Arch Virol, accepted for publication: Jan 21 2014

DOI: 10.1007/s00705-014-1999-1

Title: A unique mitovirus from Glomeromycota, the phylum of arbuscular mycorrhizal fungi

Authors: Ryoko Kitahara, Yoji Ikeda, Hanako Shimura, Chikara Masuta, and Tatsuhiro Ezawa*.

Address: Graduate School of Agriculture, Hokkaido University, Sapporo 060-8589 Japan

Authors for correspondence:

Tatsuhiro Ezawa

Graduate School of Agriculture, Hokkaido University, Sapporo 060-8589 Japan

Tel +81-11-706-3845; Fax +81-11-706-3845

Email tatsu@res.agr.hokudai.ac.jp

No. of table: N/A

No. of figure: two

No. of color figure for online publication: one (Fig. 1)

Supplementary information: three figures and one table

Abstract

Arbuscular mycorrhizal (AM) fungi that belong to the phylum Glomeromycota associate with most land plants and supply mineral nutrients to the host plants. One of the four viral segments found by deep-sequencing of dsRNA in the AM fungus *Rhizophagus clarus* strain RF1 showed similarity to mitoviruses and is characterized in this report. The genome segment is 2,895 nucleotides in length, and the largest ORF was predicted by applying either the mold mitochondrial or the universal genetic code. The ORF encodes a polypeptide of 820 amino acids with a molecular mass of 91.2 kDa and conserves the domain of the mitovirus RdRp superfamily. Accordingly, the dsRNA was designated as *R. clarus* mitovirus 1 strain RF1 (RcMV1-RF1). Mitoviruses are localized exclusively in mitochondria and thus generally employ the mold mitochondrial genetic code. The distinct codon usage of RcMV1-RF1, however, suggests that the virus is potentially able to replicate not only in mitochondria but also in the cytoplasm. RcMV1-RF1 RdRp showed the highest similarity to the putative RdRp of a mitovirus-like ssRNA found in another AM fungus, followed by RdRp of a mitovirus in an ascomycotan ectomycorrhizal fungus. The three mitoviruses found in the three mycorrhizal fungi formed a deeply branching clade that is distinct from the two major clades in the genus *Mitovirus*.

Introduction

Arbuscular mycorrhizal (AM) fungi that belong to the phylum Glomeromycota associate with most land plants and supply mineral nutrients, in particular phosphorus, to the host plants through extensive hyphal networks constructed in the soil [15]. The plant-AM fungal symbiosis occurred more than 400 million years ago, and the coincidence of the appearances of early land plants and AM associations suggests that the associations were instrumental in the colonization of land by plants [14]. Although AM fungi have been playing a significant role in terrestrial ecosystems via enhancing P-cycling in the soil, biological characteristics of the fungi have been poorly understood due to their obligate biotrophic nature.

Members of the genus *Mitovirus* in the family Narnaviridae composed of a single genome segment of positive-sense RNA that encodes only RNA-dependent RNA polymerase (RdRp) [3]. Mitoviruses are localized exclusively in mitochondria of the host fungi, except for *Thanatephorus cucumeris* mitovirus that is potentially able to replicate both in the cytosol and mitochondria [6]. The infection of mitoviruses often causes malformation of mitochondria, which leads, in the case of plant pathogenic fungi, to debilitation in virulence [18] due to attenuation of mitochondrial function [12]. Accordingly, their possibility as a biological control agent has been studied extensively [1]. The impact of mitoviruses on AM symbiosis is also of interest, but no mitovirus has been described in the Glomeromycota so far.

One technical limitation for virological study in AM fungi was the difficulty in obtaining a sufficient amount of fungal material for characterization of viral genomes. We have established an open culture system for mass production of AM fungal mycelia and initiated virological studies of the fungi recently, in which four distinct dsRNA viruses, including a new class of virus, were described for the first time in the phylum [5]. In the present study, one dsRNA that was found to be similar to mitoviruses in the previous study is characterized with reference to the members of the genus *Mitovirus*.

Provenance of the virus material

Rhizophagus clarus (Nicolson & Schenck) Walker & Schüßler strain RF1 (= *Glomus* sp. strain RF1) MAFF520086 was isolated by plant trap culture of *Petasites japonicus* subsp. *giganteus* grown in acidic soil in Hokkaido, Japan in 2005 [5] and has been maintained with sorghum and groundnut grown in a greenhouse. To obtain fungal material, the strain was grown with seedlings of *Lotus japonicus* cv. Miyakojima in the mesh bag-separated open culture system [2], and dsRNA was extracted from extraradical mycelia, purified, and electrophoresed [5]. Four dsRNA segments observed in the gel were excised from the gel, purified, and randomly amplified using the anchored-N6 primer according to Márquez et al [7]. The amplicons were directly sequenced by Roche 454 FLX GS Titanium using a 1/8-scale gasket, and assembled. Among contigs obtained in the sequencing, an ORF of a

2.5-kbp contig showed similarity to RNA-dependent RNA polymerase (RdRp) of mitoviruses. The nucleotide sequence of the coding region of 2.5-kbp dsRNA was reconfirmed by sequencing three clones for each of two >1-kbp cDNAs obtained by nested RT-PCR, and the extreme ends were determined by sequencing three clones for each of three and two RACE products of the 5′ and 3′ ends, respectively (Supplementary Table S1 and Fig. S1). The sequences were analyzed and annotated with Artemis (Sangar Institute) and has been deposited in the DDBJ under accession no. AB558120. The amino acid (aa) sequence of predicted ORF was subjected to BLASTp searches and aligned with those of other mitoviruses using MUSCLE implemented in MEGA 5 [17]. Neighbor-joining (NJ) and maximum-likelihood (ML) trees were constructed with MEGA 5 for phylogenetic analysis. Four well-characterized mitoviruses and an uncharacterized mitovirus-like ssRNA were selected for comparative sequence analysis of the dsRNA of *R. clarus* RF1: TeMV found in the ectomycorrhizal fungus *Tuber excavatum* in Germany [16], CpMV found in a hypovirulent strain of the chestnut blight fungus *Cryphonectria parasitica* in USA [10], TcMV found in a hypovirulent strain of *Th. cucumeris* in USA [6], HmMV1-18 found in the violet root rot fungus *Helicobasidium mompa* in Japan [8], and an uncharacterized mitovirus-like ssRNA found in the AM fungus *Rhizophagus* sp. strain HR1 (= *Glomus* sp. strain HR1 [2]) (RMV-like ssRNA-HR1) in Japan.

Sequence properties

The sequencing of the RACE products revealed that complete genome of the dsRNA was 2,895 nucleotides (nt) in length, which was approx. 400-nt longer than that predicted by the 454 sequencing. Between a 297-nt 5′ UTR and a 135-nt 3' UTR, the largest ORF (2,463 nt) was predicted by applying either the mold mitochondrial or the universal genetic code (Supplementary Fig. S2). The predicted ORF encodes a polypeptide of 820 amino acids (aa) with a molecular mass of 91.2 kDa and conserves the domain of mitovirus RdRp superfamily (Pfam PF05919), including the GDD motif (Fig. 1a). Accordingly, the dsRNA was designated as *R. clarus* mitovirus 1 strain RF1 (RcMV1-RF1). Generally functional RdRp in mitoviruses can be translated only if the mold mitochondrial genetic code is invoked [13]. This is because tryptophan residues in mitovirus RdRps are usually encoded either by a UGA or a UGG codon, but the former codon encodes a translation terminator in the universal genetic code (in the cytosol). In fact, 55, 52, and 84% of tryptophan residues are encoded by the UGA codon in the RdRps of TeMV, CpMV, and HmMV1-18, respectively. On the other hand, all tryptophan residues in RcMV1-RF1 RdRp are encoded by the UGG codon (Supplementary Fig. S2, TGG in cDNA) as well as those in TcMV RdRp [6] and putative RdRp of RMV-like ssRNA-HR1 (data not shown), suggesting that functional RdRp could be translated both in the cytosol and in mitochondria. The codons for all tryptophan residues within the conserved domain of the selected mitoviruses are shown in Fig. 1b. The

RdRp aa sequence of RcMV1-RF1 shows high levels of similarity to those found in the two mycorrhizal fungi throughout the ORF: 34% identity to that of RMV-like ssRNA-HR1 at 98% coverage and 28% identity to TeMV RdRp at 96% coverage. Significant similarity to TcMV RdRp in which all tryptophan residues are encoded by the UGG codon, however, was observed only within the conserved domain (43% identity at 23% coverage). The three RdRps of RcMV1-RF1, RMV-like ssRNA-HR1, and TeMV found in the mycorrhizal fungi form a subclade within the *Mitovirus* clade I [3] in the NJ-tree (Supplementary Fig. S3), although the node separating the clades I and II is poorly supported by a low-bootstrap value (28%). Whereas in the ML-tree the three viral sequences form a deeply branching clade with a bootstrap value of 99%, which is distinct from the two major clades (Fig. 2). A similar tree topology was also reported recently [4]. These observations suggest that the mitoviruses from the mycorrhizal fungi is likely to create the third distinct group in the genus.

The first member of *Mitovirus* in the Glomeromycota has been characterized in the present study. It seems likely that the distinct codon usage found in RcMV1-RF1 is a common feature of mitoviruses in AM fungi. The virus is potentially capable of replicating in the cytoplasm as well as in mitochondria. This might be an advantageous trait for horizontal transmission among the fungi, because those that belong to the same anastomosis group can exchange not only nuclei but also cytosol. Given the 400-million-year history of the close association of the fungi with plants, we also consider another possibility that ancestors of RcMV1-RF1 might be able to shuttle between the fungi and the host plant during a certain stage of their evolution. This idea is supported by the evidence that RdRps of the members in the genus *Ourmiavirus*, plant ssRNA viruses, are phylogenetically related to those of the members in the Narnaviridae [11], suggesting that mitoviruses and ourmiaviruses diverged from a common ancestor. It is thus expected that more mitoviruses employing the universal genetic code will be found in AM fungi when their sequences become available.

Acknowledgements

This study was partially supported by The Asahi Glass Foundation (CM and TE) and the Grant-in-Aid for Scientific Research (22380042) from Japan Society for the Promotion of Science (TE).

References

1. Ghabrial SA, Suzuki N (2009) Viruses of Plant Pathogenic Fungi. Ann Rev Phytopathol 47:353-384
2. Hijikata N, Murase M, Tani C, Ohtomo R, Osaki M, Ezawa T (2010) Polyphosphate has a central role in the rapid and massive accumulation of phosphorus in extraradical mycelium of an arbuscular mycorrhizal fungus. New Phytol 186:285-289
3. Hillman BI, Cai G (2013) Chapter Six - The Family Narnaviridae: Simplest of RNA Viruses. In: Ghabrial SA (ed) Advances in Virus Research. Academic Press, pp 149-176
4. Hintz WE, Carneiro JS, Kassatenko I, Varga A, James D (2013) Two novel mitoviruses from a Canadian isolate of the Dutch elm pathogen *Ophiostoma novo-ulmi* (93–1224). Virol J 10:252
5. Ikeda Y, Shimura H, Kitahara R, Masuta C, Ezawa T (2012) A novel virus-like double-stranded rna in an obligate biotroph arbuscular mycorrhizal fungus: a hidden player in mycorrhizal symbiosis. Mol Plant-Microbe Interact 25:1005-1012
6. Lakshman DK, Jian J, Tavantzis SM (1998) A double-stranded RNA element from a hypovirulent strain of *Rhizoctonia solani* occurs in DNA form and is genetically related to the pentafunctional AROM protein of the shikimate pathway. Proc Nat Acad Sci USA 95:6425-6429

7. Márquez LM, Redman RS, Rodriguez RJ, Roossinck MJ (2007) A virus in a fungus in a plant: three-way symbiosis required for thermal tolerance. Science 315:513-515

8. Osaki H, Nakamura H, Nomura K, Matsumoto N, Yoshida K (2005) Nucleotide sequence of a mitochondrial RNA virus from the plant pathogenic fungus, *Helicobasidium mompa* Tanaka. Virus Res 107:39-46

9. Poch O, Sauvaget I, Delarue M, Tordo N (1989) Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. EMBO J 8:3867-3874

10. Polashock JJ, Hillman BI (1994) A small mitochondrial double-stranded (ds) RNA element associated with a hypovirulent strain of the chestnut blight fungus and ancestrally related to yeast cytoplasmic T and W dsRNAs. Proc Nat Acad Sci USA 91:8680-8684

11. Rastgou M, Habibi MK, Izadpanah K, Masenga V, Milne RG, Wolf YI, Koonin EV, Turina M (2009) Molecular characterization of the plant virus genus *Ourmiavirus* and evidence of inter-kingdom reassortment of viral genome segments as its possible route of origin. J Gen Virol 90:2525-2535

12. Rogers HJ, Buck KW, Brasier CM (1987) A mitochondrial target for double-stranded RNA in diseased isolates of the fungus that causes Dutch elm disease. Nature 329:558-560

13. Shackelton LA, Holmes EC (2008) The role of alternative genetic codes in viral evolution and emergence. J Theoret Biol 254:128-134

14. Simon L, Bousquet J, Levesque RC, Lalonde M (1993) Origin and diversification of endomycorrhizal fungi and coincidence with vascular land plants. Nature 363:67-69
15. Smith SE, Read DJ (2008) Mycorrhizal symbiosis, 3rd edn. Academic Press
16. Stielow JB, Bratek Z, Klenk H-P, Winter S, Menzel W (2012) A novel mitovirus from the hypogeous ectomycorrhizal fungus *Tuber excavatum*. Arch Virol 157:787-790
17. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731-2739
18. Wu M, Zhang L, Li G, Jiang D, Ghabrial SA (2010) Genome characterization of a debilitation-associated mitovirus infecting the phytopathogenic fungus *Botrytis cinerea*. Virol 406:117-126

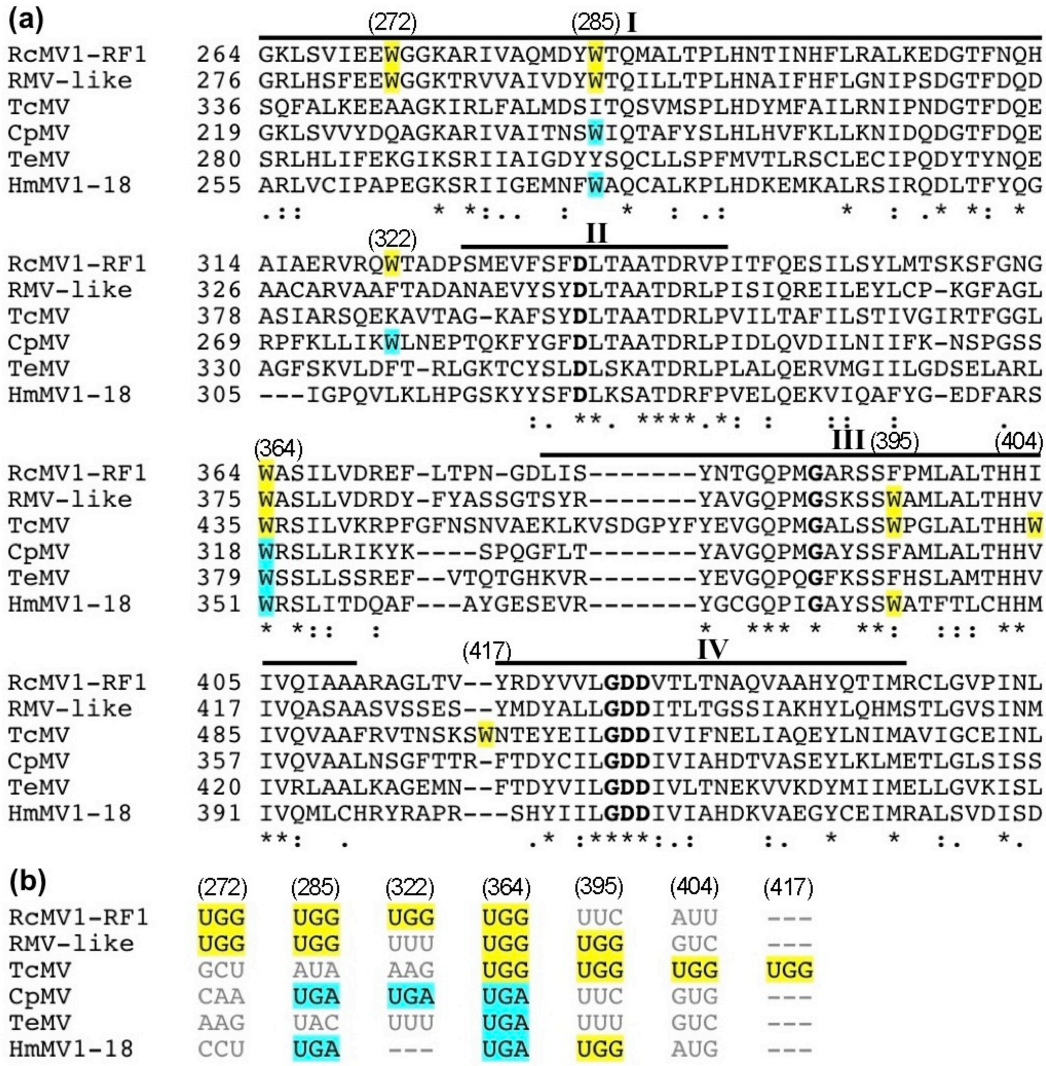


Fig. 1 a) Alignment of the conserved domain of *Mitovirus* RdRp in RcMV1-RF1. Motifs are labeled according to Poch et al. [9], and the consensus amino acids are written in bold letters. Numbers in parentheses represent the number of amino acid residue in RcMV1-RF1 RdRp. **b)** Codons for the tryptophan residues (W) corresponding to those indicated in the alignment

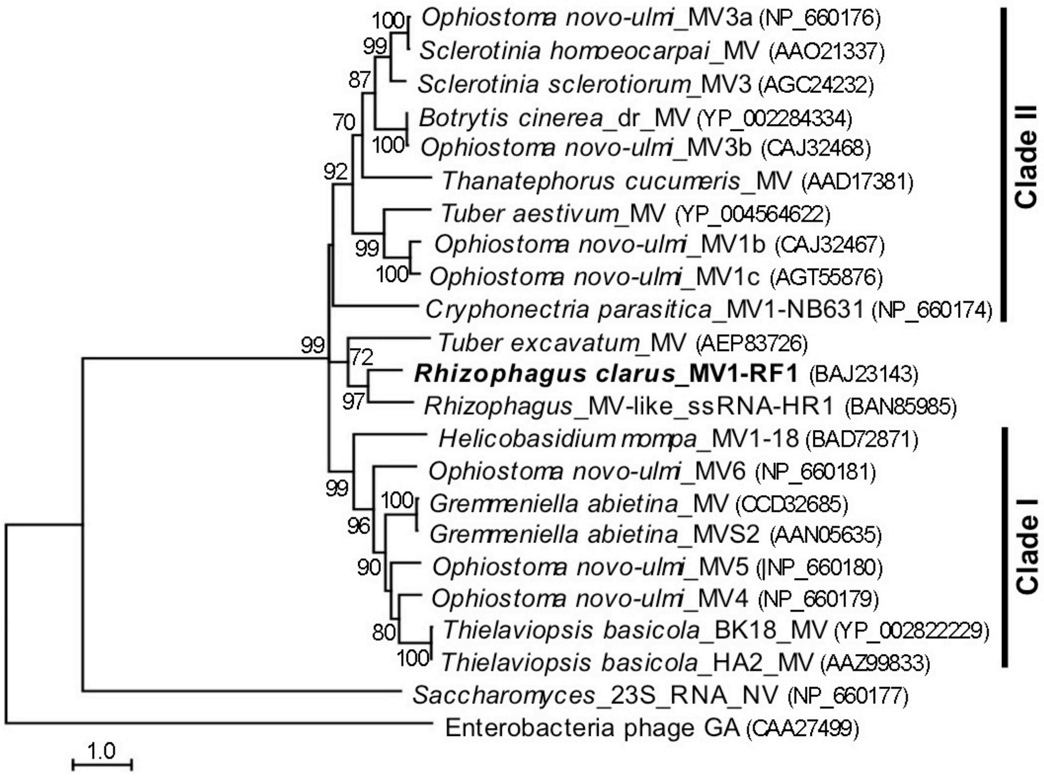
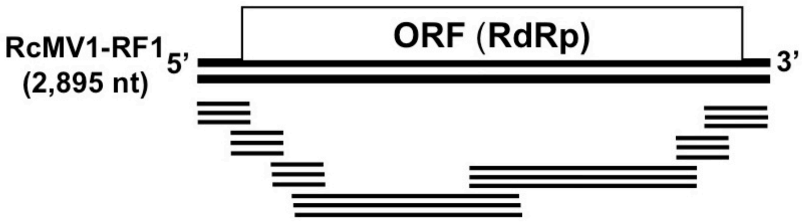


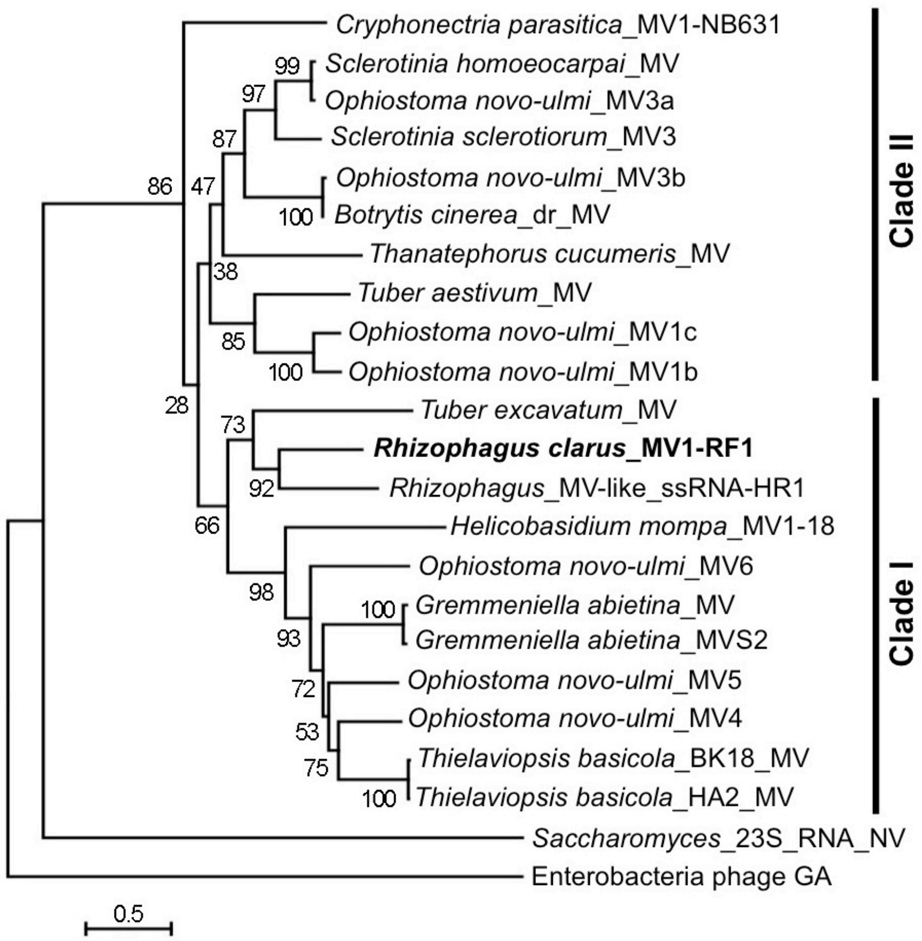
Fig. 2 Phylogenetic position of RcMV1-RF1. Maximum-likelihood tree was constructed based on the amino acid sequences of mitovirus RdRP according to the JTT matrix-based model. Percentage bootstrap values (1000 replication) are indicated at the nodes. Two major clades (I and II) in *Mitovirus* are labeled according to Hillman and Cai [3]. Accession numbers are given in parentheses



Supplementary Fig. S1 Genome structure of *R. clarus* mitovirus 1 strain RF1 (RcMV1-RF1). An ORF encoding RNA-dependent RNA polymerase (RdRp) was predicted. Relative positions and size of cDNA clones sequenced for confirmation are drawn below the genome

```
acgatgctacaagcggagaggttggacgcttgggtgataaaagccaatcattctagtagg  60
gtagtgcccgaacggatctcacaaacctagtggggttgagcgacagaccggtcggtatat  120
tgcatcgtaagacagcgggcctgggtagctttgttgacgtactgatcgccgtggagccca  180
ttatattcaaaagtagctttatagtgtaccactcttacattatcctttcccaaaattccc  240
caggaagccgcagttacctctcattcacctaacactcaatacattctacttgttgtcatg  300
                                                          M
tcaacatttaatttcaactctcctaatgaggtgagacgtgtcacacgctctaacatagct  360
S  T  F  N  F  N  S  P  N  E  V  R  R  V  T  R  S  N  I  A
agcttgtttgctgccctcatgcatctcaattctcttgtaactgtaacttattctcaacgg  420
S  L  F  A  A  L  M  H  L  N  S  L  V  T  V  T  Y  S  Q  R
cttcttgcctttctgggagcagtgtataatcgtattctcgtccttttcgatgccaacgca  480
L  L  A  F  L  G  A  V  Y  N  R  I  L  V  L  F  D  A  N  A
cgttctcttatttccgagcttaagctcgttcgccggtggtttcttgaattcatccgtaac  540
R  S  L  I  S  E  L  K  L  V  R  R  W  F  L  E  F  I  R  N
gggaacactgataacccgggtctcgaatggacaggtgggatgactctaacaactgtccg  600
G  N  T  D  N  P  G  L  E  W  D  R  W  D  D  S  N  N  C  P
gtactcttagagggtctcgacgctctttgggatgagatagaatctccatcggactctcac  660
V  L  L  E  G  L  D  A  L  W  D  E  I  E  S  P  S  D  S  H
tatacggaccacgcggcacagcttatatttactcttttatctatagaccgtatcatcgtt  720
Y  T  D  H  A  A  Q  L  I  F  T  L  L  S  I  D  R  I  I  V
gtacctgcagttcctaactattctactatagaggatggccctgtcattgaacctgtgggc  780
V  P  A  V  P  N  Y  S  T  I  E  D  G  P  V  I  E  P  V  G
tatcctactaacgaggagcttgtgaacgctctctcttctcttcatattgacccagcggcc  840
Y  P  T  N  E  E  L  V  N  A  L  S  S  L  H  I  D  P  A  A
ttcaaggccttctataaccaacaggtacacgacttcgattacgaggttctttcaactagg  900
F  K  A  F  Y  N  Q  Q  V  H  D  F  D  Y  E  V  L  S  T  R
ggccccaatggggatgctacgtggactgcccatttagacgcgagagcttgggctttggac  960
G  P  N  G  D  A  T  W  T  A  H  L  D  A  R  A  W  A  L  D
actgagcttttcagacgttttagtgcgtggctcgaggagtcacgcctgactcgtatctta  1020
T  E  L  F  R  R  F  S  A  W  L  E  E  S  R  L  T  R  I  L
cgtgacctttcgggtgtattcgttcagcagccgctgaggctatccctaaccttagtccc  1080
R  D  L  F  G  C  I  R  S  A  A  A  E  A  I  P  N  L  S  P
atacttgggaagttatctgttatcgaggagtggggcgggaaagcccgcatcgtagcccag  1140
I  L  G  K  L  S  V  I  E  E  W  G  G  K  A  R  I  V  A  Q
atggactactggacgcaaatggctctcacacctcttcataatactatcaaccatttcctt  1200
M  D  Y  W  T  Q  M  A  L  T  P  L  H  N  T  I  N  H  F  L
cgcgctcttaaggaggatggcactttcaatcaacatgccatcgcagagcgagttcgtcag  1260
R  A  L  K  E  D  G  T  F  N  Q  H  A  I  A  E  R  V  R  Q
tggacagcggatccgtcgatggaggtcttctctttcgatttaactgctgctactgatcga  1320
W  T  A  D  P  S  M  E  V  F  S  F  D  L  T  A  A  T  D  R
gtacctattactttccaggagtcaatcctttcttatcttatgacgtcgaaatcgttcggc  1380
V  P  I  T  F  Q  E  S  I  L  S  Y  L  M  T  S  K  S  F  G
aacgggtgggcctctattctggtagatagagagttccttactccaaatggcgaccttatt  1440
N  G  W  A  S  I  L  V  D  R  E  F  L  T  P  N  G  D  L  I
tcttataatactggccaaccaatggggcacggtcgtcattccctatgttggcacttacg  1500
S  Y  N  T  G  Q  P  M  G  A  R  S  S  F  P  M  L  A  L  T
catcatattatcgtgcagattgctgcagcgcgggcaggtcttaccgtatatcgggattat  1560
H  H  I  I  V  Q  I  A  A  A  R  A  G  L  T  V  Y  R  D  Y
gtcgtacttggtgacgatgttacattaactaacgctcaggtagcagcccactaccagacg  1620
V  V  L  G  D  D  V  T  L  T  N  A  Q  V  A  A  H  Y  Q  T
ataatgcggtgtcttggtgttcctatcaacctttctaagtctatagtccacgtggatgga  1680
I  M  R  C  L  G  V  P  I  N  L  S  K  S  I  V  H  V  D  G
ggcgtatcaatggccgaaatctgtaaacgtgttttcatggatggagtggagattagccgc  1740
G  V  S  M  A  E  I  C  K  R  V  F  M  D  G  V  E  I  S  R
ttcaaccccaaacttattgttaacgtcatacgtgacggtcgtttaggtcctgaccttcag  1800
F  N  P  K  L  I  V  N  V  I  R  D  G  R  L  G  P  D  L  Q
aacgatcttattatccgtgggtgggatccctctaacgaggtattctggaagttcatggct  1860
N  D  L  I  I  R  G  W  D  P  S  N  E  V  F  W  K  F  M  A
ggtctcttatctatcgataaccttactcttcttatacgccttaactgtgcacctatctct  1920
G  L  L  S  I  D  N  L  T  L  L  I  R  L  N  C  A  P  I  S
attactggtcttcttcggcaatttgcctccaactctaagttggcacaactctcagcttgg  1980
I  T  G  L  L  R  Q  F  A  S  N  S  K  L  A  Q  L  S  A  W
atcccggcctaccaggatctcaaaccggaacatctagttgagctctttacttacgtcacc  2040
I  P  A  Y  Q  D  L  K  P  E  H  L  V  E  L  F  T  Y  V  T
gcttctgaggctctgaagcgtcttgacggtatcctacgtgcggctgtcactattaacgat  2100
A  S  E  A  L  K  R  L  D  G  I  L  R  A  A  V  T  I  N  D
tctctttctatcatcgctgctgcgaacgctcaccctgaccgtattccacagtacgtacgg  2160
S  L  S  I  I  A  A  A  N  A  H  P  D  R  I  P  Q  Y  V  R
gacacttggttaggcgaaggcctaactaaggaggaacgggcacgcctggagggacttatt  2220
D  T  W  L  G  E  G  L  T  K  E  E  R  A  R  L  E  G  L  I
gcgtctacaggccccataactcctaaccatcctttagtatcagcttctcgtgcagaagct  2280
A  S  T  G  P  I  T  P  N  H  P  L  V  S  A  S  R  A  E  A
aaccgtatttcggagctgcttcatcagctcaattcacatgatactgctattatcaccagg  2340
N  R  I  S  E  L  L  H  Q  L  N  S  H  D  T  A  I  I  T  R
gccagacttggtcttctggatgtgttccgtacttctatctcttctatttggctcgacgat  2400
A  R  L  G  L  L  D  V  F  R  T  S  I  S  S  I  W  L  D  D
ggtaacatcagagcgggtgagtctcgttctatcttcacgcgtatgttaactactcttgta  2460
G  N  I  R  A  G  E  S  R  S  I  F  T  R  M  L  T  T  L  V
tccctgttcactagtgaaaagcgggtatcgaaatcaggacgtaacttatcactgtcatac  2520
S  L  F  T  S  E  K  R  V  S  K  S  G  R  N  L  S  L  S  Y
tcagttgttctaacatctttatctcgtctctggactgtagccttagacttcggcggtcag  2580
S  V  V  L  T  S  L  S  R  L  W  T  V  A  L  D  F  G  G  Q
gtaactgttaacgcacttcgtgcgaacgttactcgtgatatccacaacgctgtggataac  2640
V  T  V  N  A  L  R  A  N  V  T  R  D  I  H  N  A  V  D  N
ctgaaagccgcggaggaggctgctgttcttatctcttcttcttcagtccctactacatct  2700
L  K  A  A  E  E  A  A  V  L  I  S  S  S  S  V  P  T  T  S
cctacgcctgcgactcctaagggcatccgcagacgccgtgcattcgcacggatttcctaa  2760
P  T  P  A  T  P  K  G  I  R  R  R  R  A  F  A  R  I  S
aaatcccccccgggattctgagtatggggacttgtaaattttatcataggggggttcctt  2820
cttgcaacttagtgtagtagtaagcacagtgtaatgaatccacacaaggccggagcataa  2880
ccggggttgcaccct                                               2895
```

Supplementary Fig. S2 cDNA sequence of RcMV1-RF1 and predicted ORF according to either the mitochondrial or the universal genetic code. All tryptophan residues are written in a red letter, and their codons are highlighted



Supplementary Fig. S3 Phylogenetic position of RcMV1-RF1. The Neighbor-joining tree was constructed based on the JTT matrix-based model. Percentage bootstrap values (1000 replication) are indicated at the nodes. Two major clades (I and II) in *Mitovirus* are labeled according to Hillman and Cai [3]. Accession numbers are indicated in Fig. 2