Virus Research 265 (2019) 162–165

Contents lists available at ScienceDirect

Virus Research

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

Short communication

The complete sequences of two divergent variants of the rhabdovirus raspberry vein chlorosis virus and the design of improved primers for virus detection

Sue Jones[a], Wendy McGavin[b], Stuart MacFarlane[b,*]

[a] Information and Computational Sciences Group, The James Hutton Institute, Dundee, DD2 5DA, UK
[b] Cell and Molecular Sciences Group, The James Hutton Institute, Dundee, DD2 5DA, UK

ARTICLE INFO

Keywords:
Raspberry vein chlorosis virus
Virus detection
Rhabdovirus

ABSTRACT

The complete sequence was obtained for two variants of raspberry vein chlorosis virus (RVCV), confirming that this virus is a rhabdovirus most closely related to the cytorhabdoviruses alfalfa dwarf virus and strawberry crinkle virus. The two RVCV variants share only a 68% nucleotide sequence identity so that the previously published RT-PCR diagnostic test for this virus was not able to efficiently detect both variants. Using the new, complete sequence information several new primer sets have been designed that allow a much improved RVCV detection.

The association of vein chlorosis symptoms in raspberry with virus infection has been known for many years, where severe vein chlorosis can lead to marked stunting of cane growth (Cadman, 1952). This early paper identified the small raspberry aphid *Aphis idaei* as a vector of vein chlorosis disease but showed that the large raspberry aphid *Amphorophora idaei* (also referred to as *A. rubi.* ssp. *idaei,* and a vector of raspberry mosaic disease) did not transmit vein chlorosis disease. Since the first description of vein chlorosis disease in the UK, it has been reported in Canada, mainland Europe, Russia (U.S.S.R) and New Zealand (references described in Jennings and Jones, 1986). Later, electron microscopy work described membrane-bounded bacilliform virus-like particles of two sizes (430 × 65 nm and 125 × 30 nm) in vein chlorosis symptomatic raspberry plants (Jones et al., 1974), and similar-shaped particles of 175 × 66 nm in small raspberry aphids fed on vein chlorosis-affected raspberry plants (Murant and Roberts, 1980). Despite the apparent difference in observed particle sizes this work led to the hypothesis that raspberry vein chlorosis virus (RVCV) was likely to be a rhabdovirus. This allowed us to use degenerate rhabdovirus RT-PCR primers and an early version of deep sequencing (454) technology to obtain c. 3 kb of the RVCV sequence, leading to the design of an RVCV-specific diagnostic test (McGavin et al., 2011).

In this previous work, using RNA from a single vein chlorosis symptomatic raspberry plant (AP12) maintained in our culture collection, we identified 6 sequence contigs obtained by 454 sequencing that were related to other, published plant rhabdovirus sequences and were suspected to be from RVCV (McGavin et al., 2011). Subsequently we were able to use the remaining 454 contigs and RT-PCR amplification of intervening sequences to derive approximately 10 kb of RVCV sequence in two segments but were not able to obtain the complete sequence of this virus. In particular, this was because we found the affected plant to contain a defective form of RVCV that was missing more than 3 kb of sequence expected to encode the virus phosphoprotein, P3 (putative movement protein) and matrix protein genes. Although the AP12 plant had previously been used for successful aphid transmission and grafting experiments and so must have contained complete, infectious RVCV, using flanking primers for RT-PCR we only ever amplified sequences carrying the deletion.

Recently, we examined more plants located in Bullionfield at the James Hutton Institute that displayed vein chlorosis symptoms. However, the previously published RT-PCR test did not amplify a diagnostic RVCV fragment in all of these symptomatic plants, motivating us to investigate this virus further. RNA was extracted from symptomatic raspberry leaves and purified using a Qiagen RNeasy Plant Mini kit following the manufacturer's instructions with the addition of Ambion Plant RNA Isolation Aid at the lysis step. The purified RNA was eluted in RNAse-free water and stored at −80 °C prior to subsequent use. Four RNA samples (one from the Hutton RVCV culture collection plant (AP12), one from Bullionfield plant 126 (Moy x Latham varietal cross) collected in 2007, a second from the same plant re-sampled in 2017, and a third Bullionfield sample (plant B0, Moy x Latham cross, sampled in 2017) were mixed in equal amounts and supplied to the Glasgow Polyomics centre at the University of Glasgow, Scotland. The

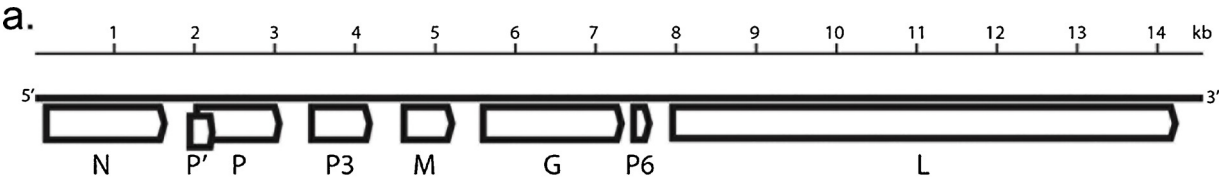
* Corresponding author.
E-mail address: Stuart.MacFarlane@hutton.ac.uk (S. MacFarlane).

https://doi.org/10.1016/j.virusres.2019.03.004
Received 12 February 2019; Received in revised form 1 March 2019; Accepted 2 March 2019
Available online 29 March 2019
0168-1702/

Table 1
Primers for co-detection of RVCV Hutt1 and RVCV Hutt2.

Forward primer	Reverse primer	Amplicon size Hutt1/Hutt2	Tm (°C)
3649-CCAACAAAGCTGATATWCCAG	3648-CCTCATCTAAGTARTCTTCCA	257/252 bp	57
3650-TCYCGACATGATTGCTCATAC	3651-GGGAACATCCTCATTATCTTY	298/298 bp	57
3652-ATMGATGCGGATATGGATGAC	3653-TCTCCTATTGCTAGAGTWGG	269/269 bp	57

a.



b.

	N	P′	P	P3	M	G	P6	L
RVCV Hutt2	75	58	74	77	73	83	81	84
SCV	43	23	32	38	30	43	42	58
ADV	42	-	39	36	35	46	40	60
TYMaV	28	-	15	23	20	33	19	45
PVA	30	-	15	17	11	21	-	41
LYMV	29	16	15	13	16	25	-	39
LNYV	29	15	16	11	18	26	-	39

c.

RVCV Hutt1 5'+	ACGAAUGAUAAAAUAAGAAAAAGAAAAAUG
RVCV Hutt2 5'+	ACGAAUGAUAAAAUAAGAAAGGAAAAAUGU
ADV 5'+	ACGAAUGAUAAAUUAAGAAAAUAAUAUGUC
SCV 5'+	ACGAAUGAUAAAGUACAAAAGAGCUUAUCU
RVCV Hutt1 5'-	ACGAAUGAUAAAAUAAGAAAAAUUAUCUGG
RVCV Hutt2 5'-	ACGAAUGAUAAAAUAAGAAAAAUUAUCUGG
ADV 5'-	ACGAAUGAUAAAUUAAGAAAAUUUUGUUGC
SCV 5'-	ACGAAUAAUAAAAUACUAGAAGUUUAUUUA

d.

Virus	termination signal	intergenic sequence
RVCV Hutt1 N/P	3'-CCAAUUUAUUUU	GAU-5'
RVCV Hutt2 N/P	3'-CCAAUUUAUUUU	GAU-5'
RVCV Hutt1 P/MP	3'-CCAAUUUAUUUU	GAU-5'
RVCV Hutt2 P/MP	3'-CCAAUUUAUUUU	GAU-5'
RVCV Hutt1 P3/M	3'-CCUAAUUAUUUU	GAU-5'
RVCV Hutt2 P3/M	3'-CCUAUUUAUUUU	GAU-5'
RVCV Hutt1 M/G	3'-CCAAUUUCUUUU	GAU-5'
RVCV Hutt2 M/G	3'-CCAAUUUCUUUU	GAU-5'
RVCV Hutt1 G/P6	3'-CCUAUUUAUUUU	GAU-5'
RVCV Hutt2 G/P6	3'-CCUAAUUAUUUU	GAU-5'
RVCV Hutt1 P6/L	3'-CCAUAUUAUUUU	GAU-5'
RVCV Hutt2 P6/L	3'-CCACAUUAUUUU	GAU-5'
RVCV Hutt1 L/t	3'-CCAAUUUCUUUU	GAU-5'
RVCV Hutt2 L/t	3'-CCAAAUUCUUUU	GAU-5'
RVCV Consensus	3'-CCAAUUUAUUUU	GAU-5'
ADV Consensus	3'-CCAAAUUAUUUU	GAU-5'
SCV Consensus	3'-CCAAAUUAUUUU	GAU-5'

Fig. 1. (a.) Genome diagram of RVCV Hutt1. Virus genes are shown as open boxes beneath the plus-strand antigenome RNA, drawn to scale. (b.) Pairwise amino acid identities of RVCV Hutt1 proteins in comparison to RVCV Hutt2 and a selection of the most closely related plant rhabdoviruses. Shaded numbers exceed the 50% cut-off for cytorhabdovirus species demarcation. (c.) Alignment of nucleotide sequences at the 5′ termini of positive-strand (+) and negative-strand (−) genome RNAs of RVCV variants, ADV and SCV. Shaded bases are different to RVCV Hutt1 5′ (+). (d.) Alignment of predicted RVCV Hutt1 and Hutt2 intergenic region sequences, on the negative-strand genomic RNA, compared to the consensus for ADV and SCV.

combined RNA sample was depleted of ribosomal RNA, converted to library form using a TruSeq kit and subjected to paired end (2 × 75bp) sequencing using an Illumina NextSeq 500 system (Table 1).

Quality control of the paired-end RNA-sequence reads was conducted using Trimmomatic (Bolger et al., 2014), which allowed the removal of Illumina sequencing adapters and low quality reads. We used sliding window trimming with parameters: LEADING:20 TRAILING:20 SLIDINGWINDOW:4:20 and a minimum trimmed length of 30 nucleic acids. To remove the host reads, trimmed reads were mapped to the genome of *Rubus occidentalis* (Black raspberry) (VanBuren et al., 2018), as the genome of red raspberry is not yet determined. The *R. occidentalis* genome (assembly v1.1) was downloaded from the Genome Database for Rosaceae (Jung et al., 2014). Mapping was conducted using Bowtie2 (Langmead and Salzberg, 2012) with score-min value "L,0,−0.2", and un-mapped reads were designated as non-host reads. The non-host reads were then de-novo assembled into contigs using Trinity (Grabherr et al., 2011) with a minimum contig length of 200 bps. A local Blast database of the contigs was then created. BlastN was used to compare the contigs with the RVCV sequence obtained previously. A Tblastn search was also made of the contigs using the 7 proteins (N,P,P3,M,G,P6,L) of the closely related rhabdovirus alfalfa dwarf virus (ADV; NC_028237.2). Contigs that aligned with an e-

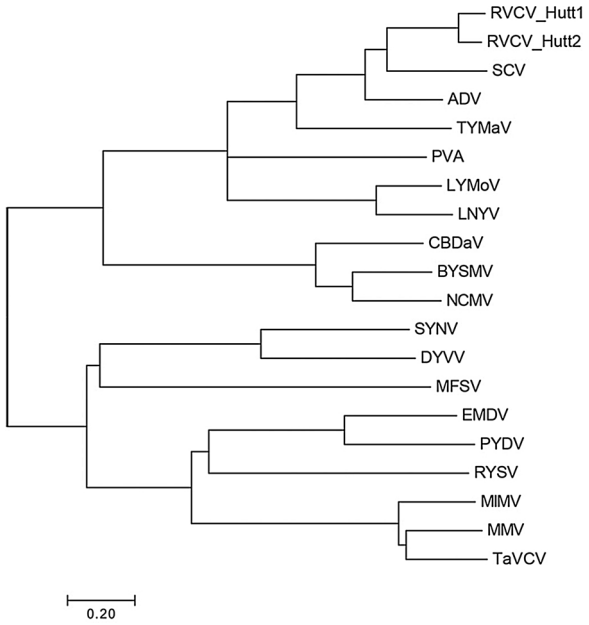


Fig. 2. Maximum likelihood phylogenetic tree comparing the complete L protein sequences of plant rhabdoviruses. The tree with the highest log likelihood is shown and is drawn to scale, with branch lengths measured in the number of substitutions per site. The phylogenetic analyses were conducted in MEGA7 (Kumar et al., 2016). Complete genome sequences used for the alignment are: MIMV (maize Iranian mosaic virus, DQ186554), TaVCV (taro vein chlorosis virus, AY674964), MMV (maize mosaic virus, AY618418), MFSV (maize fine streak virus, AY618417), PVA (persimmon virus A, NC_018381), TYMav (tomato yellow mottle associated virus, KY075646), SCV (strawberry crinkle virus, MH129615), ADV (alfalfa dwarf virus, NC_028237), LYMoV (lettuce yellow mottle virus, EF687738), LNYV (lettuce necrotic yellows virus, NC_007642), CBDaV (Colocasia bobone disease-associated rhabdovirus, NC_034551), BYSMV (barley yellow striate mosaic virus, KM213865), NCMV (northern cereal mosaic virus, NC_002251), DYVV (Datura yellow vein virus, NC_028231), SYNV (sonchus yellow net virus, L32603), EMDV (eggplant mottled dwarf virus, NC_025389), PYDV (potato yellow dwarf virus, GU734660), RYSV (rice yellow stunt virus, NC_003746).

value < 0.001 were evaluated further.

The paired-end Illumina RNA-sequencing run gave 75.9 million reads (with lengths from 35 to 75 nt and with a 42% GC content). This number was reduced to 67.5 million after quality control treatment, and further reduced to 48.6 million reads after subtraction of *R. occidentalis* genome sequences. These were assembled into 114,809 contigs of > 200 nts. A BlastN alignment identified 8 short contigs (ranging in size from 207 to 603 nts) that mapped to the c. 10 kb of RVCV sequence we had already assembled using previous sequence data and gap-filling RT-PCR. From the remaining assembled dataset we identified two long contigs (7813 nt and 6807 nt), and three shorter contigs (668 nt, 455 nt and 312 nt) that encoded proteins related to those of ADV.

Further gap-filling by RT-PCR confirmed that the three shorter contigs were located in the genome region that was missing from the previously obtained RVCV sequence, thereby enabling us to assemble a contiguous sequence of over 14 kb that contained all the genes expected for a plant rhabdovirus (Fig. 1a). Unexpectedly, the two, long sequence contigs could be joined by RT-PCR amplification of an overlapping fragment to produce a second, different c. 14.5 kb rhabdovirus sequence. Further primers were designed to enable the amplification of the entire new rhabdovirus sequence in overlapping fragments, thus confirming that the predicted sequence was correctly assembled (Supplementary Table 1). For determination of the terminal sequences of the viral genomic RNA a modified rhabdovirus partial purification procedure, based on the method described by Jackson and Wagner (1998), was used. In this modified procedure, after virus-containing material had been separated from plant debris and concentrated within a sucrose gradient, it was extracted using a Qiagen RNeasy Plant Mini kit as described above. The RNA isolated by this procedure was tested by RT-PCR to confirm the presence of the viral sequences before random amplification of cDNA ends (RACE) analysis was done using either the SMARTer RACE 5′/3′ Kit (Clontech) or by homopolymer tailing as described before (McGavin and MacFarlane, 2008).

This work revealed that the original RVCV that was previously part-sequenced has a genome of 14,512 nt whereas the newly discovered, second virus has a genome of 14,667 nt, with an overall nucleotide sequence identity of 67.9% between the two viruses (Supplementary Table 2). Pairwise comparison of the various proteins encoded by the two viruses shows amino acid identities ranging from 84 to 58% (Fig. 1b). Based on this level of relatedness the two rhabdoviruses should be considered as variants of RVCV, with the original virus now named as RVCV Hutt1 and the newly discovered virus being named as RVCV Hutt2. The RVCV Hutt1 sequence has the GenBank accession number MK240091, and RVCV Hutt2 has accession number MK257717.

We found that in the AP12 plant the deleted form of RVCV Hutt1 (missing 3430 nt) was also present and hypothesised that this shorter genome could lead to a reduced titre of both itself and full-length RVCV Hutt1 in this plant. The illumina sequence data showed that only 439 reads were derived from RVCV Hutt1 (0.0003% of total trimmed reads) whereas, 8879 Hutt2-derived reads were found (0.0066% of total). Also, the RVCV Hutt1 data had a maximum read depth of only 11 and there were many regions of the genome for which no reads were found (Supplementary Fig. 1). For RVCV Hutt2 the maximum read depth was 134 and only the extreme 5′ and 3′ terminal sequences were missing from the dataset.

Both variants have the same genome organization encoding, on the plus-sense antigenome RNA, genes for the nucleocapsid protein (N), P' (a small open reading frame [ORF] overlapping the 5′ end of the P gene), phosphoprotein (P), P3 (putative movement protein), M (matrix protein), G (glycoprotein), P6 (protein of unknown function) and L (replicase protein). Phylogenetic analysis of the L proteins from all completely sequenced plant rhabdoviruses shows that RVCV is most closely related to ADV and strawberry crinkle virus (SCV) (Fig. 2) and this relationship extends to the other viral proteins (N, P, P3, M and G; Supplementary Figs. 2 and 3). Pairwise alignments show that the L proteins of RVCV Hutt1, RVCV Hutt2, ADV and SCV share more than 50% identity, which is the species demarcation level for rhabdoviruses, although the identities of RVCV with the other ADV and SCV proteins are all below this 50% level (Fig. 1b). As with other rhabdoviruses, the terminal sequences of the genome RNAs of RVCV are inverted repeats (22 nt for RVCV Hutt1 and 20 nt for RVCV Hutt2) (Fig. 1c) (Dietzgen et al., 2017). The RVCV Hutt1 terminal sequences were identified by RACE PCR, whereas, the RVCV Hutt2 termini were amplified using primer 3579 (5′-ACGAATGATAAAATAAGAAA) derived from the RVCV Hutt1 terminal sequence. Both are identical at the extreme termini to ADV and SCV. Furthermore, rhabdoviruses carry conserved sequences in the intergenic regions (IRs), that are involved in the transcription of viral mRNAs for expression of the viral proteins (Dietzgen et al., 2017). The putative polyadenylation and intergenic non-transcribed sequences are highly conserved between all the genes of both RVCV variants, and these are almost identical to the consensus sequences described for ADV and SCV (Fig. 1d). The difference (155 nt) in genome size of the two RVCV variants results primarily from differences in their IRs, where RVCV Hutt 1 has a 123 nt ORF in the M-G IR, and RVCV Hutt 2 has 120 nt and 132 nt ORFs in the P-P3 IR and P3-M IR, respectively (Supplementary Table 1).

The previously published RVCV detection primers were designed using only c. 3 kb of the RVCV Hutt1 sequence (McGavin et al., 2011). When we used this primer pair (1531/1532) we were able to amplify a band of the expected size from the culture collection plant AP12, from three samples supplied by a commercial company (C1, C6, C8) and from extracts of *A. idaeus* aphids used for transmission studies of RVCV

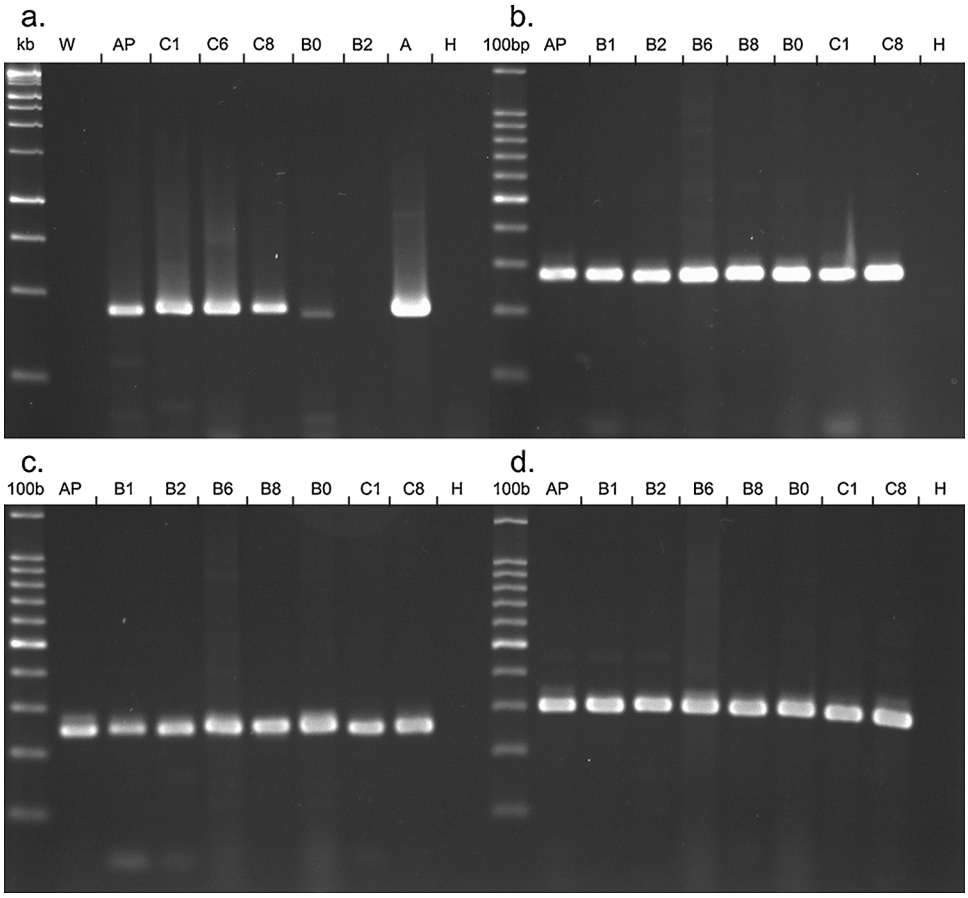


Fig. 3. RT-PCR amplification of RVCV sequences from different raspberry plant samples. (a.) original diagnostic primers 1531/1532, (b.) new degenerate primers 3648/3649 designed to detect both RVCV variants, (c.) new degenerate primers 3650/3651, (d.) new degenerate primers 3652/3653. DNA markers are (Promega) 1 kb and 100 bp. Lanes are W (water), AP (culture collection plant AP12), C1, C6, C8 (commercial samples), B0, B1, B2, B6, B8 (Bullionfield plants), A (aphid extract), H (healthy plant sample).

in AP12 (Fig. 3a; NB the aphid extract was collected in 2011 and stored frozen at −80 °C since that time). However, only a very weak band was amplified from Bullionfield sample B0 (which was included in the deep sequencing mixture) and no amplification occurred from a second Bullionfield plant (B2). New primers were designed (Sobhy and Colson, 2012) containing degeneracies to enable amplification of both RVCV Hutt1 and RVCV Hutt2 and using these primers RVCV could be detected in the AP12 plant, five different Bullionfield plants and two commercially supplied samples (Fig. 3b–d). NB sample C6 was no longer available for testing).

This work has revealed the complete genome sequence of RVCV, has identified limitations in the previously described diagnostic test and revealed the existence of a second RVCV variant with significant sequence differences to the originally described variant. Sequencing of virus variants is an essential ongoing process to enable the design of better, more inclusive diagnostic tests for virus detection.

Funding

This work was funded by the Scottish Government's Rural and Environment Science and Analytical Services division (RESAS).

Declarations of interest

None.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.virusres.2019.03.004.

References

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Ilumina sequence data. Bioinformatics 30, 2114–2120.
Cadman, C.H., 1952. Studies in Rubus virus diseases. II. Three types of vein chlorosis of raspberries. Ann. Appl. Biol. 39, 61–68.
Dietzgen, R.G., Kondo, H., Goodin, M.M., Kurath, G., Vasilakis, N., 2017. The family *Rhabdoviridae*: mono- and bi-partite negative-sense RNA viruses with diverse genome organization and common evolutionary origins. Virus Res. 227, 158–170.
Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Trinity: reconstructing a full-length transcriptome withput a genome from RNA-Seq data. Nat. Biotechnol. 29, 644–652.
Jackson, A.O., Wagner, J.D.O., 1998. Procedures for plant rhabdovirus purification, polyribosome isolation, and replicase extraction. In: In: Foster, G.D., Taylor, S.C. (Eds.), Plant Virology Protocols. Methods in Molecular Biology™, vol. 81 Humana Press.
Jennings, D.L., Jones, A.T., 1986. Immunity from raspberry vein chlorosis virus in raspberry and its potential for control of the virus through plant breeding. Ann. Appl. Biol. 108, 417–422.
Jones, A.T., Roberts, I.M., Murant, A.F., 1974. Association of different kinds of bacilliform particles with vein chlorosis and mosaic diseases of raspberry (*Rubus idaeus*). Ann. Appl. Biol. 77, 283–288.
Jung, S., Ficklin, S.P., Lee, T., Cheng, C.H., Blenda, A., Zheng, P., Yu, J., Bombarely, A., Cho, I., Ru, S., Evans, K., Peace, C., Abbott, A.G., Mueller, L.A., Olmstead, M.A., Main, D., 2014. The genome database for Rosaceae (GDR): year 10 update. Nucleic Acids Res. 42, 1237–1244.
Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874.
Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.
McGavin, W.J., MacFarlane, S.A., 2008. Rubus chlorotic mottle virus, a new sobemovirus infecting raspberry and bramble. Virus Res. 139, 10–13.
McGavin, W.J., Cock, P.J.A., MacFarlane, S.A., 2011. Partial sequence and RT-PCR diagnostic test for the plant rhabdovirus Raspberry vein chlorosis virus. Plant Pathol. 60, 462–467.
Murant, A.F., Roberts, I.M., 1980. Particles of Raspberry vein chlorosis virus in the aphid vector, *Aphis idaei*. Acta Phytopathol. Acad. Sci. Hung. 15, 103–106.
Sobhy, H., Colson, P., 2012. Gemi: PCR primers prediction from multiple alignments. Comp. Funct. Genom. 2012, 783138. https://doi.org/10.1155/2012/783138.
VanBuren, R., Wai, C.M., Colle, M., Wang, J., Sullivan, S., Bushakra, J.M., Liachko, I., Vining, K.J., Dossett, M., Finn, C.E., Jibran, R., Chagné, D., Childs, K., Edger, P.P., Mockler, T.C., Bassil, N.V., 2018. A near complete, chromosome-scale assembly of the black raspberry (*Rubus occidentalis*) genome. Gigascience 7 (8). https://doi.org/10.1093/gigascience/giy094.