Identification of viruses belonging to the family *Partitiviridae* from plant transcriptomes

Yeonhwa Jo and Won Kyong Cho*

Research Institute of Agriculture and Life Sciences, College of Agriculture and Life Sciences,
Seoul National University, Seoul 08826, Republic of Korea

*Corresponding author

Dr. Won Kyong Cho

Tel: +82-2-880-4687; Fax: +82-2-873-2317

E-mail: wonkyong@gmail.com

Running head: Identification of partitiviruses

Keywords: Viruses, *Partitiviridae*, Identification, Plant transcriptome

Abstract

Viruses in the family *Partitiviridae* consist of non-enveloped viruses with bisegmented double-stranded RNA genomes. Viruses in this family have been identified from plants and fungi. In this study, we identified several viruses belonging to the family *Partitiviridae* using plant transcriptomes. From 11 different plant species, we identified a total of 74 RNA segments representing 23 partitiviruses. Of 74 RNA segments, 28 RNA segments encode RNA-dependent RNA polymerases (RdRp) while 46 RNA segments encode coat proteins (CPs). According to ICTV demarcation for the family *Partitiviridae*, 25 RNAs encoding RdRp and 41 RNAs encoding CP were novel RNA segments. In addition, we identified eight RNA segments (three for RdRp and five for CP) belonging to the known partitiviruses. Taken together, this study provides the largest number of partitiviruses from plant transcriptomes in a single study.

Introduction

Viruses in the family *Partitiviridae* consist of non-enveloped viruses with bisegmented double-stranded (ds) RNA genomes [1]. Each RNA segment of the viruses in the family *Partitiviridae* encodes a single protein including RNA-dependent RNA polymerase (RdRp) or coat protein (CP). Members in the family *Partitiviridae* can be divided into five different genera according to the host. For example, viruses in the genera *Alphapartitivirus* and *Betapartitivirus* are identified from either plants or fungi while viruses in the genus *Gammapartitivirus* and *Deltapartitivirus* are derived from only fungi and plants, respectively [1]. In addition, viruses in the genus *Cryspovirus* are identified from protozoa. Transmission of partitiviruses is occurred by seeds (plants), cell division and sporogenesis (fungi), and oocytes (protozoa) [1]. To date, there are more than 45 partitivirus species. In particular, there are five deltapartitivirus species infecting plants including *Pepper cryptic virus 1* (PCV-1) [2], *Pepper cryptic virus 2* (PCV-2), *Fig cryptic virus* (FCV) [3], *Beet cryptic virus 2* (BCV2), and *Beet cryptic virus 3* (BCV3) [4] according to International Committee on Taxonomy of Viruses (ICTV).

To identify novel partitiviruses with dsRNA genomes, dsRNA extraction followed by next-generation sequencing (NGS) is an efficient approach. Based on those approaches, several partitiviruses have been identified. For example, Melon necrotic spot virus was identified from watermelon plant (*Citrullus lanatus* Thunb) using SOLiD NGS analysis [5]. Pittosporum cryptic virus 1 was identified from an Italian pittosporum plant by NGS with extracted dsRNA [6]. Similarly, application of NGS to extracted dsRNA has resulted in identification of Trichoderma atroviride partitivirus 1 from the fungal species *Trichoderma atroviride* [7].

Moreover, instead of traditional molecular cloning methods, *in silico* data analyses has resulted in identification of diverse dsRNA viruses from expressed sequence tag (EST) database [8] and plant transcriptomes [9-13]. Moreover, 120 mycoviruses infecting fungi have

been identified from fungal transcriptomes including several partitiviruses [14].

Recently, we carried out *in silico* data analyses to identify virus-associated contigs from plant transcriptomes. As a result, we identified a total of 74 RNA segments representing 23 partitiviruses belonging to the family *Partitiviridae* from 11 different plant species,

Materials and methods

De novo transcriptome assembly and BLAST search

We downloaded RNA-Sequencing (RNA-Seq) from the sequence read archive (SRA) database. The downloaded raw sequence data were *de novo* assembled by Trinity assembler with default parameters [15]. The assembled contigs (transcriptome) were blasted against the database containing all viral genomes for the family *Potyviridae*, which were manually collected, using TBLASTX with a cut-off E-value of $1e^{-3}$. After that, we extracted virus-associated contigs based on the TBLASTX result. To eliminate sequences derived from the plants and other contaminants, the identified virus-associated contigs were subjected to BLASTN search against NCBI's nucleotide (NT) database. Finally, we obtained clean virus-associated contigs.

Prediction of open reading frames

Of identified virus-associated contigs, the contigs more than 1 kb in length were selected for the search of open reading frames (ORFs) using ORFfinder program (<https://www.ncbi.nlm.nih.gov/orffinder/>). The identified ORF protein sequences were subjected to BLASTP search against non-redundant (NR) protein database with a cut-off E-value of $1e^{-3}$.

Virus classification

Based on BLASTP results with individual viral ORF amino acid sequence, we determined virus taxonomy. The newly identified virus was named based on the identified virus host and the homologous viral genus. In case of partitiviruses, the RNA segment encoding RdRp was named as RNA1 while other RNA segment encoding CP was named RNA2. Based on ICTV's demarcation for the family *Partitiviridae*, we determined novel viruses. For instance, amino acid sequence identity for RdRp and CP was less than 90% and 80%, respectively for species demarcation in the family *Partitiviridae*. For pairwise sequence alignment, we used sequence demarcation tool (SDT) program version 1.2 [16].

Phylogenetic analyses

To reveal phylogenetic relationship of identified viruses, we constructed phylogenetic trees using MEGA7 program [17]. We retrieved top ten viral protein sequences homologous to identified viral proteins such as RdRp and CP. The obtained amino acid sequences with the identified viral protein were together aligned by ClustalW program with default parameter implemented in MEGA7 program. The aligned protein sequences were subjected to construction of phylogenetic tree. The phylogenetic tree was inferred by using the Maximum Likelihood method based on the JTT matrix-based model with a bootstrap of 100 replicates [18].

Results

Identification of *Alloteropsis* cryptic virus 1 and 2 from *Alloteropsis semialata*

We identified six contigs associated with partitiviruses from the transcriptome of *Alloteropsis semialata* (BioProject PRJNA310121), which is a perennial grasses in the family *Poaceae* [19].

The sizes of identified contigs ranged from 1,326 bp to 1,712 bp. The six contigs encode a single ORF. Of them, two contigs encode RdRp while four contigs encode CP. Based on BLASTP search and the identified host, two novel viruses were tentatively named Alloteropsis cryptic virus 1 (AICV1) and Alloteropsis cryptic virus 2 (AICV2). Both AICV1 and AICV2 consists of three RNA fragments: RNA1 (RdRp), RNA2 (CP), and RNA3 (CP). The RNA1 of AICV1 showed sequence similarity to RdRp of Arhar cryptic virus-I (YP_009026407.1), which is an unclassified partitivirus, The RNA2 and RNA3 of AICV1 showed sequence similarity to CPs of White clover cryptic virus 1 (YP_086755.1) [20] and Citrullus lanatus cryptic virus (APT68924.1) [21], respectively. The RNA1 of AICV2 showed sequence similarity to RdRp of Raphanus sativus cryptic virus 3 (YP_002364401.1) [22] whereas RNA2 and RNA3 of AICV2 showed sequence similarity to CPs of Arhar cryptic virus-I (YP_009026398.1 and YP_009026398.2), respectively. The phylogenetic tree based on RdRp amino acid sequences showed that AICV1 RNA1 and AICV2 RNA2 were different from each other. In addition, the RNA2 and RNA3 of AICV1 in the same clade were different from those of AICV2 in the same clade according to the phylogenetic tree using CP amino acid sequences.

Identification of Amaranthus cryptic virus 1–4 from *Alloteropsis semialata*

We identified ten contigs associated with partitiviruses from the transcriptome of *Amaranthus tuberculatus* (BioProject PRJNA432348), which is a weed species in the family *Amaranthaceae* known as waterhemp [23]. The sizes of identified contigs ranged from 667 bp to 2,369 bp. The 10 contigs encode a single open reading frame (ORF). Of them, four contigs encode RNA dependent RNA polymerase (RdRp) while six contigs encode coat protein (CP). Based on BLASTP search and the identified host, the identified four novel viruses were

tentatively named Amaranthus cryptic virus 1 (AmCV1), Amaranthus cryptic virus 2 (AmCV2), Amaranthus cryptic virus 3 (AmCV3), and Amaranthus cryptic virus 4 (AmCV4). RNA1 of AmCV1 showed sequence similarity to RdRp of Persimmon cryptic virus (YP_006390091.1) whereas RNA1 of Pepper cryptic virus 2 showed sequence similarity to RdRp of Pepper cryptic virus 2 (AVL84364.1). RNA2 of AmCV1 displayed sequence similarity to CP of Pepper cryptic virus 1 (AYA43792.1) whereas RNA2 of AmCV2 displayed sequence similarity to CP of Persimmon cryptic virus (YP_006390090.1). For AmCV3, we identified two sequences (isolates Won and Cho) of RNA1 and RNA2. Both AmCV3 RNA1 isolates showed sequence similarity to RdRp of Hop trefoil cryptic virus 2 (YP_007889825.1) whereas both AmCV3 RNA2 showed sequence similarity to CP of Red clover cryptic virus 1 (AWK57379.1). RNA1 of AmCV4 showed sequence similarity to RdRp of Hop trefoil cryptic virus 2 whereas RNA2 of AmCV4 showed sequence similarity to CP of Primula malacoides virus China/Mar2007 (YP_003104769.1) [24]. The phylogenetic tree using RdRp amino sequences for AmCV1 to AmCV4 demonstrated that two distinct groups. The first group contains RNA1 of AmCV1 and AmCV2 whereas the second group includes RNA1 of AmCV3 and AmCV4. Based on phylogenetic analysis, it is likely that RNA1 of AmCV4 is a partial sequence of RNA1 for AmCV3. In case of RNA2 of AmCV1 to AmCV4, all CP sequences are different from each other.

Identification of Ambrosia cryptic virus 1–2 from *Ambrosia trifida*

We identified four contigs associated with partitiviruses from the transcriptome of *Ambrosia trifida*, which is a highly competitive annual weed known as giant ragweed (BioProject: PRJNA267208). The sizes of identified contigs ranged from 1,608 bp to 1,966 bp. The four contigs encode a single open reading frame (ORF). Of them, two contigs encode RdRp while

other two contigs encode CP. Based on BLASTP search and the identified host, the newly identified two viruses were tentatively named Ambrosia cryptic virus 1 (AmbCV1) and Ambrosia cryptic virus 2 (AmbCV2), respectively. RNA1 of AmbCV1 showed sequence similarity to RdRp of Raphanus sativus cryptic virus 4 (ATG29853.1) [25] while RNA1 of AmbCV2 showed sequence similarity to RdRp of Pepper cryptic virus 2 (AVL84364.1). RNA2 of AmbCV1 and AmbCV2 showed sequence similarity to CPs of Red clover cryptic virus 1 (AVL84364.1) and White clover cryptic virus 1 (YP_086755.1), respectively. The phylogenetic tree for AmbCV1 and AmbCV2 using RdRp sequences showed that two viruses are clearly different from each other. In addition, the phylogenetic tree using CP sequences revealed that RNA2 of AmbCV1 were distinctly related with RNA2 of AmbCV2.

Identification of Camellia cryptic virus 1 from *Camellia sinensis*

We identified three contigs associated with partitiviruses from the transcriptome of *Camellia sinensis*, which is an economically important tea plant (BioProject: PRJNA295355). Of them, BLASTP search with three contigs showed sequence similarity to camellia oleifera cryptic virus 1 (GenBank QBQ83754.1) [26], which is a member of the genus *Deltapartitivirus* in the family *Partitiviridae*. The newly identified virus was tentatively named Camellia cryptic virus 1 (CCV1) isolate Won. RNA1 (1,681 bp) of CCV1 encodes an ORF (477 aa) showing sequence similarity to RdRp of Camellia oleifera cryptic virus 1 (QBQ83753.1) with 100% coverage and 92.87% amino acid identity. RNA2 (1,318 bp) of CCV1 encodes an ORF (344 aa) showing sequence similarity to CP2 of camellia oleifera cryptic virus 1 (QBQ83755.1) with 100% coverage and 90.41% amino acid identity. RNA3 (1,410 bp) of CCV1 encodes an ORF (343 aa) showing sequence similarity to CP1 of Camellia oleifera cryptic virus 1 (QBQ83754.1) with 100% coverage and 90.96% amino acid identity. Based on BLASTP results, CCV1 is an

isolate of *Camellia oleifera* cryptic virus 1 in the genus *Deltapartitivirus* in the family *Partitiviridae*.

Identification of *Camellia* cryptic virus 2 from *Camellia reticulata*

We identified five contigs associated with partitiviruses from the transcriptome of *Camellia reticulata* (Bioproject: PRJNA297756), which is well known for its ornamental and high quality oil seeds belonging to the genus *Camellia* in the family *Theaceae* [27]. The newly identified virus was tentatively named *Camellia* cryptic virus 2 (CCV2) isolate Won. CCV2 consisted of five RNA fragments. RNA1 (1,510 bp) of CCV2 encodes an ORF (476 aa) showing sequence similarity to RdRp of Melon partitivirus (YP_009551627.1) with 99% coverage and 65.47% amino acid identity. RNA2 (1,719 bp) of CCV2 encodes an ORF (484 aa) showing sequence similarity to CP of Cucumis melo cryptic virus (QBC66122.1) [28] with 99% coverage and 36% amino acid identity. RNA3 (1,396 bp) of CCV2 encodes an ORF (343 aa) showing sequence similarity to CP1 of *Camellia oleifera* cryptic virus 1 (QBQ83754.1) [26] with 100% coverage and 91.25% amino acid identity. RNA4 (1,383 bp) of CCV2 encodes an ORF (344 aa) showing sequence similarity to CP2 of *camellia oleifera* cryptic virus 1 (QBQ83755.1) with 100% coverage and 89.24% amino acid identity. RNA5 (1,383 bp) of CCV2 encodes an ORF (396 aa) showing sequence similarity to CP of *Clohesyomyces aquaticus* partitivirus 1 (AZT88587.1) [29] with 100% coverage and 53.52% amino acid identity. The phylogenetic tree using RdRp amino acid sequences showed that CCV1 is closely related with the known *Camellia oleifera* cryptic virus 1; however, CCV2 in the same clade with Beet cryptic virus 3 was different from CCV1. The phylogenetic tree using CP amino acid sequences demonstrated that CCV1 RNA3 was closely related with COCV1 RNA2 whereas CCV1 RNA2 was closely related with COCV1 RNA3. However, RNA2 of CCV2 and RNA3

of CCV2 were different from each other. Therefore, CCV2 is a novel virus in the genus *Deltapartitivirus* in the family *Partitiviridae*.

Identification of *Dactylorhiza* cryptic virus 1–3 from *Dactylorhiza incarnata*

We identified six virus-associated contigs associated with partitiviruses from the transcriptome of *Dactylorhiza incarnata* (BioProject PRJNA317244), which is a perennial species belonging to the family *Orchidaceae* [30]. The sizes of contigs ranged from 1,333 bp to 1,590 bp. Each contig encodes an ORF. Three contigs encode RdRp while other three contigs encode CP. The three identified viruses were tentatively named *Dactylorhiza* cryptic virus 1 (DCV1), *Dactylorhiza* cryptic virus 2 (DCV2), and *Dactylorhiza* cryptic virus 3 (DCV3). RNA1 of DCV1 showed sequence similarity to RdRp of Beet cryptic virus 2 (QCF59322.1). RNA2 of DCV2 showed sequence similarity to CP of Persimmon cryptic virus (YP_006390090.1). RNA1 and RNA2 of DCV2 showed sequence similarity to RdRp of Persimmon cryptic virus (YP_006390091.1) and CP of Pepper cryptic virus 1 (BAV93049.1), respectively. RNA1 and RNA2 of DCV3 showed sequence similarity to RdRp of *Sinapis alba* cryptic virus 1 (YP_009255398.1) and CP of *Pittosporum* cryptic virus-1 (CEJ95597.2), respectively. The two different phylogenetic trees using RdRp and CP amino acid sequences, respectively, revealed that DCV1–3 belong to the different groups of partitiviruses.

Identification of *Lomandra* cryptic virus 1 from *Lomandra longifolia*

We identified three virus-associated contigs associated with partitiviruses from the transcriptome of *Lomandra longifolia*, which is an Australian native species resistant to *Phytophthora cinnamomi* (BioProject: PRJNA326999) [31]. Of them, BLASTP search with

three contigs showed sequence similarity to *Camellia oleifera* cryptic virus 1 (GenBank QBQ83754.1) [26], which is a member of the genus *Deltapartitivirus* in the family *Partitiviridae*. The newly identified virus was tentatively named *Lomandra* cryptic virus 1 (LCV1) isolate Won. RNA1 (1,314 bp) of LCV1 encodes an ORF (363 aa) showing sequence similarity to RdRp of *Camellia oleifera* cryptic virus 1 (QBQ83753.1) with 100% coverage and 74.66% amino acid identity. RNA2 (1,350 bp) of LCV1 encodes an ORF (343 aa) showing sequence similarity to CP2 of *Camellia oleifera* cryptic virus 1 (QBQ83755.1) with 95% coverage and 38.60% amino acid identity. RNA3 (1,321 bp) of LCV1 encodes an ORF (345 aa) showing sequence similarity to CP1 of *Camellia oleifera* cryptic virus 1 (QBQ83754.1) with 98% coverage and 51.47% amino acid identity. Based on BLASTP result and phylogenetic analyses, LCV1 is a novel virus in the genus *Deltapartitivirus* in the family *Partitiviridae*. In addition, this is the first report of a partitivirus from *Lomandra longifolia*.

Identification of *Helianthus* cryptic virus 1 from *Helianthus niveus*

We identified eight virus-associated contigs associated with partitiviruses from the transcriptome of *Helianthus niveus*, which is a diploid annual or perennial sunflower known as snowy sunflower in the family *Asteraceae* (BioProject: PRJNA320343) [1]. The sizes of identified contigs ranged from 1,384 bp to 2,399 bp. All contigs encode an ORF. Three contigs encode RdRp while other five contigs encode CP. The newly identified virus was named *Helianthus* cryptic virus 1 (HCV1). RNA1 of HCV1 shows sequence similarity to RdRp of Hop trefoil cryptic virus 2 (YP_007889825.1) while RNA2 of HCV1 shows sequence similarity to CP of *Primula malacoides* virus China/Mar2007 (YP_003104769.1). In addition, we identified two additional isolates: Cho and Kyong for RNA1 of HCV1 and four additional isolates: Cho, Kyong, Yeon, and Sik for RNA2 of HCV1. The phylogenetic trees using RdRp

and CP amino acid sequences showed that all isolates for HCV1 were grouped together.

Identification of *Panax* cryptic virus 1–4 from *Panax notoginseng*

We identified 15 contigs associated with partitiviruses from the two different transcriptomes of *Panax notoginseng*, which is a traditional Chinese medicine belonging to the genus *Panax* in the family *Araliaceae*. Nine contigs were identified from the first study (BioProject PRJNA393585) [32] and six contigs were identified from the second study (BioProject PRJNA472654) [33]. The sizes of identified contigs ranged from 1,013 bp to 1,836 bp. All contigs encode an ORF. From the first study, four different partitiviruses (*Panax* cryptic virus 1–4) (PCV1–4) were identified whereas three different partitiviruses (PCV1–3) were identified from the second study. The phylogenetic tree using RdRp amino acid sequences, RdRp and CP for PCV1–4 were different from each other. For example, RNA1 and RNA2 of PCV2 showed sequence similarity to RdRp (YP_006390091.1) and CP (YP_006390090.1) of Persimmon cryptic virus, respectively. Based on amino acid identity and phylogenetic analyses, we found that PCV1–3 in the second study were different isolates for PCV1–3 from the first study.

Identification of *Rhodiola* cryptic virus 1–2 from *Rhodiola rosea*

We identified 11 contigs associated with partitiviruses from the transcriptome of *Rhodiola rosea*, which is a cold-tolerant perennial plant belonging to the family *Crassulaceae* used as traditional medicine (BioProject PRJNA398393) [34]. The sizes of 11 contigs ranged from 930 bp to 2,730 bp. Each contig encodes an ORF. Two contigs encode RdRp while other nine contigs encode CP. We identified two different partitiviruses were tentatively named *Rhodiola* cryptic virus 1 (RCV1) and *Rhodiola* cryptic virus 2 (RCV2). RCV1 consists of an RdRp and six CPs while RCV2 consists of an RdRp and five CPs. RNA1 of RCV1 showed sequence

similarity to RdRp of Lysoka partiti-like virus (AWV67012.1) while RNA1 of RCV2 showed sequence similarity to RdRp of Medicago sativa deltapartitivirus 1 (ATJ00052.1). The phylogenetic tree using RdRp amino acid sequences showed that RCV1 and RCV2 belong to the different group of partitiviruses. The phylogenetic tree using CP amino acid sequences for RCV1 and RCV2 demonstrated that several groups of CPs. For example, RNA2 of RCV1 was closely related with Medicago sativa alphapartitivirus 1. Both RNA4 and RNA5 of RCV1 showed close genetic relationship with Lysoka partiti-like virus. RNA2–4 of RCV2 were clustered together with RNA6 of RCV1. In addition, RNA1 of RCV2 showed sequence similarity to CP of Raphanus sativus cryptic virus 1.

Identification of *Vicia* cryptic virus from *Vicia faba*

We identified three contigs associated with partitiviruses from the transcriptome of *Vicia faba*, which is a commercially important nitrogen-fixing legume known as faba bean belonging to the family *Fabaceae* [35]. The sizes of three contigs ranged from 1,662 bp to 1,701 bp. Three contigs encode an ORF. Two contigs showed sequence similarity to RdRp (ABN71239.1) and CP (ABN71235.1) of known *Vicia* cryptic virus (VCV), respectively [36]. However, a contig (1,697 bp) showed sequence similarity to CP of Red clover cryptic virus 1. Based on the BLAST result, we named the identified virus as *Vicia* cryptic virus isolate Won. The phylogenetic tree using RdRp amino acid sequences demonstrated that RNA1 of VCV isolate Won was closely related with other isolates of VCV. However, the phylogenetic tree using CP amino acid sequences showed that RNA2 is a new isolate of VCV; however, RNA2 is newly identified RNA segment for VCV.

Summary

In this study, we performed a large scale *in silico* data analyses to identify viruses belonging to the family *Partitiviridae*. We identified a total of 74 RNA segments associated with partitiviruses from 11 different plant species. Detailed information for identified RNA segment was provided in Table 1. We generated 20 different phylogenetic trees to reveal phylogenetic relationships of newly identified partitiviruses with known partitiviruses (Supplementary Figure). In addition, we provided BLASTP results of individual ORF for identified partitiviruses (Supplementary Table 1). Of the 11 plant species, *Panax notoginseng* (11 RNA segments representing four different partitiviruses) and *Amaranthus tuberculatus* (10 RNA segments representing four different partitiviruses) contain the largest number of RNA segments associated with partitiviruses (Table 2). Except eight RNA segments associated with *Camellia oleifera* cryptic virus 1 and *Vicia* cryptic virus, all RNA segments were novel. The phylogenetic tree using 25 RdRp amino acid sequences revealed four different groups of identified partitiviruses (Figure 1a) whereas the phylogenetic tree using 41 CP amino acid sequences showed five groups of partitiviruses. Taken together, this study provides the largest number of novel partitiviruses in a single study.

Declarations

Ethics approval and consent to participate

Not applicable

Availability of data and materials

Viral genome sequences for identified viruses were deposited in GenBank with respective accession number provided in Table 1.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korea government Ministry of Education (No. NRF-2018R1D1A1B07043597) and the support of the “Cooperative Research Program for Agriculture Science & Technology Development” (PJ01498301) conducted by the Rural Development Administration, Republic of Korea.

References

1. Vainio EJ, Chiba S, Ghabrial SA, Maiss E, Roossinck M, Sabanadzovic S, Suzuki N, Xie J, Nibert M: **ICTV virus taxonomy profile: Partitiviridae**. *Journal of General Virology* 2018, **99**(1):17-18.
2. Sabanadzovic S, Valverde RA: **Properties and detection of two cryptoviruses from pepper (*Capsicum annuum*)**. *Virus Genes* 2011, **43**(2):307-312.
3. Elbeaino T, Kubaa RA, Digiaro M, Minafra A, Martelli GP: **The complete nucleotide sequence and genome organization of Fig cryptic virus, a novel bipartite dsRNA virus infecting fig, widely distributed in the Mediterranean basin**. *Virus Genes* 2011, **42**(3):415-421.
4. Xie W, Antoniw J, White R: **Nucleotide sequence of beet cryptic virus 3 dsRNA2 which encodes a putative RNA-dependent RNA polymerase**. *Journal of General Virology* 1993, **74**(7):1467-1470.

5. Sela N, Lachman O, Reingold V, Dombrovsky A: **A new cryptic virus belonging to the family Partitiviridae was found in watermelon co-infected with Melon necrotic spot virus.** *Virus Genes* 2013, **47**(2):382-384.
6. Elbeaino T, Kubaa RA, Tuzlali HT, Digiario M: **Pittosporum cryptic virus 1: genome sequence completion using next-generation sequencing.** *Archives of Virology* 2016, **161**(7):2039-2042.
7. Chun J, Yang H-E, Kim D-H: **Identification and molecular characterization of a novel partitivirus from *Trichoderma atroviride* NCF394.** *Viruses* 2018, **10**(11):578.
8. Liu H, Fu Y, Xie J, Cheng J, Ghabrial SA, Li G, Yi X, Jiang D: **Discovery of novel dsRNA viral sequences by in silico cloning and implications for viral diversity, host range and evolution.** *PLoS One* 2012, **7**(7).
9. Jo Y, Choi H, Kim S-M, Kim S-L, Lee BC, Cho WK: **The pepper virome: natural co-infection of diverse viruses and their quasispecies.** *BMC Genomics* 2017, **18**(1):453.
10. Jo Y, Choi H, Kim S-M, Kim S-L, Lee BC, Cho WK: **Integrated analyses using RNA-Seq data reveal viral genomes, single nucleotide variations, the phylogenetic relationship, and recombination for Apple stem grooving virus.** *BMC Genomics* 2016, **17**(1):579.
11. Jo Y, Cho WK: **RNA viromes of the oriental hybrid lily cultivar “Sorbonne”.** *BMC Genomics* 2018, **19**(1):748.
12. Jo Y, Choi H, Cho JK, Yoon J-Y, Choi S-K, Cho WK: **In silico approach to reveal viral populations in grapevine cultivar Tannat using transcriptome data.** *Scientific reports* 2015, **5**:15841.
13. Kim H, Park D, Hahn Y: **Identification of novel RNA viruses in alfalfa (*Medicago sativa*): an Alphapartitivirus, a Deltapartitivirus, and a Marafivirus.** *Gene* 2018,

- 1 **638:7-12.**
- 2 14. Jo Y, Choi H, Chu H, Cho WK: **Identification of viruses from fungal transcriptomes.**
- 3 *BioRxiv* 2020.
- 4 15. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan
- 5 L, Raychowdhury R, Zeng QJNb: **Full-length transcriptome assembly from RNA-**
- 6 **Seq data without a reference genome.** 2011, **29(7):644.**
- 7 16. Muhire BM, Varsani A, Martin DP: **SDT: a virus classification tool based on pairwise**
- 8 **sequence alignment and identity calculation.** *PLoS One* 2014, **9(9).**
- 9 17. Kumar S, Stecher G, Tamura K: **MEGA7: molecular evolutionary genetics analysis**
- 10 **version 7.0 for bigger datasets.** *Molecular Biology Evolution* 2016, **33(7):1870-**
- 11 1874.
- 12 18. Jones DT, Taylor WR, Thornton JM: **The rapid generation of mutation data matrices**
- 13 **from protein sequences.** *Bioinformatics* 1992, **8(3):275-282.**
- 14 19. Dunning LT, Lundgren MR, Moreno-Villena JJ, Namaganda M, Edwards EJ, Nosil P,
- 15 Osborne CP, Christin PA: **Introgression and repeated co-option facilitated the**
- 16 **recurrent emergence of C4 photosynthesis among close relatives.** *Evolution* 2017,
- 17 **71(6):1541-1555.**
- 18 20. Boccardo G, Candresse T: **Complete sequence of the RNA2 of an isolate of White**
- 19 **clover cryptic virus 1, type species of the genus Alphacryptovirus.** *Archives of*
- 20 *Virology* 2005, **150(2):403-405.**
- 21 21. Xin M, Cao M, Liu W, Ren Y, Lu C, Wang X: **The genomic and biological**
- 22 **characterization of Citrullus lanatus cryptic virus infecting watermelon in China.**
- 23 *Virus Research* 2017, **232:106-112.**

22. Chen L, Chen J, Liu L, Yu X, Yu S, Fu T, Liu W: **Complete nucleotide sequences and genome characterization of double-stranded RNA 1 and RNA 2 in the *Raphanus sativus*-root cv. Yipinghong.** *Archives of Virology* 2006, **151**(5):849-859.
23. Kohlhase DR, O'Rourke JA, Owen MD, Graham MA: **Using RNA-seq to characterize responses to 4-hydroxyphenylpyruvate dioxygenase (HPPD) inhibitor herbicide resistance in waterhemp (*Amaranthus tuberculatus*).** *BMC Plant Biology* 2019, **19**(1):182.
24. Li L, Tian Q, Du Z, Duns GJ, Chen J: **A novel double-stranded RNA virus detected in *Primula malacoides* is a plant-isolated partitivirus closely related to partitivirus infecting fungal species.** *Archives of Virology* 2009, **154**(4):565-572.
25. Svanella-Dumas L, Theil S, Barret M, Candresse T: **Complete genomic sequence of *Raphanus sativus* cryptic virus 4 (RsCV4), a novel alphapartitivirus from radish.** *Archives of Virology* 2018, **163**(4):1097-1100.
26. Wu L-P, Du Y-M, Xiao H, Peng L, Li R: **Complete genomic sequence of tea-oil camellia deltapartitivirus 1, a novel virus from *Camellia oleifera*.** *Archives of Virology* 2020, **165**(1):227-231.
27. Yao Q-Y, Huang H, Tong Y, Xia E-H, Gao L-Z: **Transcriptome analysis identifies candidate genes related to triacylglycerol and pigment biosynthesis and photoperiodic flowering in the ornamental and oil-producing plant, *Camellia reticulata* (Theaceae).** *Frontiers in Plant Science* 2016, **7**:163.
28. Zhan B, Cao M, Wang K, Wang X, Zhou X: **Detection and Characterization of *Cucumis melo* Cryptic Virus, *Cucumis melo* Amalgavirus 1, and Melon Necrotic Spot Virus in *Cucumis melo*.** *Viruses* 2019, **11**(1):81.
29. Gilbert KB, Holcomb EE, Allscheid RL, Carrington JC: **Hiding in plain sight: New**

- 1 **virus genomes discovered via a systematic analysis of fungal public transcriptomes.**
- 2 *PLoS One* 2019, **14**(7).
- 3 30. Balao F, Trucchi E, Wolfe TM, Hao BH, Lorenzo MT, Baar J, Sedman L, Kosiol C,
- 4 Amman F, Chase MW: **Adaptive sequence evolution is driven by biotic stress in a**
- 5 **pair of orchid species (Dactylorhiza) with distinct ecological optima.** *Molecular*
- 6 *Ecology* 2017, **26**(14):3649-3662.
- 7 31. Islam M, Hussain H, Rookes J, Cahill D: **Transcriptome analysis, using RNA-Seq of**
- 8 **Lomandra longifolia roots infected with Phytophthora cinnamomi reveals the**
- 9 **complexity of the resistance response.** *Plant Biology* 2018, **20**(1):130-142.
- 10 32. Zheng Y, Chen K, Xu Z, Liao P, Zhang X, Liu L, Wei K, Liu D, Li Y-F, Sunkar R: **Small**
- 11 **RNA profiles from Panax notoginseng roots differing in sizes reveal correlation**
- 12 **between miR156 abundances and root biomass levels.** *Scientific Reports* 2017,
- 13 **7**(1):1-15.
- 14 33. Liu D, Zhao Q, Cui X, Chen R, Li X, Qiu B, Ge F: **A transcriptome analysis uncovers**
- 15 **Panax notoginseng resistance to Fusarium solani induced by methyl jasmonate.**
- 16 *Genes and Genomics* 2019, **41**(12):1383-1396.
- 17 34. Torrens-Spence MP, Pluskal T, Li F-S, Carballo V, Weng J-K: **Complete pathway**
- 18 **elucidation and heterologous reconstitution of Rhodiola salidroside biosynthesis.**
- 19 *Molecular Plant* 2018, **11**(1):205-217.
- 20 35. Webb A, Cottage A, Wood T, Khamassi K, Hobbs D, Gostkiewicz K, White M, Khazaei
- 21 H, Ali M, Street D: **A SNP-based consensus genetic map for synteny-based trait**
- 22 **targeting in faba bean (Vicia faba L.).** *Plant Biotechnology Journal* 2016, **14**(1):177-
- 23 185.
- 24 36. Blawid R, Stephan D, Maiss E: **Molecular characterization and detection of Vicia**

- 1 **cryptic virus in different Vicia faba cultivars. *Archives of Virology* 2007,**
- 2 **152(8):1477-1488.**
- 3

- 1 Table 1. Detailed information for identified RNA segments associated with partitiviruses.
- 2 Individual viral sequence was deposited in NCBI's GenBank with respective accession number (Acc. No.). The name of plant host, virus name,
- 3 size of nucleotide sequences, ORF position, and size of viral protein were provided.

Index	Acc. No.	Host	Virus name	Size of nucleotides	ORF position	Size of protein
1	MT159291	<i>Alloteropsis semialata</i>	Alloteropsis_cryptic_virus_1_RNA1_Won_1700	1712	100–1602	500
2	MH898469		Alloteropsis_cryptic_virus_1_RNA2_Won_1707	1707	103–1569	488
3	MT159292		Alloteropsis_cryptic_virus_1_RNA3_Won_1418	1418	85–1302	405
4	MT159293		Alloteropsis_cryptic_virus_2_RNA1_Won_1585	1585	80–1525	481
5	MT159294		Alloteropsis_cryptic_virus_2_RNA2_Won_1483	1483	148–1224	358
6	MT159295		Alloteropsis_cryptic_virus_2_RNA3_Won_1326	1326	65–1108	347
7	MH898471	<i>Amaranthus tuberculatus</i>	Amaranthus_cryptic_virus_1_RNA1_Won_1581	1581	74–1507	477
8	MH898472		Amaranthus_cryptic_virus_1_RNA2_Won_1561	1561	102–1400	432
9	MH898473		Amaranthus_cryptic_virus_2_RNA1_Won_1543	1543	79–1521	480
10	MH898474		Amaranthus_cryptic_virus_2_RNA2_Won_1507	1507	44–1309	421
11	MT159296		Amaranthus_cryptic_virus_3_RNA1_Won_2369	2369	69–2273	734
12	MT159298		Amaranthus_cryptic_virus_3_RNA1_Cho_878	878	69–878	270
13	MH898475		Amaranthus_cryptic_virus_3_RNA2_Won_1686	1686	68–1531	487
14	MT159299		Amaranthus_cryptic_virus_3_RNA2_Cho_1163	1163	56–1069	337
15	MT159300		Amaranthus_cryptic_virus_4_RNA1_Won_667	667	69–667	199
16	MT159297		Amaranthus_cryptic_virus_4_RNA2_Won_2211	2211	85–2142	685

17	MH898476	<i>Ambrosia trifida</i>	Ambrosia_cryptic_virus_1_RNA1_Won_1966	1966	62–1912	616
18	MH898477		Ambrosia_cryptic_virus_1_RNA2_Won_1683	1683	92–1522	476
19	MT159301		Ambrosia_cryptic_virus_2_RNA1_Won_1608	1608	94–1548	484
20	MH898480		Ambrosia_cryptic_virus_2_RNA2_Won_1700	1700	86–1552	488
21	MH898482	<i>Camellia sinensis</i>	Camellia_cryptic_virus_1_RNA1_Won_1681	1681	157–1590	477
22	MH898483		Camellia_cryptic_virus_1_RNA2_Won_1318	1318	192–1226	344
23	MH898484		Camellia_cryptic_virus_1_RNA3_Won_1410	1410	164–1195	343
24	MT159304		Camellia_cryptic_virus_2_RNA1_Won_1510	1510	81–1510	476
25	MT159305		Camellia_cryptic_virus_2_RNA2_Won_1584	1584	69–1523	484
26	MT159302		Camellia_oleifera_cryptic_virus_1_RNA2_Won_1087	1087	45–1076	343
27	MT159303		Camellia_oleifera_cryptic_virus_1_RNA3_Won_1383	1383	189–1223	344
28	MT159306		Camellia_cryptic_virus_2_RNA3_Won_1329	1329	44–1234	410
29	MH898486	<i>Dactylorhiza incarnata</i>	Dactylorhiza_cryptic_virus_1_RNA1_Won_1590	1590	150–1568	472
30	MH898488		Dactylorhiza_cryptic_virus_1_RNA2_Won_1427	1427	50–1279	409
31	MH898487		Dactylorhiza_cryptic_virus_2_RNA1_Won_1582	1582	81–1529	482
32	MH898489		Dactylorhiza_cryptic_virus_2_RNA2_Won_1498	1498	76–1356	426
33	MT159307		Dactylorhiza_cryptic_virus_3_RNA1_Won_1410	1410	88–1410	441
34	MT159308		Dactylorhiza_cryptic_virus_3_RNA2_Won_1333	1333	78–1307	409
35	MH898492	<i>Helianthus niveus</i>	Helianthus_cryptic_virus_1_RNA1_Won_2399	2399	114–2354	746
36	MT159312		Helianthus_cryptic_virus_1_RNA1_Cho_1384	1384	72–1295	407
37	MT159309		Helianthus_cryptic_virus_1_RNA1_Kyong_1384	1384	72–1295	407
38	MH898494		Helianthus_cryptic_virus_1_RNA2_Cho_2249	2249	59–2104	681
39	MT159311		Helianthus_cryptic_virus_1_RNA2_Won_1821	1821	1–1821	606
40	MT159310		Helianthus_cryptic_virus_1_RNA2_Kyong_2156	2156	59–2104	681

41	MT159314		Helianthus_cryptic_virus_1_RNA2_Yeon_1584	1584	59–1584	508
42	MT159313		Helianthus_cryptic_virus_1_RNA2_Sik_1584	1584	59–1584	508
43	MH898498	<i>Lomandra longifolia</i>	Lomandra_cryptic_virus_1_RNA1_Won_1314	1314	219–1310	370
44	MH898496		Lomandra_cryptic_virus_1_RNA2_Won_1350	1350	205–1236	343
45	MH898499		Lomandra_cryptic_virus_1_RNA3_Won_1321	1321	171–1208	345
46	MH898504	<i>Panax notoginseng</i>	Panax_cryptic_virus_1_RNA1_Won_1570	1570	94–1527	477
47	MH898506		Panax_cryptic_virus_1_RNA2_Won_1528	1528	86–1324	412
48	MH898509		Panax_cryptic_virus_2_RNA1_Won_1552	1552	88–1521	477
49	MH898510		Panax_cryptic_virus_2_RNA2_Won_1491	1491	57–1319	420
50	MT159318		Panax_cryptic_virus_3_RNA1_Won_1720	1720	8–1720	571
51	MT159319		Panax_cryptic_virus_3_RNA2_Won_1836	1836	84–1571	495
52	MT159322		Panax_cryptic_virus_4_RNA3_Won_1609	1609	74–1549	491
53	MT159320		Panax_cryptic_virus_4_RNA1_Won_1115	1115	37–1115	359
54	MT159321		Panax_cryptic_virus_4_RNA2_Won_1013	1013	71–982	303
55	MH898505		Panax_cryptic_virus_1_RNA1_Cho_1576	1576	96–1529	477
56	MH898507		Panax_cryptic_virus_1_RNA2_Cho_1536	1536	87–1325	412
57	MH898508		Panax_cryptic_virus_2_RNA1_Cho_1550	1550	85–1518	477
58	MH898511		Panax_cryptic_virus_2_RNA2_Cho_1482	1482	48–1310	420
59	MT159316	<i>Rhodiola rosea</i>	Panax_cryptic_virus_3_RNA1_Cho_1437	1437	57–1437	460
60	MT159317		Panax_cryptic_virus_3_RNA2_Cho_1804	1804	84–1571	495
61	MH898519		Rhodiola_cryptic_virus_1_RNA1_Won_1559	1559	113–1528	471
62	MT159324		Rhodiola_cryptic_virus_1_RNA2_Won_2730	2730	127–2544	805
63	MT159325		Rhodiola_cryptic_virus_1_RNA3_Won_1373	1373	74–1306	410
64	MH898521		Rhodiola_cryptic_virus_1_RNA4_Won_1580	1580	244–1320	358

65	MH898522		Rhodiola_cryptic_virus_1_RNA5_Won_1290	1290	177–1217	346
66	MT159330		Rhodiola_cryptic_virus_1_RNA6_Won_1410	1410	65–1354	429
67	MH898520		Rhodiola_cryptic_virus_2_RNA1_Won_1584	1584	80–1513	477
68	MT159326		Rhodiola_cryptic_virus_2_RNA2_Won_1361	1361	56–1306	416
69	MT159327		Rhodiola_cryptic_virus_2_RNA3_Won_1469	1469	76–1317	413
70	MT159328		Rhodiola_cryptic_virus_2_RNA4_Won_930	930	90–806	238
71	MT159329		Rhodiola_cryptic_virus_2_RNA5_Won_1037	1037	35–1037	345
72	MT159332		Vicia_cryptic_virus_RNA1_Won_1701	1701	23–1701	559
73	MT159333	<i>Vicia faba</i>	Vicia_cryptic_virus_RNA2_Won_1662	1662	107–1570	487
74	MH898526		Vicia_cryptic_virus_1_RNA3_Won_1697	1697	100–1554	484

1

1 Table 2. Summary of number of identified RNA segments, number of viruses, number of RdRp
2 and number of CP in individual plant species.

Plant species	No. of RNA	No. of viruses	No. of RdRp	No. of CP
<i>Alloteropsis semialata</i>	6	2	2	4
<i>Amaranthus tuberculatus</i>	10	4	5	5
<i>Ambrosia trifida</i>	4	2	2	2
<i>Camellia reticulata</i>	5	2	1	4
<i>Camellia sinensis</i>	3	1	1	2
<i>Dactylorhiza incarnata</i>	6	3	3	3
<i>Helianthus niveus</i>	8	1	3	5
<i>Lomandra longifolia</i>	3	1	1	2
<i>Panax notoginseng</i>	15	4	7	8
<i>Rhodiola rosea</i>	11	2	2	9
<i>Vicia faba</i>	3	1	1	2

3

4

5

6

Figure 1. Phylogenetic relationships of identified viruses based on RdRp (a) and coat protein (CP) (b) amino acid sequences. The phylogenetic trees were constructed using the Maximum Likelihood method based on the JTT matrix-based model by MEGA 7.0 with a bootstrap of 100 replicates.

(a)

1

2

Supplementary Figures

Phylogenetic relationships of identified viruses and other members of the family *Partitiviridae* based on RdRp and CP amino acid sequences. The phylogenetic trees were constructed using the Maximum Likelihood method based on the JTT matrix-based model by MEGA 7.0 with a bootstrap of 100 replicates. The orange colored boxes indicate the identified partitiviruses from this study.

Supplementary Table

Table S1. BLASTP results of identified viral proteins associated with partitiviruses.

The individual ORF proteins were subjected to BLASTP search against NCBI's non redundant protein database. The best matched protein for individual viral protein with amino acid sequence coverage, identity, and accession number were provided.