A double-stranded RNA from a *Phytophthora* species is related to the plant endornaviruses and contains a putative UDP glycosyltransferase gene

Caroline V. Hacker,1 Clive M. Brasier2 and Kenneth W. Buck1

1Department of Biological Sciences, Imperial College London, Sir Alexander Fleming Building, London SW7 2AZ, UK
2Forest Research Agency, Alice Holt Lodge, Farnham, Surrey GU10 4LH, UK

A new dsRNA was isolated from a *Phytophthora* isolate from Douglas fir. Sequence analysis showed the dsRNA to consist of 13 883 bp and to contain a single open reading frame with the potential to encode a polyprotein of 4548 aa. This polyprotein contained amino acid sequence motifs characteristic of virus RNA-dependent RNA polymerases (RdRps) in its C-terminal region and motifs characteristic of RNA helicases in its N-terminal region. These sequence motifs were related to corresponding motifs in plant viruses in the genus *Endornavirus*. In phylogenetic trees constructed from the RdRp and helicase motifs of a range of ssRNA and dsRNA viruses, the *Phytophthora* RdRp and helicase sequences clustered with those of the plant endornaviruses with good bootstrap support. The properties of the *Phytophthora* dsRNA are consistent with its being classified as the first non-plant member of the genus *Endornavirus*, for which we propose the name *Phytophthora* endornavirus 1 (PEV1). A region between the RdRp and helicase domains of the PEV1 protein had significant amino acid sequence similarity to UDP glycosyltransferases (UGTs). Two sequence motifs were identified, one characteristic of all UGTs and the other characteristic of sterol UGTs. The PEV1 UGT would be the first for an RNA virus, although ecdysteroid UGT genes have been found in many baculoviruses. The PEV1 UGT was only distantly related to baculovirus ecdysteroid UGTs, which belong to a family distinct from the sterol UGTs.

INTRODUCTION

Viruses have potential as biocontrol agents for plant-pathogenic fungi and some success has been achieved in developing *Cryphonectria* hypoviruses for the control of chestnut blight (Dawe & Nuss, 2001). *Phytophthora* is a genus consisting entirely of plant pathogens. *Phytophthora* species and other members of the family *Oomycetes* have many biological properties characteristic of fungi, but on the basis of sequence similarities are classified with diatoms and brown algae in a protist group known as the Stramenopiles (Patterson, 1989; Leipe et al., 1994; Dick, 2001) or Chromista (Cavalier-Smith, 1997). Virus-like dsRNAs in the size range 1.35–11.1 kbp have been described in isolates of *Phytophthora infestans*, the causative agent of late blight of potato (Tooley et al., 1989; Newhouse et al., 1992), but none of these has been sequenced. No dsRNAs have been reported in *Phytophthora* species, which are major pathogens of forest trees, such as *Phytophthora ramorum*, which causes sudden oak death. We have screened a number of fungal and oomycete tree pathogens for dsRNAs with potential for development as biocontrol agents. Here, we report the detection of a 13.9 kbp dsRNA in a *Phytophthora* isolate from Douglas fir and the analysis of the complete nucleotide sequence of the dsRNA.

METHODS

**Origin and culture of *Phytophthora* isolate P441.** *Phytophthora* isolate P441 was isolated from Douglas fir by E. M. Hansen and P. B. Hamm, Oregon State University, USA. The isolate was already suspected to contain dsRNA; it is considered to belong to a new and so far unnamed *Phytophthora* species that for the present is informally designated *P. taxon Douglas fir* (E. M. Hansen, personal communication). P441 was cultured in pea broth (Ristaino et al., 1998) at 20 °C for 10–14 days.

**Extraction and fractionation of nucleic acids.** Nucleic acids were extracted from the mycelium of *Phytophthora* isolate P441 and dsRNA was isolated from the total nucleic acids by lithium chloride fractionation and purified by removal of traces of DNA and ssRNA by digestion with DNase and S1 nuclease, respectively, followed by electrophoresis through a 1 % agarose gel and extraction of...
the dsRNA band with an RNaid kit (Bio 101), as described by Hong et al. (1998a). dsRNA was also analysed by polyacrylamide gel electrophoresis as described by Rogers et al. (1986). Incubations with DNase and S1 nuclease were carried out as described by Hong et al. (1998b).

**cDNA cloning, sequencing and sequence analysis.** cDNA libraries were constructed using gel-purified dsRNA either using random hexamer primers as described by Hong et al. (1998b) or by a novel RT-PCR method (M. J. Roossinck, personal communication). Nucleotide sequences were obtained by the Sanger chain-termination method (Sanger et al., 1977) using dye-terminator cycle sequencing with AmpliTaq DNA polymerase FS (ABI Prism 377).

cDNA clones between contigs were obtained using RT-PCR with sequence-specific primers. cDNA clones of the ends of the dsRNA were obtained by the rapid amplification of cDNA ends (RACE) procedure (Frohman et al., 1988) using a Gibco-BRL 5′ RACE system with the vector pGEM-T Easy (Promega) and by RNA ligase-mediated (RLM)-RACE (Schaefer, 1995). Sequences were assembled and analysed using the University of Wisconsin Genetics Computer Group programs (Devereux et al., 1984). BLASTP searches of the UniProt database were done using WU-Blast2 programs at http://www.ebi.ac.uk. TBLASTN searches of the P. ramorum and Phytophthora sojae genome databases were carried out at http://genome.jgi-psf.org. Multiple sequence alignments, construction of phylogenetic trees and bootstrap analysis were done with the CLUSTAL X program (Thompson et al., 1997). Phylogenetic trees were displayed using TREEVIEW (Page, 1996).

**RESULTS AND DISCUSSION**

**A new dsRNA in Phytophthora isolate P441 from Douglas fir**

Total nucleic acids were isolated from *Phytophthora* isolate P441, ssRNA was removed by precipitation with 2 M LiCl and dsRNA was precipitated with 5 M LiCl. Analysis by PAGE revealed a single band with a mobility considerably slower than the largest available (6–4 kbp) dsRNA marker (Fig. 1a, lane 2). Treatment of the dsRNA with DNase or S1 nuclease prior to electrophoresis had no effect on the band, which was therefore confirmed to be dsRNA (Fig. 1a, lanes 3 and 4). When the dsRNA was electrophoresed in an agarose gel, the band migrated between the 9·4 and 23·1 kbp dsDNA markers (Fig. 1b).

**Production of cDNA clones and nucleotide sequence analysis of P441 dsRNA**

cDNA clones were synthesized from denatured P441 dsRNA and sequenced as described in Methods. The complete sequence of 13 883 bp (GenBank/EMBL accession no. AJ877914) was assembled from overlapping clones. All regions were sequenced from at least two independently synthesized clones. Analysis of the coding potential of all three reading frames of both strands revealed a long open reading frame (ORF) in one strand, designated the plus strand, starting at nt 8 and ending at nt 13843, which could encode a protein of 4612 aa (calculated molecular mass 519 864 Da). However, this would imply a 5′-untranslated region (UTR) of only 7 nt. Furthermore, the methionine codon at nt 8–10 was in an unfavourable sequence context for translation initiation, with a pyrimidine (C) residue at the −3 position and an A residue at the +1 position, according to Kozak’s rules for ribosomal scanning (Kozak, 1986; Lütcke et al., 1987). There was another methionine residue starting at nt 200, which would give a 5′UTR of 199 nt and a protein of 4548 aa. The methionine codon at nt 200 was in a favourable context for translation initiation, with a purine (G) residue at the −3 position and a G residue at the +1 position. Hence, this may be the translation initiation site. Nevertheless, efficient translation of mRNAs with very short 5′UTRs (0–14 nt) has been described in the protozoan *Giardia lamblia*, which may not utilize a ribosomal scanning mechanism (Li & Wang, 2004), and further work will be needed to determine the true translation initiation site in the *Phytophthora* P441 RNA. For sequence analysis, we used the smaller 4548 aa protein. The P441 RNA 3′UTR consisted of 40 nt, including a run of nine C residues at the 3′ terminus. No other significant ORFs were found in the other two reading frames of the plus strand or the three reading frames of the minus strand of the P441 dsRNA.
**Phytophthora P441 dsRNA is related to plant viruses in the genus **Endornavirus**

A BLASTP search of the UniProt protein database using the complete amino acid sequence of the predicted P441 protein produced highly significant alignments with predicted polyprotein sequences encoded by two plant viruses in the genus *Endornavirus* (Gibbs et al., 2000, 2004), *Vicia faba endornavirus* (VfEV) (Pfeiffer, 1998) and *Oryza sativa endornavirus* (OsEV) (Moriyama et al., 1995). Seven alignments totalling 1998 aa were obtained between corresponding regions spread across most of the P441 and VfEV polyproteins (Table 1) with a composite *P* value of $3\cdot3e^{-139}$. Similarly, eight alignments totalling 1722 aa were obtained between corresponding regions spread across most of the P441 and OsEV polyproteins (Table 1) with a composite *P* value of $1\cdot7e^{-117}$. There was also significant sequence similarity between the P441 polyprotein and those of two other endornaviruses, *Oryza rufipogon endornavirus* (OrEV) (Moriyama et al., 1999) and *Phaseolus vulgaris endornavirus* (Wakarchuk & Hamilton, 1985). For the latter virus, only a small amount of sequence (630 bp, 210 aa) was available in the GenBank/EMBL database (Wakarchuk & Hamilton, 1990). Endornaviruses are endogenous dsRNA elements with genomes in the size range of 14–18 kbp and contain a single ORF spanning most of the length of the RNA (Gibbs et al., 2004). The size of the *Phytophthora* P441 dsRNA, the presence of a single ORF and the significant sequence similarity indicated that this dsRNA was related evolutionarily to the plant endornaviruses. We therefore propose that it is a novel member of the genus *Endornavirus* and suggest the name phytophthora endornavirus 1 (PEV1). This would be the first non-plant member of the genus.

In terms of size, the PEV1 dsRNA (13,883 bp) is more similar to OsEV dsRNA (13,952 bp) and OrEV dsRNA (13,936 bp) than to VfEV dsRNA (17,635 bp). In addition, like PEV1, OsEV and OrEV have methionine codons near the 5’ terminus at nt 5–7 in unfavourable sequence contexts for translation initiation by scanning and methionine codons in favourable contexts at nt 167–169. The lengths of the putative polyproteins encoded by PEV1 (4612 or 4548 aa), OsEV (4626 or 4572 aa) and OrEV (4627 or 4573 aa) were also similar (the two values are for translation initiation at the first or second methionine codon, respectively, in each case). There was no significant similarity between the N-terminal sequences of the PEV1, OsEV, OrEV or VfEV polyproteins, whether the first or second methionine codon was used for translation initiation. Alignment of the complete polyprotein amino acid sequences showed 22.3% identity and 32.1% similarity for PEV1/OsEV and 22.5% identity and 31.7% similarity for PEV1/OrEV. Sequence similarities were relatively low outside the regions identified by the BLASTP search (Table 1).

The large proteins encoded by the endornaviruses have been assumed to be polyproteins that are processed by virus-encoded proteinases by analogy with polyproteins encoded by other viruses, such as the picornaviruses (Seipelt et al., 1999). No proteinase motifs or proteolytic cleavage sites have yet been reported for the plant endornaviruses. We searched the PEV1 polyprotein sequence for conserved proteinase motifs, as described by Koonin & Dolja (1993), but were unable to identify any such motifs unequivocally. However, during the search we identified a cysteine-rich region that contained 19 cysteine residues in a sequence of 89 residues (aa 675–763). This sequence could be aligned with previously undescribed cysteine-rich regions in the polyproteins of VfEV (14 cysteines), OsEV (nine cysteines) and OrEV (nine cysteines). Whether these regions have structural or enzymic roles is not known.

There was no significant sequence similarity between the 5’UTR of PEV1 plus-strand RNA and those of VfEV, OsEV or OrEV, irrespective of whether the first or second methionine codon was used as the translation initiation site. Also there was no similarity between the 3’UTR of the plus strand of PEV1 and that of VfEV. Although the 3’UTRs of OsEV and OrEV did not terminate in a run of C residues as found in PEV1, it was noted that the 3’UTRs of these RNAs were C-rich (CCCCCTCCCCAACCCCCCG and TCCCTCCCCAACCCCCCG, respectively; Moriyama et al., 1995, 1999).

A single break has been reported in the plus strand of OsEV, OrEV and VfEV at nt 1211 (or 1256), 1197 and 2735 from the 5’ end, respectively, while the minus strand contained no break (Fukuhara et al., 1995; Moriyama et al., 1999; Pfeiffer, 1998). We examined the plus and minus

---

**Table 1.** Regions of significant similarity between the PEV1 and VfEV or OsEV amino acid sequences identified by a BLASTP search of the UniProt database

<table>
<thead>
<tr>
<th>Amino acid residues</th>
<th>Identity (%)</th>
<th>Similarity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PEV1</strong></td>
<td><strong>VfEV</strong></td>
<td></td>
</tr>
<tr>
<td>166–802</td>
<td>776–1402</td>
<td>21</td>
</tr>
<tr>
<td>977–1630</td>
<td>1572–2232</td>
<td>23</td>
</tr>
<tr>
<td>2375–2678</td>
<td>2984–3275</td>
<td>22</td>
</tr>
<tr>
<td>2930–3000</td>
<td>3506–3575</td>
<td>28</td>
</tr>
<tr>
<td>3443–3778</td>
<td>3703–4046</td>
<td>22</td>
</tr>
<tr>
<td>4052–4542</td>
<td>5230–5813</td>
<td>39</td>
</tr>
<tr>
<td><strong>PEV1</strong></td>
<td><strong>OsEV</strong></td>
<td></td>
</tr>
<tr>
<td>180–266</td>
<td>304–389</td>
<td>26</td>
</tr>
<tr>
<td>430–563</td>
<td>557–693</td>
<td>19</td>
</tr>
<tr>
<td>738–852</td>
<td>815–938</td>
<td>24</td>
</tr>
<tr>
<td>1032–1645</td>
<td>1174–1792</td>
<td>22</td>
</tr>
<tr>
<td>2424–2697</td>
<td>2349–2620</td>
<td>20</td>
</tr>
<tr>
<td>3058–3102</td>
<td>3368–3412</td>
<td>31</td>
</tr>
<tr>
<td>3573–3785</td>
<td>3647–3843</td>
<td>25</td>
</tr>
<tr>
<td>4057–4542</td>
<td>4080–4566</td>
<td>39</td>
</tr>
</tbody>
</table>
strands of PEV1 dsRNA for breaks or discontinuities using RLM-RACE. An oligonucleotide was ligated to 3'-OH groups in denatured PEV1 dsRNA. There should be three types of 3'-OH group, two from the 3' termini of the dsRNA and the third from any internal breaks in the RNA. RLM-RACE had been used to determine the sequences of the 3' termini of the RNA. When RLM-RACE was used in conjunction with a forward primer close to the 5' end of the plus strand together with a reverse primer complementary to the ligated oligonucleotide, three independently obtained clones indicated a break in the plus strand at nt 1215. Additionally, two independent clones indicated a break at nt 926, three clones indicated a break at nt 930 and two clones indicated a break at nt 933. It is noteworthy that the break at nt 1215 is similar in position to breaks in the plus strand reported for OsEV (nt 1211 and 1256) (Fukuhara...
et al., 1995) and OrEV (nt 1197) (Moriyama et al., 1999). Attempts to determine whether all or only a proportion of the PEV1 dsRNA molecules contained the breaks by denaturing gel electrophoresis were unsuccessful due to difficulties in completely denaturing the dsRNA.

**PEV1 polyprotein contains RNA-dependent RNA polymerase-like and helicase-like regions**

The region of the PEV1 polyprotein with the highest similarity to the VfEV and OsEV polyproteins, identified by the BLASTP search (Table 1), was located near the C terminus (aa 4052–4542). Inspection of this sequence showed that it contained sequence motifs, labelled A–E, characteristic of RNA-dependent RNA polymerases (RdRps) (Fig. 2a). Motifs A–D corresponded to those described by Poch et al. (1989); similar motifs have been described by Habili & Symons (1989), Bruenn (1991) and Koonin (1991). Motif E was similar to motifs described by Bruenn (1993) and Koonin (1991). When the region of the PEV1 polyprotein defined by motifs A–E was used in a BLASTP search of the UniProt database, the two highest scores were single homologous regions from the VfEV polyprotein (49 % identity, 65 % similarity, P value 3.8e−54) and OsEV polyprotein (45 % identity, 61 % similarity, P value 5.2e−48), previously identified as containing RdRp domains (Pfeiffer, 1998; Moriyama et al., 1995). The next most similar alignments were with regions of RdRps of several closteroviruses, tobamoviruses, a cucumovirus and hepatitis E virus (HEV) (P values in the range of 9.4e−5 to 9.6e−5). A neighbour-joining phylogenetic tree constructed from motifs A–E of a range of fungal and plant RNA viruses and HEV showed that PEV1 clustered with the three members of the genus *Endornavirus* with 100 % bootstrap support (Fig. 2b). Another branch with moderate (50.1 %) bootstrap support clustered PEV1 and the endornaviruses with tobamoviruses, closteroviruses, potexviruses, a cucumovirus (*Bromoviridae*) and HEV.

Another region of the PEV1 polyprotein with similarity to the VfEV and OsEV polyproteins, identified by the BLASTP search of the UniProt database, was located in the N-terminal third of the protein (aa 1032–1645). Inspection of this region showed that it contained sequence motifs, labelled I–VI (Fig. 3), characteristic of RNA helicases (Koonin & Dolja, 1993). When the region of the PEV1 polyprotein defined by motifs I–VI was used in a BLASTP search of the UniProt database, as expected the two highest scores were single homologous regions from the VfEV polyprotein (30 % identity, 52 % similarity, P value 4.3e−27) and OsEV polyprotein (27 % identity, 50 % similarity, P value 1.3e−20), previously identified as containing helicase-like domains (Pfeiffer, 1998; Moriyama et al., 1995). The next most similar alignments were with helicase-like regions of *Citrus tristeza virus* (genus *Closterovirus*) (23 % identity, 42 % similarity, P value 0.058) and *Tobacco rattle virus* (genus *Tobravirus*) (21 % identity, 38 % similarity, P value 0.11). A neighbour-joining phylogenetic tree constructed from motifs I–VI of a range of animal, fungal and plant RNA viruses showed that PEV1 clustered with the three members of the genus *Endornavirus* with 92.2 % bootstrap support (Fig. 3b).

The RdRp and helicase trees (Figs 2b and 3b) support the inclusion of PEV1 in the genus *Endornavirus*. The clustering of the RdRps of PEV1 and the endornaviruses with those of families and genera in the alpha-like virus superfamily of positive-strand RNA viruses (Buck, 1996) is consistent with previous findings of Gibbs et al. (2000) and with their suggestion that endornaviruses may have evolved...
from an alpha-like virus that has lost its capsid gene. The helicase regions of PEV1 and the endornaviruses were more distantly related to those of alpha-like viruses and did not cluster with them in the phylogenetic tree (Fig. 3b). There was no evidence for a close relationship between the RdRp or helicase of PEV1 and those of viruses in the Hypoviridae, a family of naked fungal RNA viruses with genomes in the size range of 10–13 kbp (Nuss et al., 2004). Viruses in the family Hypoviridae have previously been shown to be related to positive-strand RNA viruses in the family Potyviridae, which are members of the picorna-like virus supergroup (Buck, 1996).
PEV1 contains a putative UDP glycosyltransferase gene

A BLASTP search of the UniProt database using aa 2775–3299 of the PEV1 polyprotein produced significant alignments with many bacterial and fungal UDP-glucose:sterol glucosyltransferases. The highest score was produced from the UDP-glucose:sterol glucosyltransferase of *Ustilago maydis*, which aligned as two fragments with a composite P value of $5 \times 10^{-16}$. The shorter fragment (42% identity, 66% similarity), which was located at the N terminus of the PEV1 region (aa 2775–2794), contained an 8 aa sequence motif found in UDP-glucose:sterol glucosyltransferases from bacteria, fungi and plants (Warnecke *et al.*, 1999). The longer fragment (23% identity, 39% similarity) was located near the C terminus of this region (aa 3040–3272) and contained a 39 aa sequence motif characteristic of all UDP glycosyltransferases (UGTs) and suggested to be a UDP-sugar-binding domain (Mackenzie *et al.*, 1997). A multiple alignment of these two motifs, designated A and B, from UDP-glucose:sterol glucosyltransferases of bacteria, fungi, plants and protists is shown in Fig. 4. Out of 47 residues compared, there were 13 residues that were invariant in the 16 sequences compared and 40 of the residues in the PEV1 protein were found in at least one of the other sequences. This constitutes good evidence that PEV1 contains a UGT gene, which is probably a UDP-glucose:sterol glucosyltransferase gene.

Inspection of the sequences of the polyproteins of the plant endornaviruses revealed the presence of motifs A and B in OrEV and OsEV (Fig. 4) in analogous positions to those found in PEV1 (Fig. 5). Furthermore the entire amino acid sequences of the UDP-glucose:sterol glucosyltransferases in Fig. 4 could be aligned with analogous sequences in PEV1, OrEV and OsEV, including motifs A and B.
B, suggesting the presence of UDP-glucose:sterol glucosyltransferase genes in all three of these endornaviruses. UGTs have not previously been reported for OrEV and OsEV. Motifs A and B could not be found in the polyprotein encoded by ViEV, which could not be aligned with the putative UGTs of the other endornaviruses. Hence, it appears that ViEV lacks a UGT gene.

Glycosyltransferases have been found in several families of DNA viruses and ecdysteroid UGTs are common in baculoviruses (reviewed by Markine-Goriaynoff et al., 2004). However, the present report appears to be the first glycosyltransferase gene to be found in an RNA virus. Viral glycosyltransferases are generally considered to have been acquired from their host at some time during the evolution of the host.

Table 2. Comparison of cellular and endornavirus UGT amino acid sequences

<table>
<thead>
<tr>
<th></th>
<th>Ca</th>
<th>Pp</th>
<th>Sc</th>
<th>Um</th>
<th>Cg</th>
<th>Dd</th>
<th>Pr</th>
<th>Ps</th>
<th>At</th>
<th>Os</th>
<th>Ao</th>
<th>Rb</th>
<th>Xc</th>
<th>OrEV</th>
<th>OsEV</th>
<th>PEV1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ca</td>
<td>91.5</td>
<td>85.1</td>
<td>74.5</td>
<td>59.6</td>
<td>59.6</td>
<td>61.7</td>
<td>63.8</td>
<td>61.7</td>
<td>61.7</td>
<td>55.3</td>
<td>66.0</td>
<td>55.3</td>
<td>36.2</td>
<td>42.6</td>
<td>51.1</td>
<td></td>
</tr>
<tr>
<td>Pp</td>
<td>97.9</td>
<td>91.5</td>
<td>78.7</td>
<td>57.5</td>
<td>59.6</td>
<td>66.0</td>
<td>68.1</td>
<td>66.0</td>
<td>68.1</td>
<td>55.3</td>
<td>68.1</td>
<td>57.5</td>
<td>38.3</td>
<td>42.6</td>
<td>51.1</td>
<td></td>
</tr>
<tr>
<td>Sc</td>
<td>95.7</td>
<td>95.7</td>
<td>80.9</td>
<td>59.6</td>
<td>59.6</td>
<td>66.0</td>
<td>68.1</td>
<td>66.0</td>
<td>68.1</td>
<td>55.3</td>
<td>66.0</td>
<td>57.5</td>
<td>42.6</td>
<td>46.8</td>
<td>46.8</td>
<td></td>
</tr>
<tr>
<td>Um</td>
<td>87.2</td>
<td>89.4</td>
<td>87.2</td>
<td>57.4</td>
<td>61.7</td>
<td>68.1</td>
<td>66.0</td>
<td>66.0</td>
<td>68.1</td>
<td>55.3</td>
<td>68.1</td>
<td>55.3</td>
<td>40.4</td>
<td>42.6</td>
<td>57.5</td>
<td></td>
</tr>
<tr>
<td>Cg</td>
<td>68.1</td>
<td>70.2</td>
<td>70.2</td>
<td>68.1</td>
<td>68.1</td>
<td>72.3</td>
<td>72.3</td>
<td>68.1</td>
<td>66.0</td>
<td>48.8</td>
<td>66.0</td>
<td>57.5</td>
<td>34.0</td>
<td>40.4</td>
<td>38.3</td>
<td></td>
</tr>
<tr>
<td>Dd</td>
<td>70.2</td>
<td>70.2</td>
<td>72.3</td>
<td>70.2</td>
<td>72.3</td>
<td>68.1</td>
<td>66.0</td>
<td>66.0</td>
<td>68.1</td>
<td>44.7</td>
<td>47.7</td>
<td>49.8</td>
<td>47.7</td>
<td>46.8</td>
<td>40.4</td>
<td></td>
</tr>
<tr>
<td>Pr</td>
<td>68.1</td>
<td>68.1</td>
<td>68.1</td>
<td>68.1</td>
<td>68.1</td>
<td>72.3</td>
<td>72.3</td>
<td>68.1</td>
<td>66.0</td>
<td>48.8</td>
<td>76.6</td>
<td>68.1</td>
<td>57.5</td>
<td>34.0</td>
<td>36.2</td>
<td></td>
</tr>
<tr>
<td>Ps</td>
<td>70.2</td>
<td>70.2</td>
<td>70.2</td>
<td>66.0</td>
<td>83.0</td>
<td>74.5</td>
<td>93.6</td>
<td>80.9</td>
<td>76.6</td>
<td>68.1</td>
<td>87.7</td>
<td>59.6</td>
<td>34.0</td>
<td>36.2</td>
<td>48.9</td>
<td></td>
</tr>
<tr>
<td>At</td>
<td>70.2</td>
<td>70.2</td>
<td>70.2</td>
<td>68.1</td>
<td>78.7</td>
<td>74.5</td>
<td>80.9</td>
<td>80.9</td>
<td>76.6</td>
<td>59.6</td>
<td>59.6</td>
<td>34.0</td>
<td>36.2</td>
<td>51.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Os</td>
<td>70.2</td>
<td>72.3</td>
<td>72.3</td>
<td>72.3</td>
<td>74.5</td>
<td>76.6</td>
<td>76.6</td>
<td>76.6</td>
<td>89.4</td>
<td>42.6</td>
<td>72.3</td>
<td>57.5</td>
<td>38.3</td>
<td>40.4</td>
<td>51.1</td>
<td></td>
</tr>
<tr>
<td>Ao</td>
<td>68.1</td>
<td>68.1</td>
<td>70.2</td>
<td>63.8</td>
<td>59.6</td>
<td>55.3</td>
<td>57.4</td>
<td>59.6</td>
<td>59.6</td>
<td>55.3</td>
<td>51.1</td>
<td>51.0</td>
<td>31.9</td>
<td>55.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rb</td>
<td>74.5</td>
<td>74.5</td>
<td>74.5</td>
<td>70.2</td>
<td>76.6</td>
<td>70.2</td>
<td>76.6</td>
<td>83.0</td>
<td>78.7</td>
<td>61.7</td>
<td>55.3</td>
<td>34.0</td>
<td>36.2</td>
<td>55.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xc</td>
<td>66.0</td>
<td>63.8</td>
<td>63.8</td>
<td>59.6</td>
<td>63.8</td>
<td>57.4</td>
<td>63.8</td>
<td>61.7</td>
<td>59.6</td>
<td>61.7</td>
<td>38.3</td>
<td>44.7</td>
<td>46.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OrEV</td>
<td>46.8</td>
<td>44.7</td>
<td>44.7</td>
<td>42.6</td>
<td>51.1</td>
<td>36.2</td>
<td>36.2</td>
<td>36.2</td>
<td>40.4</td>
<td>42.6</td>
<td>44.7</td>
<td>44.7</td>
<td>78.7</td>
<td>36.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OsEV</td>
<td>53.2</td>
<td>48.9</td>
<td>48.9</td>
<td>44.7</td>
<td>53.2</td>
<td>44.7</td>
<td>38.3</td>
<td>38.3</td>
<td>38.3</td>
<td>42.6</td>
<td>44.7</td>
<td>46.8</td>
<td>91.5</td>
<td>34.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEV1</td>
<td>59.6</td>
<td>61.7</td>
<td>61.7</td>
<td>66.0</td>
<td>53.2</td>
<td>51.1</td>
<td>57.4</td>
<td>55.3</td>
<td>59.6</td>
<td>59.6</td>
<td>61.7</td>
<td>55.3</td>
<td>40.5</td>
<td>42.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
evolution of the virus and the host (Markine-Goriaynoff et al., 2004). If endornaviruses are transmitted only by intracellular routes and inherited through sexual progeny (Gibbs et al., 2004), it is possible that they have co-evolved with their hosts. Once acquired from a host, a UGT gene may also evolve with the host. Table 2 compares cellular and endornavirus UGT amino acid sequences, based on the conserved motifs A and B. The PEV1 UGT sequence was not more closely related to those of two Phytophthora species, *P. ramorum* and *P. sojae*, than it was to those of bacteria, fungi and plants. This suggests that the PEV1 UGT gene has not been acquired recently from the *Phytophthora* genome and may have been acquired by a progenitor endornavirus of an ancient host that predates the separation of bacteria, fungi and plants. This would explain the presence of a UGT gene in OsEV and OrEV in a similar position in their genomes to that in PEV1.

Virus-encoded glycosyltransferases generally have functions that are of benefit to the virus (Markine-Goriaynoff et al., 2004). For example, baculovirus ecdysteroid glucosyltransferases inactivate ecdysteroid hormones by glycosylating them, thereby preventing molting and pupation of infected insect larvae and increasing virus yield and spread (reviewed by O’Reilly, 1995). It is not yet known whether the PEV1 UGT gene produces an active enzyme, but the conservation of 13 residues that are invariant in cellular sterol UGTs suggests that the PEV1 gene is functional. Cellular sterol UGT genes in two plant-pathogenic fungi, *Colletotrichum gloeosporioides* (Kim et al., 2002) and *Magnaporthe grisea* (Sweigard et al., 1998), are required for pathogenicity. Hence, expression of the PEV1 UGT gene might modulate the pathogenicity of its *Phytophthora* host towards Douglas fir by gene silencing or other mechanisms. Homology-dependent gene silencing of a pathogenicity gene in *P. infestans* severely reduced its virulence (Latijnhouwers et al., 2004).

It is noteworthy that the PEV1 UGT sequence was more similar to those of cellular UGTs than to the UGTs of OrEV and OsEV. The RdRps that replicate RNA viruses lack proofreading mechanisms (reviewed by Buck, 1996) and therefore a UGT gene incorporated into an RNA virus has the potential to evolve more rapidly than cellular UGT genes. Hence, if both the PEV1 and OrEV (and OsEV) UGT genes have been evolving independently since the separation of protists from plants, it is likely that they would have become more divergent from each other than from the host genes. It is possible that ViEV has originated from the same UGT-containing ancestral virus as PEV1, OrEV and OsEV and that the UGT gene has been lost. This could have occurred if the UGT gene did not have a useful function to the virus in its bean host. The PEV1 UGT amino acid sequence was only very distantly related to that of baculovirus ecdysteroid UGTs, which lack motif A. The cellular sterol UGTs and baculovirus ecdysteroid UGTs have different substrate specificities and belong to different UGT families (Mackenzie et al., 1997). Hence, PEV1 and baculoviruses have most likely acquired their UGT genes after the divergence of these two families of cellular UGT genes.

**ACKNOWLEDGEMENTS**

This work was funded by the Natural Environment Research Council and the Forestry Commission.

**REFERENCES**


http://vir.sgmjournals.org


