Analysis of public domain plant transcriptomes expands the phylogenetic diversity of the family *Secoviridae*

V. Kavi Sidharthan
Institute of Forest Biodiversity, Hyderabad, India
V. Rajeswari
Tamil Nadu Agricultural University
V.K.Baranwal (vbaranwal2001@yahoo.com)
Indian Agricultural Research Institute, New Delhi, India

Short Report

Keywords: Secovirus, novel, public domain, database

Posted Date: April 25th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1588061/v1

Abstract

Secoviruses are mono-/bi-partite plant-infecting RNA viruses that cause economically important diseases in plants. In the present study, nine secoviruses tentatively named as Ananas comosus secovirus (AcSV), Artocarpus altilis secovirus (AaSV), Boehmeria nivea secovirus (BnSV), Gynostemma pentaphyllum secovirus (GpSV), Orobanche cernua secovirus (OcSV), Paris polyphylla secovirus 1 (PpSV1), Paris polyphylla secovirus 2 (PpSV2), Rhododendron delavayi secovirus (RdSV) and Yucca gloriosa secovirus (YgSV) were identified by analysis of publicly available transcriptomes of eight plant species. Coding-complete genome/genome segments of all the identified viruses encoding a polyprotein were recovered. Two of the identified viruses- AcSV and GpSV were discovered in few of the small RNA libraries of respective plant species. Putative cleavage sites were predicted in polyproteins encoded by AcSV, GpSV, PpSV2 and YgSV genome segments. Phylogenetic and sequence similarity analyses showed that AcSV, GpSV and YgSV, PpSV1 and RdSV putatively belong to the genera- *Sadwavirus* (sub genus: *Cholivirus*), *Fabavirus*, *Nepovirus* and *Waikavirus*, respectively while AaSV, BnSV and PpSV2 may represent a distinct group of viruses within the family *Secoviridae* as they could not certainly be assigned to a particular genera.

Full Text

The family *Secoviridae* contains non-enveloped icosahedral viruses that infect plants and many cause economically important diseases (Thompson et al., 2017). The positive-sense single stranded RNA genomes of secovirids are either monopartite or bipartite and are typically 9–13.7 kb in size (size of combined RNAs in case of bipartite viruses). Mostly, the genome or genome segment contains a single large Open Reading Frame (ORF) coding for a polyprotein that will be cleaved into individual proteins by 3C-like proteinases encoded by the virus. In general, secoviruses are transmitted by insects or nematodes while few members are also transmitted by seeds (Thompson et al., 2017). Currently, the family *Secoviridae* includes nine genera- *Comovirus*, *Fabavirus*, *Nepovirus*, *Cheravirus*, *Sadwavirus*, *Torradovirus*, *Sequivirus*, *Waikavirus* and *Stralarivirus* (Thompson et al., 2017; Dullemans et al., 2019) and the genus *Sadwavirus* includes three subgenera- *Satsumavirus*, *Stramovirus* and *Cholivirus* (Sanfaçon et al., 2020).

Owing to the steady increase in transcriptome sequencing projects, huge amount of data is being generated and deposited in public domain Sequence Read Archive (SRA) and Transcriptome Shotgun Assembly (TSA) databases of National Centre for Biotechnology Information (NCBI). Besides host sequences, the transcriptome data might also contain viral sequences, if the host harboured viruses during the time of sample collection. Thus, these databases act as valuable resources for comprehensive discovery of known/novel viral sequences from a wide range of hosts that would otherwise require costlier viromic studies (Bejerman et al., 2021; Sidharthan et al., 2022). Such viruses that are discovered from metagenomic datasets can be considered as *bona fide* ones and be included in International Committee on Taxonomy of Viruses (ICTV) taxonomy (Simmonds et al., 2017). We speculated that plant transcriptomes available in public domain might contain secoviral sequences and their discovery might expand the secoviral phylogenetic diversity. Thus, in the present study, we mined public domain plant transcriptomes for secoviral sequences and identified at least eight putative novel viruses.

For identification of secoviral RNA1 sequences in plant transcriptomes (Viridiplantae, taxid: 33090), tBLASTn analysis (word size: 6; e-value: 0.05; matrix: BLOSUM62) was performed using RNA1 encoded polyprotein sequences of cowpea mosaic virus (NP613283.1) and rice tungro spherical virus (NP042507.1) as queries against TSA database (accessed on March, 2022). Resulting contigs of more than 5 kb were only considered and manually examined. If more than one redundant contig was obtained from the same sample, the longest intact contig was considered as the putative RNA1 segment of the identified virus. Genome segments of same viruses identified from different samples/study were regarded as different isolates of the same virus. To obtain the second genomic segment of each putative novel virus, RNA2 encoded polyprotein sequence of respective closely related virus was used as query in tBLASTn searches and the longest intact non-redundant contig was regarded as the putative RNA2 segment. In cases where the obtained viral contigs were partial/truncated, coding-complete genome/genome segments were obtained by rnaviralSPAdes (v 3.15.4) assembly (Prjibelski et al., 2020) of datasets from where the truncated/partial contigs were derived, after trimming using Trimmomatic (v 0.36) (Bolger et al., 2014), followed by BLASTn (v 2.10.1) (Cock et al., 2015) analysis of the assembled contigs against the recovered partial/truncated viral contigs in Galaxy Australia server (https://usegalaxy.org.au/) (Afgan et al., 2016). ORFs in putative genome segments of identified viruses were predicted using NCBI ORF finder (https://www.ncbi.nlm.nih.gov/orffinder/). Molecular weight (MW) prediction and motif search was performed as described in Sidharthan and Baranwal (2021). To identify maximum sequence identity of protein sequences encoded by the recovered viral genome segments with the existing viral sequences at maximum query coverage (qcov), BLASTp analysis against NCBI 'non-redundant (nr)' database was performed. Putative cleavage sites in the polyprotein sequence encoded by identified viral genome segments were predicted by multiple sequence alignment with the polyprotein sequences of related viruses. For phylogenetic analysis, the conserved protease-polymerase (Pro-Pol) amino acid (aa) sequences of secoviruses that were retrieved from NCBI virus database (https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/) along with the corresponding sequences of identified viruses were subjected to MUSCLE alignment and phylogenetic tree construction using neighbourhood-joining (NJ) method with Poisson model and 1000 bootstrap replicates. Conserved protein domains in the identified viral sequences were visulaized using WebLogo 3 (https://weblogo.berkeley.edu/) (Crooks et al., 2004). To identify virus-positive small RNA (sRNA) libraries, BLASTn analysis (word size: 28; e-value: 0.05) of recovered viral genome sequences was performed against the available sRNA libraries of respective plant species and libraries containing at least 10 reads were considered virus-positive.

A total of nine viruses tentatively named as Ananas comosus secovirus (AcSV), Artocarpus altilis secovirus (AaSV), Boehmeria nivea secovirus (BnSV), Gynostemma pentaphyllum secovirus (GpSV), Orobanche cernua secovirus (OcSV), Paris polyphylla secovirus 1 (PpSV1), Paris polyphylla secovirus 2 (PpSV2), Rhododendron delavayi secovirus (RdSV) and Yucca gloriosa secovirus (YgSV) were identified in transcriptomes of *Ananas comosus*, *Artocarpus altilis*, *Boehmeria nivea*, *Gynostemma pentaphyllum*, *Orobanche cernua* var. *cumana, Paris polyphylla var. yunnanensis*, *Rhododendron delavayi* and *Yucca gloriosa*, respectively through tBLASTn searches. Of the identified viruses, except for RdSV that contained a single genome segment, two genome segments were recovered from the respective transcriptome libraries of the same plant sample. In total, nineteen genomic segments of the identified viruses were recovered including those of two different isolates of PpSV2 (Table 1). Genome/genome segments of all the identified viruses contained a single large ORF coding for a polyprotein (Table S1).

RNA1 of GpSV (5.7 kb) and YgSV (6.3 kb) coded for polyprotein 1 of MW 205 kDa and 221 kDa, respectively, that contained RNA helicase (HEL), 3C cysteine protease (picornain 3C) and viral RNA-dependent RNA polymerase (RdRP) motifs while polyprotein 2 encoded by RNA2 of GpSV (3.1 kb) and YgSV (5.5 kb) with MW 112 kDa and 183 kDa, respectively, contained large and small coat protein (CP) motifs. Besides, an additional movement protein (MP) motif was predicted in polyprotein 2 of GpSV (Fig. 1, Table S1). Putative cleavage sites of GpSV polyprotein 1 are Q(297)/F, Q(887)/S, Q(913)/S and Q(1122)/S while those of polyprotein 2 are Q(401)/A and Q(803)/S. Cleavage sites predicted for YgSV polyprotein 1 are Q(301)/F, Q(892)/N, Q(921)/N and Q(1125)/N and YgSV polyprotein 2 are Q(1021)/N and Q(1432)/G (Fig. 1). BLAST analysis of polyprotein sequences encoded by the genomic segments showed that GpSV shared maximum amino acid sequence similarities (75.97% at 99% qcov and 67.23% at 98% qcov for polyprotein 1 and 2, respectively) with cucurbit mild mosaic virus (CuMMV) while YgSV shared maximum sequence similarities (36.27% at 89% qcov and 24.92% at 37% qcov for polyprotein 1 and 2, respectively) with broad bean wilt virus 2 (BBWV2) (Table S1). Phylogenetic analysis grouped GpSV with CuMMV and YgSV with Prunus virus F (PrVF) (Fig. 2).

AcSV RNA1 (6.4 kb) and 2 (3.9 kb) encoded polyprotein 1 and 2, of MW 215 kDa and 127 kDa, respectively. Polyprotein 1 contained HEL, picornain 3C and RdRP domains while polyprotein 2 contained nepovirus CP domain (Fig. 1, Table S1). Proteolytic cleavage sites predicted in polyprotein 1 are Q(419)/G, Q(929)/A, Q(953)/S and Q(1179)/S while a cleavage site S(319)/G was predicted in polyprotein 2 (Fig. 1). AcSV polyprotein 1 and 2 shared a maximum of 47.70% (97% qcov) and 27.23% (99% qcov) amino acid sequence similarities with dioscorea mosaic associated virus (DMaV) and chocolate lily virus A (CLVA), respectively in BLAST analysis (Table S1). Phylogenetic analysis grouped AcSV alongwith CLVA and DMaV (Fig. 2).

OcSV RNA1 (6.9 kb) encoded for a 255 kDa polyprotein 1 with HEL and RdRP domains while the OcSV RNA2 (3.5 kb) encoded for a 117 kDa polyprotein 2 with viral MP domain (Fig. 1, Table S1). BLAST analysis revealed the maximum amino acid sequence identities of polyproteins 1 and 2 (76.61% at 99% qcov and 72.72% at 100% qcov, respectively) with Trillium govanianum cheravirus (TgCV) (Table S1). In phylogenetic analysis, OcSV and TgCV were clustered in a single sub-clade that fell apart from other cheraviruses (Fig. 2).

RNA1 (6.5 kb) and 2 (6.6 kb) of PpSV1 encoded for 232 kDa polyprotein 1 and 186 kDa polyprotein 2 with HEL, RdRP and nepovirus CP N-terminal, central, C-terminal domains, respectively (Fig. 1, Table S1). PpSV1 polyprotein 1 shared a maximum of 66.16% (99% qcov) amino acid sequence similarity with blackcurrant reversion virus (BRV) while polyprotein 2 shared 41.11% (65% qcov) sequence similarity with tomato ringspot virus (Table S1). Phylogenetic analysis grouped PpSV1 with BRV (Fig. 2).

The single large genomic segment of RdSV (12.4 kb) coded for a 422 kDa polyprotein with Waikavirus CP1, HEL, tungro spherical virus-type peptidase and RdRP domains (Fig. 1, Table S1). BLAST analysis of polyprotein sequence of RdSV showed its maximum amino acid sequence similarity of 34.71% (89% qcov) with poaceae liege virus 1 (Table S1). Phylogenetic analysis placed RdSV in a distinct sub-clade to other waikaviruses (Fig. 2).

RNA 1 of AaSV and BnSV were of length 5.8 kb and 5.9 kb, respectively while RNA 2 of AaSV and BnSV were 4.9 kb and 3.2 kb long, respectively. Genome segments of two PpSV2 isolates (designated as fruit and seed) were recovered in the present study. RNA 1 and 2 of PpSV2 isolate fruit were of length 5.6 kb and 3.2 kb, respectively while RNA 1 and 2 of PpSV2 isolate seed were 5.7 kb and 3.3 kb, respectively. RNA1-encoded polyprotein 1 of AaSV with MW 204 kDa contained HEL and RdRP motifs while that of BnSV and PpSV2 isolates with MW 213 kDa and 204 kDa, respectively contained HEL, picornain 3C and RdRP motifs. RNA2-encoded polyprotein 2 of AaSV with MW 168 kDa contained large CP domain while that of BnSV and PpSV2 isolates with MW 107 kDa and 110 kDa, respectively contained large and small CP domains (Fig. 1, Table S1). Cleavage sites predicted in polyprotein 1 of PpSV2 isolate fruit are Q(300)/W, Q(875)/S, Q(903)/Y and Q(1126)/C while those predicted in polyprotein 2 are Q(442)/A and Q(814)/T. Similarly, the predicted cleavage sites in polyprotein 1 of PpSV2 isolate seed are Q(301)/W, Q(876)/S, Q(904)/Y and Q(1127)/H and in polyprotein 2 are Q(445)/A and Q(817)/T (Fig. 1). BLAST analysis revealed that AaSV-encoded polyproteins shared 27.48% to 32.57% (at 62% to 97% qcov) amino acid sequence similarities with the corresponding sequences of Capsicum annuum fabavirus and a comoviral sequence, BnSV-encoded proteins shared 23.98% to 32.94% (at 96% to 99% qcov) amino acid sequence similarities with bean rugose mosaic virus and BBWV2 and PpSV2-encoded proteins of both the isolates shared 26.57% to 40.69% (at 95% to 99% qcov) amino acid sequence similarities with PrVF and a comoviral sequence (Table S1). Phylogenetic analysis grouped PpSV2 isolates with BnSV and both these viruses fell in a distinct clade from other comoviruses and fabaviruses while AaSV formed a distinct clade among the members of the family *Secoviridae* (Fig. 2).

The conserved GxxGxGKS motif that is found in plant picorna-like virus NTP-binding protein and three conserved domains found in picorna-like virus RdRP (Reddick et al., 1997) were determined in polyprotein of RdWV and in polyprotein 1 of all the identified bipartite viruses (Fig. S1). Plants produce small RNA (sRNA) from viral replicative forms/intermediates in response to viral infection as part of antiviral defence strategy (Pooggin, 2018). Interestingly, AcSV and GpSV reads were identified in few of the sRNA libraries of respective plant species in which the viruses were originally identified (Table S2). Also, BLASTp analysis showed that PpSV1 RNA1-encoded polyprotein sequence shared 95.80% amino acid sequence similarity (87% qcov) with one of the partial secoviral sequences available in GenBank with accession number MZ679335.1 suggesting that the sequence is of PpSV1. Further, family-level assignment of the identified viral sequences was confirmed through the GRAViTy pipeline (http://gravity.cvr.gla.ac.uk) (Aiewsakun and Simmonds, 2018).

Based on species demarcation criteria for the family *Secoviridae* (<80% and <75% amino acid sequence identities for the conserved Pro-Pol and CP, respectively) (Thompson et al., 2017), genome organization, predicted motifs, sequence similarities and phylogeny, GpSV and YgSV can be regarded as novel members of the genus *Fabavirus*, AcSV as a novel sadwavirus (sub genus: *Cholivirus*), RdSV as a novel waikavirus, PpSV1 as a novel nepovirus and AaSV, BnSV, PpSV2 as novel members of the family *Secoviridae*. It is worthy of note that two diverse secoviruses-PpSV1 and PpSV2 were identified in the same host-*P. polyphylla var. yunnanensis* wherein a stralarivirus- cnidium vein yellowing virus 2 was also identified (data not shown). On the other hand, YgSV was identified in the same sample where a novel potexvirus- Yucca gloriosa virus 1 was identified (Bejerman and Debat, 2021). Similarly, RdSV was identified in the sample wherein a novel betanucleorhabdovirus (Rhododendron delavayi virus 1) and a novel betaflexivirus (Rhododendron delavayi virus 2) were earlier identified (Bejerman et al., 2021; Bejerman and Debat, 2021).

Though OcSV shared 89.2% amino acid sequence identity with TgCV Pro-Pol sequence, OcSV CP polyprotein sequence shared <75% amino acid sequence similarity with the corresponding sequence of TgCV. Also, TgCV was identified in *Trillium govanianum* (Sidharthan et al., 2022) while OcSV was identified in *O.*

cernua var. *cumana* in the present study. Further studies on biological properties of both these viruses are needed to ascertain OcSV as a novel member of the genus *Cheravirus*. The present report will facilitate further studies on understanding the biological properties of the identified viruses.

Declarations

Acknowledgements

We thank Galaxy Australia, a service provided by the Australian Biocommons and its partners. The authors are grateful to the Head, Division of Genetics and Tree Improvement, the GCR, the Director, IFB (ICFRE), Hyderabad, the Head, Division of Plant Pathology and the Director, ICAR-IARI, New Delhi for their support. V. Kavi Sidharthan thanks ICFRE for the HRD training on Bioinformatics and Computational Biology. The authors are also thankful to the original submitters of datasets/contigs explored/used in the study.

Author contributions

V. Kavi Sidharthan: conceptualization, methodology, formal analysis and investigation, writing- original draft preparation; V. Rajeswari: formal analysis and investigation, writing- review and editing; V. K. Baranwal: conceptualization, resources, supervision, writing- review and editing.

Data availability

The viral genome sequences described in the study are available at NCBI Third Party Annotation database with accession numbers BK061318–BK061336.

Compliance with ethical standards

Conflict of interest

The authors declare that they have no conficts of interest.

Research involving human participants and/or animals

This work does not contain any animal or human participants.

Informed consent

This work does not contain any animal or human participants.

References

1. Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Čech M, et al. (2016) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Res. 44: W3–W10. doi: 10.1093/nar/gkw343
2. Aiewsakun P, Simmonds P (2018) The genomic underpinnings of eukaryotic virus taxonomy: creating a sequence-based framework for family-level virus classification. Microbiome 6(1): 1–24
3. Bejerman N, Debat H (2021) Exploring the tymovirids landscape through metatranscriptomics data. *bioRxiv*
4. Bejerman N, Dietzgen RG, Debat H (2021) Illuminating the plant rhabdovirus landscape through metatranscriptomics data. Viruses 13:1304
5. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120
6. Christ B, Xu, C, Xu M, Li FS, Wada N, Mitchell AJ … Weng JK (2019) Repeated evolution of cytochrome P450-mediated spiroketal steroid biosynthesis in plants. Nat commun 10(1): 1–11
7. Cock PJ, Chilton JM, Grüning B, Johnson JE, Soranzo N (2015) NCBI BLAST + integrated into Galaxy. Gigascience 4(1): s13742-015
8. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. Genome Res 14:1188–1190
9. Jiang Z, Zhao Q, Bai R, Yu R, Diao P, Yan T … Wuriyanghan H (2021) Host sunflower-induced silencing of parasitism-related genes confers resistance to invading *Orobanche cumana*. Plant Physiol 185(2): 424–440
10. Laricchia KM, Johnson MG, Ragone D, Williams EW, Zerega NJ, Wickett NJ (2018) A transcriptome screen for positive selection in domesticated breadfruit and its wild relatives (*Artocarpus spp.*). Am J Bot 105(5): 915–926
11. Liao D, Zhu J, Zhang M, Li XE, Sun P, Wei J, Qi J (2019) Comparative transcriptomic analysis reveals genes regulating the germination of morphophysiologically dormant *Paris polyphylla* seeds during a warm stratification. Plos one 14(2): e0212514
12. Ling LZ, Zhang SD, Zhao F, Yang JL, Song WH, Guan SM … Cheng L (2017) Transcriptome-wide identification and prediction of miRNAs and their targets in *Paris polyphylla* var. yunnanensis by high-throughput sequencing analysis. Int J Mol sci 18(1): 219
13. Liu CH, Liu Y (2017) Fruit quality and differentially expressed genes of winter-harvested pineapple in response to elevated temperature over a short postharvest period. Postharvest Biol Technol 130: 21–27.
14. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A (2020) Using SPAdes de novo assembler. Curr Protoc Bioinform 70(1):e102
15. Reddick BB, Habera LF, Law MD (1997) Nucleotide sequence and taxonomy of maize chlorotic dwarf virus within the family Sequiviridae. J Gen Virol 78(5):1165–1174

16. Sanfaçon H, Dasgupta I, Fuchs M, Karasev AV, Petrzik K, Thompson JR, …Yoshikawa N (2020) Proposed revision of the family Secoviridae taxonomy to create three subgenera,"Satsumavirus","Stramovirus" and "Cholivirus", in the genus Sadwavirus. Arch Virol 165(2): 527–533
17. Sidharthan VK, Baranwal VK (2021) Mining of the water hyssop (*Bacopa monnieri*) transcriptome reveals genome sequences of two putative novel rhabdoviruses and a solendovirus. Arch Virol 166(7): 1985–1990
18. Sidharthan VK, Nagendran K, Baranwal VK (2022) Exploration of plant transcriptomes reveals five putative novel poleroviruses and an enamovirus. Virus Genes 1–10
19. Simmonds P, Adams MJ, Benkő M, Breitbart M, Brister JR, Carstens EB, Davison AJ, Delwart E, Gorbalenya AE, Harrach B, Hull R (2017) Consensus statement: virus taxonomy in the age of metagenomics. Nat Rev Microbiol 15:161–168
20. Thompson JR, Dasgupta I, Fuchs M, Iwanami T, Karasev AV, Petrzik K, … Consortium IR (2017) ICTV virus taxonomy profile: Secoviridae. J Gen Virol 98(4): 529
21. Wang Y, Zeng Z, Li F *et al.* (2019) A genomic resource derived from the integration of genome sequences, expressed transcripts and genetic markers in ramie. BMC Genom 20: 476. https://doi.org/10.1186/s12864-019-5878-8
22. Yang Q, Liu S, Han X, Ma J, Deng W, Wang X … Xia X (2019) Integrated transcriptome and miRNA analysis uncovers molecular regulators of aerial stem-to-rhizome transition in the medical herb *Gynostemma pentaphyllum*. BMC Genom 20(1):1–15.
23. Zhang L, Xu P, Cai Y, Ma L, Li S, Li S … Wang J (2017) The draft genome assembly of R*hododendron delavayi* Franch. var. delavayi Gigascience, *6*(10): gix076.
24. Zinkgraf M, Gerttula S, Groover A (2017) Transcript profiling of a novel plant meristem, the monocot cambium. J Integr Plant Biol 59(6): 436–449

Table 1

Table 1. Summary of identification of novel secoviruses from public domain plant transcriptomes

Virus	Segment	Length (nt)	Virus Isolate	Library/contig from where the genome was derived	Plant Species	Plant Cultivar/ Isolate/Genotype	Tissue	Reference	TPA Accessio
AcSV	RNA 1	6405	Yellow Mauritius	SRR5127122	*Ananas comosus*	Yellow Mauritius	Pulp	Liu and Liu, 2017	BK0613
	RNA 2	3863							BK0613
AaSV	RNA 1	5812	USA	GGGH01067645.1	*Artocarpus altilis*	Karawa, Kea, Meisaip, Meisei, Puupuu, Samoan 2, Toneno	Leaf, perianth	Laricchia et al., 2018	BK0613
	RNA 2	4924		GGGH01066960.1					BK0613
BnSV	RNA 1	5939	Zhongzhu 1	GHEV01013589.1	*Boehmeria nivea*	Zhongzhu 1	Leaves, stems, roots, flowers	Wang et al., 2019	BK0613
	RNA 2	3192		GHEV01028031.1					BK0613
GpSV	RNA 1	5704	China	SRR9265253	*Gynostemma pentaphyllum*	na	Stem	Yang et al., 2019	BK0613
	RNA 2	3107		GHVI01071603.1					BK0613
OcSV	RNA 1	6890	China	GIQJ01019157.1	*Orobanche cernua* var. *cumana*	na	Tubercle, shoot	Jiang et al., 2021	BK0613
	RNA 2	3506		GIQJ01006599.1					BK0613
PpSV1	RNA 1	6467	Dian chonglou	GHCK01017149.1	*Paris polyphylla* var. *yunnanensis*	Dian chonglou	Root, stem, leaf, seed	Liao et al., 2019	BK0613
	RNA 2	6559		GHCK01132374.1					BK0613
PpSV2	RNA 1	5595	Fruit	GHNI01193515.1		na	Fruit	Christ et al., 2019	BK0613
	RNA 2	3224		GHNI01193508.1					BK0613
	RNA 1	5711	Seed	GFOY01008760.1		na	Seed	Ling et al., 2017	BK0613
	RNA 2	3290		GFOY01024272.1					BK0613
RdSV	RNA	12434	Kun Ming	SRR5121284	*Rhododendron delavayi*	Kun Ming; Yunnan	Flower	Zhang et al., 2017	BK0613
YgSV	RNA 1	6309	VFLL015B	GFHP01009638.1	*Yucca gloriosa*	VFLL015B	Cambium and recent derivatives	Zinkgraf et al., 2017	BK0613
	RNA 2	5477		GFHP01009639.1					BK0613

na: Information not available

Figures

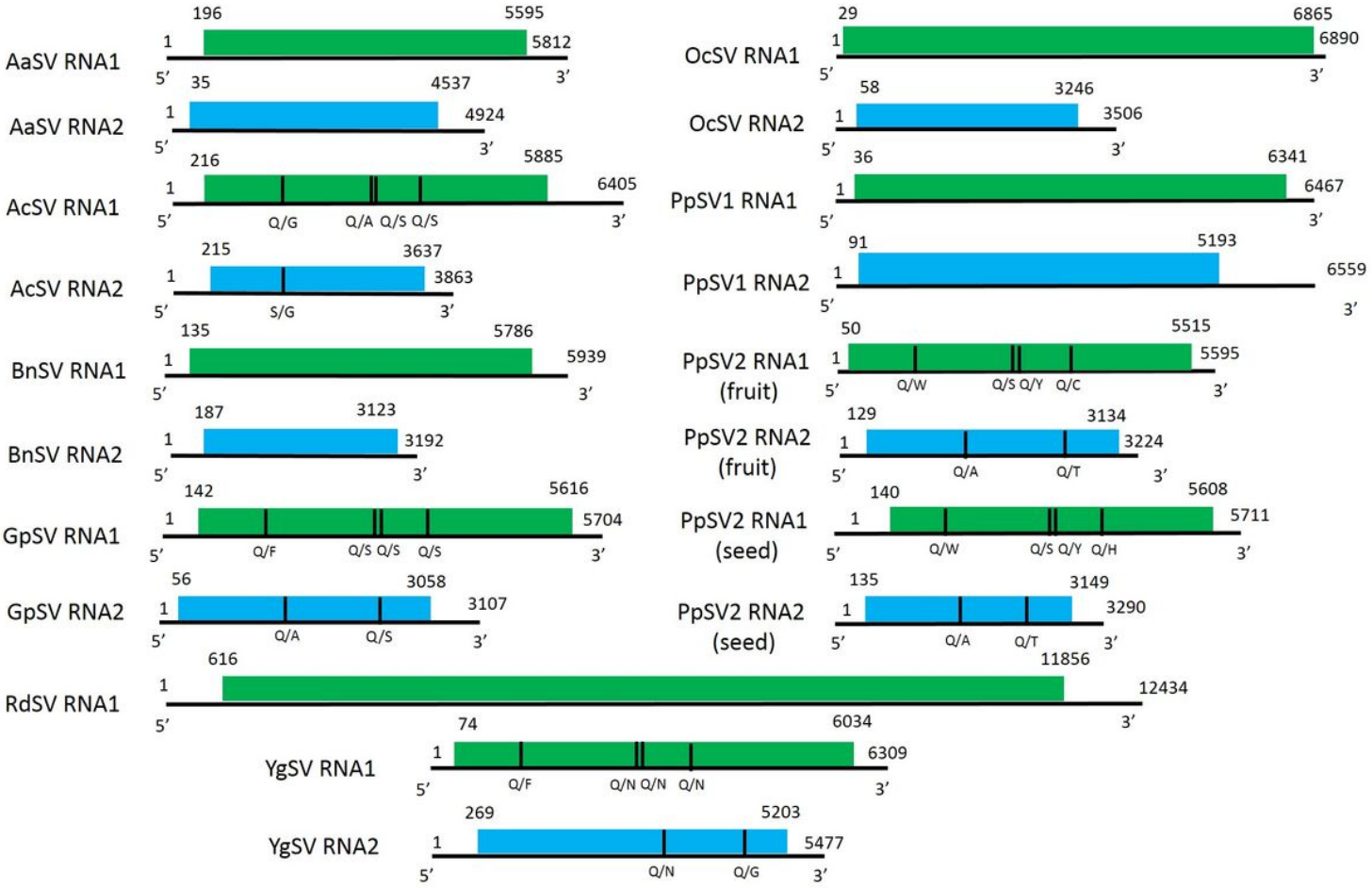

Figure 1

Genome organization of secoviruses identified in the present study. Each ORF is indicated by a different color. Cleavage sites predicted in polyproteins of few of the identified viruses are indicated

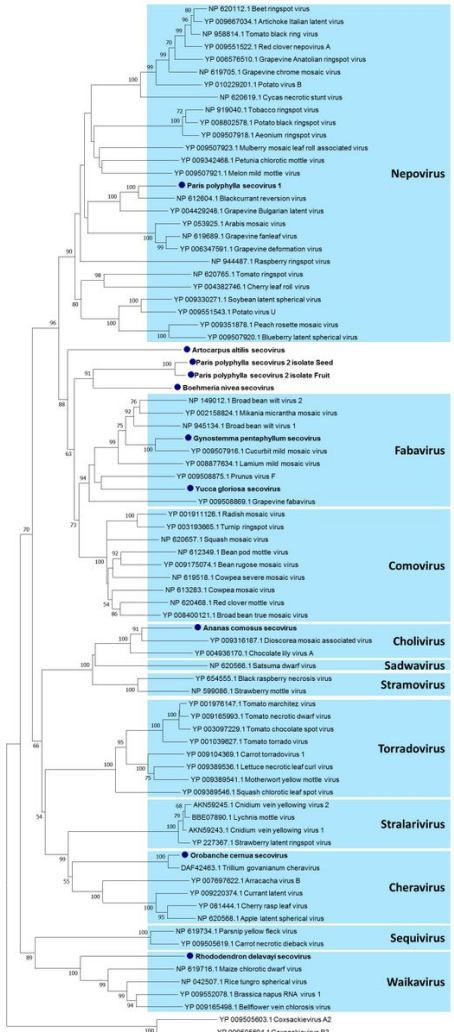

Figure 2

Phylogenetic tree showing the relationship of secoviruses identified in the present study with other secoviruses. Phylogenetic tree was constructed using Neighbourhood-Joining (NJ) method and Poisson model with 1000 bootstrap replicates. Bootstrap values >50% are only indicated. Identified viruses of the present study are shown in bold and indicated by blue circles. Enteroviral sequences were used as outgroups during phylogenetic tree construction.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ESM1.xlsx
- ESM2.docx