Peer-reviewed version available at Viruses 2018, 10, 406; doi:10.3390/v10080406

A Virus in American Blackcurrant (*Ribes americanum*) with Distinct Genome Features Reshapes Classification in the *Tymovirales*

Thanuja Thekke-Veetil[1], Thien Ho[1], Joseph D. Postman[2], Robert R. Martin[3], Ioannis E. Tzanetakis[1,#]

[1] Department of Plant Pathology, Division of Agriculture, University of Arkansas System, Fayetteville, AR 72701, United States of America
[2] National Clonal Germplasm Repository, United States Department of Agriculture, Corvallis, OR 97333, United States of America
[3] Horticultural Crops Research Unit, United States Department of Agriculture, Corvallis, OR 97331, United States of America

[#] To whom correspondence should be addressed. Tel: 1 479 575 3180; Fax: 1 479 575 7601; Email: itzaneta@uark.edu

Abstract: A novel virus with distinct genome features was discovered by high throughput sequencing in a symptomatic blackcurrant plant. The virus tentatively named as blackcurrant virus A (BCVA) has distinct genome organization and molecular features bridging genera in the order *Tymovirales*. The genome consists of 7106 nucleotides excluding the poly(A) tail. Five open reading frames were identified with the first encoding a putative viral replicase with methyl transferase (MTR), AlkB, helicase and RNA dependent RNA polymerase (RdRp) domains. The other four putative proteins exhibit no significant homology to other virus proteins. The genome organization downstream of the replicase resembles that of members of the order *Tymovirales* with an unconventional triple gene block (TGB) movement protein arrangement. Phylogenetic analysis using replicase conserved motifs loosely placed BCVA within the *Betaflexiviridae* whereas it was evolutionarily distant to existing members of the family when using the putative TGBp 1-like and coat protein sequences. Our analysis strongly suggests that BCVA is a novel virus that should be classified as a species in a new genus in the *Betaflexiviridae* or a new family in the order *Tymovirales*.

Key words: *Betaflexiviridae;* blackcurrant; blackcurrant virus A; characterization; detection

Peer-reviewed version available at Viruses 2018, 10, 406; doi:10.3390/v10080406

1. Introduction

Technological developments in high throughput sequencing (HTS) has resulted in the rapid discovery and characterization of several novel plant viruses [4; 6; 11; 24]. The USDA National Clonal Germplasm Repository (NCGR) maintains and distributes various accessions of specialty fruit and nut crop species from around the world [26]. This includes currant (*Ribes* spp.), a berry known for its potential health effects [32]. This study initiated when an American blackcurrant maintained in the USDA-NCGR, Corvallis, Oregon (*Ribes americanum*; PI 617879) showed virus-like symptoms (Figure 1). The plant was infected with two viruses [11], a waikavirus [33] and an undescribed virus in the order *Tymovirales*.

Figure 1: Symptoms observed on American blackcurrant infected with blackcurrant virus A. Infected *Ribes americanum* cultivar Gall (PI 617879, left) showing ragged leaf margins and crinkled leaf surface compared to healthy plant (right).

The order *Tymovirales* includes four families, namely *Alphaflexiviridae*, *Betaflexiviridae*, *Gammaflexiviridae* and *Tymoviridae* [13]. Members of this group are non-enveloped, flexuous, filamentous viruses infecting primarily plants and have genomes resembling eukaryotic mRNAs [2]. This manuscript describes the characterization of a novel virus tentatively named as blackcurrant virus A (BCVA). The evolutionary relationship of BCVA with other known members of the order *Tymovirales* indicated that BCVA is a virus with unique genome features, possibly representing a new genus in the *Betaflexiviridae* family or being the type representative of a yet to be described family. Discovery of this novel virus adds information to the current knowledge on this diverse group of viruses and expands the number of virus species reported in the order.

2. Materials and Methods

Double-stranded RNAs were extracted from 20 g of symptomatic leaves from currant accession PI 617879 (Figure 1) and subjected to a degenerate oligonucleotide-primed reverse transcription-PCR (DOP RT-PCR) followed by HTS and sequence assembly as described [12].

The majority of the genome including the 3′ terminal with poly-A tail was obtained during the HTS assembly. The FirstChoice® RLM-RACE Kit (ThermoFisher Scientific, USA) was used to obtain the 5′ terminus. The 5′ and 3′ RACE-RT-PCRs were performed (an additional confirmation of the 3′ end) using the BCVA specific primers and programs described in Table S1. All PCR products were sequenced for at least three-fold coverage of the regions. The complete nucleotide sequence of BCVA was deposited in GenBank under accession number MF166685.

The genome organization, putative ORFs and protein sequences were derived using the NCBI ORF Finder [37]. RNA binding amino acid (aa) residues present in BCVA proteins were identified using Pprint (Prediction of Protein RNA- Interaction; [19]) and transmembrane domains identified using TMHMM Server v. 2.0 [18]. The secondary structures of the proteins were predicted by PSI-Pred and I-TASSER [14; 39] and nuclear localization signals identified with cNLS Mapper [16]. The protein sequences of members belonging to the families in the order *Tymovirales* and family *Closteroviridae* were obtained from GenBank (Table 1) and aligned with the BCVA putative proteins using ClustalW on Bioedit [10]. Phylogenetic analyses was performed using the maximum likelihood method with 1,000 bootstrap pseudoreplicates on MEGA v.7 [20] and phylogenetic trees displayed using Treeview in MEGA v.7.

Tissue, tested positive for the virus by RT-PCR, was examined by transmission electron microscopy as follows: Leaf and root tissues were either thin sectioned using a microtomer or macerated on a glass slide in water. Sap was applied to a Formvar coated copper EM grid and incubated for 1 min at room temperature. The excess liquid was then drawn off with a piece of filter paper and the samples were stained by adding 4µl of 2% stain to the and incubating for 1 min at room temp before the excess liquid was drawn off using a piece of filter paper. Grids were stained with either uranyl acetate or phosphotungstic acid and examined in a FEI Titan ChemiSTEM TEM.

3. Results

The genome of BCVA is 7106 nucleotides (nt), excluding the poly-A tail and initiates with the pentanucleotide motif GAAAA$_{1\text{-}5}$. The genome encodes five ORFs

with untranslated regions (UTRs) of 140 and 172 nt for the 5′ and 3′ termini, respectively (Figure 2).

ORF 1 encodes for a putative protein of 207 kDa (p207; 1826 aa) presumably involved in virus replication as it contains MTR (aa 44-334; Cdd: pfam01660), AlkB homologue of 2OG-Fe(II) oxygenase (aa 712-806; pfam13532), HEL (aa 1031-1266; pfam01443) and RNA dependent RNA polymerase (RdRp; aa 1526-1732; pfam00978) motifs. The polyprotein contained five nuclear localization signals; one monopartite signal, AVRKRLRFA$_{1436-44}$ and four bipartite signals; FAKCRQLDPENLLLSEALLVND LIKWLRE$_{328-65}$, KIRLGSKDRVIGSCKDWTTKVIKISKG$_{913-39}$, GCGKTKPLMDLILKSNDN ILILVPRKRLGDS$_{1034-64}$ and PRKRLGDSWTSKMGHKKNVRV$_{1057-78}$. The entire protein shares low aa identities with other virus replicases; the highest (23%) being the orthologs in Carrot Ch virus 1 and 2 (CtChV-1 and CtChV-2, genus *Chordovirus*). The sequence identity between the RdRp conserved domains of BCVA and members of the order range from 28-57%; the highest being with members of the genera *Trichovirus* and *Prunevirus* (*Betaflexiviridae*; 55-57 %) and the lowest to members of the *Tymoviridae*, *Alpha-* and G*ammaflexiviridae* (28-32%) (data not shown). The MTR domain also showed the higher identity to members of the *Betaflexiviridae* (24-33%) compared to the other families in the order.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 April 2018 doi:10.20944/preprints201804.0279.v1

Peer-reviewed version available at Viruses 2018, 10, 406; doi:10.3390/v10080406

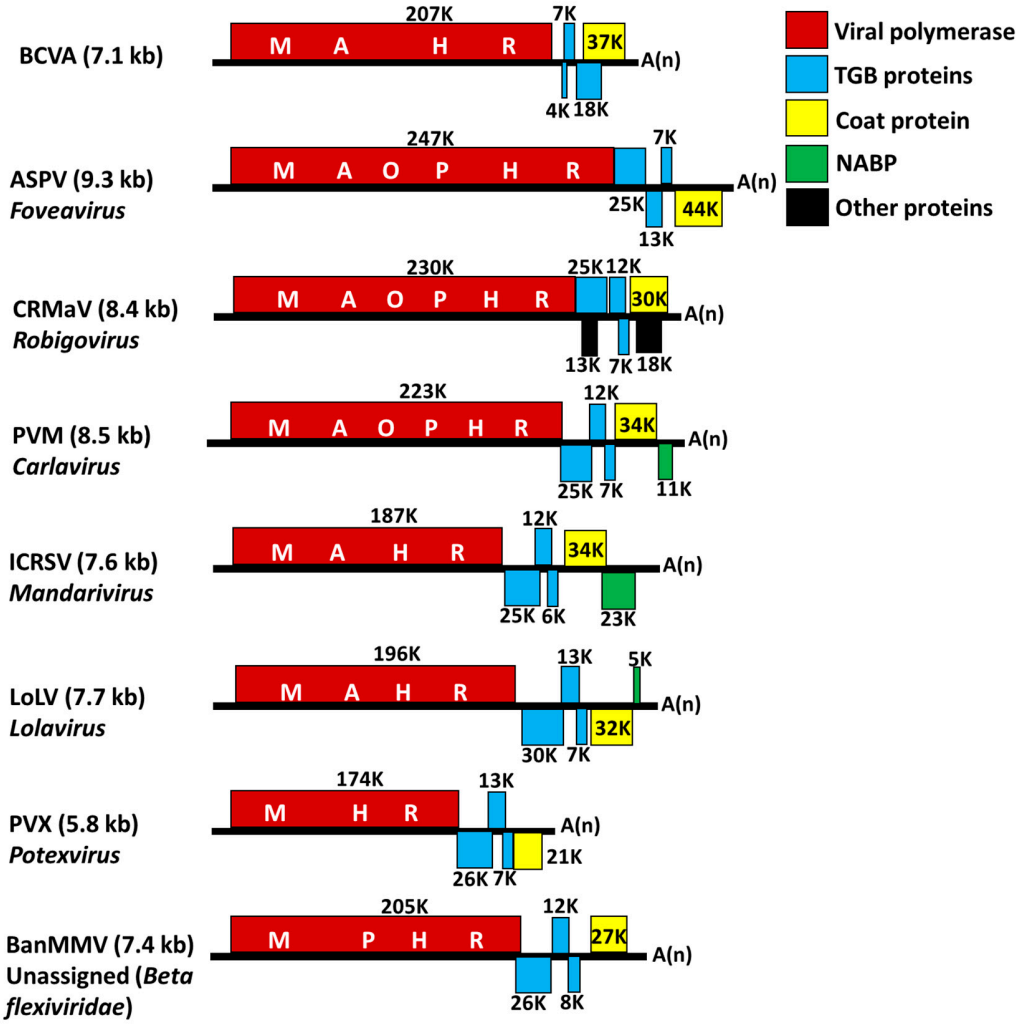


Figure 2: Genome organization of blackcurrant virus A (BCVA) in comparison to the genomes of members of *Alphaflexiviridae* and *Betaflexiviridae* with triple gene block (TGB) movement proteins. M-methyltransferase, A- AlkB, O- OTu-like peptidase, P-papain-like protease, H- helicase and R- RNA-dependent RNA polymerase. Boxes represent open reading frames. Size of the proteins encoded in the ORFs are indicated. Abbreviations: ASPV-apple stem pitting virus, CRMaV- cherry rusty mottle associated virus, PVM- potato virus M, ICRSV- indian citrus ringspot virus, LoLV- lolium latent virus, PVX- potato virus X, BanMMV- banana mild mosaic virus.

Predicted products of ORFs 2-5 showed no significant similarity to other viral proteins and for this reason HTS products were verified by Sanger sequencing performed on RT-PCR products of viral dsRNA as well as total nucleic acids extracted at different times to eliminate the possibility of contamination (Table S1). ORF 2 encodes a putative peptide of 4 kDa (p4; 39 aa) with a transmembrane domain (aa_{13-35}).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 April 2018 doi:10.20944/preprints201804.0279.v1

Peer-reviewed version available at Viruses 2018, 10, 406; doi:10.3390/v10080406

ORF 3 partially overlaps ORF 2 and encodes a peptide of 60 aa with a calculated molecular mass of 7 kDa with two predicted transmembrane domains at the N and C termini ($aa_{7-24, 34-56}$). The secondary structures of the transmembrane regions contain one and two α-helixes respectively (Figure S1). In p4, the exposed hydrophilic regions were positively charged containing four Lys residues at the N terminal and two Arg residues at the C terminal. Both hydrophilic and hydrophobic regions contained nine aa residues predicted to interact with RNA (Table 2). P7 contains only two RNA interacting aa residues and their significance in RNA binding is to be determined (Table 2).

P4

MNFKSYLLKKIKSVGIGLASSLIIYIASFVFNVLYRRSF

P7

MCYIDVAFDLVCLFICVLILVALLKLTYCNSSAFCVALALTIYSLFLNFNLLVLLYDLSR

P18

MYGYNNGIRKTSDRFSKGSVSKDKYGQRYNCGTDRLPFLVMADVSKLKIDFENATENMLF

QIVSLLLHFCVLQNIGQRKAKRGKIKKKKAAYNEYRRNKDGASSSYQGGGGLARTRDSQE

NERQVDAARDKRAEFYSDSSSTEGDGDGSGQTRNERHFV

P37

MESEKLVIVSAKVPFRRTSMAKDTTAARTDFLSSLWRMSLNSKLISKMRQRTCYSRLCHS

CCTSAYCKILARGRQREERLRRRKLPIMSTGEIRMEPLPVTREGEAWLELEIARKMKGKL

TLQETNGRNSILTVAQPKEMVMDLDRPEMRDILFNLDFTKRLIDQDVFVCSYLVKKAKRV

GVEVCTDFHCYFVDTDMTVSALLDAIEIASFFGCINSAVFEICATGSCLCKVGLRELIIE

VEKRTIEIPLKCGYHGIKHLTEVEDRQWKVLCANPLIKLEEIEEIYIFWNSLGLKNHERH

VKALLDVNGLKESTLRILGAI

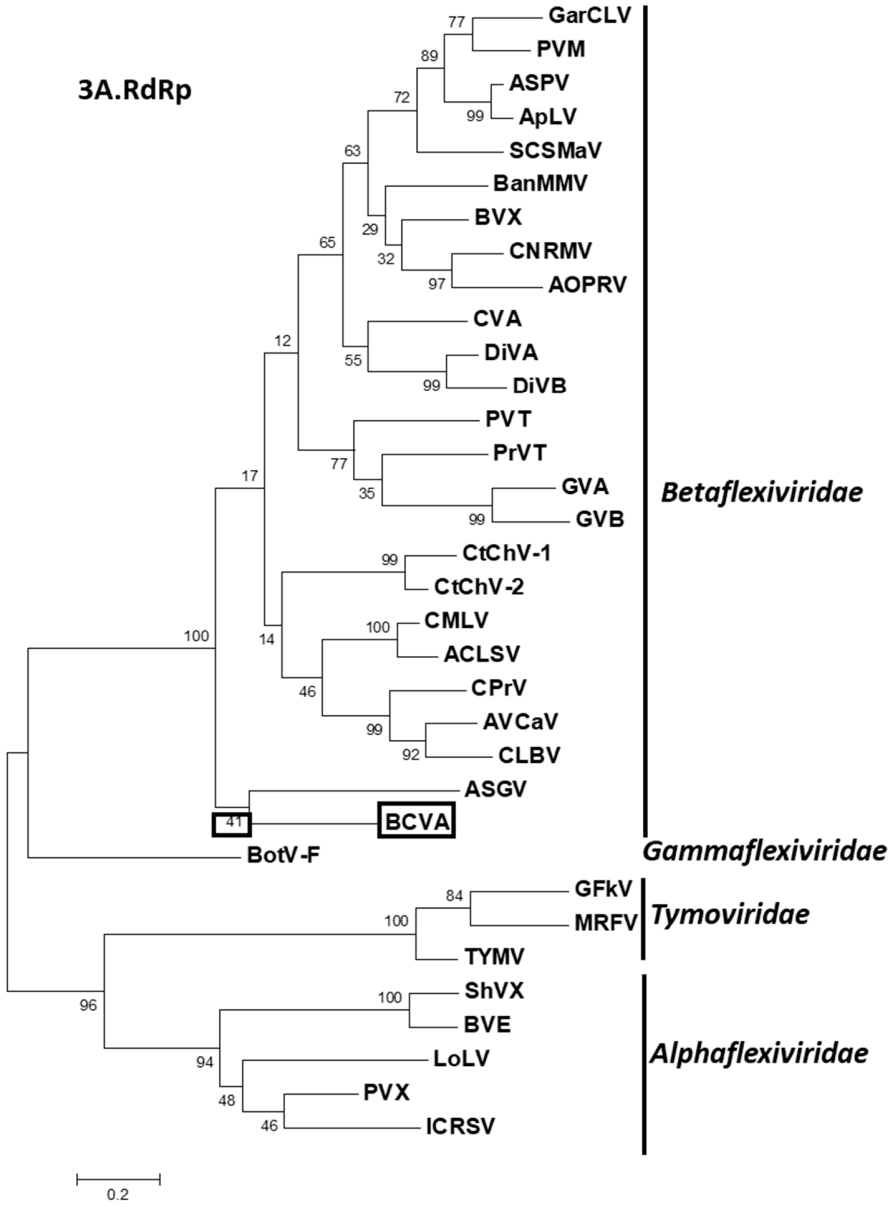
Table 2. RNA interacting amino acid residues predicted in blackcurrant virus A proteins. Amino acid residues in red color are RNA interacting while blue color are non-interacting. Shaded regions are transmembrane domain(s) in P4 and P7 while in P18 shaded region aligned with the helicase domain of ATP-dependent DNA helicase.

Separated by an intergenic region of five nt, the fourth ORF encodes a protein of 159 aa with molecular mass of 18 kDa (p18; Figure 2). This putative protein has marginal identity (24% identity, blossom 45) with a bacterial superfamily 1 (SF1) ATP dependent DNA helicase domain. The p18 secondary structure (Figure S1) containing coils, α-helixes and β-sheets with 60 % of the aa residues predicted to interact with RNA (Table 2). These aa appeared as long stretches of primarily positively charged residues (Lys and Arg) which were located in the coils present in the N- and C- termini as well as

in the α-helixes. The central region of the protein contained the longest stretch (50 RNA interacting aa) in two α-helixes which was highly rich in Lys and Arg (30 %) in a region mapped to the helicase domain (Table 2).

ORF 5 putatively encodes a 37 kDa (p37; 321 aa) protein of unknown function. This protein predicted to contain several α-helixes and coils and a single β-sheet (Figure S1). Among several RNA binding residues throughout the protein (Table 2), a stretch of 15 residues was predicted in the second and third α-helixes from the N-terminal containing seven Arg residues. P37 also has two aa residues ($Glu_{280,283}$) involved in ligand binding with calcium (Ca) (I-TASSER; [39]).

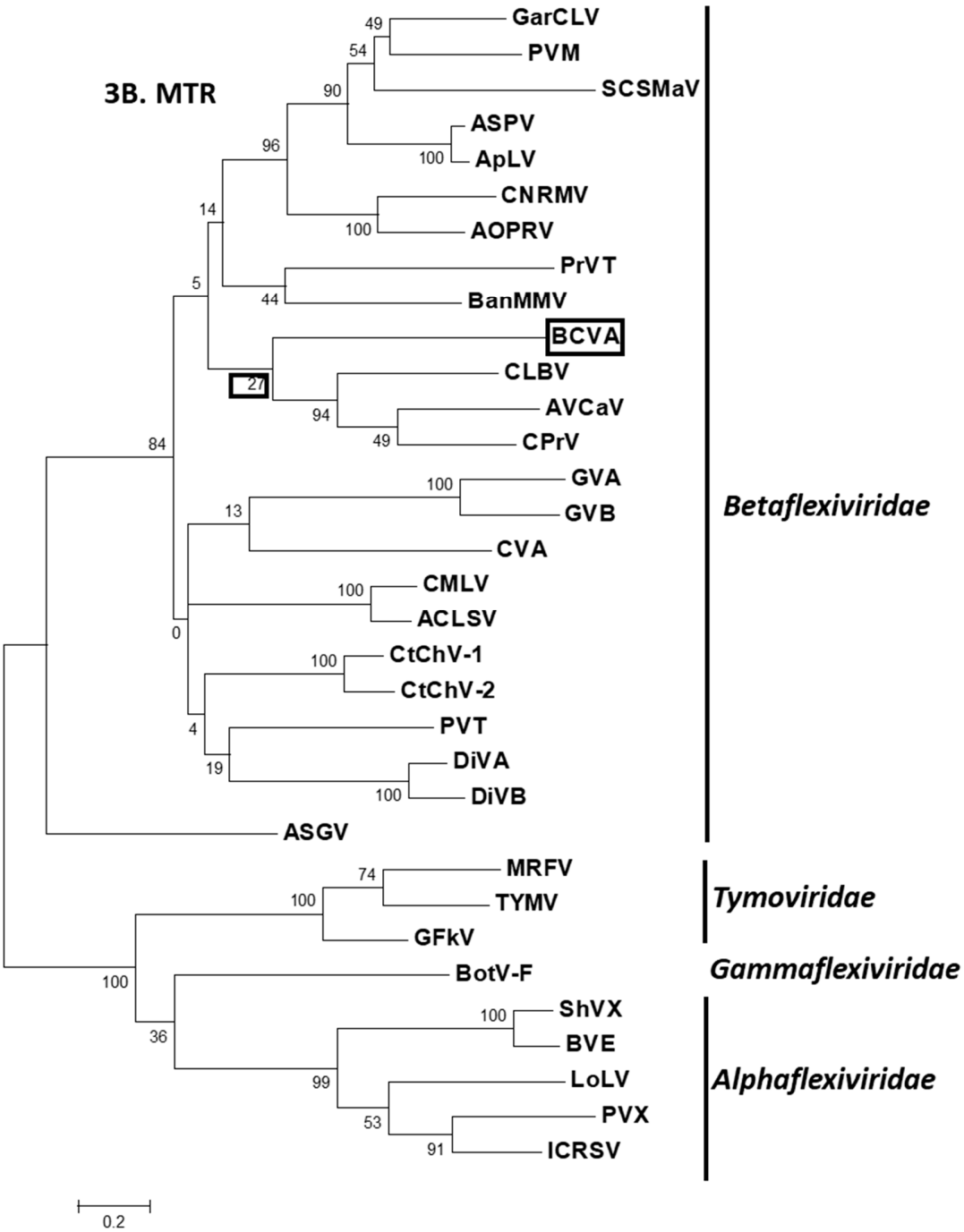
The phylogenetic analysis using the conserved domains of the RdRp and MTR (conserved in the alphavirus-like SF [15]) loosely placed BCVA with members of *Betaflexiviridae* (Figure 3A and 3B), with clustering only supported by low bootstrap values. However, the analysis on TGBp 1-like p18 and p37 capsid protein (CP) with the putative orthologs showed that BCVA is evolutionarily distant from members of the *Tymovirales* (Figure 3C and 3D). Analysis of the presumed CP included members of *Closteroviridae* as they share the structural core and evolutionary origin with tymo-members [7].



169

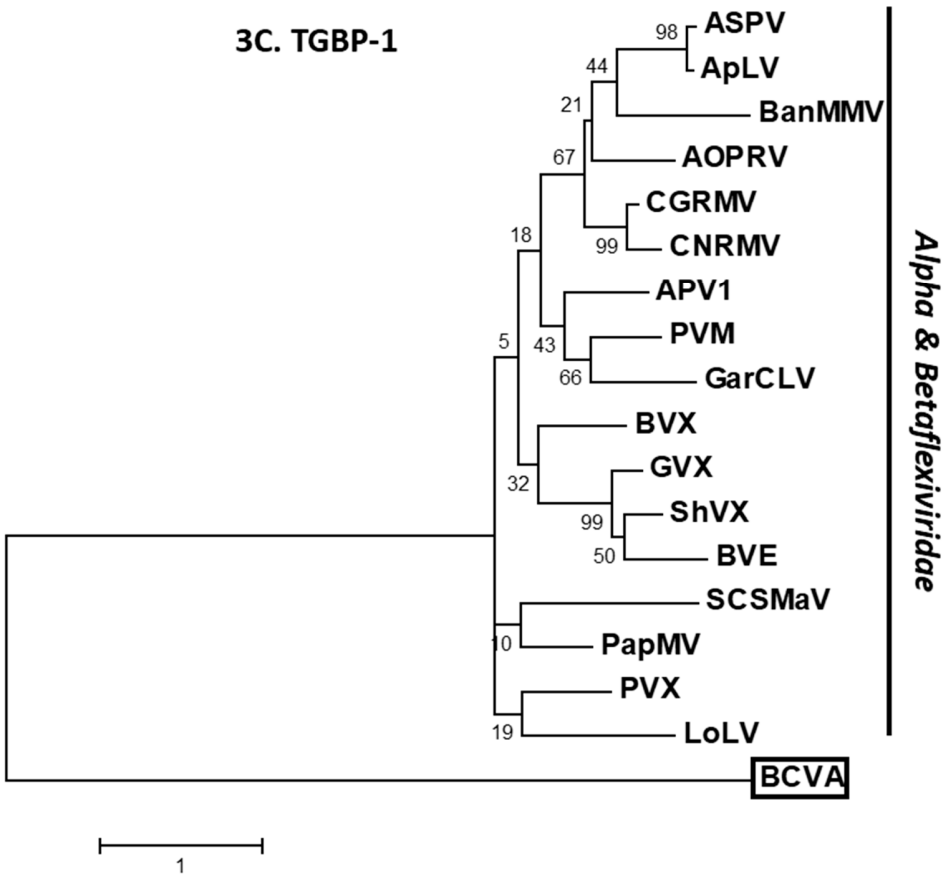
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 April 2018 doi:10.20944/preprints201804.0279.v1

Peer-reviewed version available at Viruses 2018, 10, 406; doi:10.3390/v10080406



170

171



172

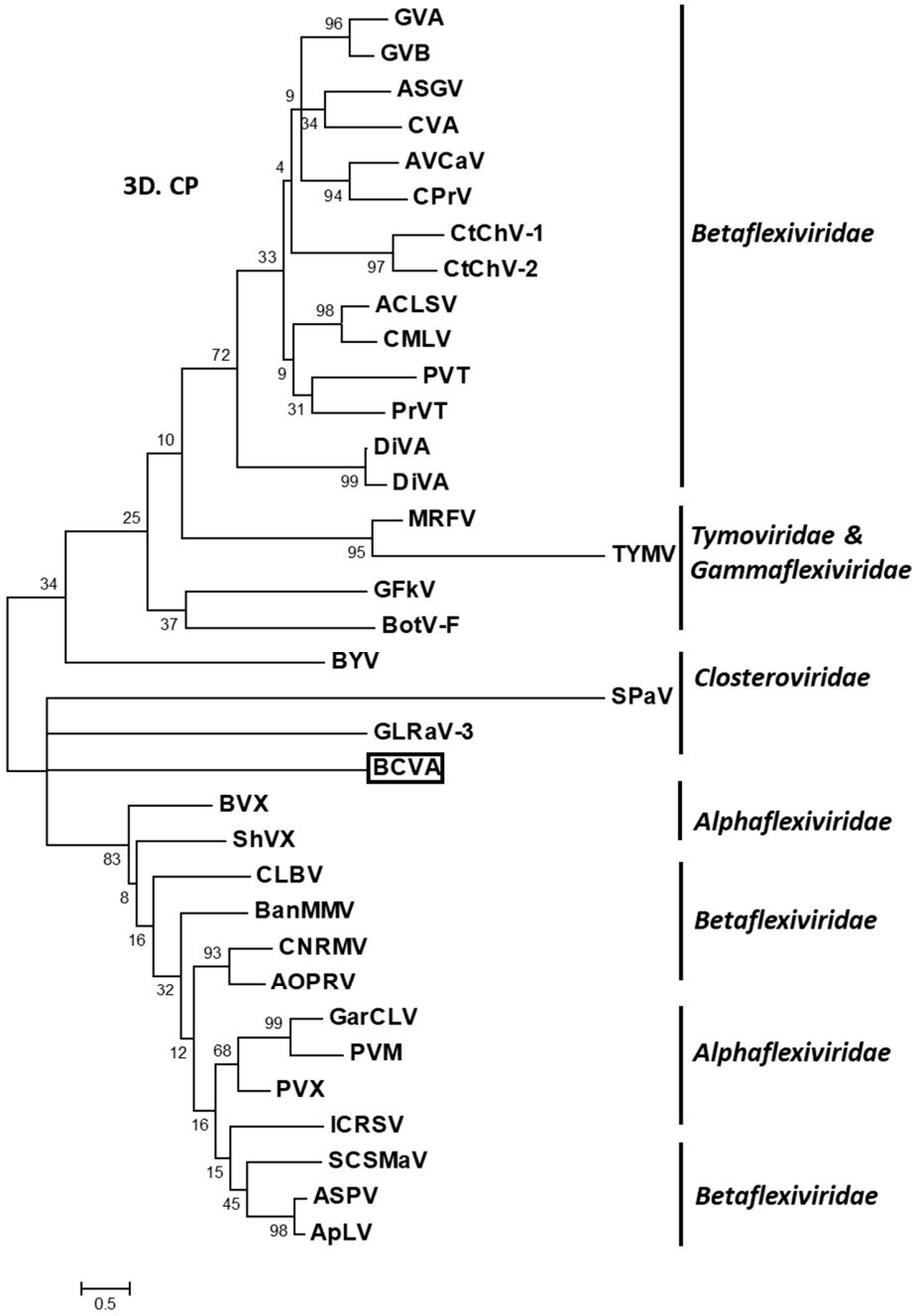


Figure 3: Phylogenetic relationship of blackcurrant virus A (BCVA) with members of the order *Tymovirales*. Coat protein (CP) analysis included representative members of *Closteroviridae*. Phylogenetic analysis was performed using the conserved domains of RNA-dependent RNA polymerase (RdRp; 3A) protein sequences, methyl transferase (MTR; 3B), triple gene block protein 1 (TGBp1; 3C) and putative CP (3D) sequences. The trees were generated by the maximum likelihood method using MEGA 7 and bootstrap

values (indicated for each branch node) were estimated using 1,000 pseudo replicates. The details of virus isolates used for the studies are provided in Table 1.

Tissue tested positive for the virus by RT-PCR was used in electron microscopy. No virus-like particles were identified nor where there any structures that belonged to an endophyte and potential host of the virus (data not shown).

4. Discussion

BCVA has features bridging genera in both the *Alphaflexiviridae* and *Betaflexiviridae* families. Because of the unique genome organization we employed electron microscopy on tissue that was tested positive for the virus to determine whether a bacterial or fungal endophyte may be the cryptic host of the virus. In the absence of any structures other than those associated with the plant cell we determined that BCVA is a plant virus. The genome initiated with the pentanucleotide sequence $GAAAA_{1\text{-}5}$, conserved in both families [21; 28; 35; 42]; thought to be essential for plus-strand RNA synthesis and protein translation [34; 38]. The genome encodes five ORFs with resemblance to the genomes of members with TGB movement proteins such as; *Potexvirus*, *Allexivirus*, *Lolavirus*, *Mandarivirus* (all belonging to *Alphaflexiviridae*) and *Foveavirus, Robigovirus, Carlavirus* and the unassigned members (in the *Betaflexiviridae*) (Figure 2). The replicase size is characteristic of a ''carlavirus-like replicase'' (>195 kDa), when compared to the "potex-like" (<195 kDa) [2; 23]. This protein contains conserved domains of MTR, HEL, AlkB and RdRp domains that are present in members of both families. The BCVA replicase showed the highest aa identity to its conserved orthologous domains (MTR and RdRp) of members of the *Betaflexiviridae*. The incongruence in the taxonomical grouping of replicase domains, p18 TGBp 1-like and p37 CP-like proteins is yet common among members of this group [22; 42].

Despite the similarity in the 5' part of the genome coding for the replicase, the BCVA genome differed significantly from genomes of members of the *Alphaflexiviridae* and *Betaflexiviridae* (Figure 2). First and foremost, the putative BCVA MP (p4, p7 and p18) and CP (p37) do not share any detectable similarity to any known viral proteins. TGB movement proteins are of two types; 'hordei-like'(rod shaped viruses of *Hordeivirus*, *Benyvirus*, *Pomovirus* and *Pecluvirus*) and potex-like (members of *Tymovirales*) and are usually overlapping [25], with the exception of TGBps reported in the genus *Robigovirus* and an unassigned member, sugarcane striate mosaic-associated virus, in which only two, TGBp 2 and 3, are overlapping.

The organization of TGB ORFs is highly conserved among plant viruses [25; 27]. Potex-like TGB proteins occurs in the order of a bigger sized protein, TGBp 1 (~25 kDa) followed by TGBp 2 and TGBp 3, smaller proteins (~12 kDa and ~7 kDa respectively; Figure 2) in which TGBp 2 has two and TGBp 3 has single transmembrane domains

bordered by charged residues. TGBp 2 and TGBp 3 are membrane proteins and localize on endomembranes and cell walls [23; 25]. TGBp 1 of these viruses usually harbors a NTP binding helicase domain of SFI and has ATPase, RNA binding and RNA helicase activity and is believed to increase the size exclusion limit of plasmodesmata [23]. Based on protein size and characteristics BCVA seems to have a potex-like TGB proteins, however the arrangement of the putative TGB ORFs (ORFs 2-4) is reversed. The RNA binding site of potexviral TGBp1 conserved positively charged residue (Arg), which is essential for interaction with RNA [25] as with the case of BCVA p18 (Table 2). The size of these three proteins were significantly different from other members (Figure 2). Although there was no significant sequence identity with its putative orthologs, the presence of transmembrane domains and the NTP binding helicase domain in p18 makes us hypothesize that these three proteins constitute a TGB movement protein block for BCVA.

ORFs downstream of the viral replicase are thought to be translated from 3'-terminal subgenomic RNAs in which TGB proteins get translated through leaky scanning mechanisms in all families of the order other than *Tymoviridae* [25]. Sequence context around the initiation codon affects the translation efficiency of mRNAs. The consensus optimal context for translation initiation in mammals is GCC(A/G)CC**AUG**G (the initiation codon in bold; [17]). In some cases, viruses also follow this rule [30]. The Kozak context of translation initiation was observed in BCVA ORF 5 although its initiation codon is 15 nt downstream of the overlapping ORF 4. This is expected if the protein is expressed given that the ribosomes first identify the start codon of p18 and p37 initiation context need to be optimal if any ribosomes are to bypass p18.

The last ORF of most *Alpha-* and *Betaflexiviridae* members is the CP (Figure 2). However, p37 of BCVA lacked conserved domains of the CP of flexuous viruses. The predicted RNA binding region of the protein is rich in Arg. Arg-rich RNA binding motif in the CP is known to involve in genome binding and subsequent packaging[5]. The p37 also has aa residues thought to be involved in Ca binding. Calcium plays an important role in the assembly and disassembly and/or replication processes in viruses [1; 8; 9]. The size of the protein is similar to the range shown by members that have TGB MP (28-45 kDa; Figure 2). The position in the genome, predicted size, Arg-rich RNA binding region and presence of Ca binding sites suggests that p37 is the BCVA coat protein.

The genomes of the members of *Tymovirales* are highly diverse with respect to the number and organization of genes indicative of major role of recombination in the evolution of this virus group. There is significant diversity in the 3' region downstream of viral replicase (30k/TGB MP, CP and NABP) suggestive of recombination events in which the ancestral viruses had a common 5' part while acquiring the 3' genes from different origins which further diversified in time [22]. Several studies also have

255 described evidence of recombination in the members of this group [3; 31; 36; 40; 41].
256 Martelli et al. [23] suggested coevolution of CP and MP in this virus group. BCVA
257 genome follows this theory in which the replicase, yet with low bootstrap values,
258 groups with the *Betaflexiviridae* whereas the remaining proteins are phylogenetically
259 distant. Changes in phylogenetic relationships for different regions of the genome of a
260 virus is an indication of recombination [29]. BCVA polymerase is similar to *Trivirinae*
261 subfamily but the rest of the genome is *Quinvirinae*- like (*Carlavirus*, *Foveavirus*,
262 *Robigovirus;* Figure 2) in which the TGBp- like proteins' organization is reverse which
263 suggests the role of recombination between these two subfamilies in the evolution of
264 this novel virus.

265 **5. Conclusion**

266 To summarize, BCVA is a unique, previously undescribed virus of tymo-like
267 lineage which cannot be assigned to any of the currently recognized taxa of plant
268 viruses. With regard to the molecular features; number of genes encoded in the genome,
269 organization, presence of conserved pentanucleotide at 5' termini, BCVA resembles
270 members of the order *Tymovirales*. Phylogenetic analysis based on conserved RdRp
271 domain, the principal determinant in the evolutionary framework of positive-strand
272 RNA viruses, and MTR domain, clearly qualify BCVA as a putative species in the
273 *Betaflexiviridae* family. However, differences in genome portions downstream of
274 polymerase (organization of TGBp-like proteins and lack of conserved domains of
275 ortholog proteins) and the distant phylogenetic relationship with other members
276 distinguishes it from both existing genera and unassigned members of the family,
277 suggesting a tentative classification. Therefore, we propose that BCVA represents a
278 new, monotypic genus in the family *Betaflexiviridae* or family (given the low replicase aa
279 identities with existing viruses) for which the names *Curravirus*/*Curraviridae* is
280 proposed.

281 **Supplementary Materials:** The following are available online, Figure S1: Predicted
282 secondary structures (PSI-Pred; [14]) of the blackcurrant virus A proteins. Table S1:
283 Primers and PCR conditions used for the amplification of the blackcurrant virus A
284 genome.
285 **Acknowledgments:** The authors acknowledge the financial support of USDA awards
286 14-8130-0392-CA and Hatch project 1002361

287 **Author Contributions:** Thanuja Thekke-Veetil (TT), Thien Ho (TH), Ioannis E. Tzanetakis
288 conceived and designed the experiments; TT and TH performed the experiments; TT and
289 TH analyzed the data; TT wrote the paper; Joseph D. Postman and Robert R. Martin did
290 the electron microscopy work. All authors approved the final version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Abdel-Meguid, S.; Yamane, S. T.; Fukuyama, K.; Rossmann, M. G. The location of calcium ions in southern bean mosaic virus. Virology 1981, 114, 81-85.
2. Adams, M.J.; Antoniw, J.F.; Bar-Joseph, M.; Brunt, A.A.; Candresse, T.; Foster, G.D.; Martelli, G.P.; Milne, R.G.; Fauquet, C.M. The new plant virus family *Flexiviridae* and assessment of molecular criteria for species demarcation. Arch. Virol. 2004, 149, 1045-1060.
3. Alabi, O.J.; Al Rwahnih, M.; Mekuria, T.A.; Naidu, R.A. Genetic diversity of grapevine virus A in Washington and California vineyards. Phytopathology 2014, 104, 548-560.
4. Barba, M.; Czosnek, H.; Hadidi, A. Historical perspective, development and applications of next-generation sequencing in plant virology. Viruses 2014, 6, 106-136.
5. Choi, Y.G.; Rao, A.L.N. Molecular studies on bromovirus capsid protein: VII. Selective packaging of BMV RNA4 by specific N-terminal arginine residues. Virology 2000, 275, 207-217.
6. Di Bello, P.L.; Laney, A.G.; Druciarek, T.; Ho, T.; Gergerich, R.C.; Keller, K.E.; Martin, R.R.; Tzanetakis, I.E. A novel emaravirus is associated with redbud yellow ringspot disease. Virus Res. 2016, 222, 41-47.
7. Dolja, V.V.; Boyko, V.P.; Agranovsky, A.A.; Koonin, E.V. Phylogeny of capsid proteins of rod-shaped and filamentous RNA plant viruses: two families with distinct patterns of sequence and probably structure conservation. Virology 1991, 184, 79-86.
8. Durham, A. C.; Hendry, D. A. Cation binding by tobacco mosaic virus. Virology 1977, 77, 510-519.
9. Gajardo, R.; Vende, P.; Poncet, D.; Cohen, J. Two proline residues are essential in the calcium-binding activity of rotavirus VP7 outer capsid protein. J. Virol. 1997, 71, 2211-2216.
10. Hall, T.A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 1999, 41, 95-98.
11. Ho, T.; Tzanetakis, I. E. Developing a virus detection and discovery pipeline using next generation sequencing. Virology 2014, 471-473, 54-60.
12. Ho, T.; Martin, R.R.; Tzanetakis, I.E. Next-generation sequencing of elite berry germplasm and data analysis using a bioinformatics pipeline for virus detection and discovery. In: Lacomme, C. (Ed.), Plant Pathology: Techniques and Protocols

(Methods in Molecular Biology), 1302. Springer, New York, USA, 2015, pp. 301-313.

13. ICTV, 2016. Virus Taxonomy 2016 Release, Order *Tymovirales*, https://talk.ictvonline.org/taxonomy/
14. Jones, D.T. Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. 1999, 292, 195-202.
15. Koonin, E.V.; Dolja.; V.V. Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. Crit. Rev. Biochem. Mol. Biol. 1993, 28, 375-430.
16. Kosugi, S.; Hasebe, M.; Tomita, M.; Yanagawa, H. Systematic identification of yeast cell cycle-dependent nucleocytoplasmic shuttling proteins by prediction of composite motifs. Proc. Natl. Acad. Sci. USA 2009, 106,10171-10176.
17. Kozak, M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 1986, 44, 283-292.
18. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L. L. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J Mol. Biol. 2001, 305, 567-580.
19. Kumar, M.; Gromiha, M.M.; Raghava, G.P.S. Prediction of RNA binding sites in a protein using SVM and PSSM profile. Proteins 2008, 71, 189-194.
20. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 2016, 33, 1870-1874.
21. Li, R.; Mock, R. Characterization of a flowering cherry strain of cherry necrotic rusty mottle virus. Arch. Virol. 2008, 153, 973-978.
22. Marais, A.; Faure, C.; Mustafayev, E.; Candresse, T. Characterization of new isolates of apricot vein clearing-associated virus and of a new *Prunus*-Infecting virus: Evidence for recombination as a driving force in *Betaflexiviridae* evolution. PloS one 2015, 10, p.e0129469.
23. Martelli, G. P.; Adams, M. J.; Kreuze, J. F.; Dolja, V. V. Family *Flexiviridae*: a case study in virion and genome plasticity. Annu. Rev. Phytopathol. 2007, 45, 73-100.
24. Massart, S.; Olmos, A.; Jijakli, H.; Candresse, T. Current impact and future directions of high throughput sequencing in plant virus diagnostics. Virus Res. 2014, 188, pp. 90-96.
25. Morozov, S. Y.; Solovyev, A. G. Triple gene block: modular design of a multifunctional machine for plant virus movement. J. Gen. Virol. 2003, 84, 1351-1366.
26. Postman, J.; Hummer, K.; Stover, E.; Krueger, R.; Forsline, P.; Grauke, L.J.; Zee, F.; Ayala-Silva, T.; Irish, B. Fruit and nut genebanks in the US national plant germplasm system. HortSci. 2006, 41, 1188-1194.

27. Rupasov, V. V.; Morozov, S. Yu.; Kanyuka, K. V.; Zavriev, S. K. Partial nucleotide sequence of potato virus M RNA shows similarities to potexviruses in gene arrangement and the encoded amino acid sequences. J. Gen. Virol. 1989, 70, 1861-1869.
28. Sabanadzovic, S.; Abou Ghanem-Sabanadzovic, N.; Tzanetakis, I.E. Blackberry virus E: an unusual flexivirus. Arch. Virol. 2011, 156, 1665-1669.
29. Salminen, M. Detecting recombination in viral sequences. In: Salemi, M, Vandamme, A.M (eds) The phylogenetic handbook: a practical approach to DNA and protein phylogeny. University Press, Cambridge, UK, 2003, pp 348–377.
30. Sánchez-Puig, J.M.; Blasco, R. AUG context and mRNA translation in vaccinia virus. Span. J. Agric. Res. 2008, 6, 73-80.
31. Singh, L.; Hallan, V.; Martin, D.P.; Ram, R.; Zaidi, A.A. Genomic sequence analysis of four new chrysanthemum virus B isolates: evidence of RNA recombination. Arch. Virol. 2012, 157, 531-537.
32. Terry, L. Health-promoting properties of fruit and vegetables, 2014, CABI, p. 432.
33. Thekke-Veetil, T.; Ho, T.; Postman, J.D.; Tzanetakis, I.E. Molecular characterization of a new member of the genus *Waikavirus.* Phytopathology (in press).
34. Verchot-Lubicz, J.; Ye, C.M.; Bamunusinghe, D. Molecular biology of potexviruses: recent advances. J. Gen. Virol. 2007, 88, 1643-1655.
35. Villamor, D. V.; Druffel, K. L.; Eastwell, K. C. Complete nucleotide sequence of a virus associated with rusty mottle disease of sweet cherry (*Prunus avium*). Arch. Virol. 2013, 158, 1805-1810.
36. Villamor, D.E.; Eastwell, K.C. Viruses associated with rusty mottle and twisted leaf diseases of sweet cherry are distinct species. Phytopathology 2013, 103, 1287-1295.
37. Wheeler, D. L.; Church, D.M.; Federhen, S.; Lash, A.E.; Madden, T.L.; Pontius, J.U.; Schuler, G.D.; Schriml, L.M.; Sequeira, E.; Tatusova, T.A.; Wagner, L. Database resources of the National Center for Biotechnology, Nucl. Acids Res. . 2003, 31, 28-33.
38. Wong, S.M.; Mahtani, P.H.; Lee, K.C.; Yu, H.H.; Tan, Y.; Neo, K.K.; Chan, Y.; Wu, M.; Chng, C.G. Cymbidium mosaic potexvirus RNA: complete nucleotide sequence and phylogenetic analysis. Arch. Virol. 1997, 142, 383-391.
39. Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. The I-TASSER Suite: Protein structure and function prediction. Nat. Methods 2015, 12, 7-8.
40. Yoon, J.Y.; Joa, J.H.; Choi, K.S.; Do, K.S.; Lim, H.C.; Chung, B.M. Genetic diversity of a natural population of apple stem pitting virus isolated from apple in Korea. Plant Pathol. J. 2014, 30, 195-199.

405 41. Zanardo, L.G.; Silva, F.N.; Lima, A.T.M.; Milanesi, D.F.; Castilho-Urquiza, G.P.;
406 Almeida, A.M.R. Molecular variability of cowpea mild mottle virus infecting
407 soybean in Brazil. Arch. Virol. 2014, 159, 727-737.
408 42. Zhang, Y.P.; Kirkpatrick, B.C.; Smart, C.D.; Uyemoto, J.K. cDNA cloning and
409 molecular characterization of cherry green ring mottle virus. J. Gen. Virol. 1998,
410 79, 2275-2281.
411

Table 1. Details of virus isolates used for the phylogenetic studies of blackcurrant virus A. RdRp-RNA dependent RNA polymerase, MTR-Methyl transferase, TGB-Triple gene block, *α-Alphaflexiviridae*, β-*Betaflexiviridae*, *γ -Gammaflexiviridae*, T-*Tymoviridae*, C- *Closteroviridae.*

Virus Acronyms	Virus Names	Genus (Family)	Genbank Accession Numbers		
			RdRp/ MTR	TGB Protein 1	Coat Protein
ACLSV	Apple chlorotic leaf spot virus	*Trichovirus* (β)	NP_040551.1	-	CAA68083.1
CMLV	Cherry mottle leaf virus	*Trichovirus* (β)	AOY07780.1	-	AOY07782.1
GVA	Grapevine virus A	*Vitivirus* (β)	AFV73358.1	-	AAL76173.1
GVB	Grapevine virus B	*Vitivirus* (β)	AIL90366.1	-	AAL40797.1
CNRMV	Cherry necrotic rusty mottle virus	*Robigovirus* (β)	NP_059937.1	ALP45967.1	BAU25796.1
AOPRV	African oil palm ringspot virus	*Robigovirus* (β)	YP_002776347.1	YP_002776348.1	AAL87671.2
CGRMV	Cherry green ring mottle virus	*Robigovirus* (β)	-	ALP45844.1	-
CtChV-1	Carrot Ch virus 1	*Chordovirus* (β)	AHA85534.1	-	YP_009104001.1
CtChV-2	Carrot Ch virus 2	*Chordovirus* (β)	AHA85531.1	-	YP_009103998.1
DiVA	Diuris virus A	*Divavirus* (β)	YP_006905850.1	-	YP_006908998.1
DiVB	Diuris virus B	*Divavirus* (β)	AFV57240.1	-	YP_006908993.1
AVCaV	Apricot vein clearing associated virus	*Prunevirus* (β)	AKN09002.1	-	AKN09014.1
CPrV	Caucasus prunus virus	*Prunevirus* (β)	AKN08994.1	-	AKN08996.1
PVT	Potato virus T	*Tepovirus* (β)	AFU55321.1	-	AFU55323.1
PrVT	Prunus virus T	*Tepovirus* (β)	YP_009051684.1	-	YP_009051686.1
ASGV	Apple stem grooving virus	*Capillovirus* (β)	APT42870.1	-	AAL40796.1
CVA	Cherry virus A	*Capillovirus* (β)	AMH87272.1	-	BAD22827.1
GarCLV	Garlic common latent virus	*Carlavirus* (β)	AGG13282.1	AEV51823.1	AGZ03888.1
PVM	Potato virus M	*Carlavirus* (β)	AHL30493.1	AGU27045.1	ADH52722.1
ASPV	Apple stem pitting virus	*Foveavirus* (β)	NP_604464.1	AEP02961.1	NP_604468.1
ApLV	Apricot latent virus	*Foveavirus* (β)	YP_004089619.1	YP_004089620.1	ADM43445.1
APV 1	Asian prunus virus 1	*Foveavirus (β)*	-	ABB42818.1	-

CLBV	Citrus leaf blotch virus	*Citrivirus* (β)	NP_624333.1	-	ALZ41789.1
BanMMV	Banana mild mosaic virus	Unassigned (β)	NP_112029.1	NP_112030.1	AAR11507.1
SCSMaV	Sugarcane striate mosaic-associated virus	Unassigned (β)	NP_624313.1	NP_624314.1	NP_624317.1
BVX	Banana virus X	Unassigned (β)	AAW50958.1	AAW50959.1	AAW50962.1
BotV-F	Botrytis virus F	*Mycoflexivirus* (γ)	NP_068549.1	-	NP_068550.1
ShVX	Shallot virus X	*Allexivirus* (α)	NP_620648.1	Q04581.1	ACV88262.1
GarVX	Garlic virus X	*Allexivirus* (α)	-	AFV61768.1	-
ICRSV	Indian citrus ringspot virus	*Mandarivirus* (α)	AAK97522.1	-	AAO72987.1
PVX	Potato virus X	*Potexvirus* (α)	BAE07083.1	BAJ12052.1	AAV27212.1
PapMV	Papaya mosaic virus	*Potexvirus (α)*	-	NP_044331.1	-
LoLV	Lolium latent virus	*Lolavirus* (α)	YP_001718499.1	YP_001718500.1	-
BVE	Blackberry virus E	Unassigned (α)	AEI17897.1	YP_004659201.1	-
GFkV	Grapevine fleck virus	*Maculavirus* (T)	NP_542612.1	-	AEK12260.1
MRFV	Maize rayado fino virus	*Marafivirus* (T)	NP_115454.1	-	AIY22520.1
TYMV	Turnip yellow mosaic virus	*Tymovirus* (T)	AMH40140.1	-	AAA46590.1
BYV	Beet yellows virus	*Closterovirus* (C)	-	-	AAA72955.1
SPaV	Strawberry pallidosis-associated virus	*Crinivirus* (C)	-	-	YP_025089.1
GLRaV-3	Grapevine leafroll-associated virus 3	*Ampelovirus* (C)	-	-	ANZ03340.1