Archives of Virology (2020) 165:257–260
https://doi.org/10.1007/s00705-019-04473-w

ANNOTATED SEQUENCE RECORD

Molecular characterization of common bean curly stunt virus: a novel recombinant geminivirus in China

Ruonan Zhang[1] · Xiaoyun Wu[1] · Xue Jiang[1] · Xiaoxia Wu[1] · Xiaoyan Luan[2] · Xiaofei Cheng[1,3]

Received: 24 August 2019 / Accepted: 17 October 2019 / Published online: 26 November 2019

Abstract

A new geminivirus was identified in common bean (*Phaseolus vulgaris*) showing severe stunt and leaf curling symptoms in Heilongjiang province of China, via sequencing and assembly of small RNAs. The genome of this geminivirus comprises 2,959 nucleotides (nt) and shares 21.77-54.97% nt sequence identity with other geminiviruses. The coat protein (CP) shares the highest amino acid (aa) sequence identity (23.5%) with that of sesame curly top virus (SeYMV; genus *Turncurtovirus*), whereas the C1 (Rep) shares the highest aa sequence identity (66.5%) with that of beet severe curly top virus (BSCTV; genus *Curtovirus*). This geminivirus neighbors the turncurtoviruses in phylogenetic trees based on the full genome sequence or the amino acid sequence of the Rep protein, but it forms a distinct clade in the phylogenetic tree based on the coat protein. Recombination analysis showed that parts of the C1 coding region of this geminivirus were recombined from a curtovirus or turncurtovirus. Based on these results, the name "common bean curly stunt virus" (CBCSV) is proposed for this virus.

Introduction

Viruses in the family *Geminiviridae* infect a large number of plants and cause significant economic losses to many crops worldwide [1]. The genome of geminiviruses typically consists of one or two circular single-stranded DNA (ssDNA) molecules of about 2.5-3.2 kilobase pairs (kb) that are encapsidated within twinned (geminate) icosahedral particles of about 22 × 38 nm in size [2]. New geminiviruses are rapidly emerging via a high mutation rate and frequent recombination [3]. Currently, the family is divided into nine genera, namely *Becurtovirus*, *Begomovirus*, *Capulavirus*, *Curtovirus*, *Eragrovirus*, *Grablovirus*, *Mastrevirus*, *Topocuvirus*, and *Turncurtovirus*, on the basis of host range, type of vector, genome organization, genome-wide pairwise sequence comparisons, and phylogenetic relationships [2].

During a survey of viral diseases of legumes in July 2018, common bean (*Phaseolus vulgaris*) plants showing virus-like symptoms were observed and collected in Harbin, Heilongjiang province, China. A small-RNA (sRNA) library was constructed from an equal-ratio leaf mixture of three common bean samples (DN-18, DN-19, and DN-20) showing severe stunt and leaf curling, vein banding, and chlorosis, respectively. The library was sequenced using the Illumina HiSeq-400 sequencing platform at Lianchuan Biotechnology Co., Ltd. (Hangzhou, China). A total of 17,060,132 reads with length between 17 and 27 nucleotides (nt) were obtained after adaptor trimming and read-quality filtering using Cutadapt 2.4 [4]. These reads were assembled using Velvet and Oases with a *k-mer* value of 17 [5, 6], and the resulting contigs were analyzed as described earlier [7]. A total of 22, 16, and 27 contigs had a high level of similarity to bean common mosaic virus (BCMV; genus *Potyvirus*), broad bean wilt virus 2 (BBWV-2; genus *Fabavirus*), and alfalfa mosaic virus (AMV; genus *Alfamovirus*) respectively. This analysis also revealed 12 contigs with a low level of

Handling Editor: F. Murilo Zerbini.

Ruonan Zhang and Xiaoyun Wu have contributed equally.

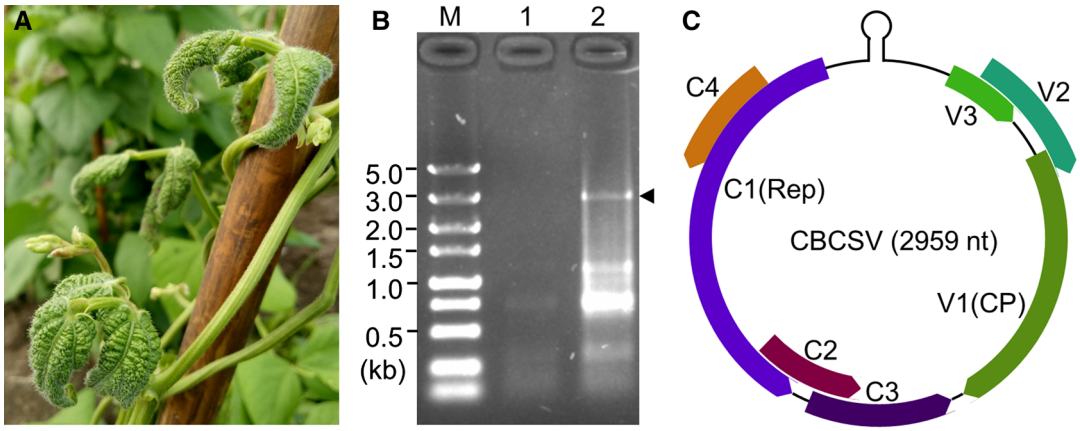
Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00705-019-04473-w) contains supplementary material, which is available to authorized users.

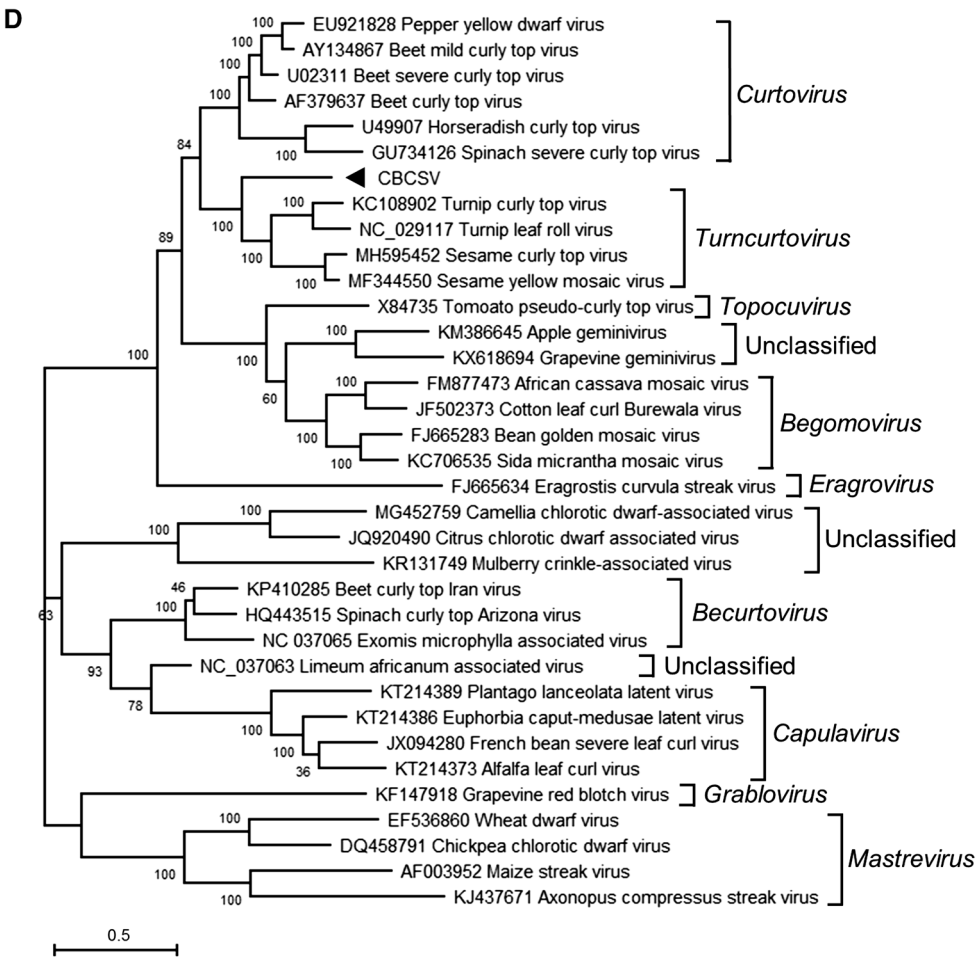
✉ Xiaofei Cheng
xfcheng@neau.edu.cn

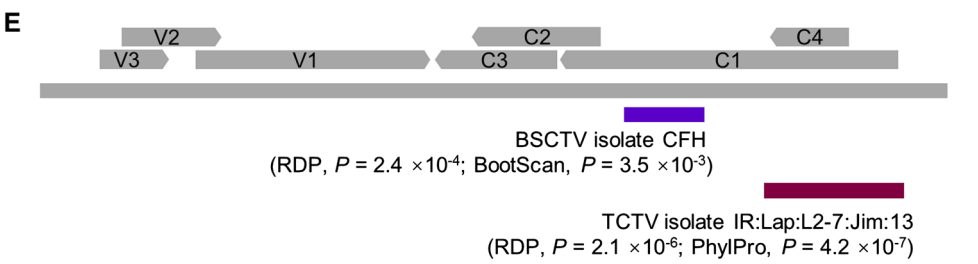
[1] College of Agriculture, Northeast Agricultural University, Harbin 150030, China

[2] Institute of Soybean, Heilongjiang Academy of Agricultural Sciences, 368 Xuefu Road, Harbin 150086, China

[3] Key Laboratory of Germplasm Enhancement, Physiology and Ecology of Food Crops in Cold Region of Chinese Education Ministry, Northeast Agricultural University, Harbin 150030, China


A
B
M
1
2
5.0
3.0
2.0
1.5
1.0
0.5
(kb)
C
C4
V2
V3
C1(Rep)
CBCSV (2959 nt)
V1(CP)
C2
C3


D
EU921828 Pepper yellow dwarf virus
AY134867 Beet mild curly top virus
U02311 Beet severe curly top virus
AF379637 Beet curly top virus
U49907 Horseradish curly top virus
GU734126 Spinach severe curly top virus
Curtovirus
CBCSV
KC108902 Turnip curly top virus
NC_029117 Turnip leaf roll virus
MH595452 Sesame curly top virus
MF344550 Sesame yellow mosaic virus
Turncurtovirus
X84735 Tomoato pseudo-curly top virus
Topocuvirus
KM386645 Apple geminivirus
KX618694 Grapevine geminivirus
Unclassified
FM877473 African cassava mosaic virus
JF502373 Cotton leaf curl Burewala virus
FJ665283 Bean golden mosaic virus
KC706535 Sida micrantha mosaic virus
Begomovirus
FJ665634 Eragrostis curvula streak virus
Eragrovirus
MG452759 Camellia chlorotic dwarf-associated virus
JQ920490 Citrus chlorotic dwarf associated virus
KR131749 Mulberry crinkle-associated virus
Unclassified
KP410285 Beet curly top Iran virus
HQ443515 Spinach curly top Arizona virus
NC 037065 Exomis microphylla associated virus
Becurtovirus
NC_037063 Limeum africanum associated virus
Unclassified
KT214389 Plantago lanceolata latent virus
KT214386 Euphorbia caput-medusae latent virus
JX094280 French bean severe leaf curl virus
KT214373 Alfalfa leaf curl virus
Capulavirus
KF147918 Grapevine red blotch virus
Grablovirus
EF536860 Wheat dwarf virus
DQ458791 Chickpea chlorotic dwarf virus
AF003952 Maize streak virus
KJ437671 Axonopus compressus streak virus
Mastrevirus
0.5


E
V2
C2
C4
V3
V1
C3
C1
BSCTV isolate CFH
(RDP, P = 2.4 ×10-4; BootScan, P = 3.5 ×10-3)
TCTV isolate IR:Lap:L2-7:Jim:13
(RDP, P = 2.1 ×10-6; PhylPro, P = 4.2 ×10-7)

◂**Fig. 1** Symptoms, genome structure, and phylogeny of CBCSV. **A.** Symptoms of CBCSV infection in a common bean plant. **B.** PCR amplification of the complete genome. M, DNA marker. Lanes 1 and 2 are healthy and symptomatic common bean leaves, respectively. **C.** Schematic diagram of the CBCSV genome. **D.** Phylogenetic tree based on full genome sequences of CBCSV and selected geminiviruses. The tree was constructed using the maximum-likelihood (ML) method in the MEGA X software [13] with the Jukes-Cantor genetic distance model and 1000 bootstrap replicates. The Kimura 2-parameter nucleotide substitution model was determined by the Model Selection function in MEGA X software. **E.** Recombination events in the CBCSV genome detected by RDP 4 software [13] using default parameters. Only recombination events with a *p*-value <0.01 were accepted, and the two recombination detection methods with the lowest *p*-values are shown

similarity (less than 65 % amino acid [aa] sequence identity) to beet curly top virus (BCTV; genus *Curtovirus*), turnip curly top virus (TCTV; genus *Turncurtovirus*), or sesame yellow mosaic virus (SeYMV; genus *Turncurtovirus*). Three sets of primers were then designed on the basis of these geminiviral contigs (Supplementary Table 1). Polymerase chain reaction (PCR) was then performed using Phanta Super-Fidelity DNA Polymerase (Vazyme Biotech, Nanjing, China) and total DNA extracted from DN-18, DN-19, or DN-20 by the CTAB method [8]. Two primer sets (Turto_F1 plus Turto_F1 and Turto_F2 plus Turto_ R2; all located in the *C3* open reading frame [ORF]) successfully amplified two fragments of about 2.7 kb from sample DN-18, but not from the other two samples, whereas primers Turto_F3 (located in *C2* ORF) and Turto_R3 (located in *C1* ORF) did not give a replicon of the predicted size (~1.1 kb) for any of the three samples. The amplicons were recovered, cloned into the pEASY-Blunt vector (Transgen, Beijing, China), and sequenced. The results showed that the two fragments were from the same new geminivirus. Therefore, the primers Turto_100F and Turto_2000R (located in *V1* and *C1*, respectively) were designed to amplify a fragment of about 1.1 kb of this geminivirus. The 5´and 3´ ends of this fragment overlapped with the two 2.7-kb fragments. We also designed a primer (Turto_1488R) that is back-to-back with Turto_F1 for amplifying the entire genome of this geminivirus. A band of about 3.0 kb was successfully amplified from DN-18, but not from the healthy common bean leaf sample (Fig. 1B). The amplified 3.0-kb fragment was inserted into the pEASY-Blunt vector. Plasmids from two independent colonies were sequenced using an ABI automated DNA sequencer (Sangon Bio., Shanghai, China). The resulting fragments from Sanger sequencing were assembled using the SeqMan program in Lasergene 7.1 (DNASTAR, Inc, Wisconsin, USA). Multiple sequence alignment using the MegAlign program in Lasergene 7.1 showed that this 3.0-kb fragment was identical to the genomic sequence assembled from the two overlapping fragments, suggesting that the amplified 3.0-kb fragment represents the full genome of the new geminivirus. We also performed RT-PCR to confirm the presence of BCMV, BBWV-2, or AMV in DN-18. The RT-PCR results showed that none of the three viruses was detected in the sample showing severe stunt and leaf curling symptoms.

The full genome of this geminivirus comprises 2,959 nucleotides (nt) (GenBank accession no. MK673513) and has the highest nucleotide (nt) sequence identity (55%) to SeYMV isolate IR/Jir/JK_10-2/14 (Table 1). The origin-of-replication sequence of this geminivirus is identical to the conserved nonanucleotide motif of the majority of geminiviruses (TAATATT/AC). Sequence analysis using Lasergene Seqbuilder 7.1.0 (DNASTAR, Inc., Madison, WI, USA) revealed three ORFs in the virion sense, namely *V1* (nt 508-1272), *V2* (nt 267-590), and *V3* (nt 193-417), and four ORFs in the complementary sense, e.g., *C1* (nt 2818-1715), *C2* (nt 1848-1429), *C3* (nt 1706-1308), and *C4* (nt 2658-2401). *C1* and *V3* are separated by a 333-nt intergenic region (IR) containing the conserved nonanucleotide motif, whereas *C3* and *V1* are separated by a 35-nt IR (Fig. 1C). *V1* encodes a 254-aa coat protein (CP) that shows very limited similarity to those of other geminiviruses (Table 1). The V2 protein is homologous to that of turncurtoviruses, whereas no homologous counterpart of V3 was found in the GenBank database. *C1* encodes a 367-aa replicase (Rep) protein that has the highest aa sequence identity (65.5%) with BSCTV (Table 1). C2, C3, and C4 are homologous to those of curtoviruses on the basis of BLASTp analysis [11]. This geminivirus was located in a separate branch adjacent to the clade of turncurtoviruses in phylogenetic trees constructed based on the full genome sequences (Fig. 1D) or amino acid sequence of the Rep proteins (Supplementary Fig. 1B) of representative geminiviruses. Interestingly, this geminivirus forms a distinct clade in the CP-based phylogenetic tree (Supplementary Fig. 1A). A recombination analysis was performed using CBCSV and the most closely related geminiviruses in the NCBI database, via PSI-BLAST (GenBank accession nos.: KC108902, MF536416, KT388064, EU921828, KX529650, AF379637, KX867037, MH595452, MH595454, MH595453, U02311, and X97203) with the detection methods RDP, Chimaera, BootScan, 3Seq, GENECONV, MaxChi, SiScan, and LARD in RDP 4 software [12]. The results showed that the N-terminal portion of *C1* (nt 2344-2829) was possibly acquired by recombination from TCTV isolate IR:Lap:L2-7:Jim:13 (GenBank accession no. MF536416), the middle part of *C1* (nt 1915-2165) was possibly derived from BSCTV isolate CFH (GenBank accession no. X97203), and the upstream portion of *V3* (nt 29-136) was possibly acquired by recombination from BCTV isolate CTS07-043 (GenBank accession no. KX867037) (Fig. 1E). These data indicate that this geminivirus is a recombinant virus that is phylogenetically related to turncurtoviruses, although it has a slightly

Table 1 Percent nucleotide and amino acid sequence identity between CBCSV and selected geminiviruses

Virus name	Genome	V1 (CP)		C1 (Rep)	
		Nucleotide	Amino acid	Nucleotide	Amino acid
CaCDaV	21.8	27.6	7.6	29.7	31.1
ACMV	39.2	28.0	12.3	61.0	52.3
CLCuBuV	37.9	29.0	1.6	60.0	52.6
BGMV	40.0	26.6	12.9	62.9	56.9
SimMV	39.6	24.9	12.8	63.7	60.7
AGV	36.9	30.6	12.0	53.7	48.6
GGVA	38.2	30.6	11.6	52.0	19.6
TPCTV	39.0	32.3	14.7	61.6	60.7
TCTV	54.5	37.5	23.1	68.3	62.5
PeYDV	50.9	38.7	18.3	68.7	62.4
BCTV	50.5	38.3	19.0	68.5	17.2
HrCTV	40.1	39.2	18.3	52.7	48.2
SSCTV	41.9	38.4	19.7	53.9	49.4
BCTIV	30.1	40.3	20.3	26.6	61.8
SCTAV	26.3	38.0	19.3	27.5	17.8
ECSV	32.1	35.3	16.4	44.9	32.6
WDV	26.3	33.6	17.3	25.1	27.5
CpCDV	27.2	32.8	18.7	30.0	20.9
MSV	26.2	33.3	16.6	24.8	18.7
ACSV	26.6	31.8	16.8	25.5	26.7
GRBaV	28.4	31.7	11.8	30.0	46.7
CCDaV	23.3	26.3	9.7	28.8	21.3
MCLV	27.7	27.5	11.8	28.1	20.0
PlLV	27.1	33.6	12.0	26.3	30.8
FbSLSV	26.0	33.3	11.7	28.1	20.9
EcmLV	26.0	31.9	12.8	27.2	31.7
ALCV	26.8	31.4	12.1	26.9	33.1
SeCTV	54.4	38.0	23.5	68.6	63.4
TLRV	54.3	38.4	21.5	67.5	62.2
SeYMV	55.0	37.9	23.1	68.7	62.8
LAAV	31.7	37.4	18.3	26.8	34.0
EMaV	28.8	35.2	18.9	25.9	23.1
BMCTV	50.3	38.4	19.0	68.4	63.8
BSCTV	52.5	38.7	19.0	72.3	65.5

The sequence information used for comparison is shown in Supplementary Table 1. Nucleotide and amino acid sequences were aligned using Clustal Omega [9], and sequence identity values were calculated using the Sequence Manipulation Suite [10]

different genome structure from that of turncurtoviruses (i.e., it contains a putative *V3* gene and a 35-nt IR between *C3* and *V1*). Because this novel geminivirus has a unique genome organization and its genomic sequence is highly divergent from those of other geminiviruses, it is difficult to assign it to one of the nine current genera of the family *Geminiviridae*. We propose to name this novel geminivirus "common bean curly stunt virus" (CBCSV).

Northeastern China is an important region for production of common bean crops. A survey has suggested that this virus is widely distributed in common bean plants in Heilongjiang province. Therefore, special attention should be paid to the damage that it may cause.

Funding This study was funded by the National Natural Science Foundation of China (grant no. 31671998) and the Natural Science Foundation of Heilongjiang Province (ZD2018002).

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest.

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

References

1. Mansoor S, Zafar Y, Briddon RW (2006) Geminivirus disease complexes: the threat is spreading. Trends Plant Sci 11:209–212
2. Zerbini FM, Briddon RW, Idris A, Martin DP, Moriones E, Navas-Castillo J, Rivera-Bustamante R, Roumagnac P, Varsani A, Consortium IR (2017) ICTV virus taxonomy profile: Geminiviridae. J Gen Virol 98:131–133
3. Padidam M, Sawyer S, Fauquet CM (1999) Possible emergence of new geminiviruses by frequent recombination. Virology 265:218–225
4. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet 17:10–12
5. Zerbino DR, Birney E (2008) Velvet: algorithms for *de novo* short read assembly using de bruijn graphs. Gen Res 18:821–829
6. Schulz MH, Zerbino DR, Vingron M, Birney E (2012) Oases: robust *de novo* RNA-seq assembly across the dynamic range of expression levels. Bioinformatics 28:1086–1092
7. Wu X, Wu X, Li W, Cheng X (2018) Molecular characterization of a divergent strain of calla lily chlorotic spot virus infecting celtuce (*Lactuca sativa* var. augustana) in China. Arch Virol 163:1375–1378
8. Allen GC, Flores-Vergara MA, Krasynanski S, Kumar S, Thompson WF (2006) A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. Nat Protoc 1:2320
9. Sievers F, Higgins DG (2018) Clustal Omega for making accurate alignments of many protein sequences. Protein Sci 27:135–148
10. Stothard P (2000) The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques 28:1102–1104
11. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410
12. Martin DP, Murell B, Golden M, Khoosal A, Muhire B (2015) RDP4: detection and analysis of recombination patterns in virus genomes. Virus Evol 1:vev003
13. Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549