Journal of General Virology (2013), **94**, 1426–1434 **DOI** 10.1099/vir.0.050393-0

Genetic characterization of Blueberry necrotic ring blotch virus, a novel RNA virus with unique genetic features

Diego F. Quito-Avila,[1] Philip M. Brannen,[2] William O. Cline,[3] Philip F. Harmon[4] and Robert R. Martin[5]

Correspondence
Diego F. Quito-Avila
quitod@onid.orst.edu

[1]Centro de Investigaciones Biotecnologicas del Ecuador (CIBE), Escuela Superior Politecnica del Litoral (ESPOL), Guayaquil, Guayas EC090150, Ecuador

[2]Department of Plant Pathology, University of Georgia, Athens, GA 30602, USA

[3]Department of Plant Pathology, North Carolina State University, Raleigh, NC 27607, USA

[4]Department of Plant Pathology, University of Florida, Gainesville, FL 32605, USA

[5]USDA-ARS, Horticultural Crops Research Unit, Corvallis OR 97331, USA

A new disorder was observed on southern highbush blueberries in several south-eastern states in the USA. Symptoms included irregularly shaped circular spots or blotches with green centres on the upper and lower surfaces of leaves. Double-stranded RNA was extracted from symptomatic leaves suggesting the presence of virus(es) possibly involved in the disease. Sequencing revealed the presence of a novel RNA virus with a ~14 kb genome divided into four RNA segments. Sequence analyses showed that the virus, for which we propose the name Blueberry necrotic ring blotch virus (BNRBV), possesses protein domains conserved across RNA viruses in the alpha-virus-like supergroup. Phylogenetic inferences using different genes placed BNRBV in a clade that includes the *Bromoviridae*, the genus *Cilevirus* (CiLV) and the recently characterized *Hibiscus green spot virus* (HGSV). Despite the strong genetic relationships found among BNRBV, *Cilevirus* and HGSV, the genome of BNRBV contains three features that distinguish it significantly from its closest relatives: (i) the presence of two helicase domains with different evolutionary pathways, (ii) the existence of three conserved nucleotide stretches located at the 3′ non-coding regions of each RNA segment and (iii) the conservation of terminal nucleotide motifs across each segment. Furthermore, CiLV and HGSV possess poly(A)-tailed bipartite and tripartite genomes, respectively, whereas BNRBV has a quadra-partite genome lacking a poly(A) tail. Based on these genetic features a new genus is proposed for the classification of BNRBV.

Received 4 December 2012
Accepted 11 March 2013

INTRODUCTION

The USA is the world's leading producer of blueberries (*Vaccinium* spp.). In 2011, more than 248 000 Tons of highbush blueberry (*Vaccinium corymbosum*) were produced (Geisler, 2012). The increasing consumer demand for this highly valued berry has resulted in the development of new cultivars adapted to areas where blueberry is not native (Moore, 1994; Strik, 2006; Bañados, 2009). This increase in acreage of blueberry has exposed the crop to new pests and pathogens leading to the emergence of novel diseases that challenge its commercial production (Martin *et al.*, 2009). Viruses have been a major concern to blueberry growers in recent years. Prevalence of viral diseases has been increased, in part, by the growers' lack of knowledge on best practices to avoid virus establishment and spread (Martin *et al.* 2009, 2012). Among the most common blueberry viruses reported in the main production areas in North America are *Blueberry shock virus*, reported in the Pacific North-west (MacDonald *et al.*, 1991), Nova Scotia, Canada and in Pennsylvania and New York, USA (Gottula *et al.*, 2012); *Blueberry scorch virus* (Martin & Bristow, 1988; Wegener et al., 2006) and the nematode transmitted viruses *Tomato ringspot virus* and *Tobacco ringspot virus*, which are found sporadically in most commercial areas in the USA (Fuchs *et al.*, 2010). In addition, blueberry mosaic and blueberry fruit drop disorders, the latter being only reported in the Pacific

The GenBank/EMBL/DDBJ accession numbers for the dsRNA sequence of Blueberry necrotic ring blotch virus are JN651148–JN651151.

Four supplementary figures and one supplementary table are available with the online version of this paper.

North-west, have been observed in several commercial fields. The presence of a negative-sense RNA virus in the genus *Ophiovirus* has been associated with mosaic disease, whereas the causal agent responsible for the fruit drop disorder is still undetermined (Martin *et al.*, 2012).

Recently, a new disorder has been observed on southern highbush blueberries in several south-eastern states of the USA. Affected plants showed irregularly shaped circular spots or blotches with green centres on the upper and lower surfaces of leaves. The symptoms were also accompanied by premature defoliation in several locations. In the state of Georgia, the initial distribution of the disease was limited to very sporadic reports in 2006 and 2007, but in 2008 it was found in multiple locations throughout the major blueberry production areas in the state. The disorder was later observed in Florida, Mississippi, South and North Carolina, and has been considered as a potential yield-limiting problem among some cultivars.

The disease was initially thought to be caused by a fungal agent based on the characteristic necrotic ring symptoms. However, tests failed to detect any fungal or bacterial pathogens from symptomatic tissue. Furthermore, ELISA and PCR results were negative for known blueberry viruses, but the presence of dsRNA from several symptomatic leaf samples suggested that a virus was associated with the disease.

In this paper, we present the complete nucleotide sequence, genome organization and phylogenetic analysis of a novel virus isolated from blueberry plants affected with necrotic ring blotch in south-eastern states. The name Blueberry necrotic ring blotch virus (BNRBV) along with the creation of a new genus for its classification is suggested, based on unique genomic features detailed in this report.

RESULTS

Double-stranded RNA pattern and genome organization

DsRNA extracted from several symptomatic leaves migrated as four RNAs with molecular masses of approximately 6, 4, 2.5 and 1.7 kbp (data not shown). This pattern was consistently observed only in dsRNA preparations from symptomatic tissue collected from different fields.

Sequencing revealed a genome comprising 14 153 nt of RNA divided into four segments (NCBI GenBank accession numbers JN651148–JN651151).

RNA 1 is 5906 nt long and contains a single ORF (frame +1) spanning nucleotides 85 (**A**UG) to 5748 (UA**A**). The putative 215 kDa protein (1887 amino acids) encoded by this ORF contains methyltransferase (MTR; aa 125–425), cysteine-protease (C-Pro; aa 643–732) and helicase (HEL; aa 725–1586) domains, which are conserved across several positive-sense (+) RNA viruses. RNA 2 is 3935 nt long with a single ORF (frame +2) spanning nucleotides 347 (**A**UG) to 3781 (UA**G**). The putative 130 kDa protein (1144 aa) encoded by this segment possesses conserved domains for HEL (aa position 225–450) at the N terminus, and RNA-dependent RNA polymerase (RdRp; aa position 700–1125) at the C terminus. RNA 3 contains 2588 nt and comprises five ORFs (ORF 3a, 3b, 3c, 3d and 3e) coding for small proteins with unknown functions. ORF 3a (frame+1), at nucleotide positions 49 (**A**UG) to 228 (UG**A**), encodes a small polypeptide (59 aa) of 7 kDa. ORF 3b (frame+1) spans nucleotides 418 (**A**UG) to 1218 (UA**A**) and codes for a 31 kDa protein (206 aa). ORF 3c (frame+1) spans nucleotides 1237 (**A**UG) to 1813 (UA**G**) encoding a 22 kDa protein (191 aa). ORF 3d (frame+3) spans nucleotides 1659 (**A**UG) to 2438 (UA**A**) and codes for a 28 kDa protein (259 aa). ORF 3e (frame+1), at nucleotide positions 2083 (**A**UG) to 2328 (UA**G**), encodes a 9 kDa protein (81 aa). The 1724 nt long RNA 4 contains a single ORF spanning nucleotides 360 (**A**UG) to 1295 (UG**A**). This ORF encodes a 34 kDa protein (311 aa) with motifs conserved in the 3A superfamily of virus movement proteins (MP; aa position 20–240) (Melcher, 1990, 2000). Fig. 1(a) shows a schematic representation of the genome organization of BNRBV.

Phylogenetic relationships

Analysis of amino acid sequences derived from several ORFs encoded by different RNA segments revealed similarities to protein domains from viruses belonging to the families *Virgaviridae* and *Bromoviridae*, but also to the newly created genus *Cilevirus* (unassigned family) (Locali-Fabris *et al.*, 2006). BLAST searches at the amino acid level on individual domains revealed that the MTR domain contained several conserved motifs with invariant 'hall-mark' residues found in members of the alpha-virus-like supergroup (Rozanov *et al.*, 1992). This domain was found to be homologous to MTRs of members of the genus *Ilarvirus* (*Bromoviridae*).

The analysis of the RdRp, located at the 3′ terminus of ORF 2, placed BNRBV in a clade that includes members of the *Virgaviridae, Citrus leprosis virus* (CiLV: the type member of the genus *Cilevirus*) and the newly characterized virus *Hibiscus green spot virus* (HGSV: proposed genus *Higrevirus*).

As we have shown (see genome organization), BNRBV possesses two HEL domains: one at the 3′ end of RNA 1 (HEL-1) and the second located at the 5′ end of RNA 2 (HEL-2). Interestingly, the two HEL domains did not cluster in the same clade. Instead, HEL-1 grouped with CiLV and HGSV, sharing a common ancestor with tobamoviruses (*Virgaviridae*), whereas HEL-2 showed a closer relationship with members of subgroup 1 of the genus *Ilarvirus* (*Bromoviridae*). Furthermore, searches for conserved motifs revealed that HEL-1 contains motifs (I, II, III and V) commonly found in members of the alphavirus-like superfamily, whereas HEL-2 showed motifs (I and II) highly conserved among members of the flavivirus-like superfamily (Kadaré & Haenni, 1997). However, motif VI of HEL-2

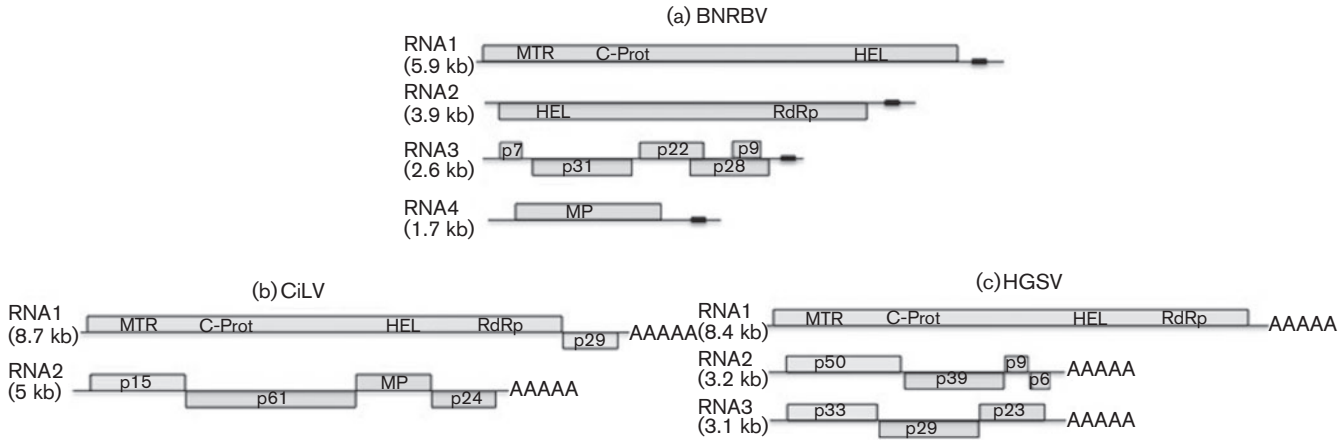

Fig. 1. Graphic representation (not to scale) of the genome organization of: (a) Blueberry necrotic ring blotch virus (BNRBV); (b) *Citrus leprosis virus* (CiLV) and (c) *Hibiscus green spot virus* (HGSV). Conserved domains are indicated inside each box. MTR: methyltransferase; C-Prot: cysteine protease; HEL: helicase; RdRp: RNA-dependent RNA polymerase; MP: movement protein. Putative proteins are indicated according to their size and position in RNA 3. Black boxes at the 3′ ends of BNRBV genome represent the presence of secondary structures, compared to the poly(A) tails present on BNRBV and CiLV.

showed a higher level of conservation to motif VI of the alphavirus-like group, which in turn, is more distant from the corresponding motif in HEL-1 (Table 1). It is worth noting that HEL-2 lacks motif V corresponding to either supergroup of viral helicases.

It seemed clear that BNRBV, along with its closest relatives (i.e. CiLV and HGSV) had resulted from recombination events between tobamoviruses and ilarviruses, which are significant representatives of two important plant virus families, *Virgaviridae* and *Bromoviridae*, respectively. Nevertheless, an evolutionary pathway that best described the most recent radiation of these viruses was still necessary. The identification of recombination break-points, which provides deeper information about the evolution of these viruses, was not possible due to the high variability at the nucleotide level among these viruses.

In an attempt to reconcile the discrepancies found across the MTR, HEL and RdRp domains, a phylogenetic tree was constructed using the corresponding amino acid sequences from representative members of the *Virgaviridae* and *Bromoviridae*. The alignments produced for each domain were concatenated and used for constructing a maximum-likelihood phylogenetic tree using the WAG + G + I amino acid substitution model (Whelan & Goldman, 2001). Two trees were generated, one using BNRBV-HEL1 and the second using BNRBV-HEL2. This analysis revealed that BNRBV and the *Bromoviridae* represent a monophyletic group when BNRBV-HEL1 was used (Fig. 2a). When BNRBV-HEL2 was used, however, it was not possible to generate a reliable tree. This was evidenced by the unsupported bootstrap values (not shown). The analysis of the MP placed BNRBV and members of the genus

Table 1. Comparison of helicase (HEL) conserved motifs between *Blueberry necrotic ring blotch virus* (BNRBV) and the two superfamilies (SF: alpha-like and flavi-like) of helicases

BNRBV-HEL-1, helicase domain from BNRBV RNA 1; BNRBV-HEL-2, helicase domain from BNRBV RNA 2. Capital letters in bold denote amino acid residues homologous to the alpha-like superfamily. Italicized capital letters denote amino acid residues homologous to the poty/flavi-like superfamily. Abbreviations: b, bulky hydrophobic residues (aliphatic or aromatic); u, bulky aliphatic residues; x, any residue.

SF	Motif					
	I	II	III	IV	V	VI
Alphavirus-like	**PG$_x$GKS/T**	**$_{bbb}$DE$_b$**	**GD$_{xx}$Q**		**T$_{bxxx}$QG$_x$T$_{bxx}$V**	**$_b$VA$_u$S/TRH**
Poty/flavivirus-like	*GSGKS/T*	$_{bbb}$*DE$_x$H*	*TATPP*	ub$_2$u*PS*	*TD$_{bx}$E$_x$G$_{uxbx}$*	*T$_{xxxxx}$QR$_x$GR$_u$GR*
BNRBV-HEL-1	**PG$_C$GKT**	$_{LLF}$**DE$_A$**	**GD$_{VM}$Q**		**T$_{IHEY}$QG$_{LEAR}$**	TGI**SRH**
BNRBV-HEL-2	*GSGKT*	$_{LLL}$*DE$_C$H*	$_A$**D$_{LS}$Q**			Y**AL**T**R**F

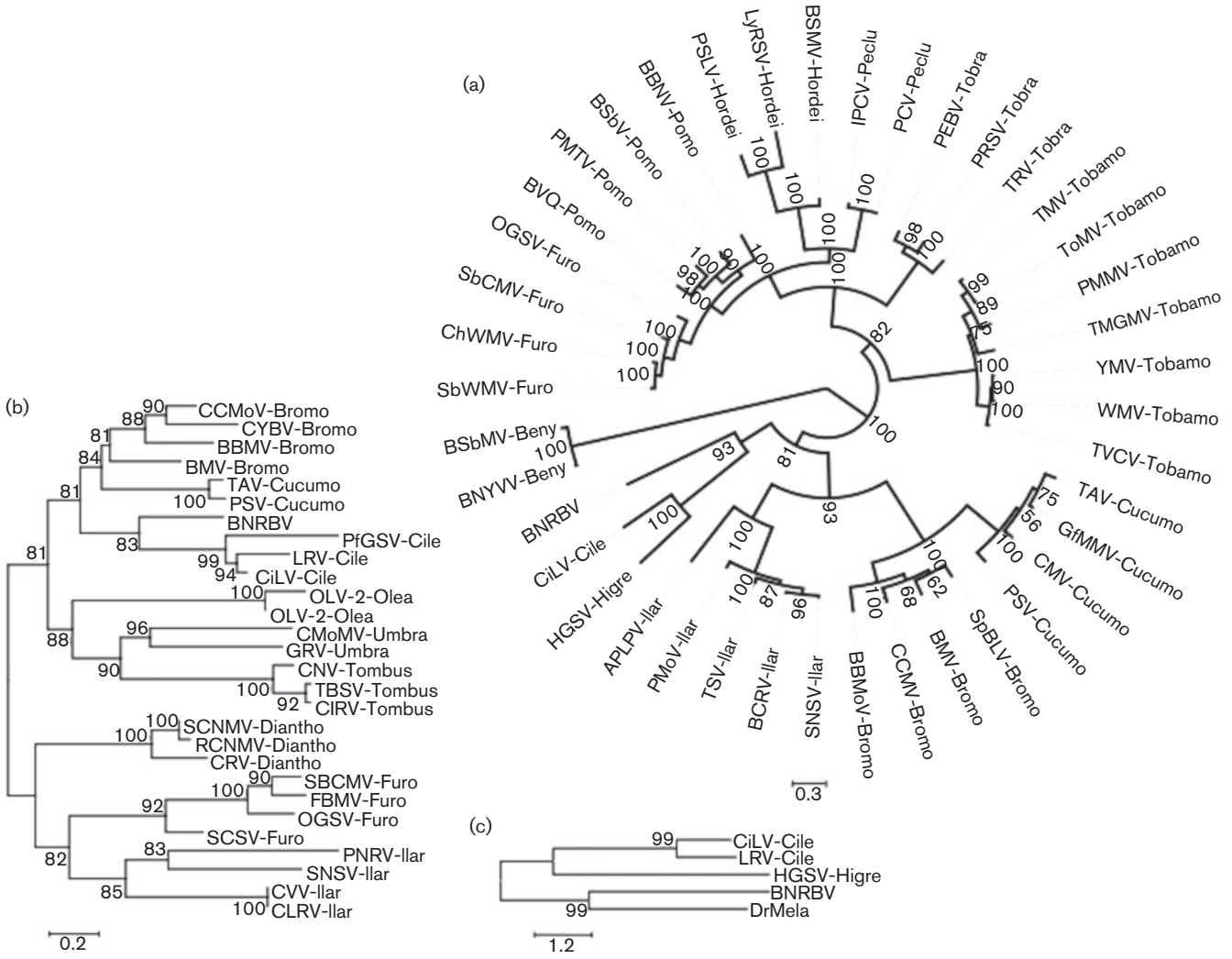

Fig. 2. Phylogenetic reconstruction. Three different inferences are shown: (a) radial tree generated by the concatenation of the methyltransferase, helicase, and RNA-dependent RNA polymerase amino acid domains; (b) rectangular tree constructed using amino acid sequences from movement proteins (MP) from representatives of the 3A-type of MP and (c) tree generated using the p22 protein of Blueberry necrotic ring blotch virus (BNRBV) and its orthologues. (Full names of the viruses and accession numbers used in the phylogenetic analysis are listed in Table S1).

Cilevirus in the same clade, which shared a common ancestor with members of the *Bromoviridae* (Fig. 2b).

Aside from the replicase (MTR-HEL-RdRp) and the MP, the only BNRBV putative protein that was found to have orthologues was that encoded by ORF 3c. Interestingly, this small protein (p22) showed similarities to proteins of comparable size from its closest relatives CiLV (p24), *Ligustrum ringspot virus* (LRV; p24) and HGSV (p23). Position-specific iterated BLAST (PSI-BLAST) searches were able to detect one additional orthologue, a deduced amino acid sequence from *Drosophila melanogaster* clone IP15837p (a clone generated by the Berkeley Drosophila Genome Project) (max. identity 29 %; e-value 6e-54). Fig. 2(c) shows phylogenetic relationships between BNRBV-p22 and its only known homologues. The function of this protein is unknown. Analysis using the PSIPRED program showed similar secondary structure patterns for BNRBV-p22 and viral homologues, suggesting a possible similar function (Fig. S1, available in JGV Online).

Non-coding terminal regions

Non-coding regions (NCRs) ranged from 84 to 359 nt and from 150 to 429 nt at the 5′ and 3′ termini, respectively. Homopolymeric tracts consisting of mainly poly(T) and poly(A) were observed along the 3′ NCRs (data not shown). The A-T content at the 3′ NCRs, including the homopolymeric tracts, was 70 % for RNA segments 1, 2 and 3, and 63 % for RNA 4. The conserved sequence

5′-CACAAAT was found at the 5′ terminus of all segments, whereas the nucleotides CG-3′ were conserved at the 3′ terminus, with RNA 1 and RNA 2 sharing two additional bases at the 3′ end (-**AT**CG-3′). No poly(A) tail was found at either RNA segment. Instead, nucleotide alignments of 3′ NCRs revealed the presence of three conserved stretches arranged in a similar fashion across the four RNA segments. RNA 3, however, showed less conservation for these nucleotide motifs (Fig. 3). Although these nucleotide motifs did show the propensity to form secondary structures of the stem–loop type (Fig. S2), none of the predicted structures had a canonical tRNA-like conformation or typical hairpin type pseudoknots, as found in other viruses (Rietveld *et al.*, 1982, 1983; Pleij *et al.*, 1985; Rudinger-Thirion *et al.*, 2006). Therefore, aminoacylation of viral tRNA-like ends in protein synthesis, as shown for *Brome mosaic virus* and other plant viruses (Dreher *et al.*, 1989; Goodwin *et al.*, 1997) might be a remote possibility. However, the RNA motifs may still play a role in genome activation through the binding of other viral proteins, e.g. CP, to the stem–loop structures, as has been the case for ilar-, alfamo- and idaeoviruses among others (Olsthoorn *et al.*, 1999; MacFarlane & McGavin, 2009).

Detection

Leaves from 63 individual plants of southern highbush blueberry that exhibited symptoms of necrotic ring blotch disease from North Carolina, Georgia and Florida were positive in the reverse transcription-PCR (RT-PCR) based detection assay developed for BNRBV. At the same time, over 200 plants that were asymptomatic and collected from several nurseries, tested negative for BNRBV using the same assay. The correlation between detection of BNRBV by RT-PCR and symptoms in plants from across several states suggests that this is indeed the causal agent of the disease.

DISCUSSION

A new blueberry virus associated with necrotic ring blotch symptoms (Fig. S3) has been identified and named Blueberry necrotic ring blotch virus (BNRBV). After the first report in 2006, the virus was detected in several symptomatic highbush blueberry cultivars in south-eastern states of the USA showing severe foliage damage and potentially lowering yields. This study presents the complete genome sequence of BNRBV and describes genomic features that support the creation of a new genus for its classification.

The genome comprised ~14 kbp of RNA divided into four segments. BLAST searches revealed that BNRBV proteins had similarity to proteins encoded by several RNA viruses belonging to the *Virgaviridae* and *Bromoviridae* – positive-sense (+) single-stranded RNA viruses.

The monophyletic origin of viral RdRps (Koonin *et al.*, 1993; Ahlquist 2006) makes this protein the starting point to determining evolutionary relationships among plant RNA viruses. In BNRBV, the RdRp, encoded by RNA 2, showed a close phylogenetic relationship to CiLV, LRV (*Cilevirus*) and HGSV (proposed genus *Higrevirus*) (Locali-Fabris *et al.*, 2006; Melzer *et al.*, 2012) (Fig. S3). BNRBV RNA1 encodes a large protein with MTR and HEL (HEL-1) domains. In agreement with the RdRp, analyses of these two domains showed a close relationship between BNRBV, CiLV and HGSV (MTR and HEL sequences were not available for LRV at the time this paper was being prepared) with motifs conserved across members of the 'alphavirus-like' superfamily. An interesting finding worth mentioning was that the phylogenetic reconstruction of the MTR placed the clade containing BNRBV, CiLV and HGSV closer to members of the *Bromoviridae*. By contrast, the analysis of the HEL-1 clustered them in a clade with members of the *Virgaviridae* (Fig. S4), suggesting a possible chimeric ancestry for these viruses, which may have occurred from recombination between species of the *Bromoviridae* and *Virgaviridae*. The analysis of the concatenated domains involved in virus replication, however, revealed that BNRBV is closer to the *Bromoviridae* (Fig. 2a). Furthermore, a putative 22 kDa protein (p22) encoded by ORF 3c of BNRBV showed 24 % identity to CiLV 24 kDa protein (p24), 12 % to LRV 24 kDa protein (p24) and 24 % to HGSV 23 kDa protein (p23). Interestingly, the deduced amino acid sequence from *D. melanogaster* clone IP15837p (Berkeley Drosophila Genome Project) showed 29 % identity with BNRBV putative p22 (e-value 6e-54), being the only non-viral orthologue found in the database (Fig. 2c). No putative function has been identified for this small polypeptide.

The analysis of the MP revealed homology to the 3A class of MPs, being related most closely to members of the *Cilevirus* (Fig. 2b). HGSV, on the other hand, was shown to have a triple gene block (TGB)-type of MP, encoded by three contiguous ORFs in RNA 2 (Melzer *et al.*, 2012). This

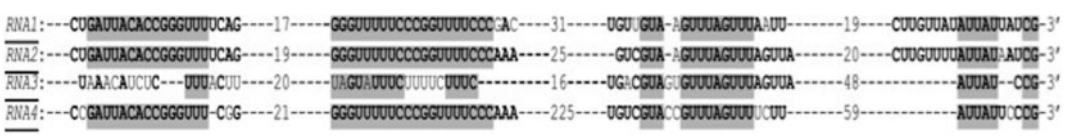

Fig. 3. Alignment of 3′ non-coding regions (NCRs) of Blueberry necrotic ring blotch virus (BNRBV). Shaded areas indicate nucleotides conserved across the four segments. Nucleotides in bold type indicate motifs conserved in at least two RNAs. Numbers denote the amount of non-conserved nucleotides between conserved stretches.

finding reveals a determinant genetic feature that distinguishes BNRBV and HGSV.

Another genetic feature that distinguishes BNRBV from its closest relatives is the presence of three stretches of nucleotides highly conserved across the 3′ NCRs of RNA 1, RNA 2 and RNA 4. These motifs are arranged in a similar fashion in each segment and were partially conserved in RNA 3 (Fig. 3). *Cilevirus* and HGSV on the other hand, do not contain these conserved nucleotide motifs. Instead, their genomes are characterized by the presence of poly(A) tails at their 3′ termini.

Analysis of RNA secondary structure on the conserved stretches did not yield a canonical tRNA-like conformation or typical hairpin type pseudoknots, found in other viruses (Rietveld *et al.*, 1982, 1983; Pleij *et al.*, 1985; Rudinger-Thirion *et al.*, 2006). However, stem–loop structures were inferred from the nucleotide stretches suggesting involvement in virus replication or transcription, as shown for other viruses (Buck, 1996; Sit *et al.*, 1998; Dreher, 1999; Olsthoorn *et al.*, 1999; Sullivan & Ahlquist 1999; Koev *et al.*, 2002; Laforest & Gehrke 2004; MacFarlane & McGavin, 2009).

Seven (I, Ia, II, III, IV, V, VI) conserved sequence motifs have been identified in viral helicases, including the 'Walker A' and 'Walker B' (motif I and motif II) implicated in NTPase/helicase activity (Gorbalenya & Koonin, 1993; Koonin *et al.*, 1993; Morozov & Solovyev, 2003). The presence of specific residues at these motifs was shown to delineate three virus superfamilies: 'alpha-like', 'poty-flavi-like' and 'picorna-like' (Kadaré & Haenni, 1997). In BNRBV, two separate HEL domains were identified. HEL-1, encoded at the 3′ end of RNA 1, contains four (I, II, III and V) of the seven motifs, with high levels of conservation with those of the 'alpha-like' group. HEL-2, encoded at the 5′ end of RNA 2, shows two motifs (I and II) highly conserved to members of the 'poty-flavi-like' group. Furthermore, residues corresponding to motif VI of the 'alpha-like' group were conserved at a higher level in HEL-2, whereas HEL-2 lacked motif V. At this point, it remains unclear whether both HEL domains are required for replication and complement each other's omitted motifs (i.e. motif VI in HEL-1 and motif V in HEL-2).

In order to elucidate the origin of both HEL domains in BNRBV, phylogenetic analyses were performed using a large group of both replicative and TGB-related helicases. Bootstrap values did not support nodes from which HEL-2 had radiated (not shown). The analysis using only the replicative helicases showed that HEL-1 was closer to the *Virgaviridae*, whereas HEL-2 was closer to the *Bromoviridae* (Fig. S4).

It should be noted that there are four important findings that rule out the possibility that the two helicases belonged to two different viruses or that they are the product of 'artificial recombination' by mixing and matching samples. First, and perhaps the most conclusive finding is that the helicases are not closely related, which suggests they are from distinct lineages. In fact, helicase conserved motifs, including 'Walker A' and 'Walker B', were present in both BNRBV helicases. Second, RNA 2 that codes the second helicase starts with the nucleotide sequence 5′-CACAAAT, which is present at the 5′ end in each of the four genomic segments of BNRBV. Third, the 3′ NCR of RNA 2 contains the three nucleotide stretches conserved at 3′ NCRs across all segments. Such conserved nucleotide stretches were not present anywhere else in the genome of BNRBV. Fourth, each helicase could be amplified from RNA extracts from BNRBV infected plants, demonstrating that there was no 'artificial recombination' during the cloning process.

Nevertheless, the only cases where RNA viruses encode two separate helicase domains have been those of viruses with the TGB movement mechanism. Therefore, the presence of a second helicase domain in the BNRBV genome is unique. To our knowledge, it represents the first case where a virus with a 3A-type of MP encodes two helicase domains. It appears that before the radiation of BNRBV, *Cilevirus* and HGSV, their common ancestor possessed both movement mechanisms, and that one became defective or extinct, resulting in a survival constraint to these viruses that led to the adaptation of a movement mechanism different from co-infecting ancestral viruses.

Furthermore, the genome of BNRBV is organized in four RNA segments (quadra-partite) whereas CiLV and HGSV are composed of bi-partite and tri-partite genomes, respectively (Locali-Fabris *et al.*, 2006; Melzer *et al.*, 2012) (Fig. 1).

In summary, differences in genome organization, the lack of poly(A) tail, the existence of conserved nucleotide elements at 3′ NCRs, and the presence of a second replicative-like helicase domain suggest that BNRBV represents a new genus, phylogenetically related but distinct from cileviruses (CiLV, LRV, *Passion fruit green spot virus*) and HGSV. The name *Blunervirus* is proposed for this new genus.

METHODS

Plant material. Symptomatic tissue (cultivars: 'Star' and 'O'Neal') was collected from several commercial fields in Alabama, Georgia, Florida, Michigan, and North and South Carolina during the summer of 2009. The virus genome sequence was determined from samples collected in Georgia and RT-PCR for virus detection was performed from individual leaves collected at multiple locations.

Double-stranded RNA extraction. Batches of 10 g of fresh leaf tissue were powdered in liquid nitrogen and dsRNA was extracted using a lithium salt-based protocol described by Tzanetakis & Martin (2008) and recovered by cellulose CF-11 chromatography (Morris & Dodds, 1979). The dsRNA was treated with RNase T1 from *Aspergillus oryzae* (Sigma-Aldrich) and DNase 1 from bovine pancreas (Sigma-Aldrich) under high salt conditions to remove single-stranded RNA and dsDNA, respectively. After the digestions, the dsRNA was washed, purified on CF-11 cellulose columns and precipitated with EtOH. The dsRNA was analysed at room temperature by electrophoresis on a 1.5 % agarose gel, stained with ethidium bromide and visualized with a UV light.

cDNA synthesis, cloning and sequencing. Purified dsRNA was denatured and used as the template for reverse transcription (RT) following the methodology described by Tzanetakis *et al.* (2005). The random primer 5′-GCCGGAGCTCTGCAGAATTCNNNNNN-3′ was used to generate first-strand cDNA (Froussard, 1992). Briefly, a mixture containing the purified dsRNA and primers was denatured with CH_3HgOH at room temperature for 30 min. Then, a second mix, containing reverse transcription buffer, DTT, dNTPs and Maxima Reverse Transcriptase (Fermentas) was added to the denaturant mix and incubated for 1 h at 50 °C. The reaction was terminated by heating at 85 °C for 10 min and subsequently amplified by PCR using the anchor primer 5′-GCCGGAGCTCTGCAGAATTC-3′ (Froussard, 1992). The PCR products were cloned into the Strataclone pSC-A-amp/kan vector (Agilent Technologies) and sequenced on an ABI 3730XL DNA analyser by Macrogen (Seoul, Korea).

Detection. Detection primers were developed to amplify a 560 bp fragment of a conserved region of the polymerase gene (F: 5′-GGTTTCGACACCTCGGCATG-3′ R: 5′-CCAGCTGCCCTTGAGA-CTTRTC-3′). Total RNA or viral dsRNA was heat denatured at 95 °C for 5 min and reverse transcribed by Maxima Reverse Transcriptase (Fermentas) in a total volume of 25 µl. After denaturing the transcriptase, 2.5 µl of cDNA was used in a total of 25 µl of PCR mix. The PCR was performed as follows: one cycle of initial denaturation at 94 °C for 4 min, 40 cycles of denaturation for 30 s at 94 °C, annealing for 30 s at 57 °C, extension for 45 s at 72 °C and a final extension for 10 min at 72 °C. PCR products were analysed by electrophoresis through ethidium bromide-stained 2 % agarose gels in 1X Tris-phosphate-EDTA (TPE) buffer.

High throughput sequencing and sequence assembly. The Illumina deep sequencing platform was implemented to obtain viral genomic reads randomly generated using the primer 5′-GCCG-GAGCTCTGCA**GGTACC**NNNNNN-3′ from dsRNA as described above (bases in bold type represent *Kpn*I restriction site). Random fragments were PCR-amplified using the anchor primer 5′-GCCGGAGCTCTGCAGGTACC-3′ and subsequently digested by *Kpn*I at 37 °C for 24 h.

Digested products were fragmented with NEBNext dsDNA Fragmentase (New England BioLabs) according to the manufacturer's instructions. NEBNext DNA Sample Prep Master Mix Set 1 (New England Biolabs) was used to prepare DNA fragments for Illumina sequencing. Illumina sequencing compatible adapters/5Phos/GATC-GGAAGAGCGGTTCAGCAGGAATGCCGAG and 5′-ACACTCTTT-CCCTACACGACGCTCTTCCGATC*T together with PCR amplification primers 5′-AATGATACGGCGACCACCGAGATCTACACTCT-TTCCCTACACGACGCTCTTCCGATC*T and 5′-CAAGCAGAAGA-CGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTT-CCGATC*T (Integrated DNA Technologies) were used. Cluster generation was performed by the Center for Genome Research and Biocomputing (CGRB) at Oregon State University (Corvallis, OR, USA) following the manufacturer's recommendations. Two 80 bp and two 160 bp single read lanes were sequenced on an Illumina GAIIx sequencer (Illumina) by the CGRB. Resulting reads were assembled using software programs Velvet (Zerbino & Birney, 2008), AbySS (Simpson *et al.*, 2009) and Edena (Hernandez *et al.*, 2008). Contigs generated from optimized assemblies from the three different programs were further assembled using CAP3 software (Huang & Madan, 1999).

Contigs were subjected to BLASTX and BLASTP searches to obtain a reference sequence. The closest virus species in the GenBank was used for sequence alignment and orientation purposes. Sequence gaps were filled in after PCR amplification using specific primers developed from previously known sequences.

The 5′ and 3′ termini were obtained by poly(A) tailing (Ambion, USA) of the 3′ ends of dsRNAs, as described by Isogai *et al.* (1998), followed by RT-PCR using primers developed to known sequences near the ends and oligo dT. To reconfirm the first (5′ end) and last (3′ end) bases of each segment, a 3′ blocked DNA oligonucleotide was ligated to both 3′ ends of the dsRNAs using T4 RNA ligase 2 (New England Biolabs) followed by RT-PCR using its complementary primer and specific primers for each end as described (Attoui *et al.*, 2000).

Phylogenetic reconstruction and protein analysis. The phylogenetic analysis comprised three steps. First, conserved motifs were identified for each protein (or domain) analysed in this study (i.e. MTR, HEL, RdRp, MP and unknown BNRBV protein p22) using BLAST searches. Second, amino acid sequences from several RNA viruses containing conserved domains for each protein were aligned using MUSCLE (Edgar, 2004) and subsequently subjected to amino acid substitution models prediction. This analysis was performed using maximum-likelihood (ML) by MEGA 5.0 (Tamura *et al.*, 2011). Models with the lowest Bayesian information criterion scores were regarded as the ones that describe the evolutionary substitution pattern the best. Third, evolutionary relationships, based on the model obtained previously, were computed by MEGA software version 5.0 (Tamura *et al.*, 2011) using the ML method with bootstrap test (1000 replicates) (Efron 1982; Felsenstein 1985; Saitou & Nei, 1987). In order to increase the reliability of the phylogenetic inference, highly divergent regions were manually removed prior to the phylogenetic analysis (Lake 1991; Swofford *et al.* 1996).

Amino acid sequences corresponding to the MTR, HEL and RdRp were concatenated and phylogenetic relationships analysed as a whole based on the best amino acid substitution model common to all three domains.

For the MP, partial or full amino acid sequences of the 3A type of MP corresponding to viruses of the genera *Dianthovirus* and *Tombusvirus* (*Tombusviridae*), *Furovirus* (*Virgaviridae*) and *Umbravirus* were included, based on distant genetic relationships found between their MPs and those from members of the *Bromoviridae* (An *et al.*, 2003; Rochon *et al.*, 2012). In addition, MPs of two tentative members of the genus *Cilevirus*: *Passion fruit green spot virus* and *Ligustrum ringspot virus* were included in the phylogenetic analysis. It is important to note that HGSV lacks an orthologous 3A type of MP; consequently, it was not included in this phylogenetic reconstruction.

GenBank accession numbers of the amino acid sequences used for the alignments and phylogeny are presented in Table S1.

Protein secondary structure predictions for BNRBV p22, CiLV p24 and HGSV p23 were made using the PSIPRED program which performs a reliable inference of the protein structure using the amino acid sequence alone. Confidence levels in the predictions, on a scale of zero to unity, are shown for each motif in the predicted structure (McGuffin *et al.*, 2000).

Analysis of non-coding regions (NCRs). In order to predict stem–loops and tRNA-like secondary structures at the 3′ terminal NCRs, the ScanSE version 1.21 software (Lowe & Eddy, 1997) and RNA fold program, included in the ViennaRNA Package 2.0 (Lorenz *et al.*, 2011) were used.

ACKNOWLEDGEMENTS

The authors acknowledge funding from the Oregon and Washington Blueberry Commissions, the British Columbia Blueberry Industry Development Council, the Southern Regional Small Fruit Consortium, the Georgia Agricultural Commodity Commission for Blueberries, the NIFA-Specialty Crops Research Initiative, the National Clean Plant Network and USDA-ARS. The authors would also like to thank Dr Nicklaus Grunwald and Javier Tabima for providing insight into phylogenetic inferences.

REFERENCES

Ahlquist, P. (2006). Parallels among positive-strand RNA viruses, reverse-transcribing viruses and double-stranded RNA viruses. *Nat Rev Microbiol* **4**, 371–382.

An, H., Melcher, U., Doss, P., Payton, M., Guenzi, A. C. & Verchot-Lubicz, J. (2003). Evidence that the 37 kDa protein of Soil-borne wheat mosaic virus is a virus movement protein. *J Gen Virol* **84**, 3153–3163.

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.

Bañados, M. (2009). Expanding blueberry production into non-traditional production areas: northern Chile and Argentina, Mexico and Spain. *Acta Hortic* **810**, 439–445.

Buck, K. W. (1996). Comparison of the replication of positive-stranded RNA viruses of plants and animals. *Adv Virus Res* **47**, 159–251.

Dreher, T. W. (1999). Functions of the 3′-untranslated regions of positive strand RNA viral genomes. *Annu Rev Phytopathol* **37**, 151–174.

Dreher, T. W., Rao, A. L. N. & Hall, T. C. (1989). Replication in vivo of mutant brome mosaic virus RNAs defective in aminoacylation. *J Mol Biol* **206**, 425–438.

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

Efron, B. (1982). The Jackknife, the Bootstrap and Other Resampling Plans. *CBMS-NSF Regional Conference Series in Applied Mathematics, Monograph 38.* Philadelphia: SIAM.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39**, 783–791.

Froussard, P. (1992). A random-PCR method (rPCR) to construct whole cDNA library from low amounts of RNA. *Nucleic Acids Res* **20**, 2900.

Fuchs, M., Abawi, G. S., Marsella-Herrick, P., Cox, R., Cox, K. D., Carroll, J. E. & Martin, R. R. (2010). Occurrence of *Tomato ringspot virus* and *Tobacco ringspot virus* in highbush blueberry in New York State. *J Plant Pathol* **92**, 451–459.

Geisler, M. (2012). Ag Marketing Resource Center, Iowa State University. http://www.agmrc.org/commodities__products/fruits/blueberries-profile/

Goodwin, J. B., Skuzeski, J. M. & Dreher, T. W. (1997). Characterization of chimeric turnip yellow mosaic virus genomes that are infectious in the absence of aminoacylation. *Virology* **230**, 113–124.

Gorbalenya, A. E. & Koonin, E. V. (1993). Helicases: amino acid sequence comparisons and structure-function relationships. *Curr Opin Struct Biol* **3**, 419–429.

Gottula, J., Cox, K., Carroll, J. & Fuchs, M. F. (2012). Blueberry shock disease. http://www.nysipm.cornell.edu/factsheets/berries/bb_shock.pdf

Hernandez, D., François, P., Farinelli, L., Osterås, M. & Schrenzel, J. (2008). De novo bacterial genome sequencing: millions of very short reads assembled on a desktop computer. *Genome Res* **18**, 802–809.

Huang, X. & Madan, A. (1999). CAP3: A DNA sequence assembly program. *Genome Res* **9**, 868–877.

Isogai, M., Uyeda, I. & Hataya, T. (1998). An efficient cloning strategy for viral double-stranded RNAs with unknown sequences. *Ann Phytopathol Soc Jpn* **64**, 244–248.

Kadaré, G. & Haenni, A.-L. (1997). Virus-encoded RNA helicases. *J Virol* **71**, 2583–2590.

Koev, G., Liu, S., Beckett, R. & Miller, W. A. (2002). The 3′ terminal structure required for replication of Barley yellow dwarf virus RNA contains an embedded 3′ end. *Virology* **292**, 114–126.

Koonin, E. V., Dolja, V. V. & Morris, T. J. (1993). Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. *Crit Rev Biochem Mol Biol* **28**, 375–430.

Laforest, S. M. & Gehrke, L. (2004). Spatial determinants of the alfalfa mosaic virus coat protein binding site. *RNA* **10**, 48–58.

Lake, J. A. (1991). The order of sequence alignment can bias the selection of tree topology. *Mol Biol Evol* **8**, 378–385.

Locali-Fabris, E. C., Freitas-Astúa, J., Souza, A. A., Takita, M. A., Astúa-Monge, G., Antonioli-Luizon, R., Rodrigues, V., Targon, M. L. P. N. & Machado, M. A. (2006). Complete nucleotide sequence, genomic organization and phylogenetic analysis of Citrus leprosis virus cytoplasmic type. *J Gen Virol* **87**, 2721–2729.

Lorenz, R., Bernhart, S. H., Höner Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F. & Hofacker, I. L. (2011). ViennaRNA Package 2.0. *Algorithms Mol Biol* **6**, 26.

Lowe, T. M. & Eddy, S. R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* **25**, 955–964.

MacDonald, S., Martin, R. & Bristow, P. (1991). Characterization of an Ilarvirus associated with a necrotic shock reaction in blueberry. *Phytopathology* **81**, 210–214.

MacFarlane, S. A. & McGavin, W. J. (2009). Genome activation by raspberry bushy dwarf virus coat protein. *J Gen Virol* **90**, 747–753.

Martin, R. R. & Bristow, P. (1988). A carlavirus associated with blueberry scorch disease. *Phytopathology* **78**, 1636–1640.

Martin, R. R., Tzanetakis, I. E., Caruso, F. L. & Polashock, J. J. (2009). Emerging and reemerging virus diseases of blueberry and cranberry. *Acta Hortic* **810**, 299–304.

Martin, R. R., Polashock, J. J. & Tzanetakis, I. E. (2012). New and emerging viruses of blueberry and cranberry. *Viruses* **4**, 2831–2852.

McGuffin, L. J., Bryson, K. & Jones, D. T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics* **16**, 404–405.

Melcher, U. (1990). Similarities between putative transport proteins of plant viruses. *J Gen Virol* **71**, 1009–1018.

Melcher, U. (2000). The '30K' superfamily of viral movement proteins. *J Gen Virol* **81**, 257–266.

Melzer, M. J., Sether, D. M., Borth, W. B. & Hu, J. S. (2012). Characterization of a virus infecting Citrus volkameriana with citrus leprosis-like symptoms. *Phytopathology* **102**, 122–127.

Moore, J. (1994). The blueberry industry of North America. *Acta Hortic* **346**, 15–26.

Morozov, S. Y. & Solovyev, A. G. (2003). Triple gene block: modular design of a multifunctional machine for plant virus movement. *J Gen Virol* **84**, 1351–1366.

Morris, T. J. & Dodds, J. A. (1979). Isolation and analysis of double-stranded RNA from virus infected-plant and fungal tissue. *Phytopathology* **69**, 854–858.

Olsthoorn, R. C. L., Mertens, S., Brederode, F. T. & Bol, J. F. (1999). A conformational switch at the 3′ end of a plant virus RNA regulates viral replication. *EMBO J* **18**, 4856–4864.

Pleij, C. W., Rietveld, K. & Bosch, L. (1985). A new principle of RNA folding based on pseudoknotting. *Nucleic Acids Res* **13**, 1717–1731.

Rietveld, K., Van Poelgeest, R., Pleij, C. W., Van Boom, J. H. & Bosch, L. (1982). The tRNA-like structure at the 3′ terminus of turnip yellow mosaic virus RNA. Differences and similarities with canonical tRNA. *Nucleic Acids Res* **10**, 1929–1946.

Rietveld, K., Pleij, C. W. & Bosch, L. (1983). Three-dimensional models of the tRNA-like 3′ termini of some plant viral RNAs. *EMBO J* **2**, 1079–1085.

Rochon, D., Lommel, S., Martelli, G. P., Rubino, L. & Russo, M. (2012). Tombusviridae. In *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses*, pp. 1111–1138. Edited by A. M. Q. King, M. J. Adams, E. B. Carstens & E. J. Lefkowitz. London: Elsevier.

Rozanov, M. N., Koonin, E. V. & Gorbalenya, A. E. (1992). Conservation of the putative methyltransferase domain: a hallmark of the 'Sindbis-like' supergroup of positive-strand RNA viruses. *J Gen Virol* **73**, 2129–2134.

Rudinger-Thirion, J., Olsthoorn, R.-C., Giegé, R. & Barends, S. (2006). Idiosyncratic behaviour of tRNA-like structures in translation of plant viral RNA genomes. *J Mol Biol* **355**, 873–878.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.

Simpson, J. T., Wong, K., Jackman, S. D., Schein, J. E., Jones, S. J. & Birol, I. (2009). ABySS: a parallel assembler for short read sequence data. *Genome Res* **19**, 1117–1123.

Sit, T. L., Vaewhongs, A. A. & Lommel, S. A. (1998). RNA-mediated trans-activation of transcription from a viral RNA. *Science* **281**, 829–832.

Strik, B. (2006). Blueberry production and research trends in North America. *Acta Hortic* **715**, 173–184.

Sullivan, M. L. & Ahlquist, P. (1999). A brome mosaic virus intergenic RNA3 replication signal functions with viral replication protein 1a to dramatically stabilize RNA in vivo. *J Virol* **73**, 2622–2632.

Swofford, D. L., Olsen, G. J., Waddell, P. J. & Hillis, D. M. (1996). Phylogenetic inference. In *Molecular Systematics*, pp. 407–514. Edited by D. M. Hillis, C. Moritz & B. K. Mable. Sunderland, MA: Sinauer Associates.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731–2739.

Tzanetakis, I. E. & Martin, R. R. (2008). A new method for extraction of double-stranded RNA from plants. *J Virol Methods* **149**, 167–170.

Tzanetakis, I. E., Keller, K. E. & Martin, R. R. (2005). The use of reverse transcriptase for efficient first- and second-strand cDNA synthesis from single- and double-stranded RNA templates. *J Virol Methods* **124**, 73–77.

Wegener, L. A., Punja, Z. K., Martin, R. R., Bernardy, M. G. & MacDonald, L. (2006). Epidemiology and strain identification of Blueberry scorch virus on highbush blueberry in British Columbia. *Can J Plant Pathol* **28**, 250–262.

Whelan, S. & Goldman, N. (2001). A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol* **18**, 691–699.

Zerbino, D. R. & Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* **18**, 821–829.