Acta virologica **62**: 304 – 309, 2018 doi:10.4149/av_2018_224

# Novel *Divavirus* (the family *Betaflexiviridae*) and *Mitovirus* (the family *Narnaviridae*) species identified in basil (*Ocimum basilicum*)

C. J. GOH[1], D. PARK[1], H. KIM[1], F. SEBASTIANI[2], Y. HAHN[1*]

[1]Department of Life Science, Research Center for Biomolecules and Biosystems, Chung-Ang University, Seoul 06974, South Korea; [2]Institute for Sustainable Plant Protection, Department of Biology, Agriculture and Food Sciences, The National Research Council of Italy, Sesto Fiorentino, Italy

*Received October 3, 2017; revised December 19, 2017; accepted July 10, 2018*

**Summary.** – Transcriptome data obtained from a plant sample often contain a large number of reads that are derived from associated RNA virus genomes that were co-isolated during RNA preparation. These virus-derived reads can be assembled into a novel plant RNA genome sequence. Here, a basil (*Ocimum basilicum*) transcriptome dataset was analyzed to identify two new RNA viruses, which were named Ocimum basilicum RNA virus 1 (ObRV1) and Ocimum basilicum RNA virus 2 (ObRV2). A phylogenetic analysis of the ObRV1 RNA-dependent RNA polymerase (RdRp) motif indicated that ObRV1 is a novel species of the genus *Divavirus* of the family *Betaflexiviridae*. ObRV1 is the fourth divavirus species to be identified. The ObRV2 RdRp motif showed sequence similarity to viruses of the genus *Mitovirus* of the family *Narnaviridae*, which infect fungal mitochondria. Although most of the known mitoviruses do not produce a functional RdRp using the plant mitochondrial genetic code, the ObRV2 encodes a full-length RdRp using both the fungal and plant mitochondrial genetic codes.

**Keywords:** basil; Ocimum basilicum RNA virus 1; Ocimum basilicum RNA virus 2; Divavirus; Mitovirus

## Introduction

Basil (*Ocimum basilicum* L.) is a popular herb that belongs to the genus *Ocimum* of the family *Lamiaceae*. It is widely cultivated in Asia, the Mediterranean region, and the United States of America for commercial and industrial purposes (Dudai *et al.*, 2002; Suppakul *et al.*, 2003; Chiang *et al.*, 2005). Basil is also used as a medicinal, antiviral, antibacterial, and antimicrobial agent, with its essential oils being known to have many beneficial compounds, such as monoterpenoids, which act against pathogens such as virus, bacteria, and fungi (Edris and Farrag, 2003; Suppakul *et al.*, 2003; Chiang *et al.*, 2005).

Transcriptome analyses of basil have been conducted to investigate its genetic features, such as the genes involved in secondary metabolism and production of unique essential oils (Rastogi *et al.*, 2014; Torre *et al.*, 2016). The plant transcriptome data often contain sequences derived from RNA viruses, which can be identified to be novel plant RNA viruses (Nouri *et al.*, 2016; Visser *et al.*, 2016; Park and Hahn, 2017a,b). In this study, a previously reported transcriptome dataset collected from red and green morphs of sweet basil was analyzed and potential sequences of two novel RNA viruses, a divavirus and a mitovirus, were identified (Torre *et al.*, 2016).

Currently, only three species are reported in the genus *Divavirus* of the family *Betaflexiviridae* (Wylie and Jones, 2011; Wylie *et al.*, 2013). The *Betaflexviridae* viruses have a linear, single-stranded, positive-sense RNA genomic segment 6.5–9.0 kb in size, which encodes 2–6 proteins, including the RNA-dependent RNA polymerase (RdRp) and several other proteins depending on the genus (Adams, 2012; Igori *et al.*, 2016).

---

[*] Corresponding author. E-mail: hahny@cau.ac.kr; phone: +82-2-820-5812.

**Abbreviations:** ObRV1 = Ocimum basilicum virus 1; ObRV2 = Ocimum basilicum virus 2; ORF = open reading frame; RdRp = RNA-dependent RNA polymerase; SRA = Sequence Read Archive

**Table 1. Summary of RNA viruses identified in this study**

| Acronym | Full name | Acc. No. | Length (nt) | ORF | Position | Length (aa) |
|---|---|---|---|---|---|---|
| ObRV1 | Ocimum basilicum RNA virus 1 | MF196913 | 6930 | Replicase/Coat protein | 44–6721 | 2255 |
| | | | | Movement protein | 5050–6219 | 389 |
| ObRV2 | Ocimum basilicum RNA virus 2 | MF196914 | 2784 | Replicase | 347–2629 | 760 |

Viruses of the family *Narnaviridae,* to which the genus *Mitovirus* belongs, have an unencapsidated, linear, single-stranded, positive-sense RNA genome 2.5–2.9 kb in size, which encodes a single open reading frame (ORF) for the replicase containing the RdRp motif (Hillman and Cai, 2013). Mitoviruses replicate in the cytosol or mitochondria of infected fungi and have been mainly isolated from plant-pathogenic fungi (Ghabrial and Suzuki, 2009; Xie and Ghabrial, 2012).

## Materials and Methods

The previously reported basil transcriptome dataset (a total of 13 gigabase pairs), which was isolated from red and green sweet basils, was analyzed in this study (Torre *et al.*, 2016). The transcriptome data is available in the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI) under the accession number SRA313233. Sickle program (version 1.33; https://github.com/najoshi/sickle) was used to collect high-quality reads with the option "-q 30 -l 55." The SPAdes Genome Assembler (version 3.10.1) was used to assemble the RNA-seq reads into contigs (Bankevich *et al.*, 2012).

The RdRp motif sequences of known RNA viruses were retrieved from the Pfam database (release 30.0; http://pfam.xfam.org). A total of 345 non-redundant RdRp motif sequences were collected from 19 RdRp families: their accession numbers are PF00602, PF00603, PF00604, PF00680, PF00946, PF00972, PF00978, PF00998, PF02123, PF03431, PF04196, PF04197, PF05788, PF05919, PF07925, PF08467, PF08716, PF08717, and PF12426.

The assembled basil transcript contigs were compared with the Pfam-derived RdRp sequences, using BLAST (Altschul *et al.*, 1990) with the parameter "-evalue e-5." Matched contigs 1000 bp or longer were selected. Annotation of contigs and identification of closely related viruses were carried out by BLAST searches of all viral genomes and proteins at the NCBI website (https://www.ncbi.nlm.nih.gov/BLAST). The basil RNA-seq reads were mapped to the viral contigs, using BWA, and variants were called using Samtools/Bcftools (Li and Durbin, 2009; Li, 2011).

For phylogenetic analysis, the RdRp motif-containing sequences were multiply aligned using the MUSCLE program (Edgar, 2004). The aligned region spanning the core RdRp motif defined by a Pfam analysis was extracted and subjected to phylogenetic analysis. A phylogenetic tree was constructed by the neighbor-joining method implemented in the ClustalW2 software (version 2.1) (Larkin *et al.*, 2007).

## Results and Discussion

Genome sequences of two new viruses (Table 1) were identified by analysis of the basil transcriptome data (SRA accession number SRA313233) (Torre *et al.*, 2016). One viral contig, named Ocimum basilicum RNA virus 1 (ObRV1), was most similar to those of divaviruses (Table 2) and considered a novel member of the genus *Divavirus*. The other contig, named Ocimum basilicum RNA virus 2 (ObRV2), showed sequence similarity to those of mitoviruses (Table 2), suggesting that it belongs to the genus *Mitovirus*.

### *Ocimum basilicum RNA virus 1 (ObRV1)*

The first of the two viral contigs discovered in the basil transcriptome was named Ocimum basilicum RNA virus 1 (ObRV1) and showed sequence similarity to those of plant RNA viruses of the genus *Divavirus* of the family *Betaflexiviridae*. There were sequence variations at 33 positions when the basil RNA-seq reads mapped to the ObRV1 genome were analyzed, indicating that the contig is a composite sequence derived from at least two closely related viruses (Supplementary Table S1).

The ObRV1 RNA genome is 6,930 base pairs (bp) long and predicted to have two overlapping ORFs. ORF1 encodes a 2,255-amino acid (aa)-long RdRp motif-containing protein and ORF2 a 389-aa-long movement protein (Table 1). A domain analysis using Pfam showed that the ObRV1 ORF1 protein contains four known domains: viral methyltransferase (PF01660) at aa 43–357, viral RNA helicase (PF01443) at aa 858–1,119, RdRp (PF00978) at aa 1,255–1,653, and trichovirus coat protein (PF05892) at aa 2,039–2,223. These genome and domain organizations are the shared features of viruses in the genera *Divavirus* and *Capillovirus* (Jelkmann, 1995; Wylie and Jones, 2011).

The ObRV1 RdRp showed 30–60% aa sequence identities with RdRps of the viruses that belong to the family *Betaflexiviridae* (Table 2). Among them, Hardenbergia virus A (HarVA), Diuris virus A (DiVA), and Diuris virus B (DiVB) were the top three similar viruses with aa sequence identities of 50–60%. These three viruses are species in the genus *Divavirus*, suggesting that ObRV1 is a new divavirus.

A phylogenetic analysis using multiply aligned RdRp motif sequences confirmed that ObRV1 is a new species of the genus *Divavirus* (Supplementary Fig. S1 and Fig. 1a).

**Table 2. Identities among the RdRp sequences of ObRV1, ObRV2, and their respective related viruses**

| | **Acronym** | **Full name** | **Acc. No.**[a] | **Sequence identity**[b] |
|---|---|---|---|---|
| ObRV1 | HarVA | Hardenbergia virus A | YP_004376201.1 | 1341/2237 (60%) |
| | DiVA | Diuris virus A | YP_006905850.1 | 1243/2248 (55%) |
| | DiVB | Diuris virus B | YP_006905848.1 | 1196/2256 (53%) |
| | CVA | Cherry virus A | ANE06570.1 | 357/949 (38%) |
| | CVA | Cherry virus A | APG53775.1 | 331/848 (39%) |
| | PCMV | Peach chlorotic mottle virus | YP_001497153.1 | 336/928 (36%) |
| | ApLV | Apricot latent virus | ADT91605.1 | 334/910 (37%) |
| | NeLV | Nerine latent virus | YP_009174681.1 | 335/1030 (33%) |
| | NSV | Narcissus symptomless virus | YP_842438.1 | 336/1030 (33%) |
| | CTLaV | Cherry twisted leaf associated virus | AHJ80270.1 | 326/840 (39%) |
| | CTLaV | Cherry twisted leaf associated virus | AHJ80314.1 | 326/859 (38%) |
| | CNRMV | Cherry necrotic rusty mottle virus | ALP45953.1 | 338/909 (37%) |
| | CRMaV | Cherry rusty mottle associated virus | AHA59466.2 | 323/846 (38%) |
| | CRMaV | Cherry rusty mottle associated virus | AHJ80324.1 | 317/839 (38%) |
| | CRMaV | Cherry rusty mottle associated virus | AKN20442.1 | 347/971 (36%) |
| | CGRMV | Cherry green ring mottle virus | AFU54620.1 | 318/825 (39%) |
| | BanMMV | Banana mild mosaic virus | NP_112029.1 | 321/870 (37%) |
| ObRV2 | BcMV1 | Botrytis cinerea mitovirus 1 | CEZ26300.1 | 165/506 (33%) |
| | FpMV4 | Fusarium poae mitovirus 4 | YP_009272901.1 | 104/224 (46%) |
| | OMV3a | Ophiostoma mitovirus 3a | NP_660176.1 | 101/241 (42%) |
| | SsMV3 | Sclerotinia sclerotiorum mitovirus 3 | YP_009182164.1 | 110/283 (39%) |
| | MpMV3 | Macrophomina phaseolina mitovirus 3 | AMM45292.1 | 100/273 (37%) |
| | RsMV-17 | Rhizoctonia solani mitovirus 17 | ANR02693.1 | 91/230 (40%) |
| | RsMV-15 | Rhizoctonia solani mitovirus 15 | ALD89120.1 | 100/276 (36%) |
| | RsMV-14 | Rhizoctonia solani mitovirus 14 | ALD89119.1 | 102/297 (34%) |
| | HetMV1 | Heterobasidion mitovirus 1 | AIF33766.2 | 158/547 (29%) |
| | COV | Clitocybe odora virus | YP_005352912.1 | 174/577 (30%) |
| | BcMV3 | Botrytis cinerea mitovirus 3 | YP_009182161.1 | 113/304 (37%) |
| | FpMV3 | Fusarium poae mitovirus 3 | YP_009272900.1 | 87/219 (40%) |
| | SlaMV5 | Soybean leaf-associated mitovirus 5 | ALM62240.1 | 87/233 (37%) |
| | RsMV-19 | Rhizoctonia solani mitovirus 19 | ANR02690.1 | 87/221 (39%) |
| | BsMV1 | Buergenerula spartinae mitovirus 1 | AHY03257.1 | 91/242 (38%) |
| | CpMV1 | Cryphonectria parasitica mitovirus 1-NB631 | NP_660174.1 | 79/216 (37%) |
| | RcMV | Rhizoctonia cerealis mitovirus | AIT71973.1 | 89/241 (37%) |
| | SsMV15 | Sclerotinia sclerotiorum mitovirus 15 | AHF48631.1 | 81/208 (39%) |
| | HfMV1 | Hymenoscyphus fraxineus mitovirus 1 | AIU44705.1 | 81/208 (39%) |
| | SsMV16 | Sclerotinia sclerotiorum mitovirus 16 | AHF48632.1 | 89/240 (37%) |
| | SsMV2 | Sclerotinia sclerotiorum mitovirus 2 | AHX84129.1 | 88/232 (38%) |
| | SlaMV3 | Soybean leaf-associated mitovirus 3 | ALM62243.1 | 76/203 (37%) |
| | OMV5 | Ophiostoma mitovirus 5 | NP_660180.1 | 82/213 (38%) |
| | SlaMV2 | Soybean leaf-associated mitovirus 2 | ALM62242.1 | 78/209 (37%) |
| | FgMV1 | Fusarium globosum mitovirus 1 | YP_009126872.1 | 110/367 (30%) |
| | TbMV | Thielaviopsis basicola mitovirus | AAZ99833.1 | 127/435 (29%) |

[a]Acc. No. for the RdRp motif-containing protein sequence. [b]Amino acid sequence identities in a format of “identical residues/aligned length (% identity).”

The ObRV1 was clustered with the three known divaviruses, with HarVA being the closest species. Currently, there are three known species of the genus *Divavirus* (Wylie and Jones, 2011; Wylie *et al.*, 2013); ObRV1 is the fourth divavirus species to be reported.

### *Ocimum basilicum RNA virus 2 (ObRV2)*

The second viral contig was named Ocimum basilicum RNA virus 2 (ObRV2) and showed sequence similarity to those of viruses of the genus *Mitovirus* of the family *Nar-*

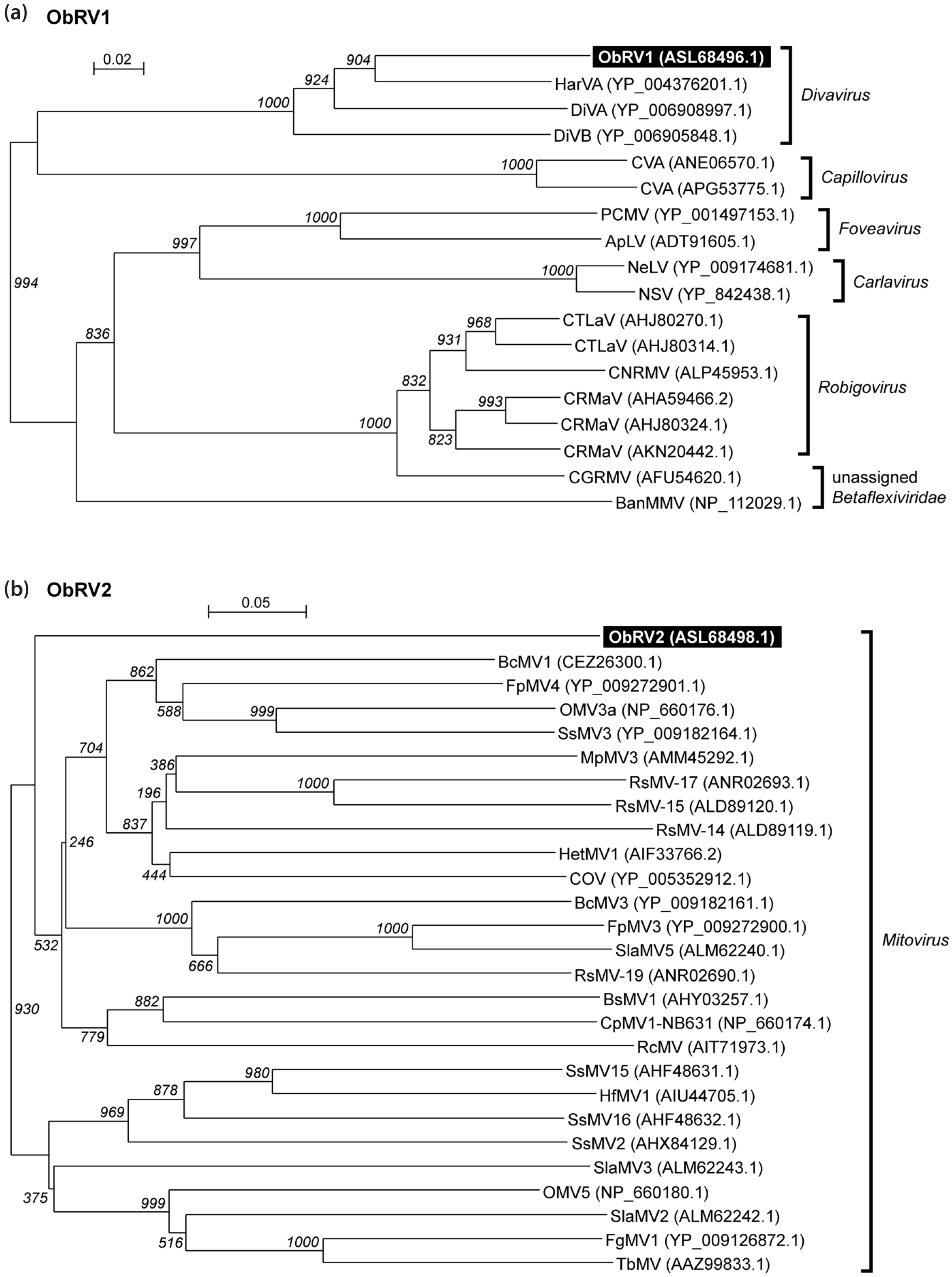

Fig. 1

**Phylogenetic analysis of newly identified basil RNA viruses**

Phylogenetic trees constructed based on the amino acid sequences of the RdRp motifs of (**a**) ObRV1 and (**b**) ObRV2 are presented. The bootstrap percentages calculated from 1,000 bootstrap replicates are shown at the nodes. See Table 2 for the full names of the viruses and Supplementary Figs. S1 (ObRV1) and S2 (ObRV2) for the multiple sequence alignments of the amino acid sequences of the RdRp motifs.

*naviridae*. Given that mitoviruses generally infect fungi, it was uncertain whether this mitovirus genome was isolated directly from basil cells or fungal cells that infected basil (Hillman and Cai, 2013).

The ObRV2 genome is 2,784-bp long and predicted to have a single ORF encoding a 760-aa-long protein when it was translated using both the fungal mitochondrial and plant (standard) genetic codes (Table 1). A Pfam analysis indicated that the ORF contains a mitovirus-type RdRp motif (PF05919) at aa 268–579. There were sequence variations at 61 positions, indicating that the ObRV2 genome contig is also a composite sequence derived from two or more clones (Supplementary Table S2).

The ObRV2 RdRp motif-containing protein showed about 30–40% similarity to those of other mitoviruses (Table 2). A phylogenetic analysis based on a multiple sequence alignment of the RdRp domain sequences placed ObRV2 as a distinct clade within the genus *Mitovirus* (Fig. 1b and Supplementary Fig. S2).

Most of the fungal mitoviruses replicate within the mitochondria of fungal cells and are unable to produce functional RdRps in plant cells because the genetic codes of fungal mitochondria and plants are different. For instance, the UGA codon, which is a stop codon in the standard genetic code, encodes for tryptophan in fungal mitochondria (Nibert, 2017). However, the ObRV2 ORF encodes a full-length RdRp using the plant nuclear and mitochondrial genetic codes, suggesting that ObRV2 may be able to replicate in both fungal and plant cells.

Interestingly, there are mitovirus-derived sequences in many plant nuclear and mitochondrial genomes, some of which are expressed as plant genes (Bruenn *et al.*, 2015; Xu *et al.*, 2015). Because the ObRV2 sequence was found in a plant transcriptome and encodes a full-length protein, it could also be derived from a mitovirus-like gene integrated in the basil nuclear or mitochondrial genome. However, most plant mitovirus-like genes are derived from the RdRp domain and are relatively shorter than those encoding the mitovirus RdRps; plant mitovirus-like genes encode for proteins about 500-aa long or shorter, while the mitovirus RdRps are 700-aa long or longer. The ObRV2 RdRp is 760-aa long. Therefore, it is highly likely that the ObRV2 sequence represents a mitovirus that infects the basil or a basil-associated fungus.

In conclusion, basil transcriptome data were analyzed to identify viral genome sequences representing two novel viral species of the genera *Divavirus* and *Mitovirus*. The method used in this study can be applied to other plant transcriptome data to discover plant viral genome sequences co-isolated with their host RNAs.

**Acknowledgment.** This research was supported by the National Research Foundation of Korea funded by the Korean Government (2017R1A1B4005866).

**Supplementary information** is available in the online version of the paper.

## References

Adams MJ, Candresse T, Hammond J, Kreuze JF, Martelli GP, Namba S, Pearson MN, Ryu KH, Saldarelli P, Yoshikawa N (2012): Family Betaflexiviridae. In King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (Eds): Virus Taxonomy. Classification and nomenclature of viruses. Ninth report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, San Diego, CA, pp. 920–941.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990): Basic local alignment search tool. J. Mol. Biol. 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012): SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477. https://doi.org/10.1089/cmb.2012.0021

Bruenn JA, Warner BE, Yerramsetty P (2015): Widespread mitovirus sequences in plant genomes. PeerJ 3, e876. https://doi.org/10.7717/peerj.876

Chiang LC, Ng LT, Cheng PW, Chiang W, Lin CC (2005): Antiviral activities of extracts and selected pure constituents of Ocimum basilicum. Clin. Exp. Pharmacol. Physiol. 32, 811–816. https://doi.org/10.1111/j.1440-1681.2005.04270.x

Dudai N, Chaimovitsh D, Reuveni R, Ravid U, Larkov O, Putievsky E (2002): Breeding of sweet basil (Ocimum basilicum) resistant to Fusarium wilt caused by Fusarium oxysporum f.sp. basilicum. J. Herbs Spices Med. Plants 9, 45–51. https://doi.org/10.1300/J044v09n02_07

Edgar RC (2004): MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797. https://doi.org/10.1093/nar/gkh340

Edris AE, Farrag ES (2003): Antifungal activity of peppermint and sweet basil essential oils and their major aroma constituents on some plant pathogenic fungi from the vapor phase. Nahrung 47, 117–121. https://doi.org/10.1002/food.200390021

Ghabrial SA, Suzuki N (2009): Viruses of plant pathogenic fungi. Annu. Rev. Phytopathol. 47, 353–384. https://doi.org/10.1146/annurev-phyto-080508-081932

Hillman BI, Cai G (2013): The family Narnaviridae: simplest of RNA viruses. Adv. Virus Res. 86, 149–176. https://doi.org/10.1016/B978-0-12-394315-6.00006-4

Igori D, Lim S, Zhao F, Baek D, Park JM, Cho HS, Kim HS, Kwon SY, Moon JS (2016): The complete sequence and genome organization of Ligustrum virus A, a novel carlavirus. Arch. Virol. 161, 3593–3596. https://doi.org/10.1007/s00705-016-3054-x

Jelkmann W (1995): Cherry virus A: cDNA cloning of dsRNA, nucleotide sequence analysis and serology reveal a new plant capillovirus in sweet cherry. J. Gen. Virol. 76 (Pt 8), 2015–2024. https://doi.org/10.1099/0022-1317-76-8-2015

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007): Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948. https://doi.org/10.1093/bioinformatics/btm404

Li H (2011): A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27, 2987–2993. https://doi.org/10.1093/bioinformatics/btr509

Li H, Durbin R (2009): Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Nibert ML (2017): Mitovirus UGA(Trp) codon usage parallels that of host mitochondria. Virology 507, 96–100. https://doi.org/10.1016/j.virol.2017.04.010

Nouri S, Salem N, Falk BW (2016): Complete genome sequence of Diaphorina citri-associated C virus, a novel putative RNA virus of the Asian citrus psyllid, Diaphorina citri. Genome Announc. 4, e00639–00616. https://doi.org/10.1128/genomeA.00639-16

Park D, Hahn Y (2017a): Genome sequence of Spinach cryptic virus 1, a new member of the genus Alphapartitivirus (family Partitiviridae), identified in spinach. J. Microbiol. Biotechnol. 27, 834–837. https://doi.org/10.4014/jmb.1611.11026

Park D, Hahn Y (2017b): Genome sequences of Spinach deltapartitivirus 1, Spinach amalgavirus 1, and Spinach latent virus identified in spinach transcriptome. J. Microbiol. Biotechnol. 27, 1324–1330. https://doi.org/10.4014/jmb.1703.03043

Rastogi S, Meena S, Bhattacharya A, Ghosh S, Shukla RK, Sangwan NS, Lal RK, Gupta MM, Lavania UC, Gupta V, Nagegowda DA, Shasany AK (2014): De novo sequencing and comparative analysis of holy and sweet basil transcriptomes. BMC Genomics 15, 588. https://doi.org/10.1186/1471-2164-15-588

Suppakul P, Miltz J, Sonneveld K, Bigger SW (2003): Antimicrobial properties of basil and its possible application in food packaging. J. Agric. Food Chem. 51, 3197–3207. https://doi.org/10.1021/jf021038t

Torre S, Tattini M, Brunetti C, Guidi L, Gori A, Marzano C, Landi M, Sebastiani F (2016): De novo assembly and comparative transcriptome analyses of red and green morphs of sweet basil grown in full sunlight. PLOS ONE 11, e0160370. https://doi.org/10.1371/journal.pone.0160370

Visser M, Bester R, Burger JT, Maree HJ (2016): Next-generation sequencing for virus detection: covering all the bases. Virol. J. 13, 85. https://doi.org/10.1186/s12985-016-0539-x

Wylie S, Jones M (2011): Hardenbergia virus A, a novel member of the family Betaflexiviridae from a wild legume in Southwest Australia. Arch. Virol. 156, 1245–1250. https://doi.org/10.1007/s00705-011-0963-6

Wylie SJ, Li H, Dixon KW, Richards H, Jones MG (2013): Exotic and indigenous viruses infect wild populations and captive collections of temperate terrestrial orchids (Diuris species) in Australia. Virus Res. 171, 22–32. https://doi.org/10.1016/j.virusres.2012.10.003

Xie J, Ghabrial SA (2012): Molecular characterizations of two mitoviruses co-infecting a hypovirulent isolate of the plant pathogenic fungus Sclerotinia sclerotiorum. Virology 428, 77–85. https://doi.org/10.1016/j.virol.2012.03.015

Xu Z, Wu S, Liu L, Cheng J, Fu Y, Jiang D, Xie J (2015): A mitovirus related to plant mitochondrial gene confers hypovirulence on the phytopathogenic fungus Sclerotinia sclerotiorum. Virus Res. 197, 127–136. https://doi.org/10.1016/j.virusres.2014.12.023

Acta virologica **62**: 304 – 309, 2018 doi:10.4149/av_2018_224

# Supplementary information

## Novel *Divavirus* (the family *Betaflexiviridae*) and *Mitovirus* (the family *Narnaviridae*) species identified in basil (*Ocimum basilicum*)

C. J. GOH[1], D. PARK[1], H. KIM[1], F. SEBASTIANI[2], Y. HAHN[1*]

[1]Department of Life Science, Research Center for Biomolecules and Biosystems, Chung-Ang University, Seoul 06974, South Korea; [2]Institute for Sustainable Plant Protection, Department of Biology, Agriculture and Food Sciences, The National Research Council of Italy, Sesto Fiorentino, Italy

*Received October 3, 2017; revised December 19, 2017; accepted July 10, 2018*

**Supplementary Table S1. Sequence variation of ObRV1**

| Position | Con[a] | Alt[b] | Quality[c] | Con # | Alt # | Con % | Alt % |
|---|---|---|---|---|---|---|---|
| 274 | A | G | 222 | 2212 | 2792 | 44.2 | 55.8 |
| 535 | G | A | 222 | 7568 | 4917 | 60.6 | 39.4 |
| 709 | T | C | 222 | 6593 | 4802 | 57.9 | 42.1 |
| 871 | G | A | 222 | 6262 | 5003 | 55.6 | 44.4 |
| 904 | A | G | 222 | 5955 | 4258 | 58.3 | 41.7 |
| 907 | A | T | 222 | 5978 | 4217 | 58.6 | 41.4 |
| 1330 | T | C | 222 | 7713 | 4836 | 61.5 | 38.5 |
| 1421 | T | C | 222 | 8349 | 5928 | 58.5 | 41.5 |
| 1423 | G | T | 222 | 8496 | 5927 | 58.9 | 41.1 |
| 1912 | G | A | 222 | 7838 | 4337 | 64.4 | 35.6 |
| 2305 | T | C | 222 | 7486 | 4512 | 62.4 | 37.6 |
| 2407 | C | T | 222 | 8855 | 5501 | 61.7 | 38.3 |
| 2758 | T | C | 222 | 7397 | 4396 | 62.7 | 37.3 |
| 2761 | A | G | 222 | 7580 | 4294 | 63.8 | 36.2 |
| 3403 | G | A | 222 | 8320 | 6555 | 55.9 | 44.1 |
| 3689 | T | C | 222 | 8390 | 6545 | 56.2 | 43.8 |
| 4036 | A | G | 222 | 9544 | 7299 | 56.7 | 43.3 |
| 4222 | G | A | 222 | 9823 | 8396 | 53.9 | 46.1 |
| 4243 | T | A | 222 | 9350 | 7276 | 56.2 | 43.8 |
| 4576 | A | G | 222 | 9959 | 7574 | 56.8 | 43.2 |
| 4645 | T | C | 222 | 10088 | 8018 | 55.7 | 44.3 |
| 4978 | A | G | 222 | 11797 | 8422 | 58.3 | 41.7 |
| 4984 | G | A | 222 | 12145 | 8704 | 58.3 | 41.7 |
| 5130 | T | C | 222 | 11593 | 8383 | 58.0 | 42.0 |
| 5253 | T | C | 222 | 15217 | 11009 | 58.0 | 42.0 |
| 5307 | C | T | 222 | 15145 | 10481 | 59.1 | 40.9 |
| 5343 | C | T | 222 | 12969 | 9216 | 58.5 | 41.5 |
| 5511 | G | A | 222 | 13535 | 9896 | 57.8 | 42.2 |
| 5539 | A | G | 222 | 12512 | 9866 | 55.9 | 44.1 |
| 5785 | C | T | 222 | 15605 | 11631 | 57.3 | 42.7 |
| 5905 | C | T | 222 | 13495 | 11168 | 54.7 | 45.3 |
| 6511 | G | A | 222 | 12695 | 8850 | 58.9 | 41.1 |
| 6551 | C | T | 222 | 15251 | 11238 | 57.6 | 42.4 |

[a]Sequence of the assembled contig; [b]Alternative sequence observed in RNA-seq reads; [c]Phred-scaled quality score for the assertion made in Alt by BCFtools.

**Supplementary Table S2. Sequence variation of ObRV2**

| Position | Con[a] | Alt[b] | Quality[c] | Con # | Alt # | Con % | Alt % |
|---|---|---|---|---|---|---|---|
| 335 | CAA | CAAA | 183 | 21 | 18 | 53.8 | 46.2 |
| 352 | A | T | 181 | 21 | 15 | 58.3 | 41.7 |
| 355 | A | T | 179 | 21 | 15 | 58.3 | 41.7 |
| 442 | C | T | 222 | 35 | 32 | 52.2 | 47.8 |
| 445 | T | C | 222 | 39 | 30 | 56.5 | 43.5 |
| 451 | A | G | 222 | 42 | 33 | 56.0 | 44.0 |
| 454 | T | C | 222 | 47 | 35 | 57.3 | 42.7 |
| 583 | C | T | 212 | 42 | 22 | 65.6 | 34.4 |
| 601 | T | C | 222 | 41 | 25 | 62.1 | 37.9 |
| 676 | C | T | 222 | 39 | 30 | 56.5 | 43.5 |
| 691 | C | T | 222 | 40 | 22 | 64.5 | 35.5 |
| 715 | C | T | 222 | 36 | 24 | 60.0 | 40.0 |
| 736 | A | G | 222 | 30 | 27 | 52.6 | 47.4 |
| 763 | G | A | 222 | 38 | 27 | 58.5 | 41.5 |
| 776 | C | G | 222 | 40 | 30 | 57.1 | 42.9 |
| 835 | C | T | 222 | 31 | 26 | 54.4 | 45.6 |
| 836 | G | C | 222 | 30 | 26 | 53.6 | 46.4 |
| 889 | G | A | 222 | 45 | 52 | 46.4 | 53.6 |
| 911 | C | A | 222 | 48 | 43 | 52.7 | 47.3 |
| 912 | G | A | 222 | 47 | 43 | 52.2 | 47.8 |
| 926 | A | C | 222 | 52 | 44 | 54.2 | 45.8 |
| 940 | C | T | 222 | 55 | 39 | 58.5 | 41.5 |
| 961 | T | C | 222 | 56 | 31 | 64.4 | 35.6 |
| 967 | G | A | 222 | 50 | 25 | 66.7 | 33.3 |
| 1006 | C | G | 219 | 34 | 21 | 61.8 | 38.2 |
| 1007 | A | C | 217 | 34 | 21 | 61.8 | 38.2 |
| 1065 | C | G | 222 | 30 | 20 | 60.0 | 40.0 |
| 1066 | C | T | 221 | 31 | 20 | 60.8 | 39.2 |
| 1085 | G | A | 222 | 32 | 26 | 55.2 | 44.8 |
| 1102 | C | T | 222 | 30 | 29 | 50.8 | 49.2 |
| 1178 | C | A | 222 | 38 | 24 | 61.3 | 38.7 |
| 1267 | A | T | 222 | 47 | 35 | 57.3 | 42.7 |
| 1276 | G | A | 222 | 49 | 33 | 59.8 | 40.2 |
| 1321 | A | G | 222 | 50 | 32 | 61.0 | 39.0 |
| 1330 | T | C | 222 | 49 | 36 | 57.6 | 42.4 |
| 1465 | T | C | 222 | 25 | 21 | 54.3 | 45.7 |
| 1486 | C | T | 12.9425 | 45 | 6 | 88.2 | 11.8 |
| 1504 | G | A | 222 | 26 | 26 | 50.0 | 50.0 |
| 1510 | T | C | 222 | 25 | 28 | 47.2 | 52.8 |
| 1594 | A | T | 222 | 23 | 21 | 52.3 | 47.7 |
| 1600 | C | T | 221 | 22 | 17 | 56.4 | 43.6 |
| 1732 | T | C | 222 | 52 | 42 | 55.3 | 44.7 |
| 1775 | G | A | 172 | 48 | 40 | 54.5 | 45.5 |
| 1777 | A | C | 157 | 49 | 40 | 55.1 | 44.9 |
| 1810 | G | A | 222 | 30 | 32 | 48.4 | 51.6 |
| 1837 | C | T | 222 | 34 | 31 | 52.3 | 47.7 |
| 1939 | T | G | 174 | 32 | 13 | 71.1 | 28.9 |
| 1996 | C | T | 203 | 26 | 28 | 48.1 | 51.9 |
| 2024 | T | G | 222 | 29 | 37 | 43.9 | 56.1 |
| 2047 | A | G | 222 | 36 | 36 | 50.0 | 50.0 |
| 2062 | G | A | 222 | 31 | 34 | 47.7 | 52.3 |
| 2074 | T | C | 222 | 35 | 37 | 48.6 | 51.4 |
| 2092 | C | T | 222 | 32 | 41 | 43.8 | 56.2 |
| 2125 | T | C | 222 | 46 | 44 | 51.1 | 48.9 |
| 2161 | G | A | 222 | 51 | 41 | 55.4 | 44.6 |
| 2170 | G | A | 222 | 49 | 46 | 51.6 | 48.4 |
| 2239 | T | C | 222 | 36 | 31 | 53.7 | 46.3 |
| 2248 | A | G | 222 | 33 | 29 | 53.2 | 46.8 |
| 2251 | C | T | 222 | 33 | 30 | 52.4 | 47.6 |
| 2266 | A | G | 222 | 31 | 32 | 49.2 | 50.8 |
| 2449 | C | T | 123 | 49 | 19 | 72.1 | 27.9 |

[a] Sequence of the assembled contig; [b] Alternative sequence observed in RNA-seq reads; [c] Phred-scaled quality score for the assertion made in Alt by BCFtools.

## Supplementary Fig. S1

### Multiple alignment of the amino acid sequences of the RdRp regions of ObRV1 and related viruses

```
ObRV1:ASL68496.1:1274-1651      PKHSNGDSLTFFAAVKKRLKFSSPQIEREKFEKVRHLGSEMFELLLEKIPLDNKNDDLMMQICVNEYIERKVSKPAGTIK    80
HarVA:YP_004376201.1:1274-1651  PKHSNQDSLTFFAAVKKRLKFSSPHVEREKFEKVRHLGSEMFDVFLSKINLDNSYNHELMQMSVNEYIEKKVSKTSNTIK    80
YP_006905850.1:1274-1651        PKHQNSDSLTFLAAVKKRLKFSSPSVERERFEKVRHLGNEMLDIFLDKIKIDNKLNSEMMARSYNEYVLKKVSKTANTIA    80
DiVB:YP_006905848.1:1274-1651   PKHSNSDTLTFFAAVKKRLKFSSPQIEREKFEKVRHLGQEMLELFLEKVKVSNHYDKQMELMSYNEYIEKKVAKTGNTIA    80
CVA:ANE06570.1:1246-1621        PRHFANDDLTFWSAVKKRLVFKNPLNNVHDFEKAKPFGKEMLDIFLKKVPLIPSFDQRMYEESISEFEEKKISKNAAMIG    80
CVA:APG53775.1:1246-1621        PRHFANDDLTFWSAVKRRLVFKNPLSNAHDFEKAKPFGKELLNIFLRKVPLKPNFDQRMYDESVSEFEEKKISKNAAMIG    80
PCMV:YP_001497153.1:1748-2123   PRHKGTDSVTFLMAVKKRLSFSQPSIESAKLRRAMPFGKFLLQTFLKYIPLCGKHDSDLMSKAVNDFEEKKLSKSAATIE    80
ApLV:ADT91605.1:1791-2166       PRHKGTDTATFLMAVKKRLSFSSPAAEHAKLRRARPFGKFLLDTFLKRIPLNCRHDCEMMSSAVHAFEEKKLSKSMATIE    80
NeLV:YP_009174681.1:1486-1861   PRHKGDDTATFLMAVKKRLRFSKPAQEVAKFRAAEPFGEYMLGVFLKHIKLNKNHEPQKMADAKREFEEKKTSKSAAVIE    80
NSV:YP_842438.1:1484-1859       PRHKGDDTATFLMAVKKRLRFSKPAQETAKFRAAEPFGEYMLGIFLKHIKLNKNHEPLKMAAAKREFEEKKTSKSAAVIE    80
CTLaV:AHJ80270.1:1639-2014      PRHKAGDTATFVMAARKRLKFSRPANERQKFLSAIPYGDTMLRVFLEKVKLKPNFDHRLFEEARNDFEEKKLQKSMATLE    80
CTLaV:AHJ80314.1:1641-2016      PRHRAGDTATFVMAARKRLKFSLPANEQQKFLSAIPYGDTMLKVFLNKVKLKPNFDYRLFEEARNDFEEKKLQKSMATLE    80
CNRMV:ALP45953.1:1640-2015      PRHKSGDTATFLMAARKRLKFSFPANERQKYLAAIPYGDTMLKVFLEKVKLKPAFDHLLFDEARNDFEEKKLGKSMATLE    80
CRMaV:AHA59466.2:1631-2006      PRHKAGDTATFVMAARKRLKFSYPARERQKFMAAIPYGASMLQVFLKRVRLQPNFDHKLFEESRSDFEEKKLQKSMATLE    80
CRMaV:AHJ80324.1:1628-2003      PRHKAGDTATFVMAARKRLKFSFPARERQKYMAAIPYGESMLQVFLKRVKLQPNFDHRLFEESRADFEEKKLQKSMATLE    80
CRMaV:AKN20442.1:1631-2006      PRHKAGDTATFVMAARKRLKFSFPAREKQKYLSAIPYGESMLQVFLKRVKLNSNFDYNLFEEARDDFEEKKLQKSMATLE    80
CGRMV:AFU54620.1:1628-2003      PRHKAGDTATFVMAARKRLKFSIPAKEGQKFRAAIPYGESMLQIFLKKVKVQPNFDHKLFAEAVSDFEEKKLTKSMATLE    80
BanMMV:NP_112029.1:1372-1747    PRHRNSDTVTFLMAVKKRLSFSQPRKEMQRYLMNKRKGEEMADAFEKFVPIKSNFSREKFIEAKMEFESKKLEKSKATIE    80
                                *.*   *  **  *...** *. *  :          *  :   :   : :    .      .   :  .*  *    :

ObRV1:ASL68496.1:1274-1651      SHSGRSDCDWKLNDVFLFIKTQLCTKYEKRFSDAKAGQTLACFSHVILNRFAAPARYIEKKLSLCLPDNYYIHQKKNFDM   160
HarVA:YP_004376201.1:1274-1651  SHSSRSEPDWKLNDVFLFMKTQLCTKYEKRFCDAKAGQTLACFSHIVLNRFAAPTRYIEKKISQCLPSNYYIHQKKNFDC   160
YP_006905850.1:1274-1651        SHSSRSEPDWKLNEIFLFMKTQLCTKFEKRFSDAKAGQTLACFSHIILNRFAAPTRYVEKKISEGLGKNFYIHQKKNFDV   160
DiVB:YP_006905848.1:1274-1651   AHSNRSEPDWNLNEIFLFMKTQLCTKYEKRFCEAKAGQTLACFSHIVLNRFAGPTRYIEKKISSYLPNNYYIHQKKNFDK   160
CVA:ANE06570.1:1246-1621        AHHDRSTTDWPINEIFLFIKSQLCTKKEKMFCDAKAGQTLACFSHLILCKFAPLNRYIEKKVTQCLPGNFYIHQKKNFDE   160
CVA:APG53775.1:1246-1621        AHHDRSTTDWPTNEIFLFIKSQLCTKKEKMFCDAKAGQTLACFSHLILCKFAPLNRYIEKKVTQSLPGNFYIHQKKNFDE   160
PCMV:YP_001497153.1:1748-2123   NHSGRSSRDWPVDKALIFMKSQLCTKFDNRFRSAKAGQTLACFQHSVLCRFAPYMRYIEAKLFNALPDRFYIHSGKNIDD   160
ApLV:ADT91605.1:1791-2166       NHSGRSCEDWPVDKALIFMKSQLCTKFDNRFRSAKAGQTLACFQHSVLCRFAPYMRYIESKVVEALPSNLYIHSGKNIEE   160
NeLV:YP_009174681.1:1486-1861   NHSGRSCRDWLIDLGFLFIKNQLCTKFEKRFADAKAAQTILCFQHEVLCRFAPYYVRYIEKKLNDALPAKYYIHSGKNIDD   160
NSV:YP_842438.1:1484-1859       NHSGRSCRDWLIDLGFLFIKNQLCTKFEKRFADAKAAQTILCFQHEVLCRFAPYVRYIEKKLNDALPAKYYIHSGKNIDD   160
CTLaV:AHJ80270.1:1639-2014      NHSGRSDPDWEVEKALIFMKSQLCTKFDNRFRDAKAGQTLACFHHNVLCRLAPYIRYIEKKVFAALPKNLYIHSGKNFDD   160
CTLaV:AHJ80314.1:1641-2016      NHSGRSDPDWEVEKALIFMKSQLCTKFDNRFRDAKAGQTLACFHHNVLCRLAPYIRYIEKKVFNSLPKNLYIHSGKNFDD   160
CNRMV:ALP45953.1:1640-2015      NHSGRSDPDWEVEKALIFMKSQLCTKFDNRYRDAKAGQTLACFHHNVLCRLAPYVRYIEKKVFKALPRNLYIHSGKNFDD   160
CRMaV:AHA59466.2:1631-2006      NHSGRSDPDWSIEKALIFMKSQLCTKFDNRFRSAKAGQTLACFHHDVLCRLAPYIRYIEKKVFKALPKNLYIHSARNFDD   160
CRMaV:AHJ80324.1:1628-2003      NHSGRSDPDWGVEKALIFMKSQLCTKFDNRFRNAKAGQTLACFHHDVLCRLAPYIRYIEKKVFKALPSNLYIHSARNFDD   160
CRMaV:AKN20442.1:1631-2006      NHSGRSDPDWNVENASIFMKSQLCTKFDNRFRDAKAGQTLACFHHNVLCRLAPYIRYIEKKVFKALPKNLYIHSGKNFDD   160
CGRMV:AFU54620.1:1628-2003      NHSGRSDPDWEKERALIFMKSQLCTKFDNRFRAAKAGQTLACFHHNVLCRLAPYIRYIEKKIFSHLPKNYYIHSGKNFDE   160
BanMMV:NP_112029.1:1372-1747    NHAQRSNREWKIDEAMIFMKSQLCTKFEKRFVEAKAGQTLACFSHIVLCRFAPYVRYMEKIVNENLPKNFYIHNGKNFDD   160
                                 *  **  :*  :   :*:*.***** :: :  ***.**: ** * :* .:*   **:*  :   *  . ***. .*::

ObRV1:ASL68496.1:1274-1651      LNAWVVRNDFSDECLESDYEAFDSSQDCLILAFEYELLKYMGWSQSLLDDYLDLKFNLGCRLGNLAVMRFTGEFGTFLFN   240
HarVA:YP_004376201.1:1274-1651  LNEWVIRNDFSESCLESDYEAFDSSQDCLILAFEYELLKYMGWKQDLLDDYLDLKFNLGCRLGNLAVMRFTGEFGTFLFN   240
YP_006905850.1:1274-1651        LNDWVVANNFDSYCLESDYEAFDSSQDCLILAFEYELLKYLGWDQSLLDDYLDLKFNLGCRLGNLAVMRFTGEFGTFLFN   240
DiVB:YP_006905848.1:1274-1651   LNEWVIENKFTGESLESDYEAFDSSQDCIILAFEYELLSYLGWNQSLLDDYLTIKFNLGCRLGNLAVMRFTGEFGTFLFN   240
CVA:ANE06570.1:1246-1621        LEKWVKSYDFSGVCTESDYEAYDASQDSYTLAFEYELMKYLGVSNSMIEDYLYLKMHLNCKLGNLAIMRFTGEFCTFLFN   240
CVA:APG53775.1:1246-1621        LEKWVKSYNFNGVCTESDYEAYDASQDSYTLAFEYELLRYLGVSNSLIEDYLYLKMHLNCKLGNLAIMRFTGEFCTFLFN   240
PCMV:YP_001497153.1:1748-2123   LGNWVKKQNFSGECTESDYEAFDASQDHFILAFEIEIMRHLGLPEGLINDYIFIKCNLGSKLGSFAIMRFTGEASTFLFN   240
ApLV:ADT91605.1:1791-2166       LADWVKVNKFNGVCTESDYEAFDASQDHFILAFELEVMKHLGLPADLIADYTFIKTHLGSKLGNFAIMRFTGEASTFLFN   240
NeLV:YP_009174681.1:1486-1861   LNDWVKDNDFSGICTESDYEAFDASQDHYIMAFEVAIMKYLGLPRDLINDYIFIKTHLGSKLGSFAIMRFSGEASTFLFN   240
NSV:YP_842438.1:1484-1859       LNDWVKANNFSGVCTESDYEAFDASQDHYIMAFEVAVMKYLGLPRDLVNDYIFIKTHLGSKLGSFAIMRFSGEASTFLFN   240
CTLaV:AHJ80270.1:1639-2014      LRDWVIAHRFSGVCTESDYEAFDSSQDANILAFEVSLMNYLNLPRDLIEDYKHLKFNTHSKLGQFAVMRFTGEAGTFLFN   240
CTLaV:AHJ80314.1:1641-2016      LRDWVIKSNFSGMCTESDYEAFDSSQDANILAFEVSLMNYLNLPRDLIEDYKHLKFNTHSKLGQFAVMRFTGEAGTFLFN   240
CNRMV:ALP45953.1:1640-2015      LQEWVIQNNFVGTCTESDYEAFDSSQDANILAFEVSLMNYLNLPRDLIEDYKHLKFNTHSKLGRFAVMRFTGEAGTFLFN   240
CRMaV:AHA59466.2:1631-2006      LRDWVIENNFTGVCTESDYEAFDSSQDANILAFEVSLMEYLRLPRDLIEDYKYLKFHTHSKLGQFAVMRFTGEAGTFLFN   240
CRMaV:AHJ80324.1:1628-2003      LRDWVIRNNFTGVCTESDYEAFDSSQDANILAFEVSLMEYLRLPRDLIEDYKYLKFHTHSKLGQFAVMRFTGEAGTFLFN   240
CRMaV:AKN20442.1:1631-2006      LRDWVIANNFTGLCTESDYEAFDSSQDANILAFEVSLMTYLKLPRDLIEDYKHLKFHTKSKLGQFAVMRFTGEAGTFLFN   240
CGRMV:AFU54620.1:1628-2003      LRDWVLKNSFIGMCTESDYXAFDSSQDSNILAFEVSLMRYLRLPRDLIEDYKYLKFNTRSKLGQFAVMRFTGEAGTSLFN   240
BanMMV:NP_112029.1:1372-1747    LNDYVKKNNFSGRCIESDYEAFDASQDSQILAFEVAIMRRMNMPQEFIDDYVWLKCNLRSKLGNMAIMRFTGEAATFLFN   240
                                *  :*    *   . **** *:*:***   :***  ::  :     :: **  :* :  ..** :*:***:**  * ***

ObRV1:ASL68496.1:1274-1651      TLANMVFTFMSYDLTGKEAICFAGDDMCCNKGIRRRTDGRFDHILNRLSLKAKAVITTEPTFCGWRLTKYGIFKKPELVL   320
HarVA:YP_004376201.1:1274-1651  TLANMVFTFMSYELNGKESICFAGDDMCCNKGIKKRIDGKFDHILKRLSLKAKAMITNEPTFCGWRLTPFGIYKKPELIL   320
YP_006905850.1:1274-1651        TLANMVFTFMTYDLNGTESICFAGDDMCCNRGIKARVDGKYDHILKRLTLKAKAVITKEPTFCGWRLTKYGIFKKPELVL   320
DiVB:YP_006905848.1:1274-1651   TLANMTFTFMSYDLDDKEAICFAGDDMCCNKPVKKRSDGRYDHILKRLTLKAKVSYTREPTFCGWRLTKHGIFKKPELVL   320
CVA:ANE06570.1:1246-1621        TLTNMLFTFMKYDVRKTHAICFAGDDMCANVRLPENLQ--HTNLLKKFSLKAKVDFTRSPTFCGWNLSRYGIVKKPELIA   318
CVA:APG53775.1:1246-1621        TLTNMLFTFMKYDVRKTHAICFAGDDMCANVRLPENHQ--YSDLLKRSSLKAKVDFTRSPTFCGWNFSRYGIVKKPELIA   318
PCMV:YP_001497153.1:1748-2123   TMANMLFTFLRYELSGHESISFAGDDMCANRRLRVRST--YKSFLEKIRLKAKVQFTNFPTFCGWGLCPEGVFKKPDLVL   318
ApLV:ADT91605.1:1791-2166       TMANMLFTFLRYDLNGKEAICFAGDDMCANSRLKVTQT--HSKFLDKIKLKAKVQFTITPTFCGWGLCEHGVFKKPDLVL   318
NeLV:YP_009174681.1:1486-1861   TMSNMLFTFMKYEINGSESICFAGDDMCASKKLRLSDE--HSSYLEKLRLKAKVCFTTRPTFCGWSLNRLGIFKKPQLVY   318
NSV:YP_842438.1:1484-1859       TMSNMLFTFMKYEITGKESICFAGDDMCASKRLRLSTV--HSSYLEKLRLKAKVCFTSRPTFCGWSLNRLGIFKKPQLVY   318
CTLaV:AHJ80270.1:1639-2014      TLANMVFTFMRYETNGREAICFAGDDMCANKLLRRKSE--FEHVLERMTLKAKVQHTSEPTFCGWRLGTFGIVKRPQLVQ   318
CTLaV:AHJ80314.1:1641-2016      TLANMVFTFMRYETNGRESICFAGDDMCANKLLRKRKE--YEHVLERMTLKAKVQHTTEPTFCGWRLGPFGIIKRPQLVQ   318
CNRMV:ALP45953.1:1640-2015      TLANMVFTFMRYETNGRESICFAGDDMCANKLLRKKGD--FEHVLNRMTLKAKVQYTTEPTFCGWRLGSFGIVKRPQLVQ   318
CRMaV:AHA59466.2:1631-2006      TLANMVFTFMRYEVNGREAICFAGDDMCANKLLRKKSE--FEHILERMTLKAKVQHTTEPTFCGWRLGNFGIIKRPQLVQ   318
CRMaV:AHJ80324.1:1628-2003      TLANMVFTFMRYEINGKEAICFAGDDMCANKLLRKKSE--FEHILDRMTLKAKVQHTTEPTFCGWRLGNFGIVKRPQLVQ   318
CRMaV:AKN20442.1:1631-2006      TLANMVFTFMRYEVNGRESICFAGDDMCANKLLRKKSE--FEHILERMTLKAKVQHTNEPTFCGWRLGNFGIIKRPQLVQ   318
CGRMV:AFU54620.1:1628-2003      TLANMVFTFMRYDLNGKEAICFAGDDMCANKRLRKKSE--FEHILDRMTLKAKVQYTTEPTFCGWRLGDFGIVKRPQLVQ   318
BanMMV:NP_112029.1:1372-1747    TMANIVFTIMAYDLKGDECILFAGDDMCMNTVRRANNS--YTHILKNLKLKAKVGITEEPTFCGWRLTIHGIYKRPQLIL   318
                                *::*: **:: *:     .* ******* .          .   *..  ****.  *  ****** :   *: *.*:*:
```

```
ObRV1:ASL68496.1:1274-1651      ERFLIAIEKNKLKDVIDSYYLECSYAYSLGERLFECFSEKDFVAHYCCIRLVHLHKDL                         378
HarVA:YP_004376201.1:1274-1651  ERFLIAIEKGRLVDVIDSYYLECSYAYSLGERLSKCFSEKDFVAHYCCVRLVHKHKSL                         378
YP_006905850.1:1274-1651        ERFLIAIEKGRLLDVIDSYYIECSYAYNLGERLFECFSEKDFSAHYCCIRIVHKNKSL                         378
DiVB:YP_006905848.1:1274-1651   ERFLIAMEKGRLKEVIDSYFIECSYAYNLGERLYECFSEKDFIAHYCCVRIVHENSSL                         378
CVA:ANE06570.1:1246-1621        ARLAVAKQKGEVNLVLDSYFLEHLYAYNKGDHLFEILSEKELEHHYNLTRFFVKHGNL                         376
CVA:APG53775.1:1246-1621        ARLAVAKQKGEINLVIDSYFLEHLYAYNKGDHLFEILSEKELEHHYNLTRFFVKNSEL                         376
PCMV:YP_001497153.1:1748-2123   ERLQIAVETNNLQNCIDNYAIEVSYAYSMGESLSKYLSEEEMDAHYNCVRFIVKHSHL                         376
ApLV:ADT91605.1:1791-2166       ERLQIARETRNLENCIDNYAIEVSCAYRMGENLNLYLSPNEMDAHYNCVRFIILHNHL                         376
NeLV:YP_009174681.1:1486-1861   ERMCIAIEKNNLQNCIDNYAIEVSYAYLMGENALVLMDEEEIHNHYMCVRTIVQNKNL                         376
NSV:YP_842438.1:1484-1859       GRMCIAIEKNNLQNCIDNYAIEVSYAYLMGENALVLMDEEEIHNHYMCVRTIVQNKNL                         376
CTLaV:AHJ80270.1:1639-2014      ERILIALEKGNFNECIDNYAIEVSYAYNLGERLVSIMSEKELDAHYFCVRTFLQHKNL                         376
CTLaV:AHJ80314.1:1641-2016      ERILIALEKGNFNECIDNYAIEVSYAYNLGERLISIMSEKELDAHYFCVRTFLQHKNL                         376
CNRMV:ALP45953.1:1640-2015      ERILIALEKGNFIECIDNYAIEVSYAYKLGERLISIMSEKELDAHYFCVRTFLQNKKL                         376
CRMaV:AHA59466.2:1631-2006      ERILIALEKGNFNECIDNYAIEVSYAYNLGERLISIMSEKELDAHYFCVRTFLQNKKL                         376
CRMaV:AHJ80324.1:1628-2003      ERILIALEKGNFHECIDNYAIEVSYAYNLGERLISIMSEKELDAHYFCVRTFLQNKKL                         376
CRMaV:AKN20442.1:1631-2006      ERILIALEKGNFHECIDNYAIEVSYAYNLGERLISIMSEKELDAHYFCVRTFLRNKNL                         376
CGRMV:AFU54620.1:1628-2003      ERILIALEKGNFHECIDNYAIEVSYAYNLGERLISIMSEKELDAHYFCVRTFLQNKCL                         376
BanMMV:NP_112029.1:1372-1747    ERFMIAIENGNLENCIDNYAIECSYAYKLGDRLVSMFSEEENCAHYILVRYIVKKRHL                         376
                                 *: :* :.  .   :*.* :*   **  *:     :. ::   **   * .  :  *
```

Asterisks (*), fully conserved residues; colons (:), residues with strongly similar properties; and periods (.), residues with weakly similar properties. See Table 2 for full names of viruses.

## Supplementary Fig. S2

### Multiple alignment of the amino acid sequences of the RdRp regions of ObRV2 and related viruses

```
ObRV2:ASL68498.1:338-532      KVRVFAIPNPIFQRFLKPLHDWEMSVLKQLDTDGTYDQLRPLHRLKGKRV----------LYSFDLSAATDMFPKVLSAS    70
BcMV1:CEZ26300.1:295-487      KVRVFAMVDIWTQSILNPLHKKIFSIIRELPTDGTFDQLKPLDRLHELSTQD--------RFSFDLSAATDRLPLALQKD    72
FpMV4:YP_009272901.1:271-459  KVRVFAITDVWTQSFLSPLHHAIFGFLKRIPMDGTFDQLKPLNALLSRGLKN--------FYSYDLSAATDRLPITLQEQ    72
OMV3a:NP_660176.1:272-465     KARVFAMADSITQSVMAPLNSWVFSKLKDLPMDGTFNQQAPLNRLVQLYQDGLLHDVE--FYSYDLSSATDRLPMAFQKQ    78
SsMV3:YP_009182164.1:264-461  KARVFAITDSITQSVMGPISDAIFKMLRQIPMDGTFNQSAPLDRLVQLSKDGAIAEKDRIFYSYDLSAATDRLPINLQKD    80
MpMV3:AMM45292.1:465-663      KVRVFAMVDAWTQWALYPIHELIFKILKDVPMDGTFDQTAPLKHVRPHSG----------FWSLDLSAATDRLPLSIQKS    70
RsMV-17:ANR02693.1:13-210     KVRVFAMVDAWTQWVLEPLHQYMFDILSHISMDGTFDQMRPVTAKAGTATA---------AYSLDLTAATDRIPMELQER    71
RsMV-15:ALD89120.1:432-629    KVRVFAMADAWTQWVLEPFHQYLFDLLRQIDMDGTFDQLRPVIKKLATVKA---------AYSLDLTAATDRLPIAIQTV    71
RsMV-14:ALD89119.1:461-658    KVRVFAMVPTWFQNLLRPLHDIIFEILSGVAQDGTFNQMKPLKGHSSRFNM---------SFSLDLTAATDRLPLIIQTA    71
HetMV1:AIF33766.2:356-558     KVRVFAMVDPLTQWALRPIHKLLFLLLRYIPMDGTFDQLKPITRLLAQKPKV--------LDSLDLSAATDRLPIAIQEW    72
COV:YP_005352912.1:336-529    KVRVFAMVDPWSQMILRPFHLGLFKILRRIRMDGTFNQLKPLERAWGFKS----------LYSMDLSSATDRLPMRIQVP    70
BcMV3:YP_009182161.1:299-499  KLRIFAIVDSITQSLLSPLHDFMFDLLKKIPNDGTFDQDLSVKRSQVKSLSSGK------AFSFDLSAATDRLPVDLTVK    74
FpMV3:YP_009272900.1:297-500  KVRVFALVDSITQSVMKPIHLGLFKVLRHLPNDGTFDQDASVTRCSEKASEAGK------AFSFDLSAATDRLPVSLTGN    74
SlaMV5:ALM62240.1:293-496     KIRVFALVDSITQSVMKPLHLGLFAVLKQLPNDGTFDQDASVTRCSIKAQQAGK------AFSFDLSAATDRLPVGLTGS    74
RsMV-19:ANR02690.1:13-215     KARLFALIDSITQSILAPLHDHLFKVLEAIPNDGTFDQDASVRRSMAKAQSTGQ------AFSFDLTAATDRLPAALTAA    74
BsMV1:AHY03257.1:312-497      KARVVALTNWWIQCMFKPLHDEIFDKLRNISQDATFDQKGSTTNFIKVNRGKT-------FYCFDLSAATDRLPIQLQSQ    73
CpMV1:NP_660174.1:339-528     KARIVAITNSWIQTAFYSLHLHVFKLLKNIDQDGTFDQERPFKLLIKWLNEPTQK-----FYGFDLTAATDRLPIDLQVD    75
RcMV:AIT71973.1:289-475       KARVVGITDAWTQMLFKPLHDEIYAKLGKIPEDGTRDQLAPVKLLLSNLKNPY-------AVSVDLSAATDRLPVELQAR    73
SsMV15:AHF48631.1:255-442     KCRPFAIFDYWSQLTLTPLHDWIYNQLRNIPNDCTFDQHKGVDKMNQIRSRKW-------FYSYDLKSATDRFPVLLQER    73
HfMV1:AIU44705.1:254-441      KSRVFAIFDYWSQTTLTPLHDWAFSTLRKIPQDCTFNQQEGVEKILRLKTRKH-------YYSYDLKSATDRFPARFQRK    73
SsMV16:AHF48632.1:195-380     KSRPFAIMDYWSQSALTPIHDLLYKLLGEIPEDCTFDQQKGVELMSRFPKSK--------LFSLDLKSATDRFPVELQRR    72
SsMV2:AHX84129.1:246-431      KSRPFAIVDYITQSALTPLHDRLYRVLKSIPQDCTFDQNKGFKDLLYGGGP---------YYSFDLTSATDRFPIFVQEM    71
SlaMV3:ALM62243.1:252-438     KSRIIAVFDYWSQTALLPLHNSLIAFLKKIHNDCTFHQGKISSLMVTFKV----------FMCYDLTAATDRFPVQLQTA    70
OMV5:NP_660180.1:253-438      KMRVIAMVDYHSQFVLKKIHNSLFNKLKLIKSDRTFTQDPIFTTPTMGHR----------FWSMDLSAATDRFPIDLQER    70
SlaMV2:ALM62242.1:235-420     KLRIVAIVDYYTQLFLKPIHEKIMNKLQNLPCDRTYTQSPLNNWKDDGNM----------FWSIDLSSATDRFPISLQRR    70
FgMV1:YP_009126872.1:236-424  KRRIIAMVDYHSQLALRKIHNGLLKLLTTLKCDRTFTQDPKHNWILDNNES---------FHSLDLSSATDRFPVKLQAR    71
TbMV:AAZ99833.1:205-396       KRRIIAMVDYHSQLVLRSIHDGLLNKLRNLPQDRTYNQDPNNAWEENKEC----------FHSLDLSSATDRFPVKLQSR    70
                              * * ..:     *  :  :       :  :  * *  *                          **.:*** :*  .

ObRV2:ASL68498.1:338-532      MLSGLFG-DEF-GAAWYDMMSNTAFRSPERLSSP--LKARVYRFTRGQPLGFYSSWPTFSLTHHMVVWLAAWRVYP---G   143
BcMV1:CEZ26300.1:295-487      ILTLLVS-PSF-AEAWGTALTGRPYKLKFG------TTEDELMYAVGQPMGALSSWGMLALTHHTIVQVAASRAGY---K   141
FpMV4:YP_009272901.1:271-459  ILSRLFG-ESF-AKAWKGLLVERPWY----------HKGIPYLYSVGQPMGALSSWGMLALTHHMIVQVAASRVGH---R   137
OMV3a:NP_660176.1:272-465     IISVLFG-SKF-AKDWATLLVGRDWY----------LKDIPYRYSVGQPMGALSSWAMLALSHHVIVQIAAMRVG----K   142
SsMV3:YP_009182164.1:264-461  ILSIYCG-ESF-ASKWSILMTDRDWYLT--------KEKRNLRYSVGQPMGALSSWAMLALTHHFIVGMAANRVG----K   146
MpMV3:AMM45292.1:465-663      LIGAILG-RES-AAAWANLLTGRTYALRTEDHNGN-ETVIKLTYAVGQPMGALSSWASLALTHHFLVQCAAWSAGFPK-W   146
RsMV-17:ANR02693.1:13-210     LVANLVN-SEF-AANWKTLLVGRKYVAFSPKR----GVSAELTYGVGQPMGALSSWSSLAITHHYMVQVSAWLAGVTPKG   145
RsMV-15:ALD89120.1:432-629    LFKHLVS-EEF-SVQWMRILVGRGYMAFSSKY----NVFKLLEYAVGQPMGALSSWASLAITHHFIVQCAAWEAGICPVG   145
RsMV-14:ALD89119.1:461-658    LMSQLIG-LDL-ATSWANLLTKIEYSINSMKF----NTYEKVKYTVGQPMGALSSWAMLALTHHLLVQISAWRVGAEKEG   145
HetMV1:AIF33766.2:356-558     IISKIFT-PSI-GKAWTTLLVGRDYRVPKRDGQTT-VLPKSVTYAVGQPMGALSSWAMLAITHHFIVQIAAWQSGVTKYG   149
COV:YP_005352912.1:336-529    LIQQLFKLTNAEAEAWRSLLVDRPYYCP--------PLNTSVMYSVGQPMGALSSWAMLAMTHHLIVQVAAWRSGFDK-K   141
BcMV3:YP_009182161.1:299-499  ILSKIFS-DEF-GTSWKQLMVNRDFFFSLTNQKDY-GAPESLRYSVGQPMGALSSWPALALTHHWILQYCSNILGR---T   148
FpMV3:YP_009272900.1:297-500  IIESLFKIPGL-SQSWQKVMVDRNFRFPLNIIENYDIEDIDYRYSVGQPMGCLSSWAGLAITHHWIMQLCSYLVTDS--W   151
SlaMV5:ALM62240.1:293-496     IIESLFQITDL-SNSWKSVMVDREFSFNDKIIKEFPELLKSYRYSVGQPMGCLSSWAGLAITHHWIMQYCSFLLKGN--W   151
RsMV-19:ANR02690.1:13-215     ILESLTG-RPI-GESWLQMMTNRNFWFNGDKATKLGISAGPYQYAVGQPMGGLSSWAGLAITHHWIVQIAALQANKG--N   150
BsMV1:AHY03257.1:312-497      ILSLL----GVRGDLWSQIISNINWH----------YRKKAVTYSVGQPMGAYSSWGMLALTHHVIVRIAAQSVG----I   135
CpMV1:NP_660174.1:339-528     ILNIIFK-NSP-GSSWRSLL-RIKYK----------SPQGFLTYAVGQPMGAYSSFAMLALTHHVIVQVAALNSGF---T   139
RcMV:AIT71973.1:289-475       ILNCL----GLPGDTWREILA-RPYE----------YMDREYVYAVGQPMGAYSSFAMLALTNHLIMYAAANKLGLVV-V   137
SsMV15:AHF48631.1:255-442     VLSLLFN-ADY-AKAWRELLTREPFRLG--------KNGELIKWGAGQPLGAKSSWAVFTLCHHLVVHIAAMR------T   137
HfMV1:AIU44705.1:254-441      VLSLIFT-TDY-ANAWLEIMTKEPFTLK--------GSETAIKWRVGQPLGAKSSWGIFTLCHHVVVNIAAIR------S   137
SsMV16:AHF48632.1:195-380     ILGLLID-PEY-AAAWSHVVNGYDFE----------YQGSPIRWSVGQPLGAKSSWALFTLSHHVVVQVAIARA-----K   135
SsMV2:AHX84129.1:246-431      VLSWLVN-EQY-ASAWVQAMVKLPFRT---------PEGTEVKFECGQPLGAKSSWAMFTISHHFVVQYCAMVL-----N   135
SlaMV3:ALM62243.1:252-438     ILAELFG-SAF-ASAWKTILVEYSFHMK--------GKGDVL-YGAGQPMGAYSSWTVFALCHHIVVQYAASLAGY---T   136
OMV5:NP_660180.1:253-438      LLSYLYG-SEI-SSAWKQLLIDRTYKT---------PEGDELHYKVGQPMGAYSSWAAFTLTHHLVVFYSARMAG----I   135
SlaMV2:ALM62242.1:235-420     LLEIAIS-KEV-ADGWNHILSTRKFET---------PEGNLISYEVGQPMGSYSSWAAFTLTHHLVVHWCAKLNG----I   135
FgMV1:YP_009126872.1:236-424  LLGEIYRDYKF-AQNWLELLLNRDYIG---------PNGEVCRYSVGQPMGAYSSWAAFTLTHHLVVAWSAYKSRK---T   138
TbMV:AAZ99833.1:205-396       LLTEMYSDPSF-GENWMNLLLNRDYLLPEEG-----LSGERLRYAVGQPMGAYSSWAAFTLSHHLVVAWCTYKSKK---V   141
                              :.          .  *   :    :                  :  ***:*  **:  ::: :* ::  .

ObRV2:ASL68498.1:338-532      KKFWDYALLGDDIVIADEAVALEYRDIMQQ-MGGVINMTKSLISHNGC-CEFAK                             195
BcMV1:CEZ26300.1:295-487      DLFLDYALLGDDICIANKAVADNYLLIMRD-LGVEINLSKSLISSTGV-VEFAK                             193
FpMV4:YP_009272901.1:271-459  TMFRDYALLGDDICIADSAVAKSYLSLMTD-YGVDINLSKSLESDIGV-AEFAK                             189
OMV3a:NP_660176.1:272-465     LPFTNYALLGDDIVIADKAVATSYHMIMTQILGVEINLSKSLVSNNSF--EFAK                             194
SsMV3:YP_009182164.1:264-461  LGFNHYALLGDDIVIADKSVADSYYMIMTEILGVQINLSKSLVSTNSF--EFAK                             198
MpMV3:AMM45292.1:465-663      KLYTNYAVLGDDVVIGDRSVALAYLRIMKS-LGVGVNTSKSLLSHRGIAFEFAK                             199
RsMV-17:ANR02693.1:13-210     EWFTDYAVLGDDLVIFNQRVKVHYLKIVAA-MGVQCGAHKSLLSPRGVAIEFAK                             198
RsMV-15:ALD89120.1:432-629    TWFTSYAVLGDDLVIFDQRVKVKYLAIVDA-LGVQCGIAKSLLSPTGTAIEFAK                             198
RsMV-14:ALD89119.1:461-658    TWFSDYAILGDDLVIFNYDVAHEYIKVIHK-IGMEIGLHKSVLSRKFTSLEFAK                             198
HetMV1:AIF33766.2:356-558     KWFTAYALLGDDLTVGNRAVSNRYQLIMIKHLGVEVGLAKSIISDKGIGLEFAK                             203
COV:YP_005352912.1:336-529    KLFRAYAVLGDDIVIFDKKVAKYYHRVLTG-LGVECNLAKSIMSHRGIGLEFAK                             194
BcMV3:YP_009182161.1:299-499  GWEENYEILGDDLVVFDSALADKYLEIAKI-LGVEINLTKSISSHDRPVFEFAK                             201
FpMV3:YP_009272900.1:297-500  DWEERYEILGDDIVIFDEELANTYLQVMEW-LGLDINLSKSIISKNKPTFEFAK                             204
SlaMV5:ALM62240.1:293-496     NWEDRYEVLGDDIVIFDTLLANQYLEVMSH-LGLEINLSKSINASNTPVFEFAK                             204
RsMV-19:ANR02690.1:13-215     SWYINYEVLGDDLVIFEADVAREYLKIMAD-LGTEINLTKSIQSPERPAFEFAK                             203
BsMV1:AHY03257.1:312-497      NKFSEYLVLGDDIVIANNEVADAYLVLMKT-LGVDINLSKSLISSQLL--EFAK                             186
CpMV1:NP_660174.1:339-528     TRFTDYCILGDDIVIAHDTVASEYLKLMET-LGLSISSGKSVISSEFT--EFAK                             190
RcMV:AIT71973.1:289-475       KGSGLYAILGDDVAISRGDLATEYNKIMHL-LGVEINPIKGFTGKIL---EFAK                             187
SsMV15:AHF48631.1:255-442     NSPLDYVLLGDDIVIRGRALATEYKRIMFQ-LGVTISDAKTHVSKDTF--EFAK                             188
HfMV1:AIU44705.1:254-441      NSDAHYIILGDDIVLRGRALATEYKRIMSN-LGVEISESKSHVSKDTF--EFAK                             188
SsMV16:AHF48632.1:195-380     ARPDCYVLLGDDIVLRGSKVATEYLKIMDG-LGVEISPAKSHISKGMF--EFAK                             186
SsMV2:AHX84129.1:246-431      IDNPRYKILGDDIVICDHALAAKYLEIMSQ-LGVEISPVKTHVSENLF--EFAK                             186
SlaMV3:ALM62243.1:252-438     SWYSNYALLGDDLVIGDEDVGHHYKIVMLS-LGVLISSNKTHEGNLLC--EFAK                             187
OMV5:NP_660180.1:253-438      KDFTNYILLGDDIVINNDKVAKYYIRTMKR-LGVELSMNKTHVSKNTY--EFAK                             186
SlaMV2:ALM62242.1:235-420     NNFSDYILLGDDIVIKNDKVAKTYMKWMNY-IGVELSESKTHVSKDTY--EFAK                             186
FgMV1:YP_009126872.1:236-424  MDFDQYIILGDDIVIKDNAIASIYRGQMMR-MGVDISIPKTHVSKDTY--EFAK                             189
TbMV:AAZ99833.1:205-396       IRSSQYIILGDDIVIKDNDIARKYIGQMSK-LGVAISMQKTHVSKDTY--EFAK                             192
                                   * :****: :    :   *        *   .  *   .      ****
```

Asterisks (*), fully conserved residues; colons (:), residues with strongly similar properties; and periods (.), residues with weakly similar properties. See Table 2 for virus names.