Cornell University

Analysis of the genomes of four viruses in the late blight pathogen

G. Cai[1], K. Myers[1], B.I. Hillman[2], and W.E. Fry[1]

[1]Cornell University, Ithaca, New York, USA; [2]Rutgers University, New Brunswick, New Jersey, USA.

Abstract: Virus-like double-strand RNAs (dsRNAs) were found in nine out of 22 isolates of *Phytophthora infestans*. Five dsRNAs were completely sequenced. Based on analysis of their sequences, these dsRNAs belonged to four different viruses, tentatively named *Phytophthora infestans* RNA virus 1, 2, 3, and 4 (PiRV1, 2, 3, and 4), respectively. PiRV1 had a bipartite genome, with RNA1 encoding an RNA-dependent RNA polymerase (RdRP) of superfamily 1 (RdRP_1, pfam00680) and RNA2 encoding a putative trypsin-like serine protease. Phylogenetic analysis based on RdRP suggested that PiRV1 represents a new family in the *picorna*-like virus superfamily. An RdRP of superfamily 4 (RdRP_4, pfam 02123) was identified in PiRV2 based on its similarity to the family profile, but it had no significant similarity with any known sequence in public databases. Also identified in PiRV2 was a putative protease motif. An RdRP_4 region was identified in PiRV3 and it was closely related to RdRPs of members of *Totiviridae*. The RdRP region in PiRV4 had only one significant hit in public database, and that was the 20S RNA replicon in *Saccharomyces cerevisiae*.

Objective: The lack of an efficient functional genomics tool is the bottleneck in molecular studies of *P. infestans* and other oomycetes. Such a system is in urgent need to exploit the recently sequenced *P. infestans* genome and accelerate our understanding of this important plant pathogen. We initiated an effort to look for extrachromosomal genetic elements in *P. infestans* that may be engineered into a vector.

Background: There have been only a very few studies on extrachromosomal genetic elements in oomycetes. In *Phytophthora* species, virus-like dsRNAs were reported in some isolates of *P. infestans* but they were not sequenced (5, 6); an autonomous linear, single-stand RNA (ssRNA) was found in a different *P. infestans* isolate and was sequenced (2), but it lacked an obvious open reading frame (ORF); and more recently, an endornavirus (*Phytophthora endornavirus 1*, PEV1) was found in an unnamed *Phytophthora* isolate from Douglas fir (1). PEV1 (13.9kb) contains a large ORF with the potential to encode a polyprotein of 4548 aa with motifs characteristic of virus RdRPs, RNA helicases, and UDP glycosyltransferases.

Results:

1. Virus-like dsRNAs found in *P. infestans*.

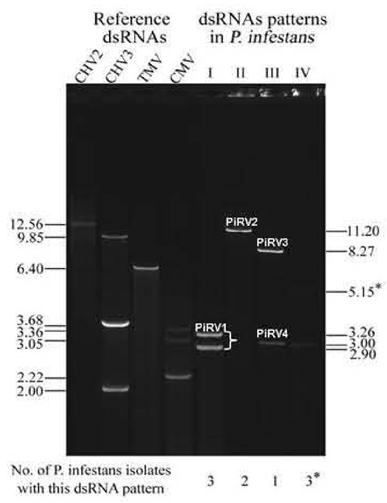

Fig. 1 Virus-like dsRNA patterns in *P. infestans* and reference viruses. CHV2, *Cryphonectria hypovirus 2* from *Cryphonectria parasitica* strain NB58; CHV3, *Cryphonectria hypovirus 3* from *C. parasitica* strain GH2; TMV, *Tobacco mosaic virus*; CMV, *Cucumber mosaic virus*. Sizes of reference dsRNAs (in kb) are labeled on the left. CHV2 and CHV3 dsRNAs have poly(A) tails of various lengths and they were assumed to be 50 bp. Sizes of *P. infestans* dsRNAs (in kb) were estimated using the appropriate reference dsRNAs and they are labeled on the right. The dsRNAs in patterns I-III were sequenced and sequence analysis showed they belong to four different viruses, tentatively named *Phytophthora infestans RNA viruses 1-4* (PiRVs 1–4), respectively. *, All three isolates with pattern IV had the 3.00 kb segment; the weak 5.15 kb segment was visible in some but not all extractions of one isolate and is considered putative.

2. Genome structure of PiRVs1-4

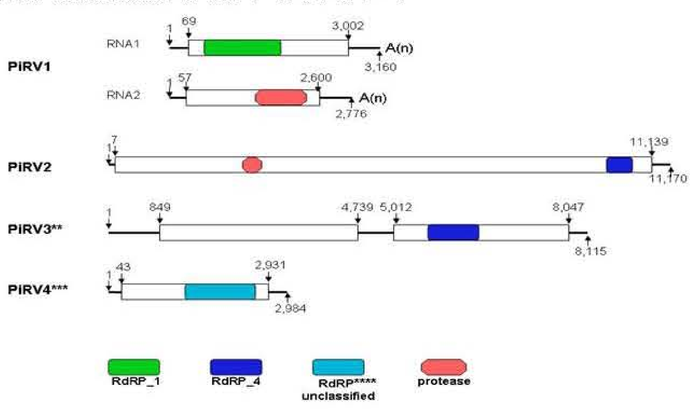

Fig.2 Genome organization of PiRVs 1-4 based on nucleotide sequence analysis. Lines represent un-translated regions (UTRs) and boxes represent ORFs. The numbers indicate nucleotide positions. The first AUG codon in each ORF was assumed to be the translation start site regardless of whether or not it was in a favorable context according to Kozak's rules (3, 4). **, sequence comparison of multiple independent clones showed that there is a string of Cs of variable length (C10, C11 or C12) starting at nt 3858; the first ORF would extend to nt 4739 only in case of C11. ***, the exact 3' terminal sequences of PiRV4 need confirmation from independent clones. ****, the RdRP region in PiRV4 has significant similarity to the RdRP in 20 S *Narnavirus* in yeast (BlastP E value 3e-06) but does not fall into any RdRP family when Blast against Pfam database (http://www.sanger.ac.uk/Software/Pfam/).

3. PiRV1 represents a new family in the *picorna*-like virus superfamily

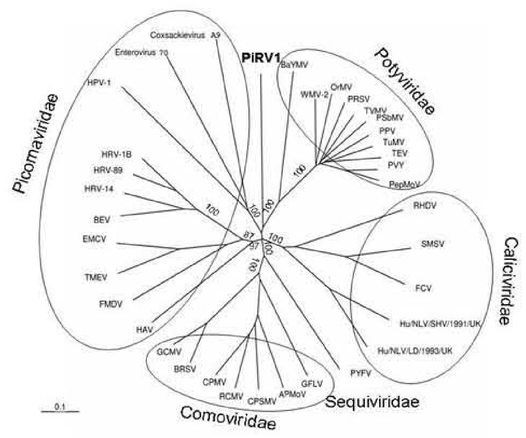

Fig. 3. Neighbor joining tree based on alignment between RdRP region of PiRV1 and that of 35 other viruses with RDRP_1 domain. Bootstrap values are in percentage based on 1000 repeats.

4. PiRV1 encoded secreted and transmembrane proteins.

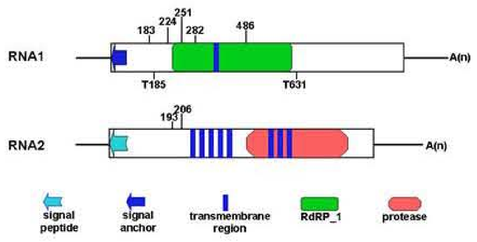

Fig. 4 Detailed genome structure of PiRV1 based on sequence analysis. The numbers above each ORF are asparagine sites that are likely subjected to N-glycosylation and those below each ORF are threonine (T) sites that are likely subjected to O-glycosylation.

5. Although PiRV2 shows no significant similarity to other viruses and sequences from other organism groups, a RdRP_4 region was identified based on a conserved domain search.

Title	Pssmid	Multi-Dom	E-value
pfam02123, RdRP_4, Viral RNA-directed RNA-polymerase. This family includes RNA-depende...	72777	No	2e-05
cd01699, RNA_dep_RNAP, RNA_dep_RNAP: RNA-dependent RNA polymerase (RdRp) is an essenti...	73158	No	4e-04

Fig. 5 Results of conserved domain search (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi?) for PiRV2

6. Preliminary analysis shows that PiRV3 is a member of *Totiviridae*; PiRV4 has only one related sequence in public databases—the 20S RNA replicon in *Saccharomyces cerevisiae*, which is related to RNA coliphage (detail not shown).

Discussion:

1. PiRV1 is likely to have single-stranded genome because it has the RdRP_1 domain and poly(A) tail. The dsRNAs likely are the replicative form of its genome.
2. In a related study, we showed that single-stranded in vitro transcripts of PiRV1 are infectious, and the virus can drive expression of GFP sandwiched between 5' and 3' UTRs of the virus;
3. It is unknown whether or not the viruses have any impact on *P. infestans* physiology and its pathogenicity on host plants. Curing of PiRV2 has been achieved and we are in progress of curing other viruses. Cured strains will be compared with wild-type strains.

References:

- C. V. Hacker, C. M. Brasier, K. W. Buck. 2005. A double-stranded RNA from a *Phytophthora* species is related to the plant endornaviruses and contains a putative UDP glycosyltransferase gene. Journal of General Virology 86:1561.
- H. S. Judelson, A. L. Fabritius. 2000. A linear RNA replicon from the oomycete *Phytophthora infestans*. Molecular and General Genetics 263:395.
- M. Kozak. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44:283.
- H. A. L¨utcke *et al.* 1987. Selection of AUG initiation codons differs in plants and animals. Embo Journal 6:43.
- J. R. Newhouse, P. W. Tooley, O. P. Smith, R. A. Fishel. 1992. Characterization of double-stranded RNA in isolates of *Phytophthora infestans* from Mexico, the Netherlands, and Peru. Phytopathology 82:164
- P. W. Tooley, A. D. Hewings, K. F. Falkenstein. 1989. Detection of double-stranded RNA in *Phytophthora infestans*. Phytopathology 79:470.