

Presence of Diverse Sugarcane Bacilliform Viruses Infecting Sugarcane in China Revealed by Pairwise Sequence Comparisons and Phylogenetic Analysis

Kashif Ahmad^{1,†}, Sheng-Ren Sun^{1,†}, Jun-Lü Chen², Mei-Ting Huang¹, Hua-Ying Fu¹, and San-Ji Gao^{1*}

¹National Engineering Research Center of Sugarcane, Fujian Agricultural and Forestry University, Fuzhou 350002, China

²Guangzhou Sugarcane Industry Research Institute, Guangzhou 510316, Guangdong, China

(Received on August 8, 2018; Revised on September 16, 2018; Accepted on October 4, 2018)

Sugarcane bacilliform viruses (SCBV), which belong to the genus *Badnavirus*, family *Caulimoviridae*, are an important DNA virus complex that infects sugarcane. To explore the genetic diversity of the sugarcane-infecting badnavirus complex in China, we tested 392 sugarcane leaf samples collected from Fujian, Yunnan, and Hainan provinces for the occurrence of SCBV by polymerase chain reaction (PCR) assays using published primers SCBV-F and SCBV-R that target the reverse transcriptase/ribonuclease H (RT/RNase H) regions of the viral genome. A total of 111 PCR-amplified fragments (726 bp) from 63 SCBV-positive samples were cloned and sequenced. A neighbor-joining phylogenetic tree was constructed based on the SCBV sequences from this study and 34 published sequences representing 18 different phylogroups or genotypes (SCBV-A to -R). All SCBV-tested isolates could be classified into 20 SCBV phylogenetic groups from SCBV-A to -T. Of nine SCBV phylogroups reported in this study, two novel phylogroups, SCBV-S and SCBV-T, that share 90.0-93.2% sequence identity and show 0.07-0.11 genetic distance with each other in the RT/RNase H region, are proposed. SCBV-S had 57.6-

92.2% sequence identity and 0.09-0.66 genetic distance, while SCBV-T had 58.4-90.0% sequence identity and 0.11-0.63 genetic distance compared with the published SCBV phylogroups. Additionally, two other *Badnavirus* species, *Sugarcane bacilliform MO virus* (SCBMOV) and *Sugarcane bacilliform IM virus* (SCBIMV), which originally clustered in phylogenetic groups SCBV-E and SCBV-F, respectively, are first reported in China. Our findings will help to understand the level of genetic heterogeneity present in the complex of *Badnavirus* species that infect sugarcane.

Keywords : badnaviruses, genetic diversity, reverse transcriptase/ribonuclease H, sequence analysis, sugarcane bacilliform virus

Handling Editor : Ju, Ho-Jong

Sugarcane (*Saccharum* spp.) is not only a key sugar-producing crop but is also a potentially renewable bioenergy crop that is grown in tropical and subtropical regions worldwide. The modern sugarcane cultivars are a complex of aneuployploids that result from commercial hybrid breeding programs involving several species and consist of approximately 80% *S. officinarum*, 10-15% *S. spontaneum* and 5-10% recombinant chromosomes (D'Hont et al. 1996). The crop is subjected to various stresses caused by biotic pathogens and abiotic conditions during the growing period. Sugarcane bacilliform viruses (SCBVs) that belong to the genus *Badnavirus* (family *Caulimoviridae*) are one of the main species that infect sugarcane. In 1985, SCBV was first identified in the infected sugarcane cultivar B34104 in Cuba, and then spread into sugarcane-growing regions worldwide (Autrey et al., 1995; Lockhart et al., 1996). SCBV is naturally transmitted by insect vectors, the

[†]These authors contributed equally to this work.

*Corresponding author.

Phone) +68-591-8826-3135, FAX) +68-591-8826-3135

E-mail) gaosanji@yahoo.com

ORCID

San-Ji Gao

<http://orcid.org/0000-0003-3442-8470>

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

mealybug species *Saccharicoccus sacchari* and *Dysmiococcus boninsis*, and can also be experimentally transmitted by crude virus-containing sap or via *Agrobacterium*-mediated inoculation (Lockhart et al., 1996). In addition, long-distance transport of SCBV-infected cane materials is an important mean of virus dissemination. The natural host range of SCBV is not limited to sugarcane, but also includes other grasses such as *Sorghum halepense*, *Brachiaria* sp., *Panicum maximum*, and *Rottboellia exaltata*; artificial SCBV infection succeeded in *Oryza sativa*, *Musa* sp., and *Sorghum vulgare* (Bouhida et al., 1993; Lockhart et al., 1996; Viswanathan et al., 1996). The visual symptoms of SCBV infection in cultivated sugarcane are stunted growth, flecking, and chlorotic stripes, but depend on the combination of host clones and virus strains (Comstock and Lockhart, 1990; Lockhart et al., 1996; Viswanathan et al., 1996). SCBV infection may lead to significant losses in biomass production in susceptible sugarcane cultivars (Lockhart and Autrey, 2000).

SCBVs that infect sugarcane are a genetically heterogeneous *Badnavirus* species complex. At present, four badnaviruses, *Sugarcane bacilliform Guadeloupe A virus* (SCBGAV), *Sugarcane bacilliform Guadeloupe D virus* (SCBGDV), *Sugarcane bacilliform MO virus* (SCBMOV), and *Sugarcane bacilliform IM virus* (SCBIMV), which originally clustered in the SCBV-A, SCBV-D, SCBV-E, and SCBV-F phylogenetic groups/genotypes, respectively, have been assigned by the International Committee on Taxonomy of Viruses (ICTV) as different species in the *Badnavirus* genus (Adams and Carstens, 2012; Adams et al., 2016; Geering and Hull, 2012). A total of 12 full length genomes were obtained for global SCBV isolates from Morocco (SCBMOV-MOR) (Bouhida et al., 1993), Australia (SCBIMV-QLD) (Geijskes et al., 2002), Guadeloupe (SCBGAV-R570, SCBGAV-R51129, and SCBGDV-Batavia) (Muller et al., 2011), India (SCBV-BT, SCBV-BRU, SCBV-BO91, SCBV-Iscam, and SCBV-BB) (Karuppaiah et al., 2013), and China (SCBV-CHN1 and SCBV-CHN2) (Sun et al., 2016). The full SCBV genome is approximately 7.3-8.0 kilobase pair (kb) in size and consists of circular double-stranded DNA that replicate through a virus-encoded reverse transcriptase (RT). SCBV retains the Badnavirus genome organization which consists of three open reading frames (ORFs) on the positive strand; ORFs 1 and 2 encode two small proteins of 176-185 and 122-135 amino acids (aa), respectively, but the function of these proteins are unknown (Geering and Hull, 2012; Sun et al., 2016). ORF 3 encodes a polyprotein of 1786-1933 amino acids which includes several functional units, such as a movement protein, aspartic protease, coat protein, RT,

and ribonuclease H (RNase H) proteins (Bouhida et al., 1993; Geering and Hull, 2012; Geijskes et al., 2002).

Various PCR assays have been used to screen for badnaviruses with three previously-developed primer pairs, SCBV-F5/SCBV-R5 (Braithwaite et al., 1995), Badna-FP/Badna-RP (Yang et al., 2003), and SCBV-F/SCBV-R (Wu et al., 2016), targeted at the RT/RNase H coding region. The three primer sets have been used in SCBV detection assays and the resulting PCR products were cloned and sequenced to explore SCBV genetic diversity (Braithwaite et al., 1995; da Silva et al., 2015; Karuppaiah et al., 2013; Muller et al., 2011; Wu et al., 2016). Currently, 18 phylogenetic groups (genotypes) of SCBV have been reported worldwide (SCBV-A to SCBV-R), and they are to some extent associated with geographical regions. Five SCBV genotypes, SCBV-A, SCBV-B, SCBV-C, SCBV-D, and SCBV-G, are present in France (Muller et al., 2011), seven SCBV genotypes, SCBV-E, SCBV-H, SCBV-I, SCBV-J, SCBV-K, SCBV-L, and SCBV-M, have been detected in India (Karuppaiah et al., 2013; Rao et al., 2014), and five SCBV genotypes, SCBV-A, SCBV-C, SCBV-F, SCBV-M, and SCBV-H, have been found in the Brazilian sugarcane germplasm collection (da Silva et al., 2015). In China, eight SCBV genotypes (SCBV-G, SCBV-H, SCBV-L, and from SCBV-N to SCBV-R) have been reported, and the two prevalent genotypes (SCBV-Q and SCBV-R) were found in sugarcane planting regions (Wu et al., 2016).

Some previous reports have described the occurrence and prevalence of SCBVs in China (Cai et al., 2009; Li et al., 2010; Wu et al., 2016; Xu et al., 2015); in the present study, two novel SCBV genotypes are proposed and two SCBV species are reported for the first time in China. Our findings will contribute to a larger understanding of the genetic diversity of SCBVs, and will also aid in the screening of sugarcane germplasm to minimize the spread of viral variants across geographical boundaries via contaminated plant material.

Materials and Methods

Leaf tissue collection. A total of 392 sugarcane leaf samples were collected from three sources; vegetative cuttings exchanged for national sugarcane regional tests in Fuzhou, Fujian province, varieties in commercial fields in Yunnan province, and a wild germplasm collection in Hainan province, China from 2015 to 2017. Initially, 114 leaf samples (three randomly-chosen leaves from each clone) were collected from 38 sugarcane clones that were originally introduced from different Chinese sugarcane

research institutes into Fuzhou, Fujian province in 2015. The second set of 216 leaf samples (six leaves per clone) were collected from the wild germplasm collection in Sanya, Hainan province in 2016, and the third set of 62 leaf samples (one leaf per cultivar in a field) was collected from three sugarcane planting regions (Baoshang, Dehong, and Lincang) in Yunnan province in 2017. The leaf samples are described in Supplementary Table 1. The leaf samples were rinsed with 75% ethanol and stored at -80°C until they were used for DNA extraction.

DNA isolation and PCR amplification. Total DNA was extracted from sugarcane leaf samples using the cetyltrimethylammonium bromide (CTAB) method (Haible et al., 2006). A set of degenerate primers SCBV-F and SCBV-R were used to amplify the RT/RNase H genomic region of SCBV (Wu et al., 2016). Total leaf DNA (200 ng/ μl) was used as a template for PCR amplification with *EX* Taq polymerase (TaKaRa, Dalian, China) under the following conditions: an initial denaturation step at 94°C for 5 min, followed by 35 cycles of 94°C for 30 s, 58°C for 30 s, 72°C for 45 s, and a final extension step at 72°C for 10 min. Amplified DNA was electrophoresed on 1.5% agarose gels, stained with ethidium bromide, and then visualized with a ChemiDoc XRS instrument (Bio-Rad Laboratories, Hercules, CA, USA) to confirm that the amplified DNA fragments were of the expected size of 726 base pair (bp).

Cloning and sequencing of PCR products. All amplified DNA fragments obtained from each leaf sample were purified using the E.Z.N.A. Gel Extraction Kit (Omega Bio-Tek Inc., Georgia, USA) and then cloned into the pMD19-T vector (TaKaRa, Dalian, China) following the manufacturer's instructions. The ligation reactions were used to transform competent *E. coli* DH5a cells (Tiangen Biotech Co. Ltd., Beijing, China). Three positive colonies per PCR product were randomly selected and the nucleotide (nt) sequence was determined in both directions by Shanghai Shenggong Biology Company (China). The sequence contigs of the three colonies from each leaf DNA sample were compared among themselves using DNAMAN (version 8). An accurate 'unique' sequence was obtained when the sequences of the three colonies from each leaf sample showed $\geq 99\%$ similarity.

DNA sequence alignment and analysis. One-hundred eleven DNA sequences of 726 bp in length were obtained in this study. Thirty-four reference sequences that are representative of 18 different worldwide SCBV genotypes (Wu et al., 2016) were downloaded from the NCBI

GenBank database. All the DNA sequences were trimmed at both ends to a length of 480 nucleotides to accommodate a considerable number of GenBank sequences in the analyses, and then were subjected to multiple alignments using CLUSTALW as implemented in MEGA6 (Tamura et al., 2013).

Recombination events were detected in the dataset of aligned RT/RH sequences with seven algorithms implemented in RDP v.3.44 software (Martin et al., 2010) using the default parameters. The neighbor-joining (NJ) phylogenetic tree, which included a corresponding sequence from an isolate of *Commelina yellow mottle virus* (ComYMV, GenBank no. NC001343) as an outgroup in the alignment, was constructed with the MEGA6 program. The robustness of the trees was determined via bootstrap analysis (1,000 replicates), and the bootstrap values were included at the internodes on the tree (Felsenstein, 1985). Pairwise sequence comparison (PASC) was carried out with the MEGA6 (Tamura et al., 2013) and BioEdit programs to estimate the genetic distances and pairwise identities of the nucleotide and amino acid sequences, respectively.

Results

PCR detection of SCBV in leaf samples. Of all the 392 sugarcane leaf samples, 16.1% (63/392) were scored as positive for SCBV with the SCBV-F/SCBV-R primer pair; 38.6% (44/114) of the *Saccharum* spp. hybrid varieties from Fujian, 8.1% (5/62) of the *Saccharum* spp. hybrid varieties from Yunnan, and 6.5% (14/216) of the wild germplasm accessions from Hainan (Table 1). Only five of 36 wild germplasm clones were detected as positive in Striped Cheribon (*S. officinarum*), Sewari (*S. robustum*), Yacheng11 (*S. spontaneum*), Guangdong29 (*S. spontaneum*), and Hainan92-79 (*Erianthus arundinaceus*). All PCR products amplified from SCBV-positive leaf DNA samples were cloned and sequenced. The SCBV sequences obtained in this study were submitted to the NCBI GenBank database (Supplementary Table 1).

Phylogenetic analysis. Because recombination may confound virus phylogenetic analyses, the aligned nucleotide sequences were subjected to recombination analysis to identify conflicting phylogenetic signals. No significant recombination events were detected using seven algorithms. Thus, the dataset of SCBV sequences and one sequence of ComYMV was used for phylogenetic analysis. The neighbor-joining phylogenetic tree revealed that all of the sequences in this study could be classified

Table 1. The occurrence of sugarcane bacilliform viruses in the Chinese sugarcane-producing regions of Fujian and Yunnan provinces and in a germplasm resource nursery in Hainan province

Province/Location	Species	No. of leaf samples tested	SCBV-positive leaf samples detected by PCR	No. of sequence variants ^a
Commercial sugarcane fields				
Fujian/Fuzhou	<i>Saccharum</i> hybrids	114	45 (39.5%)	67
Yunnan/Dehong	<i>Saccharum</i> hybrids	13	1 (7.7%)	3
Yunnan/Baoshang	<i>Saccharum</i> hybrids	8	0 (0)	0
Yunnan/Lincang	<i>Saccharum</i> hybrids	41	4 (9.8%)	10
Sugarcane germplasm resource nursery				
Hainan/Yacheng	<i>Erianthus arundinaceus</i>	36	5 (13.9%)	11
	<i>S. barberi</i>	36	4 (11.1%)	10
	<i>S. officinarum</i>	36	2 (5.6%)	4
	<i>S. robustum</i>	36	0 (0)	0
	<i>S. sinense</i>	36	0 (0)	0
	<i>S. spontaneum</i>	36	3 (8.3%)	6

^aSequences derived from three separate colonies per DNA amplicon of different isolates were aligned and considered sequence variants when they differed by at least one nucleotide.

into 20 SCBV genotypes (from SCBV-A to T) including the 18 previously-published SCBV genotypes (Fig. 1). We identified two novel SCBV genotypes, SCBV-S and SCBV-T, in this study. Twenty-four and 32 sequences from the Fuzhou isolates clustered into the SCBV-E and SCBV-F groups, respectively, and the reminding 11 sequences clustered into the SCBV-L, -N, and -Q groups. Thirteen sequences from the Yunnan isolates clustered into four groups (SCBV-L, SCBV-Q, SCBV-R, and SCBV-S). Notably, 27 sequences from the Hainan isolates infecting *S. robustum*, *S. spontaneum*, and *E. arundinaceus* were clustered into the two newly identified groups called SCBV-S and SCBV-T in our study, whereas the four isolates (HNSo1, HNSo2, HNSo3, and HNSo4) from *S. officinarum* clustered with the SCBV-G group isolates (Fig. 1). The geographical distribution of SCBV genotypes from this study and previous reports in China and other countries is shown in Fig. 2. Apart from SCBV-S and SCBV-T, another two SCBV groups (SCBV-E and SCBV-F) that were assigned to the *Badnavirus* species SCBMOV and SCBIMV by ICTV, respectively, are reported for the first time in China. Besides, SCBV-G and SCBV-N groups were firstly reported in Hainan and Fujian provinces, China, respectively.

Pairwise sequence identity between the SCBV groups.

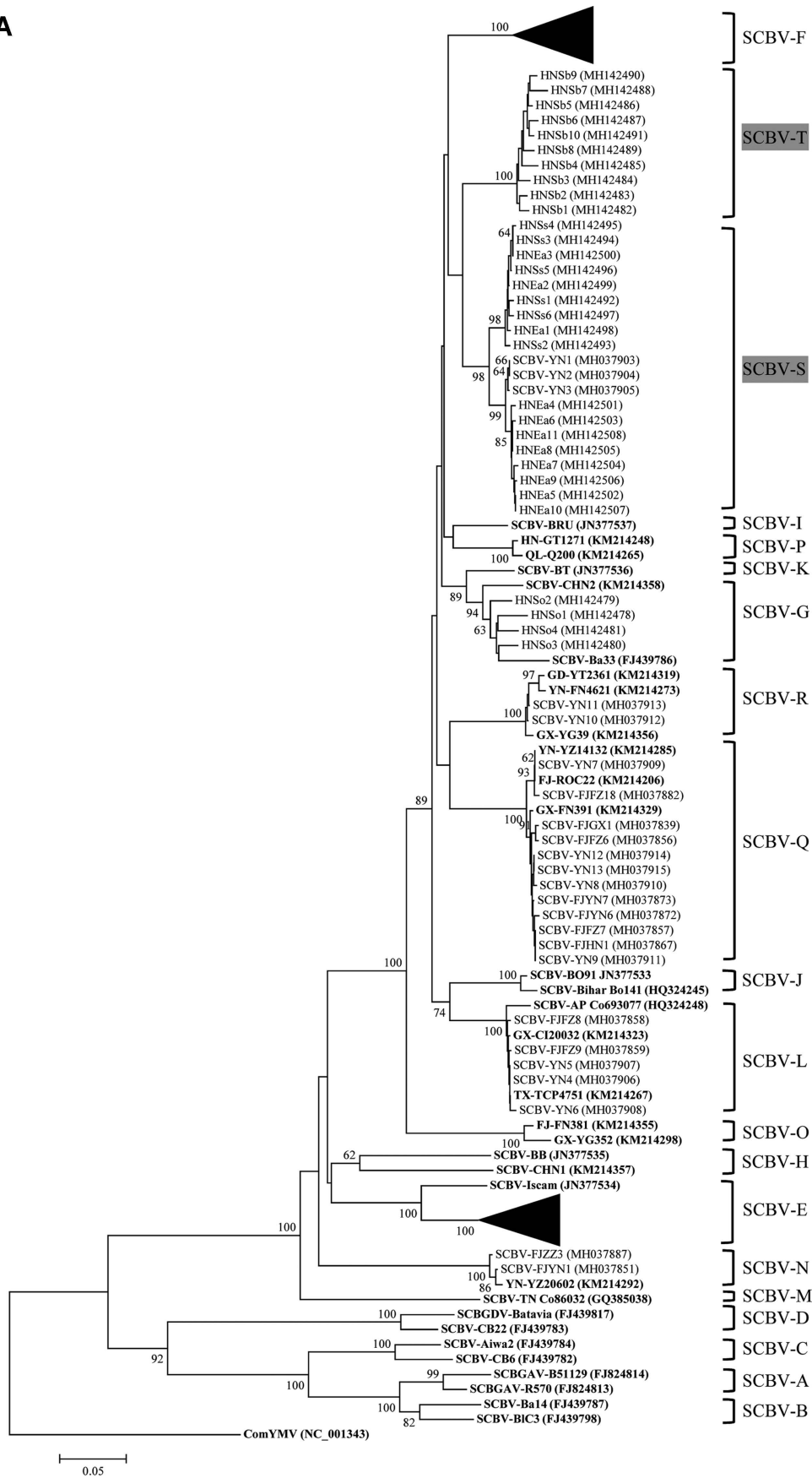
Pairwise sequence identity between the 20 phylogenetic groups (from SCBV-A to SCBV-T) ranged from 50.6-93.2% and 53.0-99.3% at the nucleotide and amino acid levels, respectively (Table S2). The SCBV-S and SCBV-T

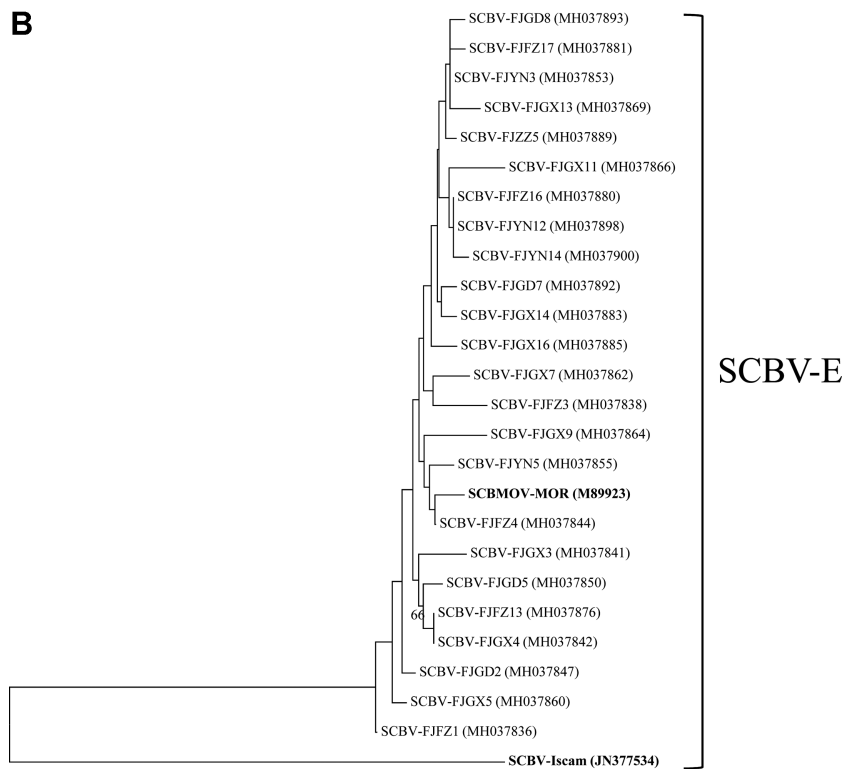
groups shared sequence identities of 90.0-93.2% (nt) and 92.7-96.9% (aa). In addition, SCBV-S shared 57.6-92.2% (nt) and 53.0-99.3% (aa) identity, whereas SCBV-T shared 58.4-90.0% (nt) and 53.0-96.9% (aa) identity with the other 18 SCBV genotypes (Table S2). The nt sequences of the Fuzhou (Fujian province) isolates shared 75.0-100% identity with each other and shared 59.8-62.4%, 59.4-62.4%, 77.0-99.8%, and 76.2-99.6% identity with four ICTV-recognized *Badnavirus* species isolates, SCBGAV, SCBGDV, SCBMOV, and SCBIMV, respectively (Table 2). The Yunnan province isolates shared 86.4-100% nt identity with one another and 59.2-61.2%, 59.2-61.4%, 76.8-79.2%, and 87.4-90.6% identity with the SCBGAV, SCBGDV, SCBMOV, and SCBIMV isolates, respectively. The Hainan province isolates shared 86.8-100% nt identity with each other and had nt identities of 58.4-61.4%, 59.4-61.6%, 76.6-79.8%, and 88.2-90.6% with the SCBGAV, SCBGDV, SCBMOV, and SCBIMV isolates, respectively.

Pairwise nucleotide sequence genetic distances between the SCBV groups.

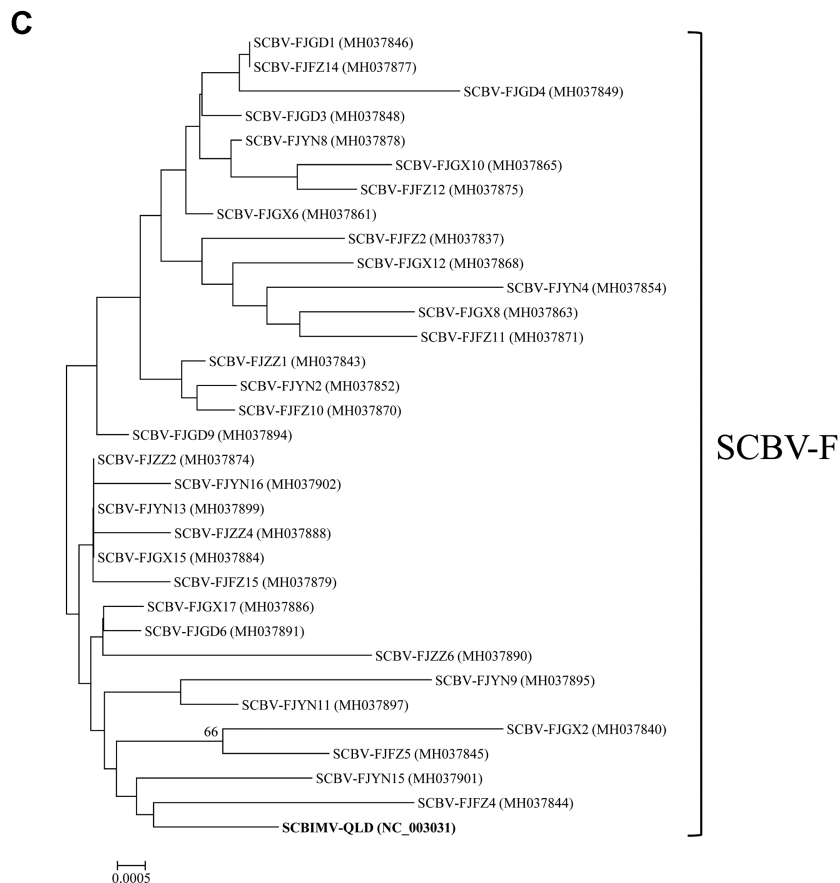
Pairwise genetic distances between the 20 phylogenetic groups ranged from 0.07-0.67 (nt) and 0-0.64 (aa) (Table S3). The genetic distance between SCBV-S and SCBV-T ranged from 0.07-0.11 and 0.03-0.08 at the nt and aa levels, respectively. SCBV-S showed genetic distances of 0.08-0.66 (nt) and 0.01-0.64 (aa), whereas SCBV-T showed distances of 0.11-0.63 (nt) and 0.03-0.64 (aa) with the other 18 SCBV genotypes (Table S3). Genetic distances between the individual Fuzhou isolates ranged from 0.00 to 0.31 (nt), while the distances between

A





SCBV-E



SCBV-F

Fig. 1. Phylogenetic analysis of sugarcane bacilliform virus (SCBV) isolates based on the partial DNA sequences of the reverse transcriptase/ribonuclease H (RT/RNase H) genomic region (480 nt). (a) The phylogeny was constructed by using the neighbor-joining (NJ) with the Tamura-Nei model. The bootstrap consensus tree was inferred from 1,000 replicates, and branches present in < 60% of the bootstrap replicates were collapsed. The analysis included 146 SCBV nucleotide sequences; 111 SCBV nucleotide sequences from this study, 34 SCBV nucleotide sequences from GenBank (bold type), and one corresponding nucleotide sequence of ComYMV (*Commelina yellow mottle virus*) as an outgroup. Newly-identified SCBV groups (genotypes) are denoted by grey boxes. Phylogenetic trees of the SCBV-E (b) and SCBV-F (c) groups were also constructed using the NJ method.

Genotype	China						Australia	Brazil	France	India	Morocco	USA
	FJ	GD	GX	HN	JX	YN						
SCBV-A												
SCBV-B												
SCBV-C												
SCBV-D												
SCBV-E												
SCBV-F												
SCBV-G												
SCBV-H												
SCBV-I												
SCBV-J												
SCBV-K												
SCBV-L												
SCBV-M												
SCBV-N												
SCBV-O												
SCBV-P												
SCBV-Q												
SCBV-R												
SCBV-S												
SCBV-T												

^aNew sugarcane bacilliform virus (SCBV) genotypes occurred in this study were shown with green boxes. The published sugarcane bacilliform virus genotypes in previous studies (Muller et al., 2011; Karuppaiah et al., 2013; Silva et al., 2015; Wu et al., 2016) were shown with grey boxes. The Chinese SCBV sequences obtained in Wu et al. (2016) and this study showed as following: FJ (Fujian) = 81, GD (Guangdong) = 12, GX (Guangxi) = 47, HN (Hainan) = 37, JX (Jiangxi) = 9, and YN (Yunnan) = 40.

Fig. 2. Geographical distribution of sugarcane bacilliform virus phylogenetic groups (genotypes) in China and other countries.

the Fuzhou isolates and the SCBGAV, SCBGDV, SCBMOV and SCBIMV isolates were 0.52-0.60, 0.52-0.59, 0.00-0.28, and 0.00-0.29, respectively (Table 2). The Yunnan province isolates showed 0.00-0.15 nt genetic distance with each other, and showed genetic distances of 0.55-0.60, 0.54-0.59, 0.25-0.29, and 0.10-0.14, respectively, with the SCBGAV, SCBGDV, SCBMOV, and SCBIMV isolates. Genetic distances within the Hainan province isolates were 0.00-0.15 (nt), and were 0.55-0.62, 0.54-0.58, 0.24-0.29, and 0.10-0.13 between the Hainan isolates and SCBGAV, SCBGDV, SCBMOV, and SCBIMV, respectively (Table 2).

Discussion

Sugarcane is extensively cultivated worldwide, and is both an important sugar-producing crop and biofuel feedstock. The global exchange of germplasm has occurred frequently in an effort to broaden the genetic base of the pool of sugarcane hybrid parents, but this has also increased the risk of long-distance virus transmission across geographical barriers resulting from the transport of SCBV-infected vegetative cuttings. Sugarcane-infecting badnaviruses present significant genomic nucleotide sequence variability and serological heterogeneity (Lockhart et al., 1996; Sun et al., 2016). Hence, in this study, we further investigated the genetic diversity of the genetically heterogeneous SCBV complex based on three sets of sugarcane leaf samples from Fujian, Yunnan, and Hainan provinces in China, even though several studies have already reported the existence

of diverse SCBVs infecting sugarcane in China (Li et al., 2010; Wu et al., 2016; Xu et al., 2015).

The RT/RNase H coding region is a common taxonomic marker for species demarcation within the genus *Badnavirus* (Geering and Hull, 2012). The demarcation criterion is > 20% nt variation in the RT/RNase H region, which was proposed by the ICTV (Geering and Hull, 2012). Also, the threshold for nt genetic distance (> 0.2) in the RT/RNase H region between species in the *Badnavirus* genus is another demarcation criteria (Bousalem et al., 2008). Genetic variation and genetic distance within the RT/RNase H regions of the two novel SCBV groups, SCBV-S and SCBV-T, were compared with the representative isolates of four *Badnavirus* species, SCBGAV, SCBGDV, SCBIMV-QLD, and SCBMOV, and showed that the SCBV-S and SCBV-T groups had > 20% nt variation when compared with SCBMOV, SCBGAV, and SCBGDV, but had only 9.4-11.8% nt variation compared to SCBIMV. Similarly, the nt genetic distances between the SCBV-S and SCBV-T group isolates were > 0.2 when compared with the four SCBV species except for SCBIMV. It is noteworthy that the SCBV-N isolate group met both of the criteria for species in the *Badnavirus* genus, indicating that SCBV-N may well be a novel SCBV species (Wu et al., 2016). The two SCBV species SCBMOV (SCBV-E group) and SCBIMV (SCBV-F group) were assigned in 2012 (Adams and Carstens, 2012), and the other SCBV species SCBGAV (SCBV-A group) and SCBGDV (SCBV-D group) were assigned more recently in 2016 by ICTV (Adams et al.,

Table 2. Nucleotide identities (% , lower left) and genetic distances (upper right) among and between sugarcane bacilliform virus (SCBV) phylogenetic groups from Fujian, Yunnan, and Hainan provinces based on the 480 nucleotide sequences of RT/RNase H regions^a

Prov- ince	Group	SCB- MOV			SCBIMV			Fujian			Yunnan			Hainan		
		SCBGAV	SCBGDV	SCBMOV	SCB-E	SCB-F	SCB-L	SCB-N	SCB-Q	SCB-L	SCB-Q	SCB-R	SCB-S	SCB-G	SCB-S	SCB-T
Fujian	SCBGAV	100 (0)	0.42	0.54	0.52-0.54	0.55-0.58	0.55	0.59-0.60	0.55-0.57	0.55	0.56-0.56	0.56-0.57	0.60	0.57-0.60	0.57-0.62	0.55-0.57
	SCBGDV	68.2	100 (0)	0.56	0.55-0.57	0.52-0.54	0.57	0.54	0.58-0.59	0.57-0.58	0.58-0.59	0.54	0.58	0.54	0.55-0.58	0.54-0.56
	SCBMOV	61.8	60.8	100 (0)	0.00-0.01	0.25-0.26	0.00-0.28	0.24	0.24-0.26	0.28	0.25	0.26	0.28-0.29	0.24-0.26	0.27-0.29	0.26-0.28
	SCBIMV	61.2	62.0	78.4	100 (0)	0.00-0.01	0.12-0.13	0.29	0.14	0.13	0.14	0.13	0.10-0.11	0.11-0.13	0.10-0.11	0.11-0.13
Yunnan	SCBV-E	61.6-62.4	60.4-61.2	99.2-99.8	78.2-78.8	99.0-100	0.25-0.27	0.27-0.28	0.24-0.25	0.24-0.26	0.27-0.28	0.24-0.26	0.25-0.26	0.27-0.29	0.23-0.27	0.25-0.29
	SCBV-F	60.4-61.2	61.6-62.4	78.2-79.0	99.0-99.6	78.0-79.2	98.6-100	0.12-0.13	0.28-0.29	0.13-0.15	0.12-0.13	0.13-0.14	0.13-0.14	0.10-0.11	0.10-0.13	0.09-0.11
	SCBV-L	61.2-61.4	59.8-59.6	77.0-77.2	88.4-88.6	76.8-77.6	88.2-88.8	99.6-100	0.26	0.14-0.15	0.00-0.01	0.14-0.15	0.13-0.14	0.09-0.10	0.13-0.15	0.08-0.10
	SCBV-N	59.8-59.6	61.4-61.6	79.6-79.8	76.2-76.4	79.4-80.2	76.2-77.0	78.0-78.2	99.4-100	0.30-0.31	0.01	0.29-0.31	0.27	0.29	0.23-0.26	0.28-0.30
Hainan	SCBV-Q	60.8-61.4	59.4-59.8	78.4-79.4	87.2-87.6	78.2-79.6	87.0-88.0	86.4-87.0	75.0-76.0	98.6-100	0.14-0.16	0.00-0.01	0.12-0.13	0.12-0.13	0.12-0.14	0.12-0.14
	SCBV-L	61.0-61.2	59.2-59.6	76.8-77.2	88.0-88.4	76.7-77.8	87.8-89	99.2-99.8	77.4-78.4	86.2-87.0	99.6-100	0.14-0.15	0.13-0.14	0.09-0.10	0.13-0.16	0.09-0.10
	SCBV-Q	60.8-61.0	59.4-59.6	78.8-79.2	87.4-87.6	78.4-79.6	87.0-88.2	86.6-87.0	75.4-76.2	98.6-100	86.4-87.2	98.8-100	0.12-0.13	0.12-0.13	0.12-0.14	0.12-0.14
	SCBV-R	60.8-61.0	61.2-61.4	78.2-78.6	88.0-88.2	78.0-78.8	87.6-88.6	87.8-88.0	77.4-77.8	88.4-89.0	87.6-88.0	88.8-89.2	99.6-100	0.12-0.13	0.13-0.14	0.10-0.13
Yunnan	SCBV-S	59.2	59.4	77.4-77.6	90.6	77.2-78.0	90.4-91.2	91.4-91.6	76.4-77.0	88.4-89.2	91.2-91.8	88.6-89.2	89.0-89.4	99.8-100	0.10-0.11	0.00-0.09
	SCBV-G	59.4-60.4	61.2-61.4	78.2-79.8	88.6-90.2	77.8-80.0	88.4-90.8	86.2-88.2	78.4-80.4	87.2-89.0	86.0-88.2	87.2-88.8	87.6-88.8	90.0-90.8	96.4-100	0.10-0.12
	SCBV-S	58.4-60.6	59.4-61.0	76.6-77.4	90.2-90.6	76.2-77.6	90.0-91.2	91.0-92.2	76.0-77.2	87.6-89.0	90.6-92.0	88.0-88.8	88.6-90.4	96.6-99.4	89.4-90.8	96.2-100
	SCBV-T	60.4-61.4	60.6-61.6	76.6-78.0	88.2-89.6	76.4-78.8	87.8-90	88.0-89.6	75.8-77.7	87.4-89.4	87.6-89.4	87.6-89.4	86.0-87.2	87.6-89.4	86.8-88.8	90.0-93.2

^aGenetic distances within SCBV phylogenetic groups are shown in parentheses.

2016). In addition, more than four SCBV species, including SCBV-C in Guadeloupe (Muller et al., 2011), SCBV-H, SCBV-I, and SCBV-J in India (Karuppaiah et al., 2013) have been proposed. However, formal recognition of these *Badnavirus* SCBV species is still pending.

In our previous studies, five genotypes of SCBV (SCBV-H, SCBV-L, SCBV-O, SCBV-Q, and SCBV-R) were reported in Fujian province, six genotypes (SCBV-G, SCBV-L, SCBV-N, SCBV-P, SCBV-Q, and SCBV-R) in Yunnan province, and three genotypes (SCBV-P, SCBV-Q, and SCBV-R) in Hainan province (Sun et al., 2016; Wu et al., 2016). However, in this study we report nine SCBV genotypes in Fujian, Yunnan, and Hainan provinces. Of the nine SCBV genotypes, we reported for the first time that SCBMOV (SCBV-E) and SCBIMV (SCBV-F) occur in China, and that SCBV-N is present in Fujian province. The two SCBV species (SCBMOV and SCBIMV) and the one genotype (SCBV-N) may have been imported from other Chinese provinces because these leaf samples were collected from sugarcane clones that were introduced into Fuzhou from other provinces as SCBV-infected vegetative cuttings imported for national sugarcane regional tests during 2013-2015. Another possibility is that they could have been infected by existing SCBVs in Fuzhou during the cultivation period. Also, the two novel genotypes (SCBV-S and SCBV-T) and group SCBV-G that once occurred in Yunnan province were detected in the wild germplasm collection in Hainan province for the first time. Most of the sequences from the wild germplasm clones clustered with the two newly-identified groups SCBV-S and SCBV-T in our study. Exceptions were HNSo1, HNSo2, HNSo3, and HNSo4, which infect *S. officinarum* and were included in the SCBV-G group, suggesting that host plants could be an important factor for SCBV genetic differentiation.

Highly divergent SCBV variants have been reported worldwide. Genetic recombination, as an evolutionary force, may have contributed to the emergence of genetic diversity in SCBV. Stronger evidence of recombination was detected within and between SCBV groups in the RT/RNase H gene coding region (Wu et al., 2016), and also at the genome level (Sun et al., 2016), and was even detected between SCBV and banana streak virus (BSV) (Sharma et al., 2015). However, only very weak recombination signals were detected in some SCBV sequences analyzed in this study. It has been reported that SCBV can naturally or experimentally infect banana, which is a natural host plant for BSV (Bouhida et al., 1993). Thus, it is possible that recombination events between SCBV and BSV could have occurred prior to the time of divergence from the

last common ancestor of the currently-known banana and sugarcane badnaviruses (Sharma et al., 2015). It is notable that some SCBV genotypes share a high degree of similarity with some BSV species (Iskra-Caruana et al., 2014; Muller et al., 2011; Wu et al., 2016). In addition to recombination, repeated bottleneck events such as horizontal vector-mediated transmission, cell-to-cell movement, and systemic spread within the host plant could also contribute to dynamic SCBV populations (Wu et al., 2016). All of the published studies suggest a complicated phylogenetic makeup of SCBV isolates worldwide, and also show that SCBV isolates are extremely diverse.

The results presented here provide a more complete picture of the prevalence, distribution, and genetic diversity of the SCBV species complex in China. In addition, our findings will play a key role in understanding the population structure of SCBV in semi-perennial and perennial host plants, and will also aid in the screening of sugarcane germplasm worldwide to reduce the spread of genetic variants of SCBV via contaminated plant material. With the continued discovery of new SCBV genotypes, it is important that the phylogenetic groups (genotypes) be identified to the level of species or strains in the genus *Badnavirus*.

Acknowledgments

This work was supported by the Sugar Crop Research System, CARS (No. CARS-170302) and the Major Science and Technology Project of Fujian Province, China (No. 2015NZ0002-2).

Conflict of Interest

There is no conflict of interest related to this research.

References

- Adams, M. J. and Carstens, E. B. 2012. Ratification vote on taxonomic proposals to the international committee on taxonomy of viruses (2012). *Arch. Virol.* 157:1411-1422.
- Adams, M. J., Lefkowitz, E. J., King, A. M., Harrach, B., Harrison, R. L., Knowles, N. J., Kropinski, A. M., Krupovic, M., Kuhn, J. H., Mushegian, A. R., Nibert, M., Sabanadzovic, S., Sanfaçon, H., Siddell, S. G., Simmonds, P., Varsani, A., Zerbini, F. M., Gorbalenya, A. E. and Davison, A. J. 2016. Ratification vote on taxonomic proposals to the international committee on taxonomy of viruses (2016). *Arch. Virol.* 161:2921-2949.
- Autrey, L., Boolell, S., Jones, P., Lockhart, B. and Nadif, A. 1995. Distribution of sugarcane bacilliform virus in various

- geographical regions. In: *Proceedings of XXI International Society of Sugarcane Technologists*, eds. by B. Napompeth and P. Wisarath, pp. 527-541. Kasetsart University, Bangkok, Thailand.
- Bouhida, M., Lockhart, B. E. L. and Olszewski, N. E. 1993. An analysis of the complete sequence of a sugarcane bacilliform virus genome infectious to banana and rice. *J. Gen. Virol.* 74:15-22.
- Bousalem, M., Douzery, E. J. and Seal, S. 2008. Taxonomy, molecular phylogeny and evolution of plant reverse transcribing viruses (family *caulimoviridae*) inferred from full-length genome and reverse transcriptase sequences. *Arch. Virol.* 153:1085-1102.
- Braithwaite, K. S., Egeskov, N. M. and Smith, G. R. 1995. Detection of sugarcane bacilliform virus using the polymerase chain reaction. *Plant Dis.* 79:792-796.
- Cai, Y. Q., Xu, D. L. and Zhou, G. H. 2009. Evidence of sugarcane bacilliform virus DNA fragment integrated into the *saccharum* inter-specific hybrids genome. *J. South China Agric. Univ.* 30:19-23 (in Chinese).
- Comstock, J. C. and Lockhart, B. E. L. 1990. Widespread occurrence of sugarcane bacilliform virus in U.S. sugarcane germ plasm collections. *Plant Dis.* 74:530.
- D'Hont, A., Grivet, L., Feldmann, P., Glaszmann, J. C., Rao, S. and Berding, N. 1996. Characterisation of the double genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. *Mol. Gen. Genet.* 250:405-413.
- da Silva, J. M., Jobim, L. J., Ramos-Sobrinho, R., Lima, J. S., Assunção, I. P., Cruz, M. M. and Lima, G. S. 2015. Incidence and species diversity of badnaviruses infecting sugarcane from a germplasm collection in Brazil. *Trop. Plant Pathol.* 40:212-217.
- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783-791.
- Geering, A. D. W. and Hull, R. 2012. Family *Caulimoviridae*. In: *Virus Taxonomy: Ninth report of the international committee on taxonomy of viruses*, eds. by A. M. Q. King, M. J. Adams, E. B. Carestens and E. J. Lefkowitz, pp. 424-443. Elsevier Academic Press, San Diego, CA, USA.
- Geijskes, R. J., Braithwaite, K. S., Dale, J. L., Harding, R. M. and Smith, G. R. 2002. Sequence analysis of an Australian isolate of sugarcane bacilliform badnavirus. *Arch. Virol.* 147:2393-2404.
- Haible, D., Kober, S. and Jeske, H. 2006. Rolling circle amplification revolutionizes diagnosis and genomics of geminiviruses. *J. Virol. Methods* 135:9-16.
- Iskra-Caruana, M. L., Chabannes, M., Duroy, P. O. and Muller, E. 2014. A possible scenario for the evolution of Banana streak virus in banana. *Virus Res.* 186:155-162.
- Karupaiyah, R., Viswanathan, R. and Kumar, V. G. 2013. Genetic diversity of sugarcane bacilliform virus isolates infecting *saccharum* spp. in India. *Virus Genes.* 46:505-516.
- Li, W. F., Huang, Y. K., Jiang, D. M., Zhang, Z. X., Zhang, B. L. and Li, S. F. 2010. Detection of sugarcane bacilliform virus isolate and its influence on yield and quality of cane in Yunan. *Acta Phytopathol. Sinica* 40:651-654 (in Chinese).
- Lockhart, B. E. L., Ireby, M. J. and Comstock, J. C. 1996. Sugarcane bacilliform virus, Sugarcane mild mosaic virus, and sugarcane yellow leaf syndrome. In: *Sugarcane Germplasm Conservation and Exchange*, eds. by B. J. Croft, C. T. Pigggin, E. S. Wallis and D. M. Hogarth, pp. 108-112. Australian Centre for International Agricultural Research (ACIAR), Brisbane, Queensland, Australia.
- Lockhart, B. E. L. and Autrey, L. J. C. 2000. Sugarcane bacilliform virus. In: *A Guide to Sugarcane Diseases*, eds. by P. Rott, R. A. Bailey, J. C. Comstock, B. J. Croft, A. S. and Saumtally, pp. 268-272. CIRAD and ISSCT Press, Montpellier, France.
- Martin, D. P., Lemey, P., Lott, M., Moulton, V., Posada, D. and Lefevre, P. 2010. RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* 26:2462-2463.
- Muller, E., Dupuy, V., Blondin, L., Bauffe, F., Daugrois, J. H., Nathalie, L. and Iskra-Caruana, M. L. 2011. High molecular variability of sugarcane bacilliform viruses in Guadeloupe implying the existence of at least three new species. *Virus Res.* 160:414-419.
- Rao, G. P., Sharma, S. K., Singh, D., Arya, M., Singh, P. and Baranwal, V. K. 2014. Genetically diverse variants of sugarcane bacilliform virus infecting sugarcane in India and evidence of a novel recombinant *Badnavirus* variant. *J. Phytopathol.* 162:779-787.
- Sharma, S. K., Vignesh, K. P., Geetanjali, A. S., Pun, K. B. and Baranwal, V. K. 2015. Subpopulation level variation of banana streak viruses in India and common evolution of banana and sugarcane badnaviruses. *Virus Genes* 50:450-465.
- Sun, S. R., Damaj, M. B., Alabi, O. J., Wu, X. B., Mirkov, T. E., Fu, H. Y., Chen, R. K. and Gao, S. J. 2016. Molecular characterization of two divergent variants of sugarcane bacilliform viruses infecting sugarcane in China. *Eur. J. Plant Pathol.* 145:375-384.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30:2725-2729.
- Viswanathan, R., Alexander, K. C. and Garg, I. D. 1996. Detection of sugarcane bacilliform virus in sugarcane germplasm. *Acta Virol.* 40:5-8.
- Wu, X. B., Alabi, O. J., Damaj, M. B., Sun, S. R., Mirkov, T. E., Fu, H. Y., Chen, R. K. and Gao, S. J. 2016. Prevalence and RT/RNase H genealogy of sugarcane bacilliform virus isolates from China. *J. Phytopathol.* 164:595-607.
- Xu, Z. Y., Zhou, L. W., Tang, Y., Lü, B. L., Gui, C. Y., Bai, R. M., Wen, R. H. and Chen, B. S. 2015. Molecular detection and distribution of sugarcane bacilliform virus in Guangxi. *J. Southern Agric.* 46:1980-1984 (in Chinese).
- Yang, I. C., Hafner, G. J., Revill, P. A., Dale, J. L. and Harding, R. M. 2003. Sequence diversity of South Pacific isolates of *Taro bacilliform virus* and the development of a PCR-based diagnostic test. *Arch. Virol.* 148:1957-1968.