010 as part of a plant virus metagenomic survey in Villenave d'Ornon (southwestern France) (Svanella-Dumas et al. 2018). The second isolate (SP) was from a cultivated lettuce collected in 2013 in Piedade County, Sao Paulo State, Brazil. This isolate was propagated in *Chenopodium quinoa* or lettuce (cv. Trocadéro) plants by sap inoculation before HTS analysis. The third LeMoV isolate, LIB3, was from a wild lettuce (*Lactuca saligna*) showing leaf yellowing and discolorations collected in 2020 in Villenave d'Ornon.

Two isolates of dandelion yellow mosaic virus (DaYMV) in mixed infection, RNA24P-1 and RNA24P-2, were identified by HTS from a dandelion plant collected in 2018 at the border of a cultivated field near Bratislava, Slovakia.

Host range study and sap inoculation of viruses

In order to characterize the novel viruses identified, leaf pieces from the original JG1 plant were ground 1:3 (wt/vol) in 0.05M potassium phosphate buffer pH 8.0. Activated charcoal

and carborundum (400 mesh) were added (100mg/ml homogenate) and the mixture was used for sap-inoculation of plants of *L. sativa* (cv. Trocadero) or of ten plant species belonging to the families *Asteraceae*, *Solanaceae*, *Chenopodiaceae* and *Amaranthaceae*. The LeMoV isolate identified in the JG2 plant was sap-transmitted to lettuce (cv. Trocadéro), *Nicotiana benthamiana* and *Gomphrena globosa* plants using the buffer described in Jadao et al. (2007). The host range study was repeated at least twice for each viral isolate. All plants were kept under greenhouse conditions and monitored for symptoms development up to one-month post-inoculation.

The presence of the viruses in upper non-inoculated leaves of the test plants was evaluated using specific RT-PCR assays for each virus using total nucleic acids (TNA) extracted from fresh leaves according to protocol 1 of Foissac et al. (2005). The sequences of the various amplification primers used [New waika-F / New waika-R (for the new waikavirus) and NSeco-varA-F / NSeco-varA-R or NSeco-varB-F / NSeco-varB-R (that discriminate between the two variants of the new sequivirus] are provided in Supplementary Table S1. These primer pairs target fragments of respectively 385, 445 or 546 nucleotides (nt) in the viral RNA-dependent RNA polymerase (RdRp) gene of each virus. The presence of LeMoV was similarly checked using primer pair LeMoV3 (Jadao et al. 2007) and LeMoV4-multi (Supplementary Table S1) which is adapted from Jadao et al. (2007) so as to improve inclusiveness and to allow amplification of isolates JG2, P22, SP described here and of previously analyzed Brazilian LeMoV isolates (DQ675190-91). This primer pair amplifies a 667 nt fragment of the RdRp gene.

For each RT-PCR assay, cDNA was first synthesized from 5µl of TNA using dT18 and pdN6 primers and the Revertaid Reverse Transcriptase (200U/µl) according to the manufacturer recommendations (Thermo Fisher Scientific, Illkirch, France). In a second step, the PCR was carried out using 3 µl of cDNA in a 25-µl reaction volume using 1 U of DyNAzymeII™ DNA

polymerase in accordance with the manufacturer recommendations (Thermo Fisher Scientific). The annealing temperature and duration of the 72°C extension step are provided in Supplementary Table S1. PCR products were analyzed by non-denaturing 1% agarose gel electrophoresis with ethidium bromide staining and, if needed, directly sequenced (Eurofins Genomics, Ebersberg bei München, Germany) to verify the specificity of the amplification reaction.

HTS analysis, genome assembly and completion of the genome of the various viral isolates

Double-stranded RNAs (dsRNAs) were purified from fresh symptomatic leaves of the original JG1, JG2, P22 and LIB3 plants as described (Marais et al. 2018). Purified dsRNAs were submited to a random amplification (Marais et al. 2018) and sequenced on Illumina platforms (2×250 or 2x150 nt). Total RNAs were purified from *N. benthamiana* leaves infected by the LeMoV-JG2 isolate using the protocol of Chang et al. (1983) and sequenced without ribodepletion an Illumina HiSeq2000 (2×150 nt). Additionally, total RNAs were extracted from fully developed dandelion leaves as mentioned in Tomasechova et al. (2020) and sequenced on an Illumina MiSeq (2×250 nt). Total RNAs extracted from the 21-015C lettuce were similarly sequenced on the Illumina iSeq100 and MiSeq platforms (2x150nt).

Following demultiplexing and quality trimming, HTS reads from the various samples were analyzed using CLC Genomics Workbench 8 and later versions (https://www.qiagenbioinformatics.com) and the VirAnnot pipeline (Lefebvre et al. 2019). *De novo* assembled contigs were annotated by BlastN and BlastX analyses against the GenBank database using a 10^{-3} *e*-value cut-off. Viral contigs were further extended by several rounds of mapping of residual reads and by integration of smaller contigs to yield scaffold covering most of the viral genomes.

The 5′ end of viral genomes was determined using a 5' RACE kit (Takara Bio Europe/Clontech, Saint-Germain-en-Laye, France) and specific reverse primers designed on the reconstructed scaffold of each virus (Supplementary Table S1). To complete the 3′ end of the genomes of the sequiviruses, that do not contain a poly-A tail, TNA were extracted and polyadenylated (Ambion/ThermoFisher Scientific, Illkirch, France). The polyadenylated RNAs were then submitted to a reverse transcription using LD-polyT as a reverse primer and the genome 3′ end was then amplified by PCR using a forward-specific internal primer designed from the genome scaffold sequenced contig (Supplementary Table S1) and the LD-prime primer, following the protocol described by Youssef et al. (2011). The same procedure was used to determine the *Waikavirus* 3' genome end, only omitting the polyadenylation step since waikavirus genomes are polyadenylated.

Internal gaps and regions of low coverage were determined or confirmed by direct sequencing of RT-PCR fragments obtained using internal primers designed from the sequences of contigs. For LeMoV-JG2, the only gap identified was in the 3' part of the genome, in a C-rich region with low coverage. Two combinations of primers (LeMoV-3Race1 / LeMoV-gap-R1 and LeMoV-3Race1 / LeMoV-gap-R2) (Supplementary Table S1), amplifying PCR products of respectively 199 and 573 nt were used to confirm the genome sequence in this short region. The nearly genomic sequence of the LeMoV isolate P22 was completed after designing six pairs of primers (Supplementary Table S1) allowing to close six small internal gaps in the HTS scaffold.

All amplified fragments were visualized on non-denaturing agarose gels and directly submitted to Sanger sequencing on both strands (Eurofins Genomics) or, when needed, sequenced after cloning in the pGEM-T Easy vector (Promega, Charbonnières-Les Bains, France).

The nearly complete genome of the LeMoV isolate SP was obtained by RT-PCR-based genome walking, using total RNAs extracted from infected *C. quinoa* leaves with the Total RNA Purification Kit (Norgen Biotek, Thorold, Canada). The primers used (Supplementary Table S1) were designed based on the genome sequence of the LeMoV P22 isolate. The one-step RT-PCR reactions were performed using 2.5 µl of RNA, 1 µM of each primer and 0.03 units of Avian myeloblastosis virus reverse transcriptase (AMV RT, Promega, Charbonnières-les-Bains, France) in a final volume of 25 µl of 1x PCR master mix (Fermentas, Vilnius, Lithuania). The annealing temperature for each primer pair is provided in Supplementary Table S1. PCR products were analyzed as described above, purified from agarose gels using the QIAquick Gel Extraction Kit (Qiagen, Courtaboeuf, France) and sequenced using an ABI 3730 DNA Analyser (Applied Biosystems, Illkirch, France) using a BigDye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems).

Sequence comparisons and phylogenetic analyses

Multiple nucleotide (nt) or deduced amino acid (aa) sequence alignments were performed using ClustalW (Thompson et al. 1994) implemented in MEGA 7 (Kumar et al. 2016). Strict identity nt or aa distances were calculated from these alignments and neighbor-joining phylogenetic trees were reconstructed using Mega 7 with bootstrap analysis (500 replicates) to evaluate branch validity.

RESULTS

Viruses detected by HTS in the French JG1 and JG2 lettuces

During the spring of 2014 and 2015, two symptomatic lettuces (JG1, cv. Shangore and JG2 cv. Guétary) showing systemic mosaic symptoms were observed during surveys. Both varieties are resistant to lettuce downy mildew (*Bremia lactucae* Regel) but only JG2 is

187 resistant to lettuce mosaic virus (LMV) and to the lettuce aphid, *Nasonovia ribisnigri*. The
188 symptoms on the JG2 lettuce included some necrotic points (Fig. 1A) that were much less
189 present on the JG1 plant. Symptomatic leaves of both plants were used to mechanically
190 inoculate lettuce plants (cv. Trocadéro), which developed systemic mosaic symptoms,
191 confirming suspicions of viral infection. In order to identify the virus(es) involved, dsRNAs
192 were purified from 0.75g of fresh leaves from the original JG1 and JG2 plants and analyzed
193 by HTS. Following demultiplexing and quality trimming, the reads were assembled into
194 contigs that were annotated by BlastN and BlastX analyses, providing evidence for a mixed
195 infection by several agents for the JG1 lettuce and for a single infection in JG2 (Table 1).
196 For JG1, the contig integrating the most reads (26,980) and showing the highest average
197 coverage (449x), showed the highest BlastN identity (49% nt identity) with bellflower vein
198 chlorosis virus (BVCV, Seo et al. 2015), a recently described member of the genus
199 *Waikavirus*. A second contig, integrating 12,327 reads for a 283x average coverage, showed
200 the closest homology (54% nt identity) with carrot chordovirus 1 (Adams et al. 2014). The
201 complete genome of this viral isolate has since been completed, allowing the description of a
202 new viral species, lettuce chordovirus 1 (LeCV1, Svanella-Dumas et al. 2018) in the family
203 *Betaflexiviridae*. The last two contigs integrating respectively 4,198 and 5,227 reads (average
204 coverage of respectively 92x and 102x) showed highest BlastN nt identities of 54% with
205 parsnip yellow fleck virus (PYFV, Turnbull-Ross et al. 1993) and 47% with carrot necrotic
206 dieback virus (CNDV, Menzel et al. 2008), two members of the *Sequivirus* genus,
207 respectively. Following the extension of these two contigs by several rounds of reads mapping
208 and aggregation of additional smaller contigs, two scaffolds of respectively 8,961 nt and
209 9,158 nt, were reconstructed, showing 79.8% nt identity. Taken together these results showed
210 that the original JG1 lettuce had a mixed infection involving LeCV1, a tentatively novel
211 *Waikavirus* and two related and tentatively novel sequiviruses.

For JG2, the largest viral contig identified from dsRNA-derived reads, was 1,996-nt long (96x average coverage, 891 reads) and showed 50% nt identity with CNDV (data not shown). Given the very incomplete assembly of the viral genome, a second HTS run was performed on this viral isolate using total RNAs purified from infected N. *benthamiana* leaves. Using this second dataset, evidence for a single viral infection was obtained. A single long viral contig was obtained and further extended by rounds of reads mapping into an 8,927-nt long scaffold integrating 119,029 viral reads (1,790x average coverage) and sharing 50% nt identity with CNDV (Table 1). Interestingly, this scaffold shared 85.4 and 86.3% nt identity, respectively (98.4 and 97.9% aa identify, respectively) in a 562 nt region shared with the two partial LeMoV sequences available (DQ675190 and DQ675191), suggesting that the viral isolate identified in the JG2 plant might represent the first identification of LeMoV in cultivated lettuce outside of south America.

Genome organization of the *Waikavirus* identified in the JG1 lettuce

Following further rounds of reads assembly and aggregation of shorter contigs, a nearly complete genome scaffold of 11,852 nt was assembled for the tentative *Waikavirus* identified in the JG1 plant. The missing terminal genome ends were determined by 5'-RACE and 3' LD-polyT / LD-prime PCR (Youssef et al. 2011) using the purified dsRNAs from the original plant as a template and specific primers designed from the scaffold sequence (Supplementary Table S1). The completed genome is 11,929 nt in length, excluding the 3'poly (A) tail (Fig. 2A) and has been deposited in GenBank under the MT348710 accession number. The genome organization is similar to that of other waikaviruses, with a single large open reading frame (ORF) of 10,389 nt encoding a polyprotein of 3,463 aa (388.5 kDa). The 5' and 3' untranslated regions (UTRs) are respectively 463 nt and 1,078 nt long. The presence of additional potential ORFs has been reported for some waikaviruses (Shen et al. 1993; Firth et al. 2008). The predicted short overlapping ORFX was identified starting at genome position

464, encoding an 88 aa protein sharing 73.3% identity with the corresponding protein of BVCV and variable conservation levels with other genus members (Supplementary Figure 1). On the other hand, the long 3' UTR of the JG1 waikavirus does not contain the small ORF2 and ORF3 described in rice tungro spherical virus (RTSV) by Thole et al. (1996). In waikaviruses and sequiviruses, the single large polyprotein encodes functional proteins released by post-translational cleavages mediated by the viral protease (Pro). From the N-terminal polyprotein end, these are a P1 protein of unknown function, the three coat protein subunits (CP1, CP2, and CP3), an NTP-binding helicase (NTB), the 3C-like protease and an RdRp. The polyprotein cleavage sites have been experimentally identified for maize chlorotic dwarf virus (MCDV) and RTSV (Reddick et al. 1997; Shen et al. 1993). They are relatively conserved between genus members and generally involve a glutamine (Q) residue at the -1 position (Sekiguchi et al. 2005; Sanfaçon et al. 2011). Considering the homologies between the lettuce virus polyprotein and those of other waikaviruses, it was possible to propose tentative cleavage sites for its polyprotein using a ClustalW-generated multiple alignment of the various polyproteins (Fig. 2 and Table 2). A comparison of the amino acid sequence at each of the proteolytic cleavage sites of the lettuce virus (hereafter referred to as LWV1-JG1) showed that glutamine was conserved at the -1 position, with different amino acids in the +1 position (Q/S, Q/T and Q/D) depending on the cleavage site. Comparisons between the LWV1-JG1 polyprotein sequence and those of other waikaviruses confirmed the presence of the expected conserved domains in the Hel-Pro-Pol module that is commonly used to characterize picorna-like viruses (Le Gall et al. 2008). This includes the A "GAPGVGKS" (aa 1806 to 1813), B "DD" (aa 1857-1858) and C "KGKHCTSKYVFSCTN" (aa 1893 to 1907) helicase motifs (Le Gall et al. 2008), the 3C-like protease catalytic triad of histidine, aspartate and cysteine (Reddick et al. 1997) present at amino acids 2670, 2707 and 2801 and the A "DYAKFDGIGSP" (aa 3149 to 3159), B "SGFSMTVIFNS" (aa 3210 to 3220) and C

262 "YGDD" (aa 3258 to 3261) motifs (Reddick et al. 1997; Le Gall et al. 2008) characterizing
263 the RdRp domain.

264 **Genome organization of the sequivirus isolates identified in the JG1 lettuce**

265 As indicated above, two long scaffolds representing sequivirus isolates could be reconstructed
266 for the JG1 plants, named here JG1-A (8,961 nt) and JG1-B (9,158 nt). The whole genome of
267 the JG1-B was completed by 5'-RACE and 3' LD-polyT / LD-prime PCR as described above
268 for the *Waikavirus* sequence, with the only exception that the genome 3' end was determined
269 after polyadenylation since sequiviruses do not possess a 3' polyA. No specific efforts were
270 made to complete the JG1-A genome sequence. The JG1-B complete genome is 10,183 nt
271 long and has been deposited in GenBank, together with the JG1-A contig sequence
272 (MT348706 and MT348705, respectively). As expected, the JG1-B genome harbors a single
273 large ORF of 9,054 nt encoding a predicted polyprotein of 3,018 aa (337.5 kDa), bordered by
274 5' and 3' UTRs of 150 nt and 980 nt, respectively (Fig. 2B). Conserved protein motifs were
275 identified at the expected locations in the JG1-B polyprotein. This includes the A
276 "GSPGVGKS" (aa 1430 to 1437), B "DD" (aa 1481 to 1482) and C "KGRTFSSKYIFSTTN"
277 (aa 1517 to 1531) helicase motifs, or the B "SGFPMTVIFNS" (aa 2555 to 2565) and C
278 "YGDD" (aa 2603 to 2606) RdRp motifs. However, polyprotein cleavage sites are much less
279 conserved between members of the *Sequivirus* genus (Sanfaçon et al. 2011), so that it was not
280 possible to propose candidates for all tentative cleavage sites of the JG1-B polyprotein (Fig.
281 2B and Table 2) from a comparison with those reported for PYFV (Turnbull-Ross et al.
282 1993). Nevertheless, four candidate sites liberating the P1 protein and the three CP subunits
283 could be proposed (N/P, Q/G, Q/G and Q/A) at aa position 383/384, 575/576, 799/800 and
284 1035/1036 of the polyprotein, respectively (Fig. 2B and Table 2). With its 79.8% overall nt
285 identity with the JG1-B sequence, the partial JG1-A isolate sequence revealed complete
286 conservation of the genome features outlined above.

Genome organization of the LeMoV isolate identified in the JG2 lettuce

The genome ends of the 8,927 nt contig for the LeMoV isolate assembled from the sequence data generated for the JG2 lettuce were determined by 5'-RACE and 3' LD-polyT / LD-prime PCR. As for the JG1-B sequivirus, it was necessary to add a poly-A tail at the 3' genome end. The completed JG2 LeMoV genome has been deposited in GenBank (MT348707) and is 9,927 nt long and has 5' and 3' UTRs of respectively 158 and 704 nt (Fig. 2C). It harbors a single large ORF of 9,066 nt encoding a predicted polyprotein of 3,022 aa (336.1 kDa), which has all the expected conserved protein motifs at the expected positions. Indeed, all aa sequences of motifs are similar to those identified for both isolates JG1-A and JG1-B except for the motif C of the helicase region for which a serine replaces an arginine at the 6th position. As for the JG1-B sequence, it was only possible to predict the four tentative cleavage sites liberating the P1 protein and the three CP subunits (Fig. 2C and Table 2). Two out of four predicted cleavage sites Q/G (577/578, between CP1 and CP2) and Q/A (1037/103, between CP3 and Helicase) are identical to those predicted for JG1-B.

Genetic diversity of JG1-A/JG1-B-like isolates and of lettuce mottle virus (LeMoV)

Two nearly complete *Sequivirus* genome sequences related to the JG1-A/JG1-B isolates were assembled from the reads obtained from a symptomatic cultivated lettuce (21-015C, cv. unknown) collected in southern France in 2022. These two contigs are respectively 10107 nt (21-015C-A, OQ446446) and 10066 nt (21-015C-B, OQ446447) and had an average coverage of respectively 460 and 343x (Table 1). No specific effort was made to complete the 5' and 3' genome UTRs. The two sequences show 23.7% nt divergence with each other and respectively 20.8% and 25% divergence with JG1-A. On the other hand, 21-015C-A shows only 8.6% divergence with JG1-B, while the corresponding value for 21-015C-B is 23.9%. Genomic organization and key protein motifs are conserved between these two new isolates

and the JG1 isolates, with only one polymorphism observed at the P1/CP1 tentative cleavage site (Table 2)

Nearly complete genome sequences were obtained for five other LeMoV isolates, P22 assembled from the reads obtained from an asymptomatic wild lettuce plant (*L. serriola*) collected in spring 2010 as part of a plant virus metagenomic survey in Villenave d'Ornon (southwestern France) (Svanella-Dumas et al. 2018), SP assembled from the reads from a cultivated lettuce collected in 2013 in Sao Paulo State, Brazil, 21-015C assembled from the reads of the cultivated lettuce mentioned above and LIB3-A and LIB3-B assembled from the reads from a wild *Lactuca saligna* collected in southwest France in 2020. The P22 contig of 9,012 nt (GenBank MT348708) integrates 2,517 reads (78x average coverage), while a large part of the SP isolate genome (8,366 nt) was determined by primer walking along the genome (GenBank MT348709). The LIB3-A (OQ446443) and LIB3-B (OQ446444) sequences are 9150 nt long while the 21-015C sequence (OQ446445) is 9807 nt long but contains a short 24 nt internal gap at the position of the C rich region in which a similar gap was closed by PCR for the JG2 sequence. No specific efforts were made to determine the missing terminal sequences for any of these five LeMoV sequences. The JG2 genomic sequence shows between 11.1% and 14.4% nt divergence (2.0-3.5% aa divergence) with these 5 novel near complete LeMoV sequences, which in turn show between 11.4% and 15% nucleotide divergence (2.2-3.9% aa divergence) with each other. All conserved protein motifs and predicted cleavage sites are identical for the various LeMoV isolates (Table 2).

Genomic organization of a DaYMV isolate identified in a Slovak dandelion

As described above, two viral contigs with homologies to sequiviruses could be assembled from the reads obtained from the dandelion RNA24P plant from Slovakia. These contigs respectively integrate 3,824 and 1,003 reads, corresponding to average coverages of 60-fold and 15-fold, respectively and showed the highest homology (49% nt identity) with CNDV

(Table 2). Further assembly efforts yielded two large scaffolds named RNA24P-1 and RNA24P-2 of 9,786 and 9,795 nt long, respectively, and sharing 77% and 81.7 % nt sequence identity (95% and 96.2 % aa), respectively, with the only available DaYMV partial sequence (DQ675189, 788 nt long). No specific efforts were made to determine the 5' and 3' terminal regions of the nearly complete RNA24P-1 and RNA24P-2 genomes and both large contigs have been deposited in GenBank (MT559283 and MT559284, respectively). These results constitute the first identification of DaYMV isolates in dandelion in Slovakia. The RNA24P-1 genome organization is similar to that of other sequiviruses, with a single large ORF of 9,102 nt encoding a polyprotein of 3,035 aa (338.7 kDa). The incomplete 5’ and 3’ UTRs are respectively 149 nt and 533 nt long (Fig. 2D), indicating that the amount of missing sequence information is likely limited. The RNA24P-2 genome has an ORF of 9,105 nt encoding a polyprotein which is one aa longer than that of RNA24P-1 in the 3C-like protease domain. All polyprotein conserved amino acid motifs are identical to those identified for LeMoV isolates, except for the helicase A motif in which a lysine replaces a serine at position 2. As for LSMV and LeMoV, it was only possible to propose candidates for the first four polyprotein cleavage sites of DaYMV (K/P, Q/G, Q/M/S and Q/G at aa positions 388/389, 580/581, 804/805 and 1041/1042, respectively; Fig. 2D and Table 2), with an M/S polymorphism between the two isolates at the +1 position of the CP2-CP3 cleavage site.

Sequence comparisons and phylogenetic relationships of the various viruses identified in the analyzed wild and cultivated lettuces

To clarify the taxonomic position of these various viruses identified above, sequence comparisons and phylogenetic analyses were performed either with the taxonomically relevant Pro-Pol region (defined as spanning the part of the polyprotein between the conserved CG protease and GDD polymerase motifs) or with the coat protein subunits for all viruses in the genera *Waikavirus* and *Sequivirus*, as well as for representative members of the

genera *Cheravirus*, *Comovirus*, *Cholivirus, Fabavirus*, *Nepovirus*, *Satsumavirus*, *Stralaravirus, Torradovirus* and for unassigned viruses in the *Secoviridae* family (Sanfaçon et al. 2020) (Fig. 3A and 3B). The trees unambiguously placed the JG1 *Waikavirus* with other genus members with strong bootstrap support and the JG1-A/JG1-B-like and LeMoV sequences within the *Sequivirus* genus, again with strong bootstrap support (Fig. 3A and 3B), confirming the assignation of these various viruses to this genus. In the case of the sequiviruses, only very short partial genome sequences were previously available for LeMoV and DaYMV (Jadao et al. 2007). Integration of these partial sequences in the phylogenetic trees confirm the identification of the JG2 isolate as LeMoV, together with the P22, SP, LIB3-A, LIB3-B and 21-015C isolates, with a tight, bootstrap-supported clustering (Fig. 3A) and shows that DaYMV unambiguously clusters away from the bootstrap-supported cluster formed by the the JG1-A, JG1-B, 21-015C-A and 21-015C-B isolates, ruling out their interpretation as DaYMV isolates (Fig. 3A).

The aa sequence identity levels between the JG1 *Waikavirus* and other waikaviruses are below 64% and 56% for the Pro-Pol region and for the CPs, respectively (Supplementary Table S2). Considering the species molecular discrimination criteria in the *Secoviridae* family (less than 80% aa sequence identity in the Pro-Pol region and less than 75% identity in the CPs) (Sanfaçon et al. 2009; Thompson et al. 2017), the isolate identified in the JG1 lettuce should be considered as belonging to a novel species, for which the name lettuce waikavirus 1 (LWV1) is proposed.

Concerning the JG1-A/B and 21-015C-A/B isolates, they share with each other 86.2-97.5% and 92.1-99.7% aa sequence identity in their Pro-Pol and CPs, respectively. They should thus be considered as isolates belonging to the same virus species. As compared with other known sequiviruses, they share at most 76.9% identity in the Pro-Pol region (with LeMoV isolates), which is below the 80% species demarcation criterion (Sanfaçon et al. 2009; Thompson et al.

386 2017) (Supplementary Table S3). However, when comparing the CPs, an aa identity value of
387 82.2-84.2% is observed with LeMoV isolates, higher than the 76% species cut-off value. This
388 suggests that other species discrimination criteria need to be taken into consideration
389 (Thompson et al. 2017), in order to be able to decide on their taxonomic situation. We
390 therefore decided to compare the experimental host range of LeMoV with that of the JG1-A
391 and JG1-B isolates.

392 **Comparative host range analysis for various sequiviral isolates**

393 Despite repeated attempts using various buffers and inoculation conditions, all efforts to
394 transmit LWV1 from the JG1 plant to lettuce (cv. Trocadéro) failed (data not shown). The
395 same was observed for LeCV1 (Svanella-Dumas et al. 2018). These observations are
396 consistent with the notion that waikaviruses are not mechanically sap-transmissible because
397 they are confined to phloem tissue (Reavy et al. 1993). On the other hand, variants JG1-A and
398 JG1-B were readily transmitted from the JG1 plant and induced systemic mosaic and star
399 chlorosis symptoms on inoculated lettuce (cv. Trocadéro) (Figure 1F). These initial
400 transmission efforts also allowed to separate the JG1-A and JG1-B isolates as demonstrated
401 using isolate-specific RT-PCR assays (Supplementary Table 1). The JG1-A and JG1-B
402 isolates, as well as the JG2 LeMoV isolate, were sap-inoculated to plants belonging to ten
403 species in four families, some of which are reported as differential hosts between LeMoV and
404 DaYMV (Jadao et al. 1997, Bos et al. 1983; Table 3). The expression of symptoms was
405 monitored over a one-month period and the presence of the various viruses was verified in
406 non-inoculated upper parts of test plants by specific RT-PCR assays. The results of these
407 experiments are presented in Table 3.
408 In summary, the JG1-A and JG1-B variants were found to accumulate in six and seven of the
409 ten tested species, respectively. For both isolates, systemic mosaic symptoms were observed
410 on lettuce (cv. Trocadero), endive (*Chicorium endivia* cv. Anjou) (Fig. 1E) and common

411 sowthistle (*Sonchus olearaceus*), while asymptomatic accumulation was observed in *N.*
412 *benthamiana, N. clevelandii and Gomphrena globosa (*very occasionally only for JG1-A).
413 JG1-B, but not JG1-A, also accumulated in *C. quinoa* in which it induced yellow systemic
414 spots (Figure 1D).
415 In parallel experiments, inoculation of the JG2 LeMoV isolate yielded results only partially
416 comparable to those reported by Jadao et al. (1997) for the Brazilian AF197 isolate since
417 contrary to AF197, JG2 caused systemic mosaic symptoms in *C. endivia* and only caused
418 symptomless infection in *G. globosa* (Table 1). Another noteworthy difference is that contrary
419 to AF197, which systemically infected and caused local lesions and systemic symptoms in *C.*
420 *quinoa*, JG2 almost never infected this host. However, a *C. quinoa* plant was identified with
421 JG2 infection in independent experiments than those reported in Table 3. The virus from this
422 plant could then readily be transmitted to *C. quinoa* but with only a drastically reduced
423 efficiency to lettuce (not shown).
424 Taken together these results show that similar to the situation with LeMoV or with DaYMV,
425 there is significant inter-isolate variability in the host range of the JG1-A and JG1-B isolates.
426 The only host that consistently separated the JG1-A and JG1-B isolates from LeMoV isolates
427 was *N. clevelandii*, which was symptomlessly infected by the former, but not by the latter.
428 Differences also exist with DaYMV, for which multiple isolates are reported to cause local
429 lesions in *C. amaranticolor* and to be unable to infect *S. oleraceus* (Table 3, Bos et al. 1983).

430 **DISCUSSION**

431 In the current study, the use of HTS data in combination with 3' and 5' RACE and
432 confirmation by RT-PCR assays for regions of low coverage allowed us to assemble complete
433 genomic sequences for three *Secoviridae* family members from symptomatic lettuces
434 cultivated in France. Similar approaches were also used to obtain near-complete genome

435 sequences for additional isolates of the same viruses from wild or cultivated lettuces in France
436 or in Brazil as well as the first near-complete genomes for yet another *Secoviridae* member,
437 DaYMV from a dandelion from Slovakia.
438 For the *Waikavirus* found in the cultivated JG1 lettuce, its genome characterization and
439 phylogenetic analysis show that it belongs to a new species, for which the name lettuce
440 waikavirus 1 is proposed. The unsuccessful results at mechanical transmission suggest that
441 like RTSV, this new virus could be confined to phloem tissues and not mechanically
442 transmissible (Reavy et al. 1993). The biology and epidemiology of waikaviruses remain
443 poorly understood, likely in part due to their phloem restriction and reliance on specific
444 arthropod vectors for efficient transmission (Reddick et al. 1997). Some waikaviruses also
445 serve as helper viruses for sequiviruses or caulimoviruses (Elnagar et al. 1976, Thompson et
446 al. 2017). However, some other sequiviruses such as DaYMV and LeMoV are reported to be
447 transmitted by aphids in the absence of a helper virus (Jadao et al. 2007). In the case of the
448 JG1 plant, given the mixed infection involving both LWV1 and the JG1-A and JG1-B
449 sequivirus isolates, it is not possible to know whether transmission complementation effects
450 exist or not between these various *Secoviridae* members. However, the identification of the
451 21-015C-A and -B isolates in a wild lettuce and in the absence of a *Waikavirus* suggests that
452 similar to LeMoV these isolates do not require the assistance of a *Waikavirus* for
453 transmission.
454 Considering the 80% cut-off molecular criterion in the Pro-Pol region defining species within
455 the family *Secoviridae* (Sanfaçon et al. 2009; Thompson et al. 2017), the JG1-A, JG1-B, 21-
456 015C-A and 21-015C-B isolates should all be considered as belonging to a single new species
457 in the genus *Sequivirus*. However, when considering the 76% cut-off value that applies to the
458 CPs, they should be considered as isolates of LeMoV. They are however unambiguously
459 different from DaYMV, the only sequivirus known to date from lettuce in Europe. At the

same time, the JG2, P22, LIB3-A, LIB3-B and 21-015C isolates from cultivated and wild lettuces are unambiguously European LeMoV isolates that tightly cluster with LeMoV sequences from South America (Figure 3A & B). The host range analysis identified a single consistent difference between LeMoV isolates and the JG1-A and JG1-B isolates, with the former being unable to infect *N. clevelandii* (Table 3). Considering together the phylogenetic analyses (Fig. 3), the species molecular criterion met for the Pol-Pro region, the differences in the predicted cleavage sites liberating the 3 CPs (Fig. 2 and Table 2) and the host range difference, we suggest that the JG1-A/B and 21-015C-A/B isolates could all be considered as belonging to a distinct species, for which the name lettuce star mosaic virus (LSMV) is proposed. LSMV is clearly pathogenic in lettuce, inducing star mosaic symptoms that may turn to pin-point or star necrosis and are quite similar to those caused by LeMoV, while more punctuated that the mosaics typically induced by lettuce mosaic virus (LMV). It should be noted that in the absence of fully conclusive diagnostics at the time, older reports of the presence of DaYMV in French lettuce crops (Blancard et al. 2003) might, in fact, have been related to LeMoV or LSMV.

Previous surveys of lettuce crops in France suggest that LSMV or LeMoV are rare, even though wild lettuce may possibly act as a reservoir in the absence of lettuce crops (hence the HTS analysis of the JG1 and JG2 plants following the observation of unusual symptoms). This situation is, at least for LeMoV, in contrast to that reported in the early 2000s in Brazil where LeMoV was considered the first or second most important lettuce virus, with a reported prevalence of 10.1%-26.1% under open field conditions in Sao Paulo state (Krause-Sakate et al. 2008; De Marchi et al. 2012). However, this situation changed during the last years, with LeMoV being now more rarely found and the main viral constraint now being thrips transmitted orthotospoviruses (RKS, unpublished data). Identifying the causes underlying these differences epidemiological trends, which might be related to differences in agricultural

485 practices, varietal choice, insect vector populations or other yet unidentified factors will
486 require further investigations. The availability of complete genomic sequences for LeMoV,
487 LSMV and DaYMV, the characterization and genome sequencing of LWV1 and the RT-PCR
488 assays developed here should provide useful tools for such studies and for a broader and more
489 thorough estimation of the distribution, prevalence and impact of these viruses on lettuce
490 crops worldwide.

491 **Acknowledgments:** M.G. and L.P. acknowledge support from the Scientific Grant Agency of
492 the Ministry of Education and Slovak Academy of Sciences (VEGA 2/0030/20) and the
493 Slovak Research and Development Agency (APVV-18-0005). The authors thank the INRAE
494 GenoToul Platform (Toulouse, France) for high throughput sequencing, and Muriel Bahut
495 from the ANAN technical platform of SFR QUASAV (Anger, France) for technical
496 assistance with libraries preparations and HTS sequencing. The help of T. Mauduit and C.
497 Higelin (UMR NFP, INRAE Bordeaux) for taking care of plants in the greenhouse is
498 gratefully acknowledged. The authors are also most grateful to the following people from
499 ANSES Unité BVO (Angers, France): Michèle Visage for the first biological and molecular
500 characterization of the viral isolates in the 21-015C plant, Benoit Remenant for his
501 bioinformatics support and Mathieu Rolland for valuable discussions.
502

LITERATURE CITED

Adams, I.P., Skelton, A., Macarthur, R., Hodges, T., Hinds, H., Flint, L., Nath, P.D., Boonham, N. and Fox, A. 2014. Carrot yellow leaf virus is associated with carrot internal necrosis. PLoS ONE 9:e109125.

Blancard, D., Lot, H. and Maisonneuve, B. 2003. Maladies des salades : identifier, connaître et maîtriser. INRA Editions, Paris, France, pp 311-312.

Bos, L., Huijberts, N., Huttinga, H. and Maat, D.Z. 1983. Further characterization of dandelion yellow mosaic virus from lettuce and dandelion. Nether. J. Plant Pathol. 89:207-222.

Candresse, T., Marais, A., Faure, C. and Gentit, P. 2013. Association of little cherry virus 1 (LChV1) with the Shirofugen stunt disease and characterization of the genome of a divergent LChV1 isolate. Phytopathology 103:293-298.

Chang, S., Puryear, J. and Cairney J. 1993. A simple and efficient method for isolating RNA from pine trees. Plant Mol. Biol. Rep. 11:113-116.

Ciuffo, M., Mammella, M., Vallino, M., Caciagli, P. and Turina, M. 2016. Molecular identification and biogical characterization of a new potyvirus in lettuce. Arch. Virol. 161:2549-2554.

De Marchi, B.R., de Almeida Spadotti D.M., Leite de Oliveira, M., Martinello Sanches, M., Krause-Sakate, R. and Pavan, M.A. 2012. Levantamento revela a predominância do Lettuce mottle virus em três regiões produtoras de alface no Estado de São Paulo. Summa Phytopathol. 38:245-247.

de Vries, I.M. 1997. Origin and domestication of *Lactuca sativa* L. Genet. Resour. Crop Evol. 44:165-174.

Dullemans, A.M., Botermans, M., de Kock, M.J.D., de Krom, C.E, van der Lee, T.A.J., Roenhorst, J.W., Stulemeijer, I.J.E., Verbeek, M., Westenberg, M. and van der Vlugt, R.A.A. 2020. Creation of a new genus in the family *Secoviridae* substantiated by sequence variation of newly identified strawberry latent ringspot virus isolates. Arch. Virol. 165:21-31.

Elnagar, S. and Murant, A.F. 1976. The role of the helper virus, anthriscus yellows, in the transmission of parsnip yellow fleck virus by the aphid *Cavariella aegopodii*. Ann. Appl. Biol. 84:169-181.

FAO 2019. The world lettuce economy. http://www.fao.org/faostat/en/#data

Firth, A.E. and Atkins, J.F. 2008. Bioinformatic analysis suggests that a conserved ORF in the waikaviruses encodes an overlapping gene. Arch. Virol. 153:1379-1383.

Foissac, X., Svanella-Dumas, L., Gentit, P., Dulucq, M.J., Marais, A. and Candresse, T. 2005. Polyvalent degenerate oligonucleotides reverse transcription-polymerase chain reaction: a polyvalent detection and characterization tool for trichoviruses, capilloviruses, and foveaviruses. Phytopathology 95:617-625.

Jadao, A.S., Krause-Sakate, R., Liberti, D., Pavan, M.A., Echer, M.M., Svanella-Dumas, L., Zerbini, F.M., Candresse, T. and Le Gall, O. 2007. Further characterization of two sequiviruses infecting lettuce and development of specific RT-PCR primers. Arch. Virol. 152:999-1007.

Kassanis, B. 1944. A virus attacking lettuce and dandelion. Nature 154:6.

Kassanis, B. 1947. Studies on dandelion yellow mosaic and other virus diseases of lettuce. Ann. Appl. Biol. 34:412-421.

548 Koloniuk, I. and Fránová, J. 2018. Complete nucleotide sequence and genome organization of
549 red clover associated virus 1 (RCaV1), a putative member of the genus *Waikavirus* (family
550 *Secoviridae*, order *Picornavirales*). Arch. Virol. 163:3447-3449.

551 Krause-Sakate, R., Firmino, A.C., Jadao, A.S., Pavan, M.A., da Silva, N., Hanai, S.M., Anbo,
552 R.H. and Nietzsche, T. 2008. Ocorrência generalizada do Lettuce mottle virus em três
553 regiões produtoras de alface comercial do Estado de São Paulo. Summa Phytopathol.
554 34:88-89.

555 Krause-Sakate, R., Jadao, A.S., Firmino, A.C., Pavan, M.A., Zerbini, F.M., Rosales, I.M.,
556 Bustamante, P. and Le Gall, O. 2005. First report of a lettuce-infecting sequivirus in Chile.
557 Plant Disease 89:1129.

558 Kumar, S., Stecher, G. and Tamura, K. 2016. MEGA7: molecular evolutionary genetics
559 analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33:1870-1874.

560 Le Fustec, M. 2020. Laitues d'hiver 2019-2020: une fin de campagne marquée par l'impact de
561 la crise sanitaire sur la demande et les dernières récoltes. Agreste Infos Rapides 2020-048.
562 http://www.agreste.agriculture.gouv.fr

563 Le Gall, O., Sanfaçon, H., Ikegami, M., Iwanami, T., Jones, T., Karasev, A., Lehto, K.,
564 Wellink, J., Wetzel, T. and Yoshikawa, N. 2007. *Cheravirus* and *Sadwavirus*: two
565 unassigned genera of plant positive-sense single-stranded RNA viruses formerly
566 considered atypical members of the genus *Nepovirus* (family *Comoviridae*). Arch. Virol.
567 152:1767-1774.

568 Lebeda, A., Kristkova, E., Kitner, M., Mieslerova, B., Jemelkova, M. and Pink, D.A.C. 2014.
569 Wild *Lactuca* species, their genetic diversity, resistance to diseases and pests, and
570 exploitation in lettuce breeding. Eur. J. Plant Pathol. 138:597-640.

571 Lefebvre, M., Theil, S., Ma, Y. and Candresse, T. 2019. The VirAnnot pipeline: a resource for
572 automated viral diversity estimation and operational taxonomy units assignation for virome
573 sequencing data. Phytobiomes J. 3:256-259.
574 Maliogka, V.I., Minafra, A., Saldarelli, P., Ruiz-García, A.B., Glasa, M., Katis, N. and
575 Olmos, A. 2018. Recent advances on detection and characterization of fruit tree viruses
576 using high-throughput sequencing technologies. Viruses 10:436.
577 Marais, A., Faure, C., Bergey, B. and Candresse, T. 2018. Viral double-stranded RNAs
578 (dsRNAs) from plants: alternative nucleic acid substrates for high-throughput sequencing.
579 Meth. Mol. Biol. 1746:45-53.
580 Maree, H.J., Fox, A., al Rwahnih, M., Boonham, N. and Candresse, T. 2018. Application of
581 HTS for routine plant virus diagnostics: state of the art and challenges. Front. Plant Sci.
582 27:1082.
583 Menzel, W. and Vetten, H.J. 2008. Complete nucleotide sequence of an isolate of the
584 Anthriscus strain of Parsnip yellow fleck virus. Arch. Virol. 153:2173-2175.
585 Moreno, A. and Fereres, A. 2012. Virus diseases in lettuce in the Mediterranean basin. Adv.
586 Virus Res 84:247-248.
587 Pavan, M.A., Krause-Sakate, R., Silva, N., Zerbini, F.M. and Le Gall, O. 2008. Virus diseases
588 of lettuce in Brazil. Plant Viruses. Global Science Books, Viçosa, pp. 35-40.
589 Petrzik, K., Sanfaçon, H., Tzanetakis, I., van der Vlugt, R., Wetzel, T., Yoshikawa, N.,
590 Lefkowitz, E.J., Davison, A.J., Siddell, S.G., Simmonds, P., Adams, M.J., Smith, D.B.,
591 Orton, R.J. and Knowles, N.J. 2017. ICTV virus taxonomy profile: *Secoviridae*. J. Gen.
592 Virol. 98:529-531.
593 Reavy, B., Mayo, M.A., Turnbull-Ross, A.D. and Murant, A.F. 1993. Parsnip yellow fleck
594 and rice tungro sperical viruses resemble picornaviruses and represent two genera in a
595 proposed new plant picornavirus family (*Sequiviridae*). Arch. Virol. 131:441-446.

596 Reddick, B.B., Habera, L.F. and Law, M.D. 1997. Nucleotide sequence and taxonomy of
597 maize chlorotic dwarf virus within the family *Sequiviridae*. J. Gen. Virol. 78:1165-1174.

598 Sanfaçon, H., Dasgupta, I., Fuchs, M., Karasev, A., Petrzik, K., Thompson, J., Tzanetakis, I.,
599 van der Vlugt, R., Wetzel, T. and Yoshikawa, N. 2020. Proposed revision of the family
600 *Secoviridae* taxonomy to create three subgenera, "*Satsumavirus*", "*Stramovirus*" and
601 "*Cholivirus*", in the genus *Sadwavirus*. Arch. Virol. 165:527-533.

602 Sanfaçon, H., Iwanami, T., Karasev, A.V., van der Vlugt, R., Wellink, J., Wetzel, T. and
603 Yoshikawa, N. 2011. Family *Secoviridae*. In "Virus Taxonomy: Ninth Report of the
604 International Committee on Taxonomy of Viruses. King, A.M.Q., Adams, M.J., Carstens,
605 E.B. and Lefkowitz, E.J. Eds, Elsevier, Academic Press, pp 881-899.

606 Sanfaçon, H., Wellink, J., Le Gall, O., Karasev, A., van der Vlugt, R. and Wetzel, T. 2009.
607 *Secoviridae*: a proposed family of plant viruses within the order *Picornavirales* that
608 combines the families *Sequiviridae* and *Comoviridae*, the unassigned genera *Cheravirus*
609 and *Sadwavirus*, and the proposed genus *Torradovirus*. Arch. Virol. 154:899-907.

610 Seo, J.K., Kwak, H.R. and Lee, Y.J. 2015. Complete genome sequence of bellflower vein
611 chlorosis virus, a novel putative member of the genus *Waikavirus*. Arch. Virol. 160:3139-
612 3142.

613 Shen, P., Kaniewska, M., Smith, C. and Beachy, R.N. 1993. Nucleotide sequence and
614 genomic organization of rice tungro spherical virus. Virology 193:621-630.

615 Svanella-Dumas, L., Tsarmpopoulos, I., Marais, A., Theil, S., Faure, C., Gaudin, J. and
616 Candresse, T. (2018) Complete genome sequence of lettuce chordovirus 1 isolated from
617 cultivated lettuce in France. Arch. Virol. 163:2543-2545.

618 Thole, V. and Hull, R. 1996. Rice tungro spherical virus: nucleotide sequence of the 3'
619 genomic half and studies on the two small 3' open reading frames. Virus Genes 13:239-
620 246.

621 Thole, V. and Hull, R. 2002. Characterization of a protein from rice tungro spherical virus
622 with serine proteinase-like activity. J. Gen. Virol. 83:3179-3186.

623 Thompson, J., Higgins, D.G. and Gibson, T.J. 1994. CLUSTALW: improving the sensitivity
624 of progressive multiple sequence alignment through sequence weighting, position-specific
625 gap, penalties and weight matrix choice. Nucl. Acids Res. 22:4673-4680.

626 Thompson, J.R., Dasgupta, I., Fuchs, M., Iwanami, T., Karasev, A.V., Petrzik, K., Sanfaçon,
627 H., Tzanetakis, I., van der Vlugt, RA.A., Wetzel, T., Yoshikawa, N. and ICTV Report
628 Consortium 2017. ICTV Virus Taxonomy Profile: *Secoviridae*. J. Gen. Virol. 74:555-561.

629 Tomašechová, J., Hančinský, R., Predajňa, L., Kraic, J., Mihálik, D., Šoltys, K., Vávrová, S.,
630 Böhmer, M., Sabanadzovic, S. and Glasa, M. 2020. High-throughput sequencing reveals
631 bell pepper endornavirus infection in pepper (*Capsicum annum*) in Slovakia and enables its
632 further molecular characterization. Plants, 9:41.

633 Turnbull-Ross, A.D., Mayo, M.A., Reavy, B. and Murant, A.F. 1993. Sequence analysis of
634 the parsnip yellow fleck virus polyprotein: evidence of affinities with picornaviruses. J.
635 Gen. Virol. 74:555-561.

636 Turnbull-Ross, A.D., Reavy, B., Mayo, M.A. and Murant, A.F. 1992. The nucleotide
637 sequence of parsnip yellow fleck virus: a plant picorna-like virus. J. Gen. Virol. 73:3203-
638 3211.

639 Verbeek, M., Dullemans, A.M., van Raaij, H.M., Verhoeven, J.T.J., van der Vlugt, R.A.A.
640 2014. Lettuce necrotic leaf curl virus, a new plant virus infecting lettuce and a proposed
641 member of the genus *Torradovirus*. Arch. Virol. 159:801-805.

642 Villamor, D.E.V., Ho, T., al Rwahnih, M., Martin, R.R. and Tzanetakis, I.E. 2019. High
643 throughput sequencing for plant virus detection and discovery. Phytopathology 109:716-
644 725.

645 Wu, Q., Ding, S.W., Zhang, Y. and Zhu, S. 2015. Identification of viruses and viroids by
646 next-generation sequencing and homology-dependent and homology-independent
647 algorithms. Annu. Rev. Phytopathol. 53:425-444.

648

TABLE 1. List of viral isolates detected in various cultivated and wild lettuces and in the RNA24P dandelion plant.

Plant	Virus	Mapped reads	% Mapped reads	Average Coverage	BlastX best hit / (% of identity)
JG1 (a)	LWV1	26,980	33.7%	449x	Bellflower vein chlorosis virus (49%)
	LeCV1	12,327	15.4%	283x	Carrot chordovirus 1 (54%)
	LSMV-JG1-A	4,198	5.2%	92x	Parsnip yellow fleck virus (54%)
	LSMV-JG1-B	5,227	6.5%	102x	Carrot necrotic dieback virus (47%)
JG2 (b)	LeMoV	119,027	0.38%	1,790x	Carrot necrotic dieback virus (50%)
P22 (a)	LeMoV	2,517	29.3%	78x	Parsnip yellow fleck virus (47%)
LIB3 (a)	LeMoV LIB3-A	800.864	29.3%	9,107x	na
	LeMoV LIB3-B	365.279	13.2%	4,157x	na
21-015C (b)	LeMoV	143,557	2.6%	2,024x	na
	LSMV 21-015C-A	33,503	0.6%	460x	na
	LSMV 21-015C-B	24,895	0.4%	343x	na
RNA24P (b)	DaYMV-RNA24P-1	3,824	60.3%	60x	Carrot necrotic dieback virus (49%)
	DaYMV-RNA24P-2	1,003	15.6%	15x	Carrot necrotic dieback virus (49%)

For each viral contig, number of reads assigned to each isolate, the proportion of mapped reads and the average coverage are provided, together with information about the best BlastX hit. The results were obtained after HTS analysis of (a) double-stranded RNAs or (b) total RNAs from infected plants. LWV1: lettuce waikavirus 1, LeCV1: lettuce chordovirus 1, LSMV: lettuce star mosaic virus, LeMoV: lettuce mottle virus, DaYMV: dandelion yellow mosaic virus. na: does not apply

TABLE 2. Polyprotein dipeptides cleavage sites for viruses in the genera *Waikavirus* and *Sequivirus* and either (a) published or (b) predicted in the present work using a ClustalW-generated multiple alignment of the viral polyproteins.

Virus sequences per species in this study	Hypothetical P1 / CP1	CP1 / CP2	CP2 / CP3	CP3 / NTP-Protein Binding	NTP-Protein Binding / Pro#	Pro / RdRp#
RTSV NC_001632 (a)	Q/A	Q/S	Q/D	Q/M	Q/D	Q/A
MCDV NC_003626 (a)	Q/S	Q/M	Q/V	Q/M	Q/V	S/P
BVCV NC_027915 (a)	Q/S	Q/S	Q/G	K/D	Q/D	E/A
RcaV1 MH325329 (a)	Q/N	Q/N	Q/L	Q/M	Q/E	A/A
BraV1 NC_040586 (b)	Q/S	Q/A	Q/A	Q/M	Q/N	T/A
PWV LC_488189 (b)	Q/T	Q/A	Q/S	Q/K	Q/D	S/V
LWV1 MT348710 (b)	Q/S (700/701)	Q/S (922/923)	Q/S (1124/1125)	Q/T (1423/1424)	Q/D (2536/2537)	Q/S (2873/2874)
PYFV NC_003628 (a)	S/P	N/A	Q/A	Q/A	S/P	Q/S
CNDV NC_038320 (b)	G/S	Q/A	Q/S	Q/G	?	?
LeMoV-JG2 MT348707 (b)	E/P (385/386)	Q/G (577/578)	Q/N (801/802)	Q/A(1037/1038)	?	?
LeMoV-P22 MT348708 (b)	E/P	Q/G	Q/N	Q/A	?	?
LeMoV-SP MT348709 (b)	E/P	Q/G	Q/N	Q/A	?	?
LeMoV LIB3-A OQ446443 (b)	E/P	Q/G	Q/N	Q/A	?	?
LeMoV LIB3-B OQ446444 (b)	E/P	Q/G	Q/N	Q/A	?	?
LeMoV 21-015C OQ446445 (b)	E/P	Q/G	Q/N	Q/A	?	?
LSMV-JG1-A MT348705 (b)	N/P	Q/G	Q/G	Q/A	?	?
LSMV-JG1-B MT348706 (b)	N/P (383/384)	Q/G (575/576)	Q/G (799/800)	Q/A (1035/1036)	?	?
LSMV-21-015C-A OQ446446 (b)	N/P	Q/G	Q/G	Q/A	?	?
LSMV-21-015C-B OQ446447 (b)	D/P	Q/G	Q/A	Q/A	?	?
DaYMV-RNA24P-1 MT559283 (b)	K/P (388/389)	Q/G (580/581)	Q/M (804/805)	Q/G (1041/1042)	?	?
DaYMV-RNA24P-2 MT559284 (b)	K/P	Q/G	Q/S	Q/G	?	-

Amino acid positions of the cleavage sites on the polyprotein are provided in parentheses () for one representative isolate sequenced here of each species. #: a question mark indicates viruses for which it was not possible to identify a tentative dipeptide cleavage site.

TABLE 3. Infection and symptomatology of sequiviral isolates in various hosts following mechanical transmission.

	JG1-AT348705	**JG1-BT348706**	**LeMoV-JG2**	**LeMoV-AF197(a)**	**DaYMV(b)**
Chicorium endivia cv. Anjou	S **(9/11)**	S **(14/43)**	S **(33/60)**	-	S
Lactuca sativa cv. Trocadéro	S **(22/24)**	S **(13/18)**	S **(29/32)**	S	S
Sonchus olearaceus	S **(5/12)**	S (**13/29)**	S **(30/36)**	S	-
Nicotiana rustica	- (0/12)	- (0/33)	- (0/36)	-	(ll)(s)
Nicotiana glutinosa	- (0/12)	- (0/33)	- (0/48)	-	(ll)
Nicotiana benthamiana	s **(6/12)**	s **(27/36)**	s (**29/36)**	ll, s	s
Nicotiona clevelandii	s (**18/56)**	s **(11/67)**	- (0/48)	-	(ll)(s)
Chenopodium amaranticolor	- (0/11)	- (0/32)	- (0/48)	-	LL, (S)
Chenopodium quinoa	- (0/12)	**S (7/32)**	- (0/36)	LL, S	LL, (S)
Gomphrena globosa	s **(1/11)**	s **(11/55)**	s **(14/56)**	LL, S	(LL), (s)

The number of infected over inoculated plants is indicated in parentheses. Also shown are infection and symptoms reported by Jadao et al. (2007) for lettuce mottle virus (LeMoV) isolate AF197 (a) and reported in Bos et al. (1983) for several isolates of dandelion yellow mosaic virus (DaYMV) (b). -, no infection; S, systemic symptoms; LL, local lesions; s, latent systemic infection; ll, latent local infection. Symptoms noted in parentheses reflect inter-isolate variability.

LEGENDS TO THE FIGURES

Fig. 1. Symptoms caused by lettuce mottle virus (LeMoV) and lettuce star mosaic virus (LSMV) in various host plants. **A**, Necrotic mosaic symptoms observed on the original field-grown JG2 lettuce (cv. Guétary). **B**, Systemic necrotic lesions of LeMoV JG2 isolate on *Chicorum endivia* cv. Anjou 13 days post-inoculation (dpi). **C**, Systemic mosaic symptoms of LeMoV JG2 isolate on lettuce cv. Trocadéro at 13 dpi. The leaf on the left is from an uninoculated control plant. **D**, Ssymptoms of lettuce star mosaic virus (LSMV) isolate JG1-B on *Chenopodium quinoa* at 36 dpi. **E**, Symptoms of lettuce star mosaic virus (LSMV) isolate JG1-A on *C. endivia* cv. Anjou at 19 dpi. **F**, Symptoms of LSMV JG1-A (left) and JG1-B (right) in lettuce cv. Trocadéro.

Fig. 2. Schematic representation of the genomic organization drawn to scale of **A,** lettuce waikavirus 1 (LWV1) isolate JG1. **B,** lettuce star mosaic virus (LSMV) isolate JG1-B. **C,** lettuce mottle virus (LeMoV) isolate JG2. D, dandelion yellow mosaic virus (DaYMV) isolate RNA24P-1, partial. The short horizontal lines show the 5' and 3' untranslated regions, respectively. The large box is the long open reading frame with the coordinates of the beginning and the end mentioned. Predicted cleavage sites and their positions are shown as triangles under predicted mature proteins and uncertain cleavage sites by a question mark. P1, hypothetical protein 1; CP1, CP2, CP3, coat proteins 1, 2, and 3; NTB, NTP-binding helicase; PRO, 3C-like protease; RdRp, RNA-dependent RNA polymerase are indicated within the polyprotein. Predicted ORFX is indicated by a box for LWV1-JG1.

Fig. 3. Unrooted phylogenetic trees reconstructed using **A,** the Pro-Pol region, and **B,** the coat protein amino acid sequences from the members of the genera *Waikavirus* and *Sequivirus*, one representative species of other genera in family *Secoviridae*, and unassigned plant viruses showing sequence homology with the members of *Secoviridae*. The sequences of the viruses determined in the present work are indicated in bold font with black dots. The trees were constructed using the neighbor-joining method and the bootstrap values (500 replicates) greater than 70% are indicated next to the branches. The scale bar represents 10% aa divergence. The abbreviations and the accession numbers of the viruses used in this work are mentioned in Supplementary TABLE S4.

Figure 1

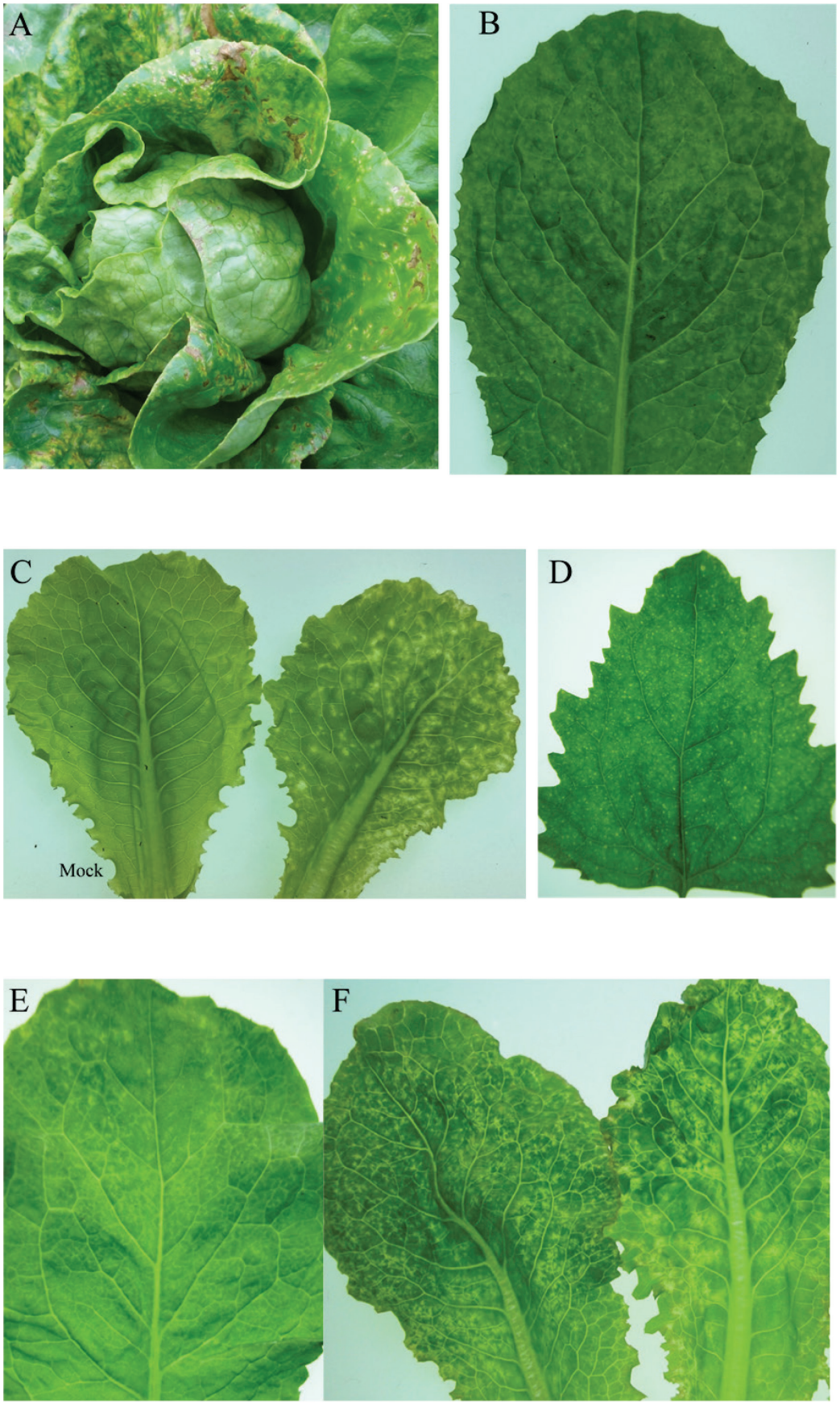
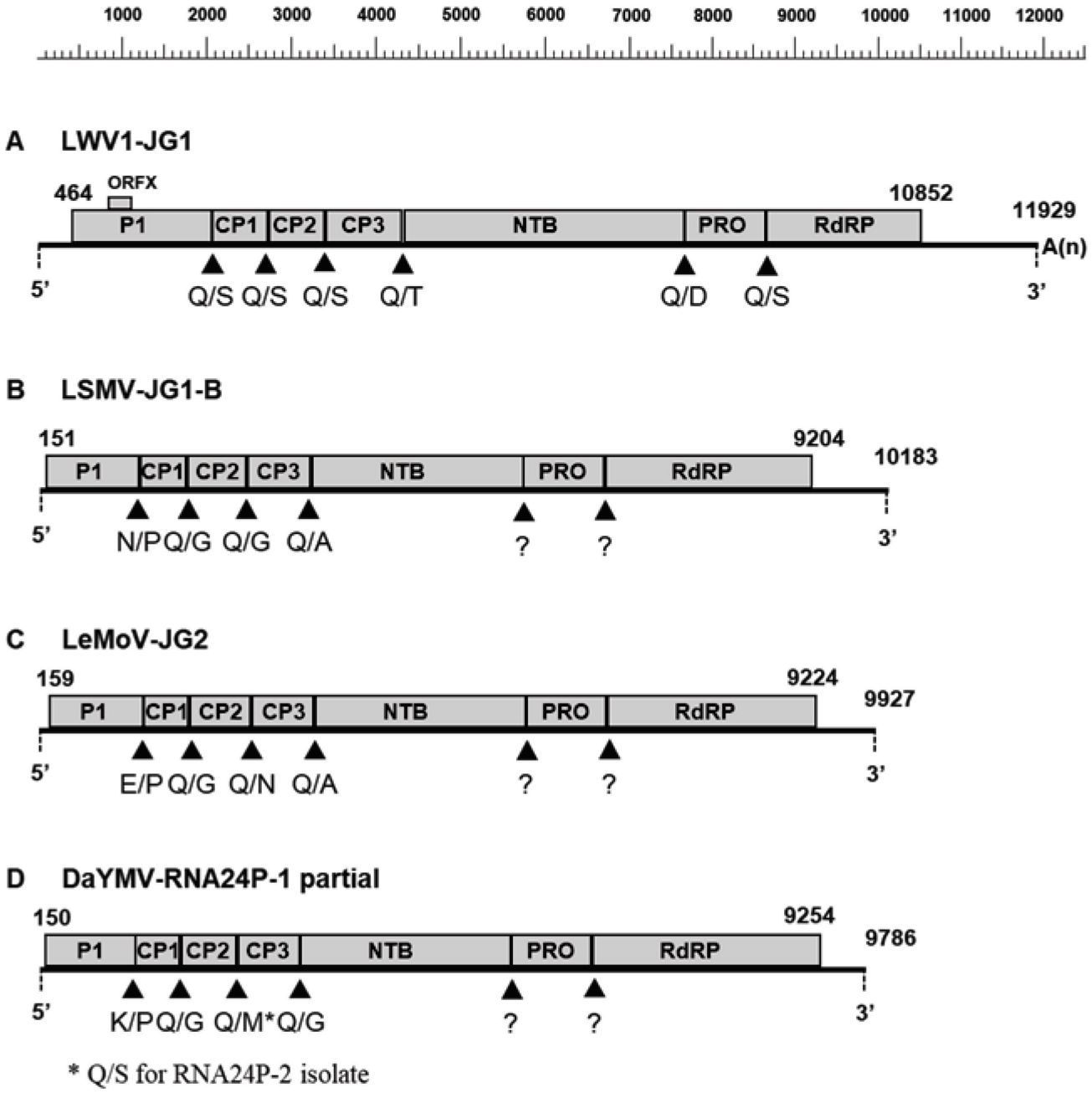
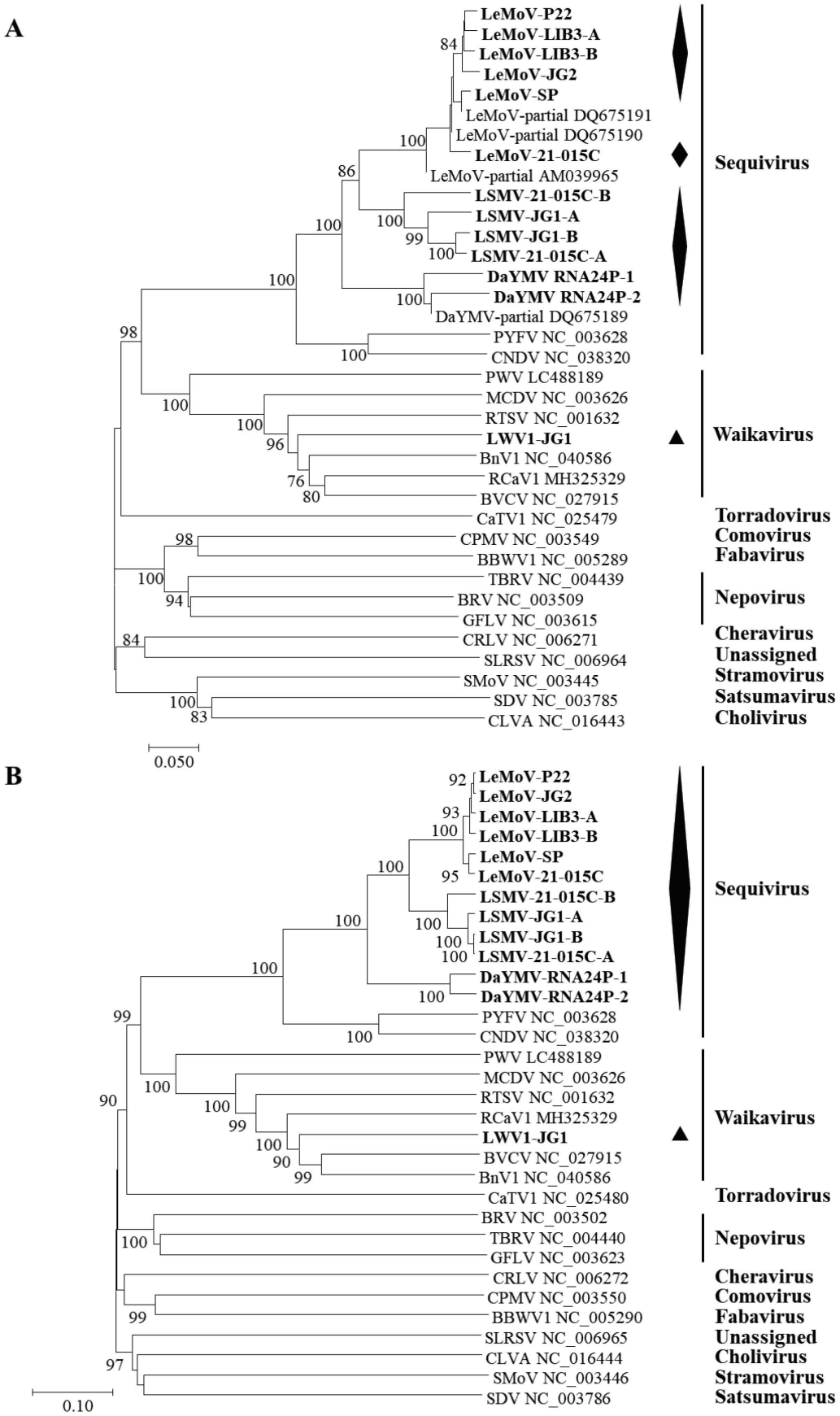


Figure 2



* Q/S for RNA24P-2 isolate

Figure 3


A
LeMoV-P22
LeMoV-LIB3-A
LeMoV-LIB3-B
LeMoV-JG2
LeMoV-SP
LeMoV-partial DQ675191
LeMoV-partial DQ675190
LeMoV-21-015C
LeMoV-partial AM039965
LSMV-21-015C-B
LSMV-JG1-A
LSMV-JG1-B
LSMV-21-015C-A
DaYMV RNA24P-1
DaYMV RNA24P-2
DaYMV-partial DQ675189
PYFV NC_003628
CNDV NC_038320
PWV LC488189
MCDV NC_003626
RTSV NC_001632
LWV1-JG1
BnV1 NC_040586
RCaV1 MH325329
BVCV NC_027915
CaTV1 NC_025479
CPMV NC_003549
BBWV1 NC_005289
TBRV NC_004439
BRV NC_003509
GFLV NC_003615
CRLV NC_006271
SLRSV NC_006964
SMoV NC_003445
SDV NC_003785
CLVA NC_016443
0.050
Sequivirus
Waikavirus
Torradovirus
Comovirus
Fabavirus
Nepovirus
Cheravirus
Unassigned
Stramovirus
Satsumavirus
Cholivirus
B
LeMoV-P22
LeMoV-JG2
LeMoV-LIB3-A
LeMoV-LIB3-B
LeMoV-SP
LeMoV-21-015C
LSMV-21-015C-B
LSMV-JG1-A
LSMV-JG1-B
LSMV-21-015C-A
DaYMV-RNA24P-1
DaYMV-RNA24P-2
PYFV NC_003628
CNDV NC_038320
PWV LC488189
MCDV NC_003626
RTSV NC_001632
RCaV1 MH325329
LWV1-JG1
BVCV NC_027915
BnV1 NC_040586
CaTV1 NC_025480
BRV NC_003502
TBRV NC_004440
GFLV NC_003623
CRLV NC_006272
CPMV NC_003550
BBWV1 NC_005290
SLRSV NC_006965
CLVA NC_016444
SMoV NC_003446
SDV NC_003786
0.10
Sequivirus
Waikavirus
Torradovirus
Nepovirus
Cheravirus
Comovirus
Fabavirus
Unassigned
Cholivirus
Stramovirus
Satsumavirus

Supplementary table titles and Supplementary figure captions

Supplementary TABLE S1. Primers used for the detection and the molecular characterization of lettuce mottle virus (LeMoV) isolates JG2, P22 and SP, of lettuce star mosaic virus (LSMV) isolates JG1-A and JG1-B and of lettuce waikavirus 1 (LWV1) isolate JG1.

Supplementary TABLE S2. Pairwise percentages of amino acid identity for the CP (lower diagonal) and Pro-Pol (upper diagonal) regions of the polyproteins of Waikavirus genus members.

Supplementary TABLE S3. Pairwise percentages of identity for the CP (lower diagonal) and Prot-Pol (upper diagonal) regions of the polyproteins of Sequivirus genus members.

Supplementary TABLE S4. Description of the viral isolates used for the phylogenetic studies

Supplementary Figure 1. Multiple alignment of the predicted short overlapping ORFX from all members of the genus Waikavirus.