Virology 435 (2013) 341–349

Contents lists available at SciVerse ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

A new virus from the plant pathogenic oomycete *Phytophthora infestans* with an 8 kb dsRNA genome: The sixth member of a proposed new virus genus

Guohong Cai [a], Jason F. Krychiw [a], Kevin Myers [b], William E. Fry [b], Bradley I. Hillman [a,*]

[a] Department of Plant Biology and Pathology, Rutgers, The State University of New Jersey, New Brunswick, NJ 08901, USA
[b] Department of Plant Pathology and Plant-Microbe Biology, Cornell University, Ithaca, NY 14853, USA

ARTICLE INFO

Article history:
Received 25 June 2012
Returned to author for revisions
25 July 2012
Accepted 6 October 2012
Available online 10 November 2012

Keywords:
Oomycete
Phytophthora infestans
Potato late blight
RNA virus
Fungal virus
New virus genus

ABSTRACT

A virus designated *Phytophthora infestans* RNA virus 3 (PiRV-3) was characterized from an isolate of *P. infestans* that was co-infected with a second previously described virus, PiRV-4, a member of the virus family *Narnaviridae* (Cai et al., 2012). The genome of PiRV-3 is 8112 nt and one strand, designated the positive strand, has two deduced overlapping open reading frames linked by a potential frameshift sequence. The first open reading frame (ORF1) is predicted to encode a protein of unknown function, and ORF2 is predicted to encode an RNA-dependent RNA polymerase (RdRp) most closely related to six unclassified dsRNA viruses of filamentous fungi. The genome organizations of five of the related viruses are similar to PiRV-3, indicating taxonomic linkage among those viruses. We suggest that PiRV-3 and related viruses should be collected into a new virus taxon.

Introduction

Phytophthora infestans is the causal agent of potato and tomato late blight and the pathogen that led to the Irish potato famine in the 1840s (Fry, 2008). It continues to be a threat to potatoes and tomatoes worldwide and causes billions of dollars in yield loss and control cost (Goodwin, 1997). *P. infestans* is a member and a model organism in the oomycetes, a group with many highly destructive plant pathogens, such as downy mildews, the sudden oak death pathogen *Phytophthora ramorum*, and *Phytophthora capsici*, which causes Phytophthora blight on many vegetables. Oomycetes have similar morphology and occupy the same habitats as fungi, but the two groups are evolutionally distant. Oomycetes are related to brown algae and diatoms and belong to the Stramenopila clade in the super group Chromalveolata (Adl et al., 2005).

In a screening for RNA viruses in *P. infestans*, we found five double-stranded RNA (dsRNA) segments from multiple isolates estimated to be 2.9–11.2 kbp (Cai et al., 2009). The five dsRNAs have been characterized by sequence analysis and biological properties and found to represent the genomes of four independent RNA viruses, designated *P. infestans* RNA virus (PiRV) 1 through 4. The PiRV-1 genome consists of two positive-sense RNAs of 3160 nt and 2776 nt, respectively, each with a poly(A) tail, and is most closely related to the *Astroviridae* family of animal-infecting positive-sense RNA viruses. PiRV-2 contains the largest of the five dsRNAs, an 11,170 bp dsRNA that is not closely related to any known virus and appears to represent a novel dsRNA virus group (Cai et al., unpublished). The smallest of the five dsRNA segments, a 3 kb element, constitutes the genome of PiRV-4, a member of the *Narnaviridae* family of positive-sense, ssRNA viruses (Cai et al., 2012). Here we report the sequence and characterization of PiRV-3, a virus that shares a similar genome structure and phylogenetic relationships with several recently characterized fungal viruses. Characterization here of PiRV-3 bolsters the case for description of a new virus family or genus that includes these viruses.

Results and discussion

PiRV-3 sequence and genome organization

Two dsRNA segments were found in a single *P. infestans* isolate from an infected tomato sample collected in Florida in 2005, FLa2005 (Fig. 1). Preliminary sequence analysis of the two elements indicated that they belong to two distinct viruses. The smaller 3 kbp dsRNA was also isolated independently in other *P. infestans* isolates and was found to represent the genome of a virus within the family *Narnaviridae* named PiRV-4 (Cai et al.,

* Corresponding author. Fax: +1 866 365 7736.
E-mail addresses: cai@aesop.rutgers.edu (G. Cai), hillman@aesop.rutgers.edu (B.I. Hillman).

0042-6822/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.virol.2012.10.012

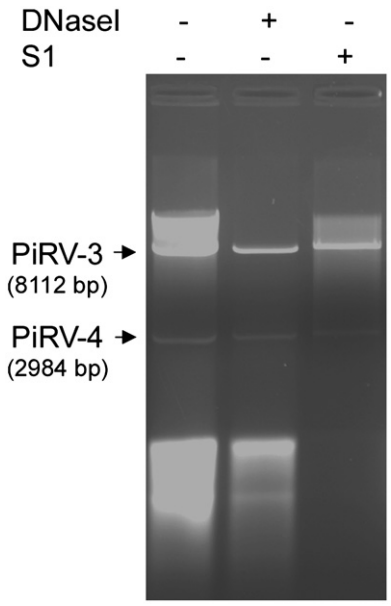

Fig. 1. Agarose-gel electrophoresis of dsRNAs extracted from isolate FLa2005 and treated with DNaseI or S1 nuclease.

2012). Here we focus on characterization of the larger dsRNA, representing the genome of PiRV-3.

The PiRV-3 dsRNA was 8112 nt with a GC content of 54.7%. Two long ORFs were identified on one of the strands, designated the plus strand. ORF1 is predicted to initiate at nt 849 following a long 5′-leader sequence that contains 10 additional AUG codons. None of the proteins deduced from this leader sequence is predicted to have a size greater than 41 aa. ORF1 terminates at nt 4739 and is predicted to encode a protein of 1296 aa with a calculated molecular mass of 140.4 kDa. A poly(C) track of variable length was found within ORF1 starting at position 3858. Sequencing of 28 independent clones covering this region revealed that 1, 7, 16, and 4 clones had the poly(C) sequences C_9, C_{10}, C_{11}, and C_{12}, respectively. The predominant form, C_{11}, is predicted to maintain the reading frame of ORF1. With the alternative poly(C) lengths, translation termination would be predicted following addition of 18 aa in the case of C_{10} or 3 aa in the case of C_9 or C_{12}. Use of high-fidelity Taq polymerase for these experiments and direct PCR product sequencing results revealing sequence heterogeneity (not shown) also supported the finding of variability of the internal poly(C) tract. Inspection of the sequence of *Phlebiopsis gigantea* virus 2 (PgV2) (Kozlakidis et al., 2009), to which PiRV-3 has considerable similarity (see below), also revealed a poly(C) tract in a similar location, but no variability in length was reported. ORF2, if counted from the first in-frame AUG codon (5012–8047), does not overlap with ORF1 and is predicted to encode a protein of 1011 aa with calculated molecular mass 113.2 kDa. When in-frame codons before the start codon are counted, ORF2 (4709–8047) has a 31 bp overlap with ORF1, and a candidate slippery sequence for −1 translational frameshift, G TTA AAC, was found immediately before the stop codon of ORF1 (see more analysis below), suggesting these two ORFs may be expressed as a fused protein. No other apparent ORF greater than 351 bp was found in either strand.

ORF1 encodes a protein of unknown function

BLAST searches of the GenBank non-redundant database revealed that ORFs 1 and 2 were both predicted to encode proteins with counterparts in an undescribed group of RNA viruses from fungi. The deduced translation product of ORF1 was most similar to five deduced protein sequences from recently characterized viruses: the ORF1-encoded proteins of PgV2 (Kozlakidis et al., 2009), the *Fusarium virguliforme* RNA viruses 1 and 2 (FvRV1 and 2; Acc.# JN671444 and JN671443), *Fusarium graminearum* virus 3 (FgV3) (Yu et al., 2009), and Grapevine associated totivirus-2 (GaTV-2) (Al Rwahnih et al., 2011) (Table 1). GaTV-2 was found in high-throughput sequencing of grapevine-associated dsRNAs and likely originated from a grapevine-associated fungus (Al Rwahnih et al., 2011). No other sequence had an E-value less than 1. When the ORF1-encoded sequences of the latter five viruses were used as queries, BLASTP searches returned similar results (Table 1; other data not shown). No significant matches were found for these proteins in Pfam (Bateman et al., 2004), PROSITE (Hulo et al., 2006), or Conserved Domain (Marchler-Bauer and Bryant, 2004) protein databases, indicating these six proteins form a novel protein family that has no detectable similarity to other known proteins. A neighbor-joining tree based on alignment of the ORF1-deduced protein sequences showed that FgV3 and GaTV-2 formed a strongly supported clade, as did FvRV1 and FvRV2. These two clades were further grouped together with 99% bootstrap support (Fig. 2A).

In the notes to the GenBank submission of FvRV1 and 2, ORF1 is speculated to be a structural/gag protein; however, virus particle purifications have not yet confirmed that prediction (Dr. Les Domier, personal communication). Our efforts thus far using various protocols to purify virus particles of PiRV-3 have yielded no fractions containing virus particles (data not shown). Similarly, Kozlakidis et al. (2009) also failed to purify virus particles from PgV2-infected tissue.

Kozlakidis et al. (2009) noted that ORF1 of *P. gigantea* virus 1 (PgV1) and PgV2 were rich in alanine and proline, and Spear et al. (2010) observed that PgV1, PgV2, FgV3, and two viruses from plant-feeding insects, *Spissistilus festinus* virus 1 (SpFV1) and *Circulifer tenellus* virus 1 (CiTV1), contained high levels of proline and alanine, and to a lesser extent, serine in their ORF1 region. These observations were based on the absolute number and frequency of these amino acids in those proteins. As amino acids do not have equal distribution in proteomes, we calculated amino acid frequencies in members of this new protein family and compared them with the average amino acids composition in 105 protein families in eight species, including archaea, eukaryote and eubacteria (Brooks et al., 2002) (Table S1). Amino acid compositions in viral proteins are highly correlated to that in their hosts (Bogatyreva et al., 2006). Based on this comparison, proline, serine, tryptophan, histidine and alanine are over-represented, and lysine, isoleucine, glutamic acid and methionine are under-represented in this protein family (amino acids listed in the order of over- or under-representation, respectively).

Spear et al. (Spear et al., 2010) observed that the ORF1-encoded proline-rich proteins in FgV3 and PgV1 contain the PSAP sequences and the one in PgV2 contains PPDY sequences. These sequences matched the late domain motifs PS/TAP and PPXY, respectively. Three late domain motifs have been characterized, YXXL, PPXY, and PS/TAP, and they have been shown to mediate protein–protein interactions and play a critical role in the pinching off of virus particles from plasma membrane for a wide range of viruses (Chen and Lamb, 2008; Freed, 2002; Wirblich et al., 2006). However, these short motifs have a high possibility of random occurrence. Examination of alignments among ORF1-encoded proteins of PiRV-3, PgV2, FvRV1, FvRV2, FgV3 and GaTV-2 showed that neither PSAP in FgV3 nor PPDY in PgV2 were in conserved regions (not shown). Instead, we found three conserved regions (I–III) around the middle section of these proteins. Region I contains two potential YXXL motifs and region III contains a PXAP motif, similar to the PS/TAP late-domain motif. Previous study showed potential relaxations of the PS/TAP motif

Table 1
Sequence similarities among ORF1-encoded proteins of PiRV-3, PgV2, FvRV1, FvRV2, FgV3 and GaTV-2.

	PiRV-3	PgV2	FvRV1	FvRV2	FgV3
PgV2	332/24/40/2e-14[a]				
FvRV1	450/24/40/5e-10	578/23/39/7e-20			
FvRV2	367/23/39/6e-07	529/22/39/4e-18	1283/48/63/0.0		
FgV3	464/21/38/4e-06	1061/24/39/8e-38	1213/28/45/2e-97	974/29/47/4e-97	
GaTV-2	562/22/39/7e-05	427/27/41/4e-18	1057/31/47/6e-106	795/31/50/7e-94	1314/44/61/0.0

[a] Viral sequences from the row were used as queries in BLASTP searches of GenBank non-redundant protein database. Values are number of query amino acids aligned/identity in percentile/similarity in percentile/E value.

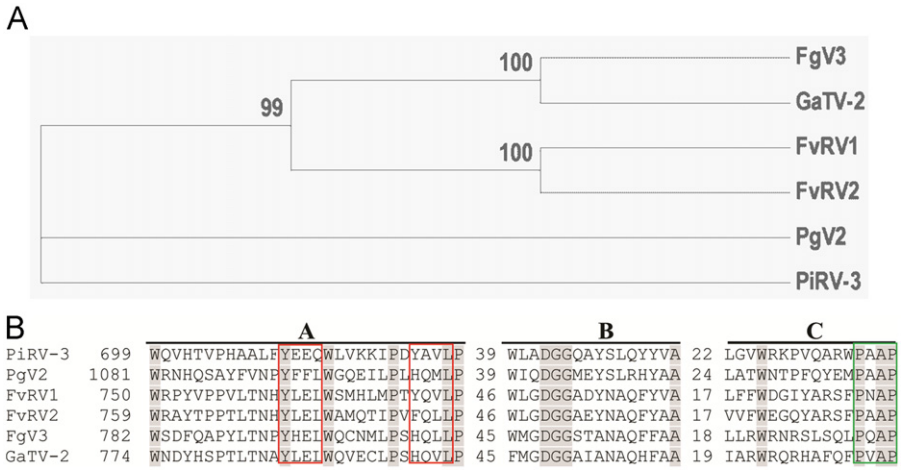

Fig. 2. A. Neighbor-joining tree based on alignments of ORF1-deduced protein sequences of PiRV-3, PgV2, FvRV1, FvRV2, FgV3 and GaTV-2. Numbers above the branches are bootstrap support values in percentile in 1000 replicates. Branches with less than 50% bootstrap support were collapsed. **B**. Presentation of partial alignments showing three conserved motifs (I–III) around the middle section of ORF1-encoded proteins. Numbers represent the number of amino acids before and between the regions. Identical residues are shadowed in gray. Sequences with strong similarity to late domain motifs YXXL and PS/TAP (Freed, 2002; Wirblich et al., 2006) are boxed in red and green, respectively.

(Wirblich et al., 2006). Between regions I and III, region II has a DGG motif that is absolutely conserved in all viruses (Fig. 2B). None of these viruses has been shown to contain virus particles, and it is not been shown that fungal or *Phytophthora* viruses have an extracellular component to their infection cycles. The function(s) of these proteins remain to be investigated.

ORF2 encodes an RNA-dependent RNA polymerase

An RdRp domain with greatest similarity to RdRps of dsRNA viruses (RdRP_4, Cdd:pfam02123) was predicted in the middle section of ORF2. In BLASTP searches, the proteins in GenBank with most significant similarities to PiRV-3 RdRp were RdRps of PgV2 (2e-70), FgV3 (5e-65), FvRV-2 (4E-59), *Diplodia scrobiculata* RNA virus 1 (DsRV1, 1e-57), (De Wet et al., 2011) and FvRV1 (3e-54). Only a partial sequence of the GaTV-2 RdRp region was available and it had E-value 8e-26. Many other members of the dsRNA virus families *Chrysoviridae* and *Totiviridae* showed similarity to PiRV-3 RdRp but with much less significant E-values (at most 5e-16). Representative sequences from BLASTP results were retrieved and aligned. Similarities among RdRps in PiRV-3, PgV2, FgV3, FvRV-2, DsRV1 and FvRV1 extended beyond the eight conserved motifs in the RdRps of dsRNA viruses in lower eukaryotes (Bruenn, 1993), with fourteen conserved regions identified (Fig. 3). A neighbor-joining tree based on RdRp alignments showed that PiRV-3, PgV2, FvRV1, FvRV2, FgV3, GaTV-2 and DsRV1 formed a strongly supported clade outside of the *Totiviridae* and *Chrysoviridae* (Fig. 4). Within this clade, the tree topology is in agreement with that based on ORF1—FvRV1 was grouped with FvRV2, as did FgV3 and GaTV-2, and these four viruses were further grouped together. This analysis also grouped CiTV1, Cucurbit yellows-associated virus (CuYaV), and SpFV1 into a strongly supported clade, as previously reported (Spear et al., 2010).

Interestingly, a region close to the N-terminus of ORF2 (aa 97–142 if counting from the first in-frame AUG codon) showed weak similarity to Cdd:pfam12818 (5.05e-03), which is a hidden Markov model representing a family of tegument proteins from the large, enveloped, vertebrate-infecting double-stranded DNA herpesviruses. However, no actual sequences in the GenBank non-redundant database have detectable similarity to this region, and this domain was not detected in other viruses in the same clade as PiRV-3. The function of this region remains to be investigated.

ORF1 and ORF2 in PiRV-3, PgV2, FvRV1, FvRV-2, FgV3 and GaTV-2 are likely expressed as a fusion protein

In addition to similarity in coding regions, PiRV-3, PgV2, FvRV1, FvRV-2, FgV3 and GaTV-2 likely share similar expression strategies. In each of these six viruses, ORF2 is in the −1 reading frame relative to ORF1 (Fig. 4 and not shown). Although the first AUG codon in ORF2 is downstream of the stop codon in ORF1, there is a short overlap between ORF1 and ORF2 if in-frame codons before the first AUG codon in ORF2 are considered. Inspection of the overlapping region identified a potential slippery sequence for −1 translational frameshift immediately before the ORF1 stop codon in each virus (Fig. 5A). In PiRV-3, PgV2, FgV3 and GaTV-2, these slippery sequences differ from the canonical X-XXY-YYZ motif (Jacks et al., 1988). Instead, they have the consensus of G-DWA-AAC (D: A, G or T; W: A or T). The nucleotide immediately before the predicted slippery sequences, cytosine,

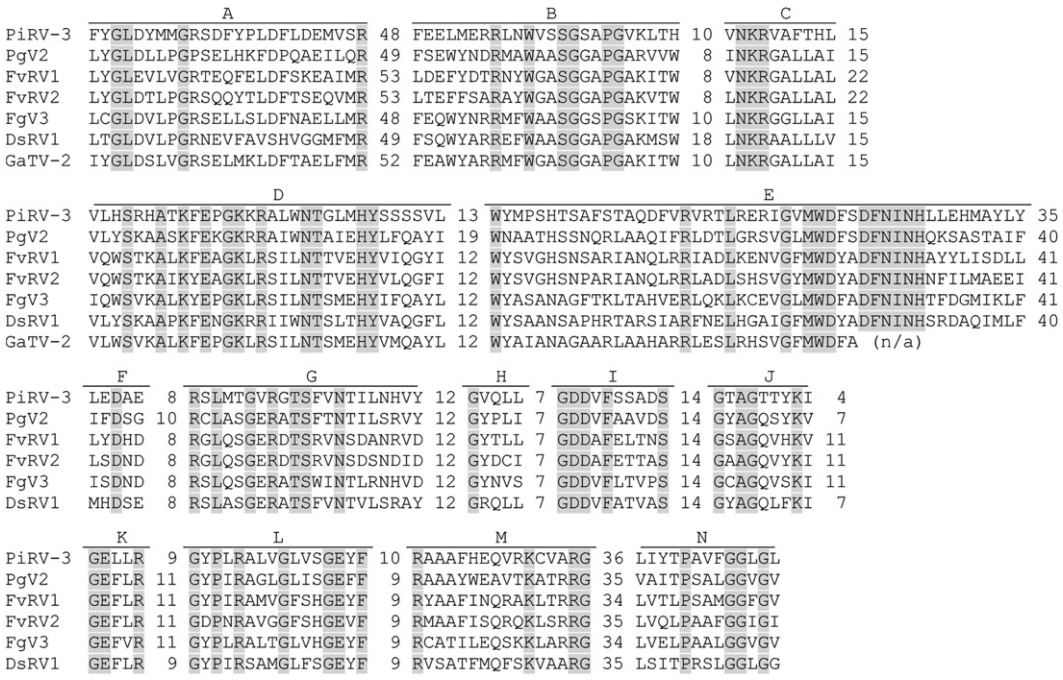

Fig. 3. Alignment of the RdRps in PiRV-3, PgV2, FvRV1, FvRV2, FgV3, GaTV-2 and DsRV1 revealed 14 conserved regions (A–N). Sequences were aligned with MUSCLE program (Edgar, 2004) implemented in MEGA5 (Tamura et al., 2011). Only partial sequence for GaTV-2 is known. Regions A, B, D, E, G, I, K and L correspond to the eight conserved motifs in the RdRps of dsRNA viruses in lower eukaryotes (Bruenn, 1993).

was also strictly conserved. Previous studies have shown that the XXX part of the canonical motif is not strictly required. For example, *Equine arteritis virus* uses GUUAAAC for frameshifting to produce ORF1a/ORF1b fusion protein (den Boon et al., 1991); *Mouse mammary tumor virus* uses GGAUUUA at the *pro*/*pol* junction (Jacks et al., 1987; Moore et al., 1987); and T7 phage utilizes GGUAAAC to produce two products from gene 10 (Condron et al.,1991).

In the majority of −1 translational frameshift cases, an H-type pseudoknot structure was predicted or confirmed shortly downstream of the slippery sequence and this pseudoknot structure is believed to assist in pausing translating ribosomes and increase the frequency of frameshifting (see Brierley et al. (2007), Dam et al. (1990), Giedroc et al. (2000) for reviews). In PiRV-3, PgV2, FgV3 and GaTV-2, a strong H-type pseudoknot structure was predicted 2–12 nt downstream of the predicted slippery sequence. In FvRV1 and FvRV2, which have the "perfect" slippery sequence, a weaker H-type pseudoknot was found further downstream (Fig. 5B). The presence of a good slippery sequence can stimulate low level translational frameshift (approximately 1%) (Giedroc et al., 2000). In *Human immunodeficiency virus 1*, a stem loop, rather than a pseudoknot, shortly downstream of the slippery sequence was found to stimulate −1 frameshifting (Parkin et al., 1992). In FvRV1 and FvRV2, a strong stem loop was predicted after the slippery sequences (Fig. 5C).

The case for a new virus genus

Based on genome organization, similarity in coding sequences and lack of a conventional virion, Spear et al. (2010) proposed that SpFV1 and CiTV1 formed a new virus genus not affiliated with any known virus family (new genus 1), with CuYaV, for which only partial sequence is known, as a tentative member. They also proposed that PgV2, FgV3 and DsRV1 form a second new genus (new genus 2) and that these two new genera constitute a new family of dsRNA viruses based on RdRp phylogeny. Our analyses support the establishments of new genera 1 and 2 and further demonstrate that PiRV-3, FvRV1 and FvRV2 also belong to new genus 2. GaTV-2, whose sequences were obtained from high throughput sequencing of grapevine-associated dsRNAs with substantial portion of the 3′ end unknown and whose host is uncertain, is considered as a tentative member of new genus 2 and not further discussed here. On the other hand, the different genome organization of DsRV1 argues against its inclusion within the new genus 2, and the contention that new genera 1 and 2 form a new virus family requires further examination.

In addition to similarities in coding regions and the predicted expression strategy of ORF2, members of new genus 2 share other similarities, such as genome size, host, and the potential expression strategy of ORF1. PiRV-3, PgV2, FvRV1, FvRV2 and FgV3 have non-segmented dsRNA genomes of 8–10 kbp (Fig. 4 and not shown). The 5′end of PgV2 was not determined but based on estimation from gel electrophoresis, PgV2 dsRNA is approximately 10 kbp. The hosts of PgV2, FvRV1 and FvRV2, and FgV3, *P. gigantea*, *F. virguliforme* and *F. graminearum*, respectively, are filamentous fungi. *P. infestans*, host of PiRV-3, is an oomycete, and thus very distantly related to fungi (Adl et al., 2005). However, fungi and oomycetes share similar morphologies including filamentous mycelia and spores, and occupy similar habitats. Properties of viruses that infect fungal and oomycete hosts are often similar, including that both often have dsRNA genomes and often lack traditional virus particles.

Members of new genus 2 have long 5′ UTRs, with 7–15 AUG codons upstream from the first AUG codon in ORF1. In each instance, some of the AUG codons are in contexts predicted to be favorable for translation initiation (data not shown) (Kozak, 1986; Lütcke et al., 1987). As such, it is unlikely that they use a cap-dependent scanning mechanism for translation initiation of ORF1. A potential mechanism for ORF1 expression is through an internal ribosomal entry site (IRES), which allows for cap-independent internal initiation of translation (see Doudna and Sarnow (2007), Filbin and Kieft (2009) for reviews). IRES elements have considerable diversity in primary sequences and secondary structures, and have proven to be difficult for computational prediction using

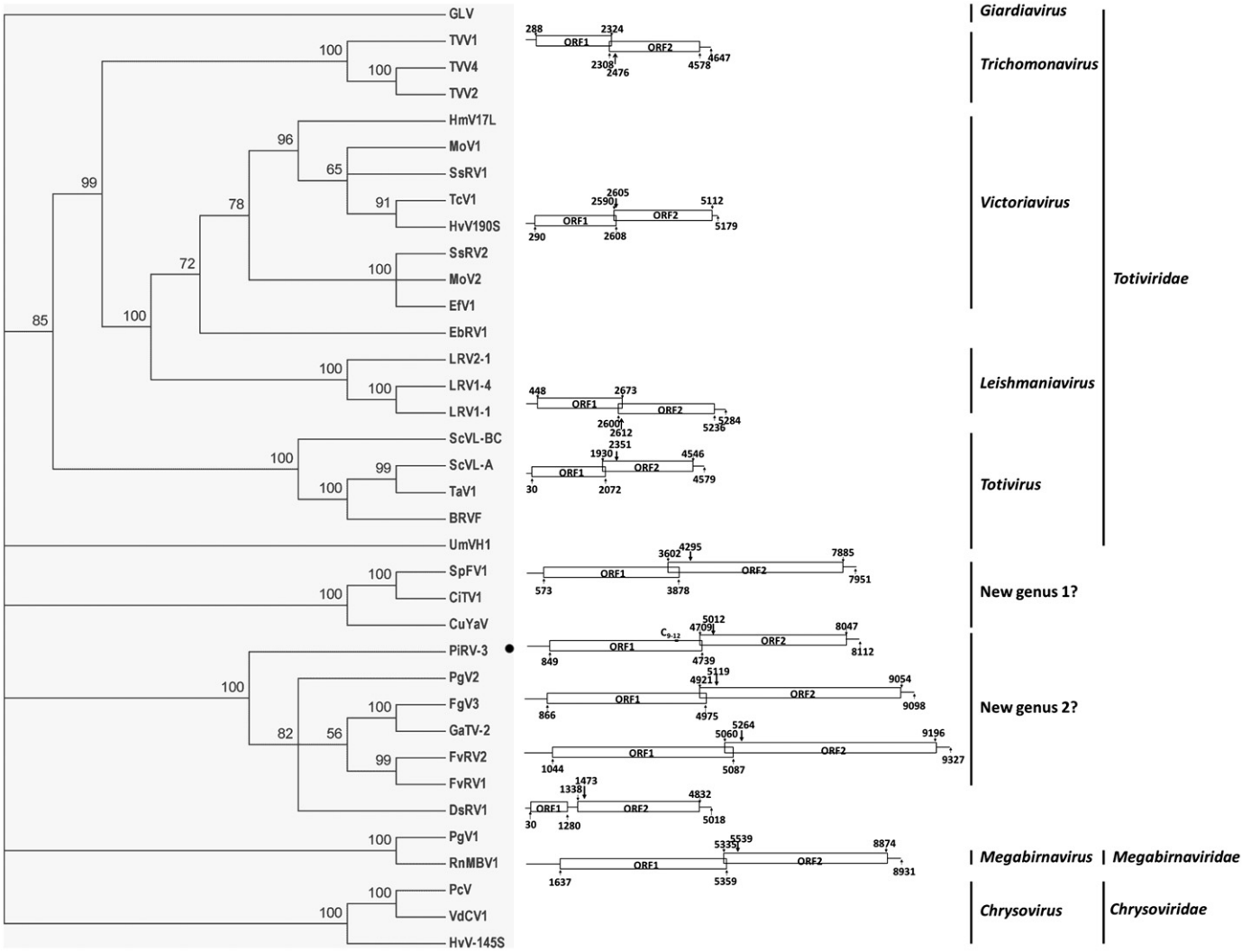

Fig. 4. Neighbor-joining tree based on alignments of RdRp sequences of PiRV-3 and related viruses. Numbers above the branches indicate bootstrap support in percentile in 1000 replicates. Branches with less than 50% support were condensed. Genome organizations of representative viruses are shown. Horizontal bars represent ORFs, with pointed ends indicating that terminal sequences are unknown; when two bars in the same virus are not at the same height, the ORF represented by the higher bar is in −1 reading frame relative to the lower bar; lines represent UTRs; numbers indicate nucleotide positions on the genomes; and bold arrows point to the positions of the first in-frame AUG codon in ORF2 of each virus. Virus names (acronym, GenBank accession) included in analysis were: Black raspberry virus F (BRVF, NC_009890), *Circulifer tenellus* virus 1 (CiTV1, NC_014360), Cucurbit yellows-associated virus (CuYaV, X92203), *Diplodia scrobiculata* RNA virus 1 (DsRV1, NC_013699), *Eimeria brunetti* RNA virus 1 (EbRV1, NC_002701), *Epichloe festucae* virus 1 (EfV1, AM261427), *Fusarium graminearum* virus 3 (FgV3, NC_013469), *Fusarium virguliforme* RNA virus 1 (FvRV1, JN671444), *Fusarium virguliforme* RNA virus 2 (FvRV2, JN671443), *Giardia lamblia* virus (GLV, NC_003555), Grapevine associated totivirus-2 (GaTV-2, GU108594), *Helicobasidium mompa* virus 17 L (HmV17L, NC_005074), *Helminthosporium victoriae* 145S virus (HvV-145S, NC_005978), *Helminthosporium victoriae* virus 190S (HvV190S, NC_003607), Leishmania RNA virus 1–1 (LRV1–1, M92355), Leishmania RNA virus 1–4 (LRV1–4, NC_003601), Leishmania RNA virus 2–1 (LRV2–1, NC_002064), *Magnaporthe oryzae* virus 1 (MoV1, NC_006367), *M. oryzae* virus 2 (MoV2, NC_010246), *Penicillium chrysogenum* virus (PcV, NC_007539), *Phlebiopsis gigantea* mycovirus dsRNA 1 (PgV1, NC_013999), *P. gigantea* mycovirus dsRNA2 (PgV2, AM111097), *Rosellinia necatrix* megabirnavirus 1 (RnMBV1, NC_013462), *Saccharomyces cerevisiae* virus L-A (ScVL-A, NC_003745), *S. cerevisiae* virus L-BC (ScVL-BC, NC_001641), *Sphaeropsis sapinea* RNA virus 1 (SsRV1, NC_001963), *S. sapinea* RNA virus 2 (SsRV2, NC_001964), *Spissistilus festinus* virus 1 (SpFV1, NC_014359), *Tolypocladium cylindrosporum* virus 1 (TcV1, NC_014823), *Trichomonas vaginalis* virus 1 (TvV1, NC_003824), *T. vaginalis* virus 2 (TvV2, AF127178), *T. vaginalis* virus 4 (TvV4, HQ607522), Tuber aestivum virus 1 (TaV1, HQ158596), *Ustilago maydis* virus H1 (UmVH1, NC_003823), and *Verticillium dahlia* chrysovirus 1 (VdCV1, HM004067).

primary sequences (Baird et al., 2006; Filbin and Kieft, 2009). Wu et al. (Wu et al., 2009) designed the program IRSS, a computational tool that combines RNA secondary structure prediction and alignment, to predict IRES elements. They claimed 72.3% sensitivity, but only if the predicted IRES belongs to one of four classes of known elements. Using this tool, an IRES was predicted in the 5′ UTR in PiRV-3, FvRV1 and FvRV2, but not in FgV3. The expression mechanism of ORF1 requires further investigation.

DsRV1 is likely a recombinant genome

DsRV1 is similar to PiRV-3 and other members of new genus 2 in the RdRp region, but it differs from those viruses in several ways: (1) its genome is only 5018 bp, much smaller than members in this proposed genus, mainly because of a shorter ORF1; (2) ORF1 and ORF2 are in the same reading frame (Fig. 4); (3) it has a short 5′ UTR and the start codon in ORF1 is the first AUG codon on the positive strand, suggesting that ORF1 is expressed from ribosomal scanning; and (4) the predicted ORF1-encoded protein has no detectable similarity to that of other viruses in the genus; instead, it is closely related to class E vacuolar protein sorting machinery protein HSE1 from ascomycete fungi (BLASTP identity up to 71%, E value 76–116). The host of DsRV1, *D. scrobiculata*, is also an ascomycete, and it therefore is likely that the DsRV1 genome resulted from recombination between a member of proposed new genus 2 and host sequence.

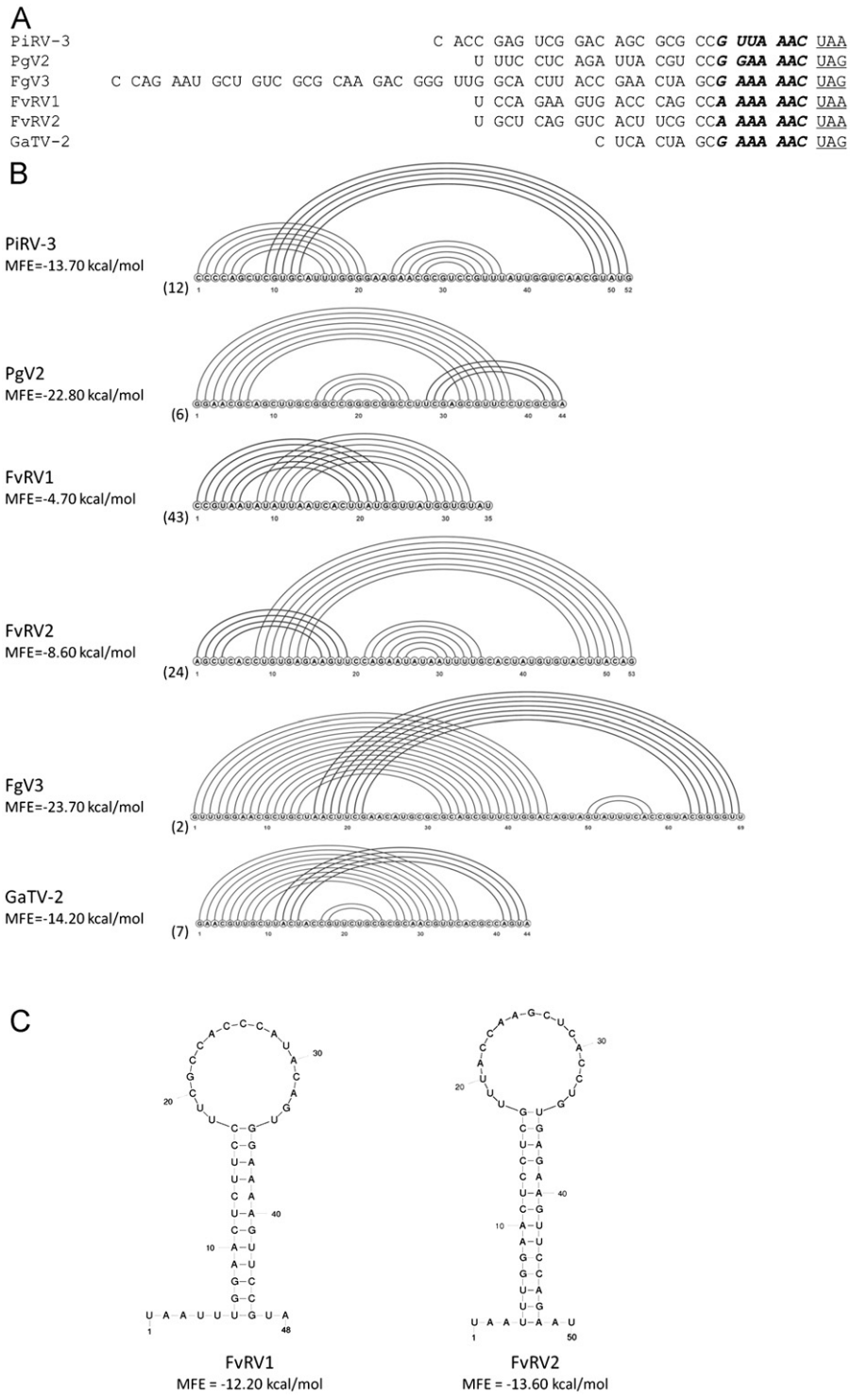

Fig. 5. ORF1 and ORF2 in PiRV3, PgV2, FvRV1, FvRV2, FgV3 and GaTV-2 are likely expressed as fusion protein. **A**. Overlapping regions between ORF1 and ORF2. Sequences are presented in the reading frame of ORF1, with stop codons underlined. Potential slippery sequences for −1 translational frameshift are bold and italic. Conserved nucleotides are labeled with the symbol "*" at the bottom. **B**. Pseudoknot structures downstream of the identified slippery sequences using HPknotter (Huang et al., 2005). Numbers in parenthesis indicate the number of bases between the end of slippery sequences and the beginning of the pseudoknot. **C**. Stem-loop structure predicted by Mfold (Zuker, 2003) in the sequences immediately downstream of the slippery sequences in FvRV1 and FvRV2. MFE, minimal free energy.

Interestingly, these fungal proteins are predicted to be associated with membranes and protein sorting; thus a role in membrane-associated viral replication of DsRV-1 ORF1-deduced protein is easy to envision. Important to our study and the elucidation of this proposed new genus 2, the observation that the DsRV-1 ORF1 product likely does not encode a capsid protein bolsters the case for a virus group devoid of a capsid protein and traditional virus particles.

Do new genera 1 and 2 belong to the same family?

Based on RdRp phylogeny, Spear et al. (Spear et al., 2010) proposed a new virus family consisting of the new genera 1 and 2. The bootstrap support for grouping these two genera together was less than 60%, however, and our analysis did not support the grouping of these two genera based on RdRp phylogeny. These

two genera have similar genome sizes and organizations (Fig. 4 and not shown) and both appear to lack a conventional capsid. The ORF1 sequences of viruses in these two proposed genera are predicted to encode proteins rich in proline and alanine; however, proline was concentrated at the N-terminus in members of proposed new genus 1 but not in members of new genus 2 (Spear et al., 2010). In BLASTP searches, there was no detectable similarity between the ORF1-encoded proteins in these two proposed genera. We suggest that a study group of the International Committee for the Taxonomy of Viruses be assembled to proceed with a taxonomic proposal for these viruses.

PiRV-3 curing and phenotype

Using a combination of Ribavirin treatment and hyphal-tipping, PiRV-3 but not PiRV-4 was cured from isolate FLa2005 and the cured isolate was named FLa2005/PiRV-3C. The curing was confirmed by dsRNA extraction and RT-PCR (Fig. 6A–C). Using quantitative RT-PCR, we did not detect statistically significant differences in PiRV-4 levels between FLa2005 and FLa2005/PiRV-3C (not shown). These two isogenic isolates grew at similar speed on rye agar, and neither produced many sporangia (not shown). The only phenotypic difference that was consistently observed was colony morphology. The doubly infected isolate FLa2005 produced aerial mycelium that was evenly distributed across the plates, while FLa2005/PiRV-3C produced a zone, usually circular, with dense aerial mycelium (Fig. 6D). FLa2005 was the only isolate in our collection bearing PiRV-3 and we have been unable to rid it of PiRV-4. Furthermore, we have been unable to move PiRV-3 to an uninfected isolate of *P. infestans* by hyphal fusion. Therefore we cannot compare the phenotype of an isolate infected with only PiRV-3 directly to its isogenic, uninfected counterpart.

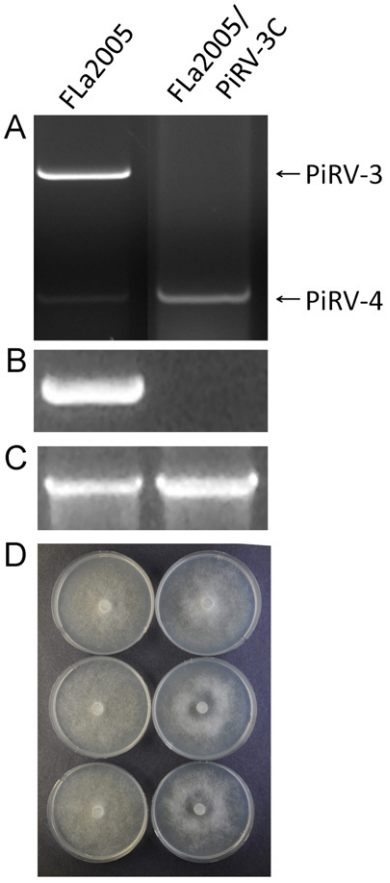

Fig. 6. Curing and phenotype of PiRV-3. **A**. Agarose electrophoresis of dsRNA extractions from FLa2005 and FLa2005/PiRV-3C. **B**. RT-PCR detection of PiRV-3. **C**. RT-PCR detection of PiRV-4. **D**. Colony morphology of these two isogenic isolates on rye agar.

Materials and methods

Isolate culture and dsRNA extraction

Isolate FLa2005 was maintained on rye agar and grown in pea broth at 18 °C in dark for 2–3 weeks for mycelium production. Approximately 4 g (fresh weight) mycelium was used to extract dsRNA, which was purified with CF-11 cellulose as previously described (Cai et al., 2009). DNaseI and S1 nuclease were used to remove trace DNA and single strand RNA, respectively (Cai et al., 2009).

dsRNA sequencing

DsRNAs from isolate FLa2005 were separated on agarose gels, individual segments were extracted from the gel and the 8.3 kb segment was used as template for cDNA library generation as previously described (Bohlander et al., 1992; Cai et al., 2009). The cDNAs were cloned into PCR4-TOPO vector (Invitrogen) following the manufacturer's instructions. Individual clones were extracted and sequenced using the Big Dye Terminator chemistry (Applied Biosystems). Sequences were assembled into contigs using Sequencher version 5.0 (Gene Codes Corp.). Sequence-specific primers were then designed and used in RT-PCR reactions to bridge the contigs. RLM-RACE was used to obtain terminal sequences as follows: a 5′ phosphorylated and 3′ blocked adapter (5′ PO4-AGGTCTCGTAGACCGTGCACC-NH2–3') was ligated to the 3′ end of each strand with T4 RNA ligase (Promega) (Attoui et al., 2000; Lambden et al., 1992), and a primer with reverse-complementarity to the adapter (GGTGCACGGTCTACGAGACCT) was used in combination with sequence-specific primers to amplify terminal sequences. The amplification products were cloned and sequenced as described above, with each position covered by three or more independent clones. The complete sequence of PiRV-3 were submitted to GenBank and given accession number JN603241.

Bioinformatic analysis

ORF1- and ORF2- encoded protein sequences were used in BLAST searches of various databases noted in "results" section using default settings. Alignments of related sequences were conducted using the MUSCLE program (Edgar, 2004) implemented in MEGA5 (Tamura et al., 2011), using "Neighbor-Joining" for the first two clustering iterations. For ORF2, the N- and C-termini diverse regions were removed after initial alignment and the remaining sequences were re-aligned as described. After alignment, Neighbor-joining trees were constructed using MEGA5 with 1000 bootstrap steps.

To predict pseudoknot structures, 200 nt sequences immediately downstream of the predicted slippery sequences were used as input in HPknotter (Huang et al., 2005) using the pknotsRG kernel (Reeder et al., 2007). The JAVA applet VARNA (Darty et al., 2009) was used for better visualization of the results. Stem loop structure was predicted using Mfold (Zuker, 2003) and confirmed with RNAfold (Bompfünewerer et al., 2008).

Virus curing and phenotype

FLa2005 was grown on a 9-cm diameter rye agar plate containing 10 μg/ml Ribavirin (SigmaAldrich) at 18 °C in the dark (Herrero and Zabalgogeazcoa, 2011). When the colony covered approximately two-thirds of the plate, hyphal tips were taken from the edge with the aid of a dissecting microscope and transferred to a new plate. The process was repeated. Periodically total RNA was extracted from mycelium using RNeasy plant mini kit (Qiagen) according to manufacturer's directions. The RNA was treated with DNaseI (Invitrogen), denatured in 15% DMSO at 99 °C for 1 min, chilled on ice for 2 min, and then used as template in RT-PCR reactions to detect PiRV-3 and PiRV-4 using the superscript III one-step RT-PCR system (Invitrogen). For PiRV-3, primer pair CGGTGAGATATTGGTAGAACGGAA (nt 6511–6534) and GGAATCGTAGACGTGCGAGGTA (nt 7615–7594) was used; for PiRV-4, primer pair GTCGTTCCACACTTCCACAATCGTG (nt 185–209) and GGTTGACGCAGGAAATCTTCTG (nt 2475–2454) was used. After five transfers, PiRV-3 was not detected by RT-PCR. DsRNA was then extracted as described above to confirm virus curing. The cured isolate was named FLa2005/PiRV-3C.

To examine phenotype changes, agar plugs were taken from the colony edge of actively growing FLa2005 and FLa2005/PiRV-3C on rye agar using #5 cork borer, and transferred to the center of new 9-cm diameter rye agar plates. The plates were incubated at 18 °C in the dark. Colony diameter was measured at 5, 7 and 9 (10 in the third experiment) days after inoculation by averaging two measurements at 90 degree difference. The Student t-test was used to detect differences in growth rate. Two weeks after inoculation, the plates were photographed. Sporangia were harvested by flooding the colonies with 10 ml distilled water and rubbing with a glass rod, and sporangia concentration was determined using a hemocytometer. The experiments were conducted three times with three plates per isolate in the first two experiments and five plates per isolate in the third experiment.

Acknowledgments

The authors thank the New Jersey Agricultural Experiment Station and College of Agriculture and Life Sciences at Cornell University for partial funding for this project. The authors thank Les Domier for information on the *Fusarium virguliforme* virus prior to publication.

Appendix A. Supplementary Information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.virol.2012.10.012.

References

Adl, S.M., Simpson, A.G.B., Farmer, M.A., Andersen, R.A., Anderson, O.R., Barta, J.R., Bowser, S.S., Brugerolle, G., Fensome, R.A., Fredericq, S., James, T.Y., Karpov, S., Kugrens, P., Krug, J., Lane, C.E., Lewis, L.A., Lodge, J., Lynn, D.H., Mann, D.G., McCourt, R.M., Mendoza, L., Moestrup, O., Mozley-Standridge, S.E., Nerad, T.A., Shearer, C.A., Smirnov, A.V., Spiegel, F.W., Taylor, M., 2005. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. J. Eukaryot. Microbiol. 52, 399–451.

Al Rwahnih, M., Daubert, S., Úrbez-Torres, J., Cordero, F., Rowhani, A., 2011. Deep sequencing evidence from single grapevine plants reveals a virome dominated by mycoviruses. Arch. Virol. 156, 397–403.

Attoui, H., Billoir, F., Cantaloube, J.F., Biagini, P., de Micco, P., de Lamballerie, X., 2000. Strategies for the sequence determination of viral dsRNA genomes. J.Virol. Methods 89, 147–158.

Baird, S.D., Turcotte, M., Korneluk, R.G., Holcik, M., 2006. Searching for IRES. RNA 12, 1755–1785.

Bateman, A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E.L.L., Studholme, D.J., Yeats, C., Eddy, S.R., 2004. The Pfam protein families database. Nucl. Acids Res. 32 (suppl_1), D138–141.

Bogatyreva, N.S., Finkelstein, A.V., Galzitskaya, O.V., 2006. Trend of amino acid composition of proteins of different taxa. Journal of Bioinformatics & Computational Biology 4, 597–608.

Bohlander, S.K., Espinosa, R., Lebeau, M.M., Rowley, J.D., Diaz, M.O., 1992. A method for the rapid sequence-independent amplification of microdissected chromosomal material. Genomics 13, 1322–1324.

Bompfünewerer, A., Backofen, R., Bernhart, S., Hertel, J., Hofacker, I., Stadler, P., Will, S., 2008. Variations on RNA folding and alignment: lessons from Benasque. J. Math. Biol. 56, 129–144.

Brierley, I., Pennell, S., Gilbert, R.J.C., 2007. Viral RNA pseudoknots: versatile motifs in gene expression and replication. Nat. Rev. Microbiol. 5, 598–610.

Brooks, D.J., Fresco, J.R., Lesk, A.M., Singh, M., 2002. Evolution of amino acid frequencies in proteins over deep time: inferred order of introduction of amino acids into the genetic code. Mol. Biol. Evol. 19, 1645–1655.

Bruenn, J.A., 1993. A closely related group of RNA-dependent RNA polymerases from double-stranded RNA viruses. Nucl. Acids Res. 21, 5667–5669.

Cai, G., Myers, K., Fry, W. E., Hillman, B.I. unpublished results.

Cai, G., Myers, K., Fry, W., Hillman, B., 2012. A member of the virus family Narnaviridae, from the plant pathogenic oomycete *Phytophthora infestans*. Arch. Virol. 157, 165–169.

Cai, G., Myers, K., Hillman, B.I., Fry, W.E., 2009. A novel virus of the late blight pathogen, *Phytophthora infestans*, with two RNA segments and a supergroup 1 RNA-dependent RNA polymerase. Virology 392, 52–61.

Chen, B.J., Lamb, R.A., 2008. Mechanisms for enveloped virus budding: can some viruses do without an ESCRT? Virology 372, 221–232.

Condron, B.G., Atkins, J.F., Gesteland, R.F., 1991. Frameshifting in gene 10 of bacteriophage T7. J. Bacteriol. 173, 6998–7003.

Dam, E.B., Pleij, C.W.A., Bosch, L., 1990. RNA pseudoknots: translational frameshifting and readthrough on viral RNAs. Virus Genes 4, 121–136.

Darty, K., Denise, A., Ponty, Y., 2009. VARNA: Interactive drawing and editing of the RNA secondary structure. Bioinformatics 25, 1974–1975.

De Wet, J., Bihon, W., Preisig, O., Wingfield, B.D., Wingfield, M.J., 2011. Characterization of a novel dsRNA element in the pine endophytic fungus *Diplodia scobiculata*. Arch. Virol. 156, 1199–1208.

den Boon, J.A., Snijder, E.J., Chirnside, E.D., de Vries, A.A., Horzinek, M.C., Spaan, W.J., 1991. Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. J. Virol. 65, 2910–2920.

Doudna, J.A., Sarnow, P., 2007. Translation initiation by viral internal ribosome entry sites. In: Mathews, M.B., Sonenberg, N., Hershey, J.W.B. (Eds.), Translational Control in Biology and Medicine. Cold Spring Harbor Laboratory Press, pp. 129–153.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.

Filbin, M.E., Kieft, J.S., 2009. Toward a structural understanding of IRES RNA function. Curr. Opin. Struct. Biol. 19, 267–276.

Freed, E.O., 2002. Viral late domains. J. Virol. 76, 4679–4687.

Fry, W., 2008. Phytophthora infestans: the plant (and R gene) destroyer. Mol. Plant Pathol. 9, 385–402.

Giedroc, D.P., Theimer, C.A., Nixon, P.L., 2000. Structure, stability and function of RNA pseudoknots involved in stimulating ribosomal frameshifting. J. Mol. Biol. 298, 167–185.

Goodwin, S.B., 1997. The population genetics of Phytophthora. Phytopathology 87, 462–473.

Herrero, N., Zabalgogeazcoa, I., 2011. Mycoviruses infecting the endophytic and entomopathogenic fungus *Tolypocladium cylindrosporum*. Virus Res. 160, 409–413.

Huang, C.-H., Lu, C.L., Chiu, H.-T., 2005. A heuristic approach for detecting RNA H-type pseudoknots. Bioinformatics 21, 3501–3508.

Hulo, N., Bairoch, A., Bulliard, V., Cerutti, L., De Castro, E., Langendijk-Genevaux, P.S., Pagni, M., Sigrist, C.J.A., 2006. The PROSITE database. Nucleic Acids Res. 34 (suppl_1), D227–230.

Jacks, T., Madhani, H.D., Masiarz, F.R., Varmus, H.E., 1988. Signals for ribosomal frameshifting in the Rous sarcoma virus gag-pol region. Cell 55, 447–458.

Jacks, T., Townsley, K., Varmus, H.E., Majors, J., 1987. Two efficient ribosomal frameshifting events are required for synthesis of mouse mammary tumor virus gag-related polyproteins. Proc. Natl. Acad. Sci. USA 84, 4298–4302.

Kozak, M., 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44, 283–292.

Kozlakidis, Z., Hacker, C.V., Bradley, D., Jamal, A., Phoon, X., Webber, J., Brasier, C.M., Buck, K.W., Coutts, R.H.A., 2009. Molecular characterisation of two novel double-stranded RNA elements from Phlebiopsis gigantea. Virus Genes 39, 132–136.

Lütcke, H.A., Chow, K.C., Mickel, F.S., Moss, K.A., Kern, H.F., Scheele, G.A., 1987. Selection of AUG initiation codons differs in plants and animals. EMBO J. 6, 43–48.

Lambden, P.R., Cooke, S.J., Caul, E.O., Clarke, I.N., 1992. Cloning of noncultivatable human rotavirus by single primer amplification. J. Virol. 66, 1817–1822.

Marchler-Bauer, A., Bryant, S.H., 2004. CD-Search: protein domain annotations on the fly. Nucleic Acids Res. 32 (suppl_2), W327–331.

Moore, R., Dixon, M., Smith, R., Peters, G., Dickson, C., 1987. Complete nucleotide sequence of a milk-transmitted mouse mammary tumor virus: two frameshift

suppression events are required for translation of gag and pol. J. Virol. 61 (2), 480–490.
Parkin, N.T., Chamorro, M., Varmus, H.E., 1992. Human immunodeficiency virus type 1 gag-pol frameshifting is dependent on downstream mRNA secondary structure: demonstration by expression in vivo. J. Virol. 66, 5147–5151.
Reeder, J., Steffen, P., Giegerich, R., 2007. pknotsRG: RNA pseudoknot folding including near-optimal structures and sliding windows. Nucleic Acids Res. 35 (suppl 2), W320–W324.
Spear, A., Sisterson, M.S., Yokomi, R., Stenger, D.C., 2010. Plant-feeding insects harbor double-stranded RNA viruses encoding a novel proline–alanine rich protein and a polymerase distantly related to that of fungal viruses. Virology 404, 304–311.
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 Method. Mol. Biol. Evol. 28, 2731–2739.
Wirblich, C., Bhattacharya, B., Roy, P., 2006. Nonstructural protein 3 of bluetongue virus assists virus release by recruiting ESCRT-I protein Tsg101. J. Virol. 80, 460–473.
Wu, T.-Y., Hsieh, C.-C., Hong, J.-J., Chen, C.-Y., Tsai, Y.-S., 2009. IRSS: a web-based tool for automatic layout and analysis of IRES secondary structure prediction and searching system in silico. BMC Bioinformatics 10, 160.
Yu, J., Kwon, S.-J., Lee, K.-M., Son, M., Kim, K.-H., 2009. Complete nucleotide sequence of double-stranded RNA viruses from *Fusarium graminearum* strain DK3. Arch. Virol. 154, 1855–1858.
Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 31, 3406–3415.