Molecular Characterization of a Novel Wheat-infecting Virus in Family *Betaflexiviridae*

Jianxiang Wu (✉ wujx@zju.edu.cn)

Zhejiang University <https://orcid.org/0000-0002-7611-7833>

Shuai Fu

Zhejiang University College of Agriculture and Biotechnology

Mengzhu He

Zhejiang University College of Agriculture and Biotechnology

Bingjian Sun

Gansu Agricultural University College of Plant Protection

Xueping Zhou

Zhejiang University College of Agriculture and Biotechnology

Research Article

Keywords: Wheat infecting virus, molecular characterization, Betaflexiviridae

Posted Date: April 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-405043/v1>

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

[Read Full License](#)

Version of Record: A version of this preprint was published at Archives of Virology on July 23rd, 2021.
See the published version at <https://doi.org/10.1007/s00705-021-05175-y>.

Abstract

Wheat plants showing yellow mosaic in leaves and stunting were observed and collected from wheat fields in the Henan Province, China. Analyses of these plants through transmission electron microscopy showed that these plants contained two filamentous virus-like particles of 200–500 nm and 1000–1300 nm long, respectively. RNA-seq result unveiled a co-infection of wheat yellow mosaic virus (WYMV) and an unknown wheat-infecting virus. The complete genome sequence of the unknown virus is 8,410 nucleotide long, excluding its 3' end poly (A) tail. This unknown virus has six open reading frames (ORFs). The ORF1 encodes a putative viral replication-associated protein, and the ORF2, 3 and 4 encode the triple gene block (TGB) proteins. The ORF5 and 6 encode the capsid protein (CP) and a protein with unknown function, respectively. Phylogenetic relationship analyses showed that this novel virus is evolutionarily related to viruses in the subfamily *Quinvirinae*, family *Betaflexiviridae*. It is, however, distinctly different from the viruses in other genera. Based on the species and genus demarcation criteria set by the International Committee on Taxonomy of Viruses (ICTV), we tentatively name this novel virus as wheat yellow stunt-associated betaflexivirus (WYSaBV). We also propose WYSaBV as a new member in a new genus in the family *Betaflexiviridae*.

Main Text

Virus diseases often cause serious yield losses in wheat fields in China and in many other countries [6]. Based on the virus taxonomy released by the ICTV in 2019 (<http://ictv.global/report>), the order *Tymovirales* has five families: *Alphaflexiviridae*, *Betaflexiviridae*, *Gammaflexiviridae*, *Deltaflexiviridae*, and *Tymoviridae*. Family *Betaflexiviridae* is the largest family with two subfamilies (*Quinvirinae* and *Trivirinae*), containing 13 genera and 108 species. Subfamily *Quinvirinae* and *Trivirinae* are distinguished by the types of their viral movement proteins. For example, viruses in the subfamily *Quinvirinae* encode triple gene block (TGB) movement proteins, while the viruses in the subfamily *Trivirinae* encode the 30K-like movement proteins [1]. Genus *Foveavirus*, *Carlavirus* and *Robigovirus* belong to the subfamily *Quinvirinae*, while genus *Capillovirus*, *Chordovirus*, *Citrivirus*, *Divavirus*, *Prunevirus*, *Ravavirus*, *Tepovirus*, *Trichovirus*, *Vitivirus*, and *Wamavirus* are classified in the subfamily *Trivirinae* [5]. Viruses in the family *Betaflexiviridae* contain monopartite positive-sense single-stranded RNA genome with 6.5-9.0 kilobases (kb) and 3' terminal polyadenylated tail. Viruses in some genera are capped at the 5' ends. Virions of viruses in the family *Betaflexiviridae* are non-enveloped flexuous filamentous of 600-1000 nm long, or even longer, and 12-13 nm in diameter [1].

Recently, RNA-seq technology has been widely used to identify viruses in field-collected plant samples and resulted in identification of many novel plant-infecting viruses [10, 11]. In 2017, we sampled some wheat plants showing yellow mosaic in leaves and stunting from wheat fields in the Kaifeng region, Henan province, China (Fig.1 a, b). Analyses of these samples under a transmission electron microscope found the presence of two different flexuous filamentous virus-like particles of 200-500 × 13 nm (Fig.1 c, d) and 1000-1300 × 13 nm, respectively (Fig.1 e, f).

To characterize these two virus-like particles, total RNA was extracted from the sampled leaf tissues using the EASYspin Plus Complex Plant RNA Kit (Aidlab Biotech, Beijing, China) followed by RNA-seq. The RNA library was synthesized after removing ribosomal RNAs using the TruSeq RNA Sample Prep Kit (Illumina, San Diego, CA, USA), and then sequenced on the Illumina HiSeq X-ten platform (Biomarker Technologies, Beijing, China). Adapters of the paired-end raw reads were trimmed and the low quality reads were filtered using the CLC Genomics Workbench 9.5 software (QIAGEN, Germantown, DEU). The clean reads were first mapped to the wheat genome to identify and remove host sequences, and then assembled *de novo* into larger contigs using the Trinity program [2, 4, 8]. Sequences representing viral RNAs were reconstructed through the manual alignment using the closely related contigs. The assembled contigs were then used to BLASTx and BLASTp search the NCBI databases to identify sequences that shared high nucleotide (nt) sequence similarities and amino acid (aa) sequence identities. The 5' end and 3' end sequences of the assembled contigs were determined through the 5' end and 3' end RACE PCR as instructed (Roche Diagnostic, Milan, Italy). The amplified products were inserted individually into the pLB cloning vector (Tiangen, Beijing, China), and then transformed into *Escherichia coli* DH5 cells followed by DNA sequencing. The resulting nucleotide and the predicted protein sequences were analyzed using the Snap program V4.1.9 software (GSL Biotech, San Diego, CA). Multiple alignments using nucleotide and amino acid sequences were done using the MEGA X software as described with default settings [3].

The result of BLASTx search identified a contig of 7621 nt that covered 99.8% (7621 nt out of 7635 nt) of WYMV genome RNA1 and shared 99% nt sequence similarity with the WYMV genomic RNA1 (FJ361766.1). The BLASTx search also identified a contig of 3634 nt that covered 99.6% (3634 nt out of 3650 nt) of WYMV genome RNA2, and shared 99% nt sequence similarity with the WYMV genomic RNA2 (FJ361769.1). This result supported that the result obtained through transmission electron microscopy (Fig. 1 c, d), and indicated that WYMV was a causal virus of the diseased wheat plants. In addition, another contig of 8403 nt was found to share the highest amino acid sequence identity (43.78%) with the replication-associated protein of peach chlorotic mottle virus (PCMV). Identification of this novel contig indicated the presence of an unknown virus in the infected wheat plants, and this virus might be a member in the family *Betaflexiviridae*. To confirm this finding, we amplified the full-length novel virus sequence through RT-PCR followed by 5' end and 3' end RACEs. The sequencing result showed that the novel virus genomic RNA contained 8140 nt, excluding its poly (A) tail (GenBank accession: MW771279). ORFs of the viral genomic RNA were then predicted using the ORF Finder software (www.ncbi.nlm.nih.gov/orffinder), and the proteins were predicted using the CDD software [7]. The results uncovered that this novel virus had six ORFs (Fig. 2 a), a 59 nt 5' end untranslated region (UTR), and an 11 nt 3' end UTR. The ORF1 started from the AUG start codon at the nt position 60 and ended at the TAA stop codon at the nt position 6213. This ORF1 was predicted to encode a putative RNA-dependent RNA polymerase (RdRp) with 2070 aa residues and an estimated molecular mass (Mr) of 233.8 kDa. This putative RdRp contained four conserved domains: a methyltransferase domain (aa position 44-357, pfam01660), a peptidase domain (aa position 1080-1167, cl05111), a helicase domain (aa position 1254-1506, cl26263), and an RNA-dependent RNA polymerase domain (aa position 1678-1976, cl03049). Although sequence alignment analysis showed that the RdRp aa sequence of this novel virus shared the

highest aa sequence identity (60.9%) with the RdRp of banana virus X (BVX) (Table 1) followed by the RdRp of PCMV (37.0%), the novel virus ORF1 is 6213 nt long, while the ORF1s of BVX and PCMV are 906 nt and 6090 nt long, respectively (Fig. 2). Further analyses of ORF2, ORF3 and ORF4 that encode triple gene block (TGB1, TGB2, and TGB3) proteins showed that the novel virus ORF2-4 were similar to the ORF2-4 of other viruses in the subfamily *Quinvirinae* [1]. For example, the ORF2 of the novel virus was predicted to start from the AUG codon at the nt position 6286 and end at the UAG stop codon at the nt position 6957 (Fig.2). This ORF2 was predicted to encode a 24.5 kDa polypeptide containing a conserved helicase domain (aa position 24-218, pfam01443), and shared the highest aa sequence identity (42.5%) with the TGB1 of PCMV (Fig. 2, Table 1). The ORF3 of the novel virus was predicted to start from the nt position 6947 and end at the nt position 7309. This ORF3 was predicted to encode a 13.4 kDa protein containing a conserved viral movement protein domain (aa position 5-104, cl03157), and shared the highest aa sequence identity (43.2%) with the TGP3 of elderberry carlavirus A (EBCVA)(Fig. 2, Table 1). The ORF4 of the novel virus was predicted to start from the nt position 7218 and end at the nt position 7475. This ORF4 was predicted to encode a 9.2 kDa protein that shared the highest aa sequence identity (33.8%) with the TGB4 of butterbur mosaic virus (BMV) (Table 1). The ORF5 of the novel virus was predicted to start from the nt position 7487 and end at the nt position 8278 (Fig. 2). This ORF5 was predicted to encode a 28.5 kDa protein with a conserved flexi CP domain (aa 83-220, cl02836). This protein shared the highest aa sequence identity (33.5%) with the CP of PCMV (33.5%) (Fig.2, Table 1). The ORF6 of the novel virus is partially overlapped with its ORF5 and was predicted to start from the nt position 7941 and end at the nt position 8399. This ORF was predicted to encode a 17.4 kDa protein with unknown function, and have no significant sequence homology with other known proteins. Alignment using genomic sequences of this novel virus and nine representative members in the subfamily *Quinvirinae* showed that the genome of this novel virus shared the highest nt sequence similarity (45.6%) with the genome of PCMV followed by BVX (44.9%) (Table 1).

To investigate the infectivity of this novel virus, leaves of the plants showing typical yellow mosaic and stunting were ground in 0.1 M phosphate buffer and then mechanically inoculated to several assay plants, including wheat, *Nicotiana benthamiana*, *N. occidentalis*, *N. tabacum* cv. Samsun, *Chenopodium quinoa*, *C. amaranticolor* and *Gomphrena globosa* grown inside a greenhouse at 25°C. All the inoculated plants remained symptomless during the study and the novel virus was not detected in them through RT-PCR at 30 days post inoculation.

Phylogenetic analyses using whole genomic sequences of the novel virus and the representative viruses in five different families in the order *Tymovirales* indicated that the novel virus was more closely phylogenetically related to the members in the family *Betaflexiviridae* (Fig. 3 a). The phylogenetic trees constructed using the whole genome sequences, and the RdRp or CP sequences of the novel virus and the representative viruses in the family *Betaflexiviridae* showed that this novel virus should be placed in the subfamily *Quinvirinae*, but is distinct from the members in other existing genera (Fig. 3 b, c, d). Alinement analyses showed that the viral genome and protein sequence similarities between the novel virus and the other members in the subfamily *Quinvirinae* ranged from 41.8% to 45.6%, and 14.7% to 60.9%, respectively (Table 1). For RdRp and CP, the aa sequence identities between the novel virus and the

other members ranged from 22.9% to 37%. Though BVX showed 60.9% aa sequence identities with the novel virus RdRp, but BVX RdRp sequence was remarkable shorten with only 301 aa in length, compared with the novel virus with 2070 aa in length. While, the novel virus CP showed only 33.5% aa sequence identities with BVX CP (Table 1). These genome, RdRp, CP sequence differences not only exceed the cut-off values for species discrimination (i.e., 72% for the nt and 80% for the aa sequence), but also the cut-off values for genus discrimination (i.e., 45% for the nt and aa sequences) [1]. Consequently, we consider that the novel virus as a new species in the subfamily *Quinvirinae*. Because the differences in virus genome size, genome structure, and phylogenetic relationship with other members in the family *Betaflexiviridae*, we were unable to place this novel virus in any existing genus in the family *Betaflexiviridae*. Therefore, we tentatively named this novel virus as wheat yellow stunt-associated virus (WYSaBV) and proposed to place it in a new genus in the family *Betaflexiviridae*.

Some questions need to be further investigated in future, including (i) Where is WYSaBV originated? (ii) What types of damages can single infection of WYSaBV on wheat cause in wheat plant? (iii) How is WYSaBV transmitted in the field? (iv) Considering that WYSaBV and WYMV mix-infection on wheat, and WYMV is transmitted by a soilborne fungus-like organism *Polymyxa graminis* [9]. Can *P. graminis* also transmit WYSaBV in the field? Nonetheless, identification of WYSaBV on wheat plants in this study indicates the need of caution in order to develop an effective management strategy for virus disease control in wheat fields in China.

Declarations

Acknowledgements

This research was supported by grants from National Natural Science Foundation of China for Youth Scholars (31801701), China Postdoctoral Science Foundation (2019T120521 and 2018M642449). We thank Dr. Xinshun Ding (retire, New York, USA) for his help on the language of this manuscript.

Compliance with ethical standards

Conflict of interest

All authors declare they have no conflict of interest.

Ethical approval

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

References

1. Adams MJ, Antoniw JF, Bar-Joseph M, Brunt AA, Candresse T, Foster GD, Martelli GP, Milne RG, Zavriev SK, Fauquet CM (2004) The new plant virus family Flexiviridae and assessment of molecular

- criteria for species demarcation. Archives of virology 149:1045–1060
2. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature biotechnology 29:644–652
 3. Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Molecular biology evolution 35:1547–1549
 4. Lai K, Berkman PJ, Lorenc MT, Duran C, Smits L, Manoli S, Stiller J, Edwards D (2012) WheatGenome.info: an integrated database and portal for wheat genome information. Plant Cell Physiol 53:e2
 5. Lefkowitz EJ, Dempsey DM, Hendrickson RC, Orton RJ, Siddell SG, Smith DB (2018) Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). Nucleic acids research 46:D708–Dd717
 6. Liu Y, Khine MO, Zhang P, Fu Y, Wang X (2020) Incidence and Distribution of Insect-Transmitted Cereal Viruses in Wheat in China from 2007 to 2019. Plant Dis 104:1407–1414
 7. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic acids research 39:D225–D229
 8. Montenegro JD, Golicz AA, Bayer PE, Hurgobin B, Lee H, Chan CK, Visendi P, Lai K, Doležal J, Batley J, Edwards D (2017) The pangenome of hexaploid bread wheat. The Plant journal: for cell molecular biology 90:1007–1013
 9. Ohki T, Netsu O, Kojima H, Sakai J, Onuki M, Maoka T, Shirako Y, Sasaya T (2014) Biological and genetic diversity of Wheat yellow mosaic virus (Genus Bymovirus). Phytopathology 104:313–319
 10. Shi M, Lin XD, Tian JH, Chen LJ, Chen X, Li CX, Qin XC, Li J, Cao JP, Eden JS, Buchmann J, Wang W, Xu J, Holmes EC, Zhang YZ (2016) Redefining the invertebrate RNA virosphere. Nature 540:539–543
 11. Wu Q, Ding SW, Zhang Y, Zhu S (2015) Identification of viruses and viroids by next-generation sequencing and homology-dependent and homology-independent algorithms. Annu Rev Phytopathol 53:425–444

Tables

Table 1 Genomic RNA and viral protein sequence similarities between WYSaBV and the representative members in the subfamily *Quinvirinae*

Genus	Virus	Genome (nt)	(ORF1) RdRp (aa)	(ORF2) TGB1 (aa)	(ORF3) TGB2 (aa)	(ORF4) TGB3 (aa)	(ORF5) CP(aa)
<i>Foveavirus</i>	ASPV	44.2	35.5	40.3	35.3	27.5	31.6
<i>Foveavirus</i>	PCMV	45.6	37.0	42.5	34.8	26.5	33.5
<i>Carlavirus</i>	PVM	42.8	33.1	35.7	29.4	20.3	25.5
<i>Carlavirus</i>	BMV	43.2	33.4	32.7	36.4	33.8	22.9
Unclassified <i>Carlavirus</i>	EBCVA	43.4	33.6	35.3	43.2	21.7	27.0
<i>Robigovirus</i>	CRMaV	42.7	30.4	33.8	26.3	19.4	28.9
Unassigned	BVX	44.9	60.9	28.4	40.9	19.4	23.5
Unassigned	SSMaV	42.4	31.9	30.5	24.3	14.7	29.1
Unassigned	BMMV	41.8	31.5	33.0	29.5	14.9	26.6

Virus names and abbreviations, and their GenBank accession numbers used in these analyses are shown as: apple stem pitting virus (ASPV, NC 003462.2), peach chlorotic mottle virus (PCMV, NC 009892.1), potato virus M (PVM, NC 001361.2), elderberry carlavirus A (EBCVA, NC 029085.1), cherry rusty mottle associated virus (CRMaV, NC 020996.1), banana virus X (BVX, AY710267.1), sugarcane striate mosaic-associated virus (SSMaV, NC 003870.1), and butterbur mosaic virus (BMV, NC 013527.1).

Figures

Figure 1

Symptom of an infected wheat plant and virus-like particles found in the infected leaf extract. (a) A field-collected wheat plant showing yellow mosaic and stunting symptoms. (b) A wheat leaf showing typical yellow mosaic symptom. (c, d) Transmission electron micrographs showing negatively-stained filamentous virus-like particles of approximately 200-500 × 13 nm in size. (e, f) Transmission electron micrographs showing negatively-stained filamentous virus-like particles of approximately 1000-1300 × 13 nm in size. Bar = 200 nm.

Figure 2

Genome organizations of WYSaBV and six representative members in the subfamily Quinvirinae. Schematic diagrams representing the genome organizations of WYSaBV and six representative members in the genus *Foveavirus*, *Carlavirus*, *Robigovirus*, and three unassigned genera in the subfamily Quinvirinae, family Betafexiviridae. The ORF encoding viral replication associated protein is shown in gray. The ORFs encoding the triple gene block proteins are shown in red. The ORF encoding the capsid protein is shown in blue, and the ORF encoding the RNA binding protein (RBP) of PMV is shown in purple. The sixth ORF of WYSaBV is shown in green.

Figure 3

Phylogenetic analyses using the genomic sequences, or the RdRp and CP sequences of WYSaBV and representative members in several genera in the order Tymovirales or family Betaflexiviridae. Phylogenetic trees constructed based on the genomic sequences, or the RdRp and CP sequences of WYSaBV and representative members belonging to several genera in the order Tymovirales or family Betaflexiviridae. Evolutionary relationship among these viruses were determined using the Neighbor-Joining method with

1000 bootstrap replications. Clusters with the bootstrap values less than 50 are not indicated. The phylogenetic trees are drawn to scale and the length of branches in the same clusters represent the evolutionary distances computed using the Poisson correction method. Evolutionary analyses were conducted using the MEGA X software.

Supplementary Files

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

- [genomicsequence.txt](#)