

Identification and complete genome sequence of a novel secovirus infecting *Broussonetia kaempferi* in China

Jinhong Yang

Ankang University

Jun Ling

Ankang University

Wei Jiang

Ankang University

Weiqing Kong

`weiqing_kongwq@126.com`

Ankang University

Short Report

Keywords:

Posted Date: April 28th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9535272/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Archives of Virology on July 30th, 2026.
See the published version at <https://doi.org/10.1007/s00705-026-06704-3>.

Abstract

In this study, a novel secovirid, tentatively named *Broussonetia kaempferi* secovirus (BkSV), was identified from a *Broussonetia kaempferi* transcriptome dataset. Its bipartite positive-sense RNA genome comprises two segments (6,300 nt and 3,564 nt), each encoding a single large polyprotein. The RNA1 polyprotein contains six conserved domains associated with replication, and the RNA2 polyprotein is cleaved into movement and coat proteins. Amino acid identities of Pro-Pol and CP relative to known members of Comovirinae were below species demarcation criteria. Phylogenetic analysis confirmed that BkSV represents a new species in the genus *Mersevirus*, family Secoviridae.

Full Text

The family Secoviridae comprises a diverse array of economically important plant viruses. These are non-enveloped, icosahedral viruses 25–30 nm in diameter that cause substantial yield losses in crops, fruits, and ornamental plants worldwide [1]. Genomically, secovirids possess linear, positive-sense, single-stranded RNA genomes of 9–13.7 kb. These genomes may be either monopartite (consisting of a single RNA component encoding a large polyprotein) or bipartite (comprising RNA1 and RNA2, each encoding a large polyprotein and encapsidated in separate virions). These polyproteins are subsequently processed into mature functional proteins by virus-encoded 3C-like proteases [2]. In addition, a poly(A) tract is present at their 3' terminus, and a specialized viral protein (VPg) is covalently attached to their 5' terminus [3].

With advances in high-throughput sequencing, many novel viruses have been identified, expanding viral diversity and prompting taxonomic revisions [4]. Based on host range, genome organization, and transmission vectors, the genus *Mersevirus* was recently established in the family Secoviridae, which currently comprises ten genera: four within the subfamily Comovirinae (*Comovirus*, *Fabavirus*, *Mersevirus*, *Nepovirus*) and six unassigned genera (*Cheravirus*, *Sadwavirus*, *Sequivirus*, *Stralarivirus*, *Torradovirus*, *Waikavirus*) [5]. The genus *Mersevirus* currently includes four viruses: *Boehmeria nivea* secovirus (BnSV), jujube-associated virus 1 (JaSV), *Mercurialis* secovirus 1 (MSV1), and *Paris polyphylla* secovirus 2 (PpSV2). Members of the genus *Mersevirus* can be distinguished from other genera in the subfamily Comovirinae by the presence of a protein containing a Ham1 domain at the N-terminus of the polyprotein, which is predicted to possess inosine triphosphate pyrophosphatase (ITPase) activity [6]. Phylogenetic analyses based on the conserved Pro–Pol region and the coat protein further support this genus as a distinct monophyletic clade [5].

Broussonetia species possess not only high economic and medicinal value but also promising applications in ecological restoration and landscaping [7]. In this study, a novel secovirid belonging to the genus *Mersevirus* was identified from *Broussonetia kaempferi* leaves exhibiting yellow spot symptoms (Fig. S1). The complete genome of the virus was characterized, and it was tentatively designated *Broussonetia kaempferi* secovirus (BkSV).

Leaf samples were collected from the research base of the Shaanxi Key Laboratory of Sericulture, Ankang University, Shaanxi Province, China. Total RNA was extracted using TRIzol reagent according to the manufacturer's instructions. mRNA was enriched using oligo(dT)-coated magnetic beads, followed by cDNA library construction and paired-end 150 bp (PE150) sequencing on the DNBSEQ-T7RS platform at Genepioneer Biotechnologies (Nanjing, China). A total of 21,957,350 raw reads were generated, with a Q20 score of 98.36% and a Q30 score of 94.77% (NCBI SRA: SRX32968257). Adapter trimming, quality filtering, and length trimming were performed using fastp [8]. The resulting clean reads were assembled using SPAdes [9], yielding 22,822 contigs with an N50 of 2,168 bp. Open reading frames (ORFs) and deduced proteins from all contigs were predicted using ORFfinder (<https://www.ncbi.nlm.nih.gov/orffinder>). Viral sequences were identified using two approaches: BLASTp searches against the NCBI viral protein database and pHMM-based detection with VirBot [10]. In total, two potential viral contigs were identified, with lengths and depths of 6,300 bp (467×) and 3,564 bp (606×), respectively. These two sequences were subsequently mapped to the NGS DNA dataset of the same sample (NCBI SRA: SRX32968258), and no matching reads were detected, indicating that the identified viruses are RNA viruses.

Primers were designed to amplify the 5' and 3' termini of the two contigs (Table S1), and 5'-RACE and 3'-RACE were performed using the TSO method (NEB, M0466S). However, no terminal sequence extension was obtained. Further analysis using ORFinder revealed that both sequences contain a single ORF, encoding large polyproteins P1 (2,063 aa) and P2 (993 aa), respectively. The 3'-UTRs of RNA1 and RNA2 are 50 nt and 81 nt in length, respectively, and both contain a poly(A) tail. The 5'-UTRs are 61 nts and 501 nt, respectively, and share a relatively conserved homologous sequence (Fig. 1). These results indicate that the two contigs, designated RNA1 and RNA2, represent the complete viral genomic sequences (GenBank number: PV528063, PV528064).

Subsequently, online BLASTp analysis showed that most viruses sharing relatively high sequence similarity with P1 and P2 belong to the genus *Mersevirus*. An exception was *Apis florea*-associated secovirus 1 (NCBI BioSample: SAMN51756889), which was identified from *Apis florea*. However, this virus has not yet been officially recognized by the International Committee on Taxonomy of Viruses (ICTV). Although some secoviruses can be transmitted by *Apis florea* [11], we could not verify whether this honeybee serves as a transmission vector, as no *Apis florea* individuals were detected in our field sites over nearly one year. Given that all Secoviridae members are plant viruses, this sequence was excluded from subsequent analyses.

To further analyze their evolutionary relationships and structural characteristics, the homologous P1 and P2 protein sequences of 37 viruses from genera *Mersevirus*, *Comovirus* and *Fabavirus* were downloaded (Table S2) and aligned using Muscle software [12]. Conserved domains were predicted via the CDD tool (<https://www.ncbi.nlm.nih.gov/cdd>) and truncated based on the characteristics of proteinase cleavage sites. The results demonstrated that P1 was cleaved into six mature proteins by the virus-encoded proteinase: protease cofactor (Co-Pro, 314 aa), putative nucleoside triphosphate (NTP)-binding protein and helicase (NTB, 586 aa), VPg (85 aa), 3C-like proteinase (3CL-Pro, 196 aa), RNA-directed RNA

polymerase (Pol, 593 aa), and Ham1 (288 aa) with putative cleavage sites Q₃₁₄/F₃₁₅, Q₉₀₀/S₉₀₁, Q₉₈₅/S₉₈₆, Q₁₁₈₁/H₁₁₈₂, and Q₁₇₇₄/G₁₇₇₅, respectively (Fig. 1). P2 was cleaved into three domains, two capsid proteins (CPs) located at the C-terminal region and a putative movement protein (MP) upstream of the CP domain. The putative cleavage sites were Q₃₈₂/N₃₈₃ and Q₇₆₃/N₇₆₄, respectively.

Amino acid identity (AAI) analysis of the Pro-Pol region was performed using SDT software [13]. The results showed the highest AAI values with members of the genus *Mersevirus*: 78.80% with BnSV and 77.30% with MSV1, whereas AAIs relative to other viruses ranged from 42.00% to 54.30%. SDT analysis of the CP region revealed identities of 59.60% with MSV1 and 56.30% with BnSV, followed by 30.80% with PpSV2 and 27.90% with JaSV; the highest identity to other secovirids was 27.70% with Cherry virus F (Fig. 2). In the family Secoviridae, species demarcation criteria established by the ICTV specify less than 80% amino acid sequence identity in the Pro-Pol region and less than 75% identity in the CP. So, this virus represents a novel species in the genus *Mersevirus*. The AAI values for CP were lower than those for Pro-Pol, indicating rapid evolution driven by their structural constraints and adaptation to host-imposed selection pressures [14].

To illustrate the phylogenetic position of BkSV, maximum likelihood (ML) and Bayesian inference (BI) phylogenetic trees were constructed based on the Pro-Pol region of 37 viruses using IQ-TREE 3 [15] and MrBayes 3.2 [16], respectively. One nepovirus from each of group A–C in the genus *Nepovirus* was used as an outgroup. The two phylogenetic trees were largely congruent, with all isolates forming three major clades in agreement with ICTV taxonomy. In both trees, BkSV clustered closely with BnSV and MSV1 and fell within the genus *Mersevirus* (Fig. 3), supporting the classification as a member of the genus *Mersevirus*.

Declarations

Funding This work was supported by the Scientific Research Program Funded by Education Department of Shaanxi Provincial Government (24JP003, 25JP003) and the Scientific Research Program Funded by Ankang Municipal Bureau of Science and Technology (AK2025-NY-24).

Author information

Authors and Affiliations

Shaanxi Key Laboratory of Sericulture, Ankang University, Shaanxi 725099, China

Jinhong Yang, Jun Ling, Wei Jiang, Weiqing Kong

Contributions

All authors contributed to the study conception. Weiqing Kong and Jinhong Yang contributed to the study design; Jinhong Yang, Jun Ling and Wei Jiang performed sample collection, material preparation

and experiments. Jinhong Yang and Weiqing Kong conducted data analysis. The first draft of the manuscript was written by Jinhong Yang. All authors reviewed the manuscript.

Corresponding author

Correspondence to Weiqing Kong.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Sidharthan VK, Rajeswari V, Baranwal VK (2022) Analysis of public domain plant transcriptomes expands the phylogenetic diversity of the family Secoviridae. *Virus Genes* 58:598–604
2. Sanfaçon H (2025) 3C-like proteases at the interface of plant-virus-vector interactions: Focus on potyvirus Nla proteases and secovirus proteases. *Virology* 602:110299
3. Mayo MA, Fritsch C (1994) A possible consensus sequence for VPg of viruses in the family Comoviridae. *FEBS Lett* 354:129–130
4. Zhang YZ, Chen YM, Wang W, Qin XC, Holmes EC (2019) Expanding the RNA virosphere by unbiased metagenomics. *Annu Rev Virol* 6:119–139
5. Fuchs M, Hily JM, Petrzik K, Sanfaçon H, Thompson JR, van der Vlugt R, Wetzels T (2022) ICTV virus taxonomy profile: Secoviridae 2022. *J Gen Virol* 103:001807
6. Mahillon M, Brodard J, Dubuis N, Gugerli P, Blouin AG, Schumpp O (2024) Mixed infection of ITPase-encoding potyvirus and secovirus in *Mercurialis perennis*: evidences for a convergent euphorbia-specific viral counterstrike. *Virol J* 21:6
7. Wang GW, Huang BK, Qin LP (2012) The genus *Broussonetia*: a review of its phytochemistry and pharmacology. *Phytother Res* 26:1–10
8. Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890
9. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477
10. Chen G, Tang X, Shi M, Sun Y (2023) VirBot: an RNA viral contig detector for metagenomic data. *Bioinformatics* 39:btad093

11. Reingold V, Eliyahu A, Luria N, Leibman D, Sela N, Lachman O, Smith E, Mandelik Y, Sadeh A, Dombrovsky A (2024) A distinct Arabidopsis latent virus 1 isolate was found in wild Brassica hirta plants and bees, suggesting the potential involvement of pollinators in virus spread. *Plants* 13:671
12. Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113
13. Muhire BM, Varsani A, Martin DP (2014) SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS One* 9:e108277
14. Thompson JR, Kamath N, Perry KL (2014) An evolutionary analysis of the Secoviridae family of viruses. *PLoS One* 9:e106305
15. Wong TK, Ly-Trong N, Ren H, Baños H, Roger AJ, Susko E, Bielow C, De Maio N, Goldman N, Hahn MW, Huttley G, Lanfear R, Minh BQ (2025) IQ-TREE 3: phylogenomic inference software using complex evolutionary models. *ecoevorXiv*
16. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61:539–542

Figures

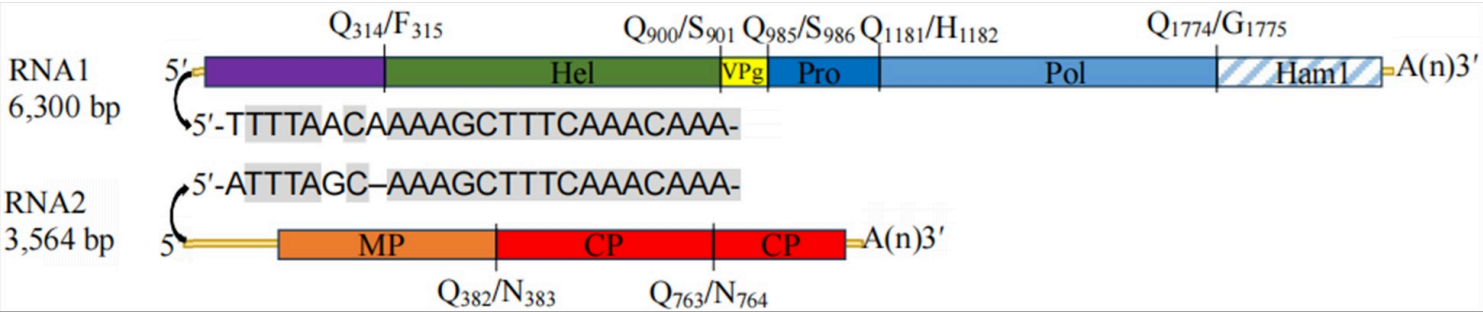


Figure 1

Genome organization of *Broussonetia kaempferisecovirus* (BkSV). Open reading frames (ORFs) of RNA1 and RNA2 are shown as boxes. For both RNAs, conserved sequences in the 5'-UTRs are shaded gray, and poly(A) tails at the 3'-end are indicated as (An). Putative proteinase cleavage sites identified by sequence alignment are marked with solid vertical lines.

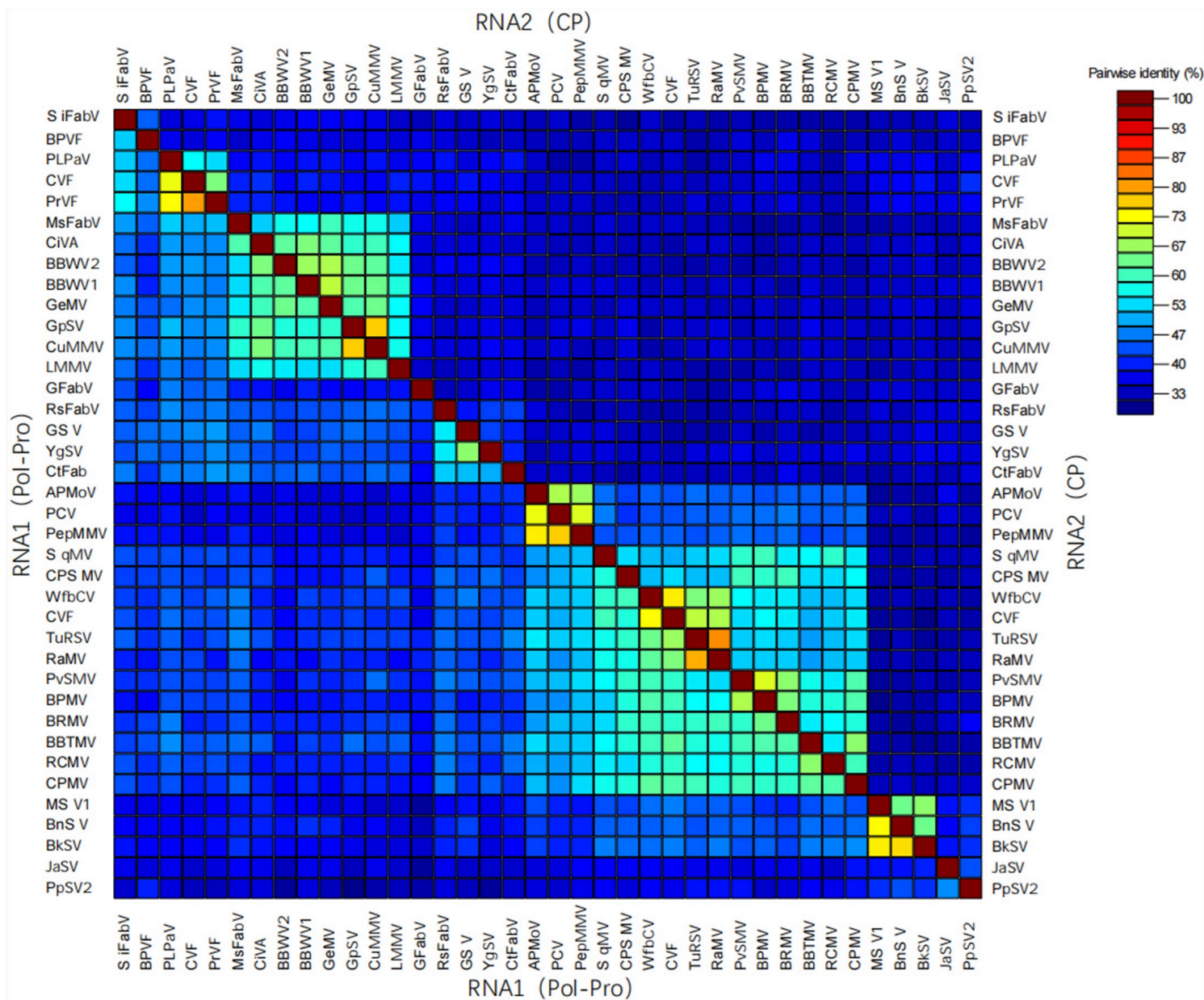


Figure 2

Amino acid identity matrix of the Pro-Pol and CP regions among 37 viruses from the genera *Mersevirus*, *Comovirus*, and *Fabavirus*.

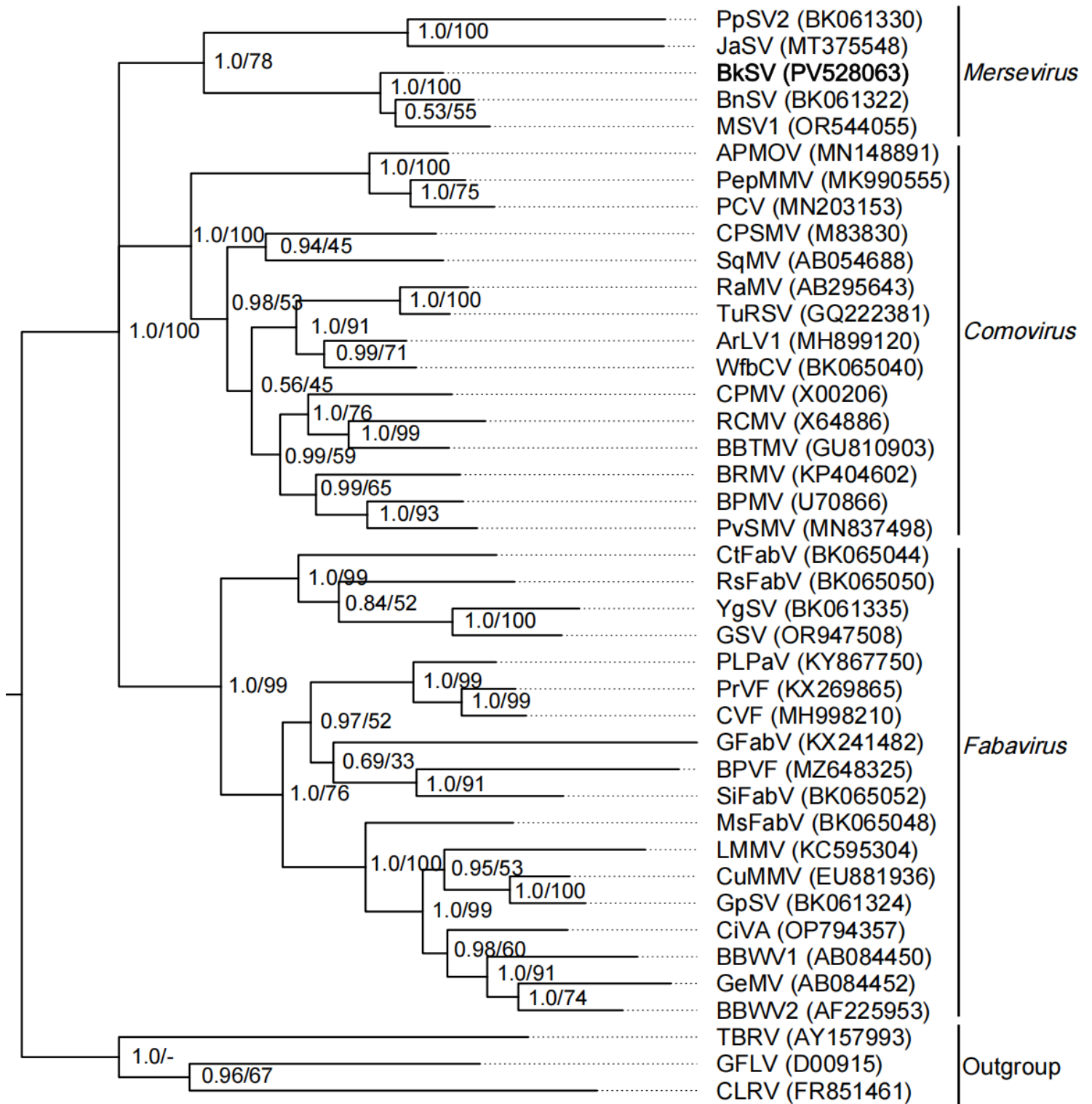


Figure 3

Phylogenetic trees inferred from the Pro-Pol region of 37 viruses. The trees were constructed using maximum likelihood (ML) and Bayesian inference (BI) methods, respectively. Numbers at nodes indicate posterior probability (BI) and bootstrap support (ML) values.

Supplementary Files

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

- [TableS.xlsx](#)
- [LeavesofBroussonetiakaempferiexhibitingyellowspotsymptoms.jpg](#)