A novel putative virus of *Gremmeniella abietina* type B (Ascomycota: Helotiaceae) has a composite genome with endornavirus affinities

Tero T. Tuomivirta,¹ Juha Kaitera² and Jarkko Hantula¹

¹Finnish Forest Research Institute, Vantaa Research Unit, Jokiniemenkuja 1, PO Box 18, FI-01301 Vantaa, Finland
²Finnish Forest Research Institute, Muhos Research Unit, Kirkkosaarentie 7, FI-91500 Muhos, Finland

Ascospore and mycelial isolates of *Gremmeniella abietina* type B were found to contain three different dsRNA molecules with approximate lengths of 11, 5 and 3 kb. The 11 kb dsRNA encoded the genome of a putative virus and is named *Gremmeniella abietina* type B RNA virus XL (GaBRV-XL). GaBRV-XL probably exists in an unencapsulated state. We identified two distinct dsRNAs (10 374 and 10 375 bp) of GaBRV-XL, both of which coded for the same putative polyprotein (3249 amino acids) and contained four regions similar to putative viral methyltransferases, DExH box helicases, viral RNA helicase 1 and RNA-dependent RNA polymerases. While a cysteine-rich region with several CxCC motifs in GaBRV-XL was similar to that of putative endornaviruses, cluster analyses of conserved regions revealed GaBRV-XL to be distinct from a broad range of viral taxa but most closely related to *Discula destructiva* virus 3. Collectively, these findings suggest that GaBRV-XL represents a novel virus group related to endornaviruses.

INTRODUCTION

Recently recognized by the International Committee on Taxonomy of Viruses (ICTVdB Management, 2006), endornaviruses are usually cryptic and non-enveloped plant and fungal dsRNA viruses that spread efficiently through mitotic and meiotic cells but not through grafts (Fukuhara et al., 1995; Moriyama et al., 1999; Wakarchuk & Hamilton, 1990; Pfeiffer, 1998). Almost all of the 14 000-18 000 nucleotides comprising endornavirus genomes code for a putative polyprotein and exhibit a characteristic nick near the 5’ end of the coding strand. Similar putative viruses have been found in the Stramenopila (*Phytophthora*; Hacker et al., 2005) and in the basidiomycete *Helicobasidium mompa* (Fukuhara et al., 2005; Osaki et al., 2006), both of which are pathogens of plants.

*Gremmeniella abietina* var. *abietina* is the causal agent of scleroderris canker on coniferous trees. In northern Europe, this ascomycete is commonly recognized as a species complex (e.g. Uotila et al., 2000) with two types (A and B) that mainly infect Scots pine (*Pinus sylvestris*). Fungal viruses of the *G. abietina* species complex have been studied thoroughly in type A and are known to include putative members of the virus families *Narnaviridae, Totiviridae* and *Partitiviridae*, some of which can co-infect a single fungal isolate (Tuomivirta et al., 2002; Tuomivirta & Hantula, 2005).

The goals of this study include a survey of dsRNA molecules in *G. abietina* type B, molecular characterization of the most common dsRNA type and its comparison to representatives of similar viral taxa.

METHODS

*G. abietina* type B strains. *G. abietina* type B strains (see Supplementary Table S1, available in JGV Online) were identified by the random amplified microsatellite (RAMS) technique of Hantula & Müller (1997). Branches from *P. sylvestris* and *Pinus contorta* <2 m in height (<20 years old) with symptoms of scleroderris canker were collected in artificially or naturally regenerated stands between June 1994 and June 1995 in northern Finland. Disease symptoms included death of lateral branches over several internodes, cankers, yellow-green woody tissues, pycnidia and apothecia. Fungi were isolated either from pycnidia or from infected wood adjacent to apothecia.

Nucleic acid isolation and electrophoresis. *G. abietina* type B isolates were grown at 20 °C on modified orange serum agar covered with a cellophane membrane (Müller et al., 1994). dsRNA isolation
was based on the protocol of Morris & Dodds (1979) as modified by Tuomivirta et al. (2002) and Tuomivirta & Hantula (2003). dsRNA molecules bind specifically to CF-11 cellulose (Whatman) in the presence of 15% ethanol and, in order to test for dsRNA encapsidation, the protocol was also conducted with non-phenol or chloroform extractions in osmotic stabilizer buffer (0.6 M NaCl in 0.1 M phosphate buffer, pH 6.0; Phillips, 1993). dsRNA recovery was quantified by electrophoresis according to Tuomivirta et al. (2002).

Buoyant densities were determined according to Tuomivirta & Hantula (2005) and each of 10 CsCl gradient fractions was tested for the presence of dsRNA through agarose gel electrophoresis. cDNA synthesis and sequencing followed Tuomivirta & Hantula (2005) except that dsRNA from isolates AU58 (Müller & Uotila, 1997) and E46 (Table 1) were ligated using T4 RNA ligase (Fermentas) to 5'-phosphorylated and 3'-inactivated adaptors T4II (5'-GCATTCCGACCCGGGTT-3') and T4III (5'-AGAGACCCAGTCTGAGCTCCA- GT-3'), respectively. Also, eight (AU58) and 10 (E46) specific primers (available on request) were designed for cloning or direct sequencing.

Sequences were compiled, aligned and analysed using Vector NTI Advance 10 (Invitrogen). Complete nucleic acid and putative amino acid sequences were screened using the protein BLAST (Altschul et al., 1997) search engine including CD-Search (Marchler-Bauer & Bryant, 2004) of the National Center for Biotechnology Information (NCBI) and the InterProScan (Zdobnov & Apweiler, 2001) protein signature search engine of the European Bioinformatics Institute (EBI). Alignment algorithms were from CLUSTAL W (Thompson et al., 1994). Phylogenetic analyses of resulting alignments were conducted according to the neighbour-joining method as applied in MEGA 3.1 (Kumar et al., 2004), Bayesian inference under a mixed amino acid model (ngen = 10,000, samplefreq = 10, burnin = 250) in MrBayes 3.1.2 (Huelsenbeck et al., 2001; Ronquist & Huelsenbeck, 2003) and fast maximum-likelihood using PhyML (Guindon et al., 2005). We used ProtTest (Abascal et al., 2005) to identify and apply the best-fit models of sequence evolution in the PhyML analysis. Finally, we used TreeView 1.6.6 (Page, 1996) to inspect the resultant dendrograms and their branch support values.

Table 1. Viral amino acid sequences used in this study

Details of other sequences used in the construction of Supplementary Fig. S4 are given in Supplementary Table S2.
RESULTS

Host and physicochemical properties of dsRNA

RAMS fingerprint analysis of all 31 isolates (Supplementary Table S1) confirmed them to be *G. abietina* type B (not shown). We detected three different dsRNA molecules among eight isolates and double infections were apparent in two isolates; E46 contained dsRNAs of 11 and 5 kb, whereas AU58 contained dsRNAs of 11 and 3 kb (not shown). Electrophoretic mobility of the 5 and 3 kb dsRNAs was identical to that of putative fungal totiviruses and mitoviruses from *G. abietina* type A (Tuomivirta & Hantula, 2005).

Ultracentrifugation of isolate E46 identified four fractions containing dsRNA (not shown). Most of the 11 kb dsRNAs were recovered from the nucleic acid pellet, but small amounts also occurred in the uppermost (1: \(d = 1.26\) g ml\(^{-1}\)) and lowermost (10: \(d = 1.505\) g ml\(^{-1}\)) fractions. The 5 kb dsRNA was found only in fraction 8 (\(d = 1.387\) g ml\(^{-1}\)) and with a similar mean density (\(d = 1.405\) g ml\(^{-1}\)) to a totivirus isolated from *G. abietina* type A (Tuomivirta & Hantula, 2005).

Using the cellulose-binding protocol, we successfully isolated the 11 kb dsRNA from a mycelial culture of isolate E46 without phenol extraction (Fig. 1). However, we were unable to recover the 5 kb dsRNA using the same approach. The 11 kb dsRNAs, probably of viral origin, were named as *Gremmeniella abietina* type B RNA virus XL (GaBRV-XL).

Production of cDNA and nucleotide sequence analysis of GaBRV-XL from isolates AU58 and E46

Complete sequences of the 11 kb dsRNAs of isolates AU58 (GaBRV-XL1) and E46 (GaBRV-XL2) were assembled from 109 and 86 sequencing reactions of clones and direct sequencing of RT-PCR products, respectively. Their fully compiled and contiguous lengths were 10 375 and 10 374 bp, respectively. The sequences were 97% similar and both contained a single, 10 287 bp open reading frame (ORF) starting at nt 24 according to a universal translational table. The ORF spanned 99.2% of the dsRNA and encoded a protein of 3249 aa with an approximate molecular mass of 383 kDa. The protein sequences were 98.7% identical and 99.4% similar (Fig. 2).

GaBRV-XL codes for a viral methyltransferase, two helicases and an RNA-dependent RNA polymerase

Based on amino acid sequence analysis, the GaBRV-XL dsRNAs encoded two strains of a novel putative virus. Both CD-Search and InterProScan inferred four conserved regions. The first was a putative viral methyltransferase (CD-Search: pfam01660; expect value 2E-14) region, ranging from aa 265 to 573. The second was a putative DEXDc, DEAD-like helicases superfamily (CD-Search: smart00487; 1E-08) region, aa 1218–1376. The third was a putative viral RNA helicase 1 (CD-Search: pfam01443; 6E-08) region, aa 1908–2025. The last region was RNA_dep_RNApol2 (CD-Search: pfam00978; 2E-14), aa 3090–3254.

The putative methyltransferase region was most similar to sequences of *Soil-borne cereal mosaic virus* (Diao et al., 1999a; Koenig et al., 1999), *Soil-borne wheat mosaic virus* (Shirako & Wilson, 1993; Ratti et al., 2004) and *Oat golden stripe virus* (Diao et al., 1999a) (unassigned members of *Furovirus*), *Fragaria chiloensis latent virus* (Tzanetakis & Martin, 2005) (*Bromoviridae: Ilarvirus*), Chinese wheat mosaic virus (Yang et al., 2001; Diao et al., 1999b) (unclassified *Furovirus*) and *Pepper mild mottle virus* (Velasco et al., 2002; GenBank accession no. BAD90599) (unassigned *Tobamovirus*), and the expect value ranged from 3.5E-06 to 6E-03. A selection of putative methyltransferases was aligned with that of GaBRV-XL.

---

**Fig. 1.** Detergent-free isolation of dsRNA from isolate E46 by different treatments. The starting material between chloroform+CF-11 and CF-11 treatments was equivalent. The estimated length of dsRNA is marked on the left.

**Fig. 2.** Genomic organization of GaBRV-XL1. The putative polyprotein of 3249 amino acids contained five conserved regions with respect to other virus families: viral methyltransferase (VMet; aa 265–573), cysteine-rich region (CRR; aa 520–1140), DExH box helicase (DEXH; aa 1218–1376), viral RNA helicase 1 (VHel1; aa 1908–2025) and RNA_dep_RNApol2 (RdRp; aa 3090–3254).
Discula destructiva similar to a partially sequenced 12 kb dsRNA identified as a putative endornavirus of bell pepper (Valverde & Gutierrez, 2007) with an expect value of 4E-4, which may represent a previously unassigned member of the Ampelovirus genus (ranging from 1E-17 to 8E-3) and for ssRNA viruses Grapevine roostock stem lesion-associated virus (NCBI reference sequence NP_835244) (an unclassified member of the Closteroviridae, genus Closterovirus) and Little cherry virus 2 (Rott & Jelkmann, 2005) (Closteroviridae. Ampelovirus).

Complete or partial polymerase amino acid sequences of GaBRV-XL1-2, DvV-3 and the putative endornaviruses listed above were aligned (Supplementary Fig. S4) and subjected to phylogenetic analysis. The alignment contained all eight conserved motifs of the RNA-dependent RNA polymerases of positive-strand RNA viruses and related double-stranded viruses (Koonin & Dolja, 1993). In addition, a new and unique motif was identified, IXa.

Putative endornaviruses contain a cysteine-rich region with several CxCC signatures

In the BLAST search, aa 545–1143 of the PEV1 polyprotein matched significantly with a region extending between positions 520 and 1140 (expect value 2E-4) of GaBRV-XL2. While similarities with other putative endornaviruses were not detected in this region, we noted that the cysteine-rich region first reported by Hacker et al. (2005) showed some degree of sequence similarity among putative endornaviruses. The cysteine-rich regions in ORV (aa 856–900; 20% cysteine) and OSV (855–899; 20%) each contained two, whereas VFV (1265–1362; 16.3%), PEV1 (675–763; 21.4%) and HmEV1-670 (1198–1269; 22.2%) and both strains of GaBRV-XL (747–846; 17%) each contained three CxCC signatures. One of the observed signatures was in the form CxCG in all viruses.

Phylogenetic analysis of the GaBRV-XL genome

The four conserved regions of GaBRV-XL were each separately aligned with sequences from similar viruses and exposed to phylogenetic analysis (Table 1, Fig. 3). The substitution model applied in MrBayes and PhyML was based on maximum-likelihood (PhyML) methods were identical and inferred a distant relationship between GaBRV-XL and other putative endornaviruses. Phylogenies based on the viral RNA helicase 1 and RNA_dep_RNApol2 regions were decisive, fully resolved and well supported by bootstrap and posterior probability values. While we note that taxonomic sampling was different for all alignments, phylogenetic resolution was weak for alignments of the viral methyl-
In summary, our phylogenetic analyses suggested GaBRV-XL to be distinct and clearly separated from other recognized viral groups.

**DISCUSSION**

We have reported the occurrence of three different dsRNA molecule types in the fungus *G. abietina* type B. The largest type, GaBRV-XL, is a sexually transmitted, cryptic, monopartite, linear dsRNA virus, somewhat similar to assigned and unassigned members of genus *Endornavirus*. Based on centrifugation trials and dsRNA extraction tests, in line with known endornaviruses, it does not exhibit a protein capsid. However, the putative totivirus infecting the same isolate seems to have one.

The genomic structure of GaBRV-XL appears to be a composite of five conserved regions. While physicochemical properties and sequences of the RNA_dep_RNApol2, viral RNA helicase 1 and cysteine-rich regions resemble those of known endornaviruses, the DExH box helicase and viral methyltransferase regions appear to be restricted to GaBRV-XL. The genome of GaBRV-XL is further distinguished from almost all known endornaviruses in that it lacks the UDP glycosyltransferase motifs (Hacker *et al.*, 2005), and RT-PCR results suggest that the nick in its coding strand is missing.

The putative methyltransferase region in GaBRV-XL is based on sequences deposited in the EBI (InterPro: IPR002588) and is widely found in a range of ssRNA viruses, including hordei-, tobra-, tomoabo-, bromo-, clostero- and caliciviruses. The methyltransferase region is involved in mRNA capping to increase stability. In eukaryotes, mRNA is transcribed in the nucleus, and cytoplasmic viruses therefore encode their own methyltransferase. The methyltransferase region is commonly found in the so-called Alphavirus supergroup and this region has also been detected, e.g. in a fungal reovirus (Hillman *et al.*, 2004) and in a mycovirus infecting *Sclerotinia sclerotiorum* (Xie *et al.*, 2006). The occurrence of the CAP structure in GaBRV-XL was not examined in this study.

The GaBRV-XL genome contained two different helicase regions with unknown roles. Two helicases in a single viral genome have been reported in a group of plant positive-strand ssRNA viruses comprising the potex-, carla-, hordei- and furoviruses, where one helicase is involved in RNA replication and the other in movement of the virus in plant tissue (Kadare & Haenni, 1997).

RNA helicases of the DEAD box and related DExD/H proteins play various roles in RNA metabolism such as nuclear transcription, pre-mRNA splicing, ribosome biogenesis, nucleocytoplasmic transport, translation, RNA decay and organellar gene expression (Tanner & Linder, 2001). In addition, DExD/H box RNA helicases can also displace proteins from RNA (Schwer, 2001).

The viral RNA helicase 1 regions in endornaviruses and foveaviruses are closely related to the alpha-like supergroup (Gibbs *et al.*, 2000; van der Heijden & Bol, 2002) and are thought to be involved in duplex unwinding during viral RNA replication (Gomez de Cedron *et al.*, 1999). A common feature of helicases is that NTP hydrolysis is usually coupled to the unwinding reaction of nucleic acids (Kadare & Haenni, 1997). Motif I is a Walker A NTP-binding motif that binds phosphates on NTP (Tanner & Linder, 2001), and it represents a critical biochemical function. Thus, that motif I is highly similar in the helicase regions of GaBRV-XL may be a convergent consequence of their similar function in the same virus–host system rather than evidence of an evolutionary relationship.

Endornaviruses are generally thought to encode a polyprotein that can be digested by protease. However, protease domains or activity have yet to be reported among endornaviruses. A candidate for a protease domain could...
be the somewhat conserved cysteine-rich region. Cysteine is unique among naturally occurring amino acids in that it contains a thiol group which is capable of reversible oxidation/reduction reactions in the active sites of thiol proteases (papain-like proteases). However, the CxCC signature has also been linked to heavy metal binding (Giritch et al., 1998; Lu et al., 2003), and cysteine itself forms disulfide bonds between adjacent residues to stabilize the tertiary structure of functional proteins.

The differences described above combined with the phylogenetic analyses conducted on the putative methyltransferase, DExH box helicase, viral RNA helicase 1 and RNA_{dep-RNApol2} regions suggest that GaBRV-XL is related to but distinct from known endornaviruses and distantly related or even unrelated to ssRNA viruses. However, the partial sequence of the RNA_{dep-RNApol2} region from Discalla destructiva virus 3 suggests that similar viruses may occur in fungi.

GaBRV-XL and putative endornaviruses discussed in this study are somewhat closely related and may have evolved from an alpha-like virus, as suggested by Gibbs et al. (2000) based on their analysis of polymerase and viral RNA helicase regions. However, the distinction of the GaBRV-XL genome from known endornaviruses suggests a complex evolutionary history for these non-enveloped viruses. Complete sequences of Discalla destructiva virus 3, other putative endornaviruses (Fukuhara et al., 2005; Coutts, 2005; Valverde & Gutierrez, 2007) and the putative XL virus of G. abietina type A (11 kb in electrophoresis; T. T. Tuomivirta and J. Hantula, unpublished) may help us to understand the evolutionary origins and genomics of this interesting group of viruses.

ACKNOWLEDGEMENTS

Ms Marianne Boström (isolate AU58), Ms Tiina Nurmi (isolate E46) and Ms Marja-Leena Santanen are acknowledged for their skilful technical assistance. Dr Antti Uotila is thanked for providing isolate AU58. Dr Michael Hardman helped with editing and language. This research was funded by the Academy of Finland.

REFERENCES


