Scientific Research Publishing

Advances in Bioscience and Biotechnology, 2023, 14, 120-141
https://www.scirp.org/journal/abb
ISSN Online: 2156-8502
ISSN Print: 2156-8456

# Genetic Diversity Analysis of Badnaviruses Infecting Banana in Burkina Faso

**Bakary Ouattara[1,2], Drissa Sérémé[1*], Léon Wêndé-M'Minèré Nitiema[1], Issa Wonni[3], Kadidia Koita[2]**

[1]Laboratoire de Protection et Défense des Cultures, Institut de l'Environnement et de Recherches Agricoles (INERA), Centre National de Recherche Scientifique et Technologique (CNRST), Ouagadougou, Burkina Faso
[2]Laboratoire de Biosciences, Université Joseph KI-ZERBO, Ouagadougou, Burkina Faso
[3]Laboratoire Central d'Horticulture, Institut de l'Environnement et de Recherches Agricoles (INERA), Centre National de Recherche Scientifique et Technologique (CNRST), Bobo Dioulasso, Burkina Faso
Email: *sereme@gmail.com

**How to cite this paper:** Ouattara, B., Sérémé, D., Nitiema, L.W.-M., Wonni, I. and Koita, K. (2023) Genetic Diversity Analysis of Badnaviruses Infecting Banana in Burkina Faso. *Advances in Bioscience and Biotechnology*, **14**, 120-141.
https://doi.org/10.4236/abb.2023.143008

**Received:** January 15, 2023
**Accepted:** March 27, 2023
**Published:** March 30, 2023



## Abstract

Badnaviruses are serious plant pararetroviruses affecting banana and causes serious economic losses to banana production worldwide. This study aims to examine the variability of BSV and SCBV nature infecting banana in Burkina Faso. Polymerase Chain Reaction (PCR) used the Badna FP/RP specific primers for the RT/RNase H regions present in badnaviruses. The PCR yielded about 579 bp amplicons from banana infected by BSV and SCBV. The 38 BSV isolates recorded low nucleotide identity ranging from 58.9% - 98.1%. Based on percentage nucleotide sequence identity and phylogenetic analyse, BSV isolates were categorized into four groups: 1, 2, 3 and 4. Group 4 shared 76.9% - 100% identity with BSOL species. However, Groups 1 and 3 recorded a low identity ranging, from 76.8% - 79.2%, 68.8% - 79.7% with BSCV, and 72.8% - 79.0% between Group 2 and BSOLV. Groups 1, 2 and 3 were assigned to a potentially new BSV species. The two SCBV isolates recorded a low nucleotide identity of 68% among themselves indicating high diversity. In addition, SCBV_Cd and SCBV_CE showed high nucleotide identity 91.3% and 58.7% with SCBV_C and SCBV, when they were compared to all published SCBV genotypes. In addition, phylogenetic analysis revealed the segregation of SCBV isolates into two genotypes, SCBV_Cd in C and SCBV_CE segregated in a new genotype namely Z. Recombination analyses showed weak signatures of recombination among some of the BSV and SCBV sequences.

## Keywords

Banana Streak Virus, Sugarcane Baciliform Virus, RT/RNase H, Polymerase Chain Reaction, Diversity

*Corresponding author.

## 1. Introduction

Banana (*Musa* spp.) is one of the important commercial horticultural crops. Banana plants are hosts to several badnaviruses named Banana Streak Virus (BSV) and Sugarcane Baciliform Virus (SCBV) leads to a loss in the production. The genomic structure of the Badnavirus consists of three Open Reading Frames (ORFs) in the positive strand [1] except in Taro Bacilliform Virus (TaBV), Cacao Swollen Shoot Virus (CSSV) and Citrus Yellow Mosaic Virus (CYMV) which have four, five and six ORFs, respectively. ORF I and ORF II encode two small proteins of unknown function of 20.8 and 14.5 kDa, respectively. ORF III, the largest ORF, encodes a polyprotein of 208 kDa consisting of the putative cell-to-cell Movement Protein (MP), the Coat Protein (CP), the Aspartic Protease (AP) and the viral replicase, which has Reverse Transcriptase (RT) and Ribonuclease H (RNase H) functions [2] [3]. Bioinformatic analysis revealed that the sequence between RT and RNase H (RT/RNase H) contains highly conserved domains in the badnaviruses genome. Thus far, this gene is used as marker for BSV detection and phylogenetic analysis [4]. Various PCR assays have been used to screen for badnaviruses with three previously-developed primer pairs.

Banana streak disease was first reported in Côte d'Ivoire in 1966 [5]. The virus is transmitted through vegetatively propagated plant material and semi- persistently by mealybugs, but not mechanically or through soil [6]. According to the classification of the genus Badnavirus by the International Committee on Taxonomy of Viruses (ICTV), the nucleotide sequence similarity of less than 80% or the amino acid sequence similarity of less than 89% is considered as a new species [7]. At present, nine BSV species of BSV are identified by ICTV. In addition, six other BSV species have not been classified. Moreover, BSV genome may integrate into the banana genome, and it can be activated to produce infectious virions under various abiotic and biotic stresses, including genetic crosses, and lead to infectious viral particles [8] [9]. Several episomal BSV sequences, namely BSOLV, BSMYV, BSGFV and BSIMV, have been shown to have integrated counterparts, termed endogenous BSV (eBSV), in the *Musa* genome [8] [10] [11]. BSV is a kind of pararetroviruses (EPRVs) that use a virus-encoded Reverse Transcriptase (RT) to reverse viral RNA (vRNA) into viral DNA, completing the viral DNA replication process [12].

Sugarcane bacilliform virus is an exception among badnaviruses which has been proven to be transmitted to two different families, viz., *Poaceae* and *Musaceae* [13] [14] [15]. Recently, natural infection of banana by SCBV has been reported in China [16] and Burkina Faso [17]. SCBV is found in major sugarcane growing regions worldwide. SCBV has a high level of genome diversity based on the RT/RNase H region [18]. Two species of SCBV, Sugarcane Bacilliform IM Virus (SCBIMV) and Sugarcane Bacilliform MO Virus (SCBMOV), originating from Australia and Morocco [19] [20], are recognized by ICTV. In addition, recently two new species, Sugarcane Bacilliform Guadeloupe A Virus

(SCBGAV) and Sugarcane Bacilliform Guadeloupe D Virus (SCBGAV) have been identified from Guadeloupe [21]. Currently, 20 phylogenetic groups (genotypes) of SCBV (SCBV-A to SCBV-T) have been reported worldwide [22] [23], and they are to some extent associated with geographical regions. Recently, five novel genotypes (SCBV-U, SCBV-V, SCBV-W, SCBV-X and SCBV-Y) are being proposed [24].

SCBV is serologically closely related to BSV [25]. Phylogenetic analysis has shown that some BSV and SCBV isolates cluster in the same branch [26], suggesting no clear separation between SCBV and BSV isolates. According to [27], in the evolutionary history of badnaviruses, banana was the ancestor host plant of BSV as well as some badnavirus owing to viral host shift.

Understanding the genetic diversity of BSV and SCBV should help in controlling diseases that cause. Many authors studied BSV diversity in Egypt [15], RDC [28], Uganda [29], Guadeloupe [21], etc. and SCBV diversity in China [16] infecting banana revealing high nucleotide divergence. To date, there has been only one study on the prevalence of BSV and SCBV isolates in the major banana-produced areas in Burkina Faso [17]. These reports indicate that BSV and SCBV were detected in seven and two out of nine survey regions, respectively. However, the level of variability is unknown. Moreover, control strategies are not easy to develop due to the absence of information on virus isolates and variability. Our major objectives were to identify the variability of BSV and SCBV isolates in Burkina Faso.

## 2. Material and Methods

### 2.1. Virus Isolates

Thirty-eight BSV and two SCBV isolates, obtained from banana-infected plants from Boucle of Mouhoun, Hauts Bassins, Cascades, Centre, Central West, North and Central East regions, were selected from a previous survey [17]. Samples were collected from the main banana growing areas in Burkina Faso across the two agroclimatic zones from July to October 2018 and September to December 2020. It is the period when viral symptom expression is optimal. Samples were kept on ice before being transferred to the laboratory where they were stored at −80˚C for further analysis.

### 2.2. Extraction of Total DNA

DNA was extracted according to the method of [30] with some modifications [17]. The recovered DNA pellet was dissolved in 30 µl of sterile water and stored at −20˚C for later use.

### 2.3. RT/RNase H Amplification

PCR test was applied to determine the diversity of the BSV in the collected leaf samples. PCR using forward primer Badna FP

(5'-ATGCCITTYGGIAARAAYGCICC-3') and reverse primer Badna RP (5'-CCAYTTRCAIACISCICCCCAICC-3') was carried out to amplify a 579 bp DNA fragment of partial reverse transcriptase/ribonuclease H (RT/RNase H) of BSV and SCBV [22] [31]. PCR reaction mixture contained 1 µl of each primer (10 µM), 1 µl of 10 mM dNTPs mix, 5 µl of 10× Taq DNA polymerase reaction buffer, 2.5 µl of 25 mM $MgCl_2$, 0.25 µl of 5 U/µl Taq DNA polymerase (Cat No. M8301, Affimetrix, USA) and 2 µl of total DNA in a final volume of 50 µl. The thermal cycles of PCR were 35 cycles at 94˚C for 30 sec, 53˚C for 30 sec, 72˚C for 1 min and a final extension on 72˚C for 10 min.

DNA fragments were routinely detected by gel electrophoresis. Ten microliters of the completed PCR were loaded onto a 1% (p/v) agarose gel in 0.5× Tris-Borate-EDTA (TBE), run at 100 V for 30 min, and the gel was stained with ethidium bromide.

All BSV amplicons were directly sequenced at Macrogen (Amsterdam, The Netherlands).

## 2.4. Sequences Data Analysis

All sequences contigs were edited and assembled using BioEtdit software to generate RT/RNase H sequences. Sequences were compared with other respective viral sequences of the NCBI database using BLAST (BLAST, http://www.ncbi.nlm.nih.gov/blast). Nucleotide identities between among query and retrieved sequences were determined using SDT v1 software [32]. Relative Synonymous Codon Usage (RSCU) was determined using Mega-6 [33].

Phylogenetic relationships between the Burkina Faso BSV and SCBV isolates and between our isolates and isolates of the main strains in world were determined using the Maximum-Likelihood (ML) method implemented in MEGA-6. The Tamura 3-parameter model with rate variation across sites estimated from a gamma distribution (T92 + G) was selected as the best fitting model. Robustness of trees was determined by bootstrap sampling of multiple sequence alignment with 1000 replications. The phylogenetic tree was visualized and edited using the MEGA-6 software.

## 2.5. Viral Recombination Analysis

The recombination detection program RDP4 v4.100 [34] was used to analyse BSV and SCBV sequences from this study for recombination events. The program RDP4 detects recombination breakpoints accurately and presents a friendly graphical interface for assessing various attributes in a recombination analysis process. The program simultaneously uses a range of different recombination detection methods to detect recombination events within aligned sequences. These methods include the RDP4 BURT method, the BOOTSCANNING method [35], the GENECONV method [36], the Maximum Chi-Square (MAXCHI) method [37], the CHIMAERA method [37], the Sister Scanning (SISCAN) method [38] and the 3SEQ method [39].

## 3. Results

### 3.1. Agarose Gel Visualization of PCR Product

PCR amplification of the RT/RNase H region using gene specific primers resulted in the amplification of ~579 bp fragment from banana infected with BSV and SCBV but no amplification was observed in healthy samples (Figure 1). The RT/RNase H specific amplified products were eluted and sequenced. RT/RNase H sequences were obtained from all 38 BSV and 2 SCBV positive samples used.

### 3.2. Nucleotide Identity

BLAST search analysis revealed that the sequences of BSV and SCBV isolates proved to be identical at the nucleotide level from 73.25% to 99.60% and 89.69% to 93.24%, respectively, with those from other parts of the world.

Nucleotide analysis of the RT/RNase H region sequences showed a wide virus identity (Figure 2). Nucleotide (nt) sequence identities in RT/RNase H regions ranged from 58.9% to 98.1% among the 38 BSV isolates from Burkina Faso. The highest percentage of nt identity was shown between the isolates BSV_BM5 and BSV_HB7, while BSV_CE was the most distant from the isolate BSV_HB5.

As shown in Table 1, four groups of BSV sequences designated Group 1 (13 isolates), Group 2 (5 isolates), Group 3 (8 isolates) and Group 4 (12 isolates) were differentiated. The nucleotide sequence identity ranged from 87.5% to 100%, 88.1% to 100%, 75.9% to 100% and 83.3% to 100% among isolates from Group 1, Group 2, Group 3 and Group 4 respectively. Among them, identity between Group 3 and Group 4 had the high homology 60.9% to 74.8%, while Group 2 and Group 3 showed a low nucleotide sequence similarity ranging from 69.0% to 79.5%.

Interestingly, our four Burkina Faso BSV group were critical analysis of the sequence identities of with the other BSV species originating from different parts of world (Table 1). Analysis of RT/RNase H regions of four BSV groups from Burkina Faso shared 62.8% - 97.8% nt identities compared to 15 BSV species with published complete genomes. The highest nucleotide sequence identity

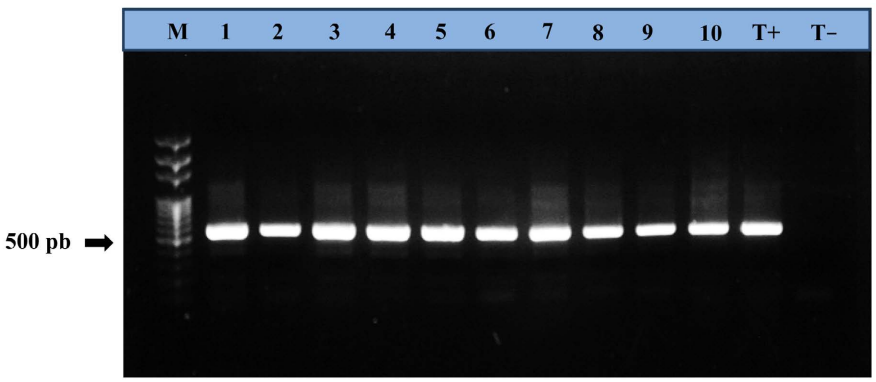


Figure 1. Agarose gel (1%) electrophoresis of PCR amplified fragments using primers for the RT/RNase H region; M: 100 bp molecular weight marker (Solis Biodyne); Lane 1 - 8, BSV positive samples; Lane 9 - 10, SCBV positive samples; T+: positive control, T−: negative control.

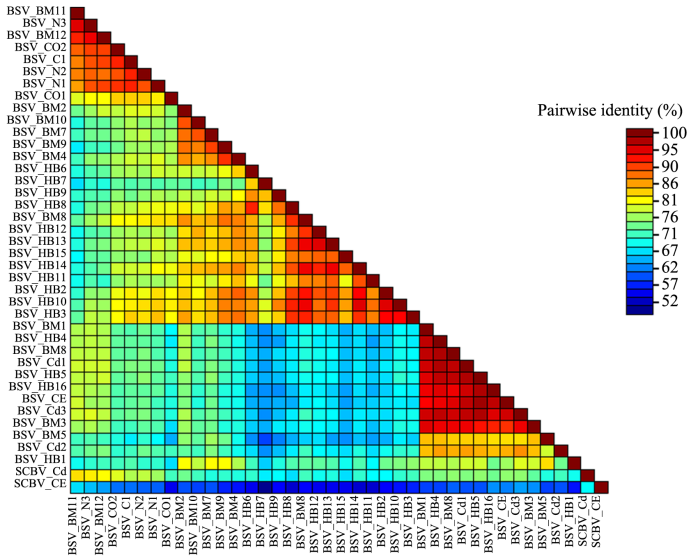


**Figure 2.** Graphical representation of percentage pairwise nucleotide sequence identity of the RT/RNase H region. The color scale bar indicates the percentage identity of 40 Burkina Faso BSV and SCBV isolates.

**Table 1.** RT/RNase H nucleotide identity between BSV groups and other BSV species from GenBank.

| | Nucleotide identity (%) | | | |
|---|---|---|---|---|
| | **Group 1** | **Group 2** | **Group 3** | **Group 4** |
| Group 1 | 87.5 - 100 | - | - | - |
| Group 2 | 67.9 - 81.2 | 88.1 - 100 | - | - |
| Group 3 | 71.7 - 80.5 | 69.0 - 86.5 | 75.9 - 100 | |
| Group 4 | 68.3 - 80.9 | 68.2 - 79 | 60.9 - 74.8 | 83.3 - 100 |
| BSOLV | 69.2 - 75.8 | 72.8 - 79.0 | 61.8 - 71.6 | 76.9 - 100 |
| BSImV | 68.1 - 71 | 65.3 - 69.4 | 6.1 - 69.8 | 65.2 - 71.2 |
| BSMyV | 66.5 - 67.9 | 64.8 - 66.6 | 60.7 - 68.4 | 60 - 67.5 |
| BSGFV | 62.3 - 66.3 | 64.5 - 67.6 | 55.9 - 66.8 | 61.3 - 63.9 |
| BSCavV | 67.6 - 69.8 | 70.7 - 77.5 | 63.1 - 70.3 | 65.1 - 73.8 |
| BSAcVNV | 65.2 - 67.2 | 66.3 - 69.9 | 61.6 - 68.4 | 63.9 - 65.1 |
| BSUIV | 63.7 - 65.1 | 62.7 - 65.8 | 55.5 - 66.9 | 58.7 - 63.4 |
| BSUgAV | 69.0 - 70.8 | 71.0 - 73.8 | 61.3 - 71.2 | 69.2 - 75.0 |
| BSULV | 61.2 - 62.7 | 61.1 - 63.7 | 58.9 - 64.7 | 60.6 - 63.4 |
| BSUMV | 61.1 - 64.8 | 59.8 - 63.0 | 58.8 - 63.7 | 58.2 - 65.7 |
| BSESV | 38.5 - 45.9 | 38.9 - 45.8 | 40.8 - 46.7 | 35.5 - 41.1 |
| BSEGV | 43.6 - 47.5 | 44.2 - 51.1 | 41.0 - 45.2 | 44.6 - 49.6 |
| BSCV | 76.8 - 79.2 | 68.7 - 75.1 | 68.8 - 79.7 | 66.5 - 70.3 |
| BSCuV | 69.4 - 73.2 | 68.8 - 78.1 | 69.4 - 73.2 | 68.8 - 79.2 |
| BSKAV | 66.2 - 70.8 | 72.5 - 78.2 | 72.9 - 63.1 | 66.5 - 80.0 |

ranged from 76.8% to 79.2%, 72.8% to 79.0%, 68.8% to 79.7% and 76.9% to 100% was observed between Group 1-BSCV, Group 2-BSOLV, Group 3-BSCV and Group 4-BSOLV, respectively. BSESV species recorded the lower levels of identity to Group 1 (38.5% to 45.9%), Group 2 (38.9% to 45.8%), Group 3 (40.8% to 46.7%) and Group 4 (35.5% to 41.1%).

Similarly, comparison of the nucleotide sequence between Burkina Faso SCBV isolates with all SCBV genotypes and species from GenBank showed the higher sequence variability (Table 2). Our two Burkina Faso SCBV isolates had a sequence identity of 68% among themselves. Nucleotide sequences of the isolates SCBV_C and SCBV_CE showed the lower homology with a SCBV_C genotype of 91.3% and 58.7%, respectively. Low nucleotide sequence similarity ranging from 59.2% and 50.4% with SCBV_C and SCBV_CE was recorded with SCBV_Y and SCBV_K species, respectively.

### 3.3. Relative Synonymous Codon Usage (RSCU)

Generally, the RSCU value of >1.05 considered as overexpressed while <0.4 is an under represented codon. Table 3 summarizes the codon usage bias in the four BSV groups. There is a high variability in RSCU between the four groups of BSV with one or several codons overexpressed for each. The four groups of BSV shared the several same overexpressed codons in some amino acids. Following codons are also well expressed in all BSV group viz. UUG (Leucine), UCA (Serine), ACA (Threonine), UAU (Tyrosine), AAA (Lysine), GAA (Glutamine),

Table 2. RT/RNase H nucleotide identity between SCBV isolates and other SBCV genotypes/species from GenBank.

| | Nucleotide identity (%) | | | | |
|---|---|---|---|---|---|
| | SCBV_Cd | SCBV_CE | | SCBV_Cd | SCBV_CE |
| SCBGAV_A | 79.2 | 64.7 | SCBV_N | 62.7 | 53.8 |
| SCBV_B | 79.2 | 63.7 | SCBV_O | 61.1 | 54.1 |
| SCBV_C | 91.3 | 68.7 | SCBV_P | 60.5 | 53.2 |
| SCBGDV_D | 67.9 | 60.6 | SCBV_Q | 61.7 | 54.7 |
| SCBMOV_E | 62.9 | 52.8 | SCBV_R | 59.8 | 51.4 |
| SCBIMV_F | 62.9 | 54 | SCBV_S | 60.7 | 51 |
| SCBV_G | 61.5 | 51.8 | SCBV_T | 60.9 | 54.8 |
| SCBV_H | 61.5 | 53.4 | SCBV_U | 64.2 | 53 |
| SCBV_I | 61.3 | 52.8 | SCBV_V | 62.2 | 53.1 |
| SCBV_J | 62.3 | 52.2 | SCBV_W | 63.4 | 54.7 |
| SCBV_K | 60.9 | 50.4 | SCBV_X | 61.3 | 55.5 |
| SCBV_L | 63.2 | 51.4 | SCBV_Y | 59.2 | 52.9 |
| SCBV_M | 65.8 | 53.8 | | | |

**Table 3.** Relative synonymous codon usage in the BSV groups and SCBV isolates.

| Amino acid | RSCU | | | | | | |
|---|---|---|---|---|---|---|---|
| | Codon | Group 1 | Group 2 | Group 3 | Group 4 | SCBV_Cd | SCBV_CE |
| Init. | AUG | 1 | 1 | 1 | 1 | 1 | 1 |
| Leu | UUA | 1.08 | 0.33 | 0.28 | 1.14 | 1.29 | 3 |
| | UUG | 1.82 | 1.12 | 1.44 | 1.6 | 0 | 1.47 |
| | CUU | 0.74 | 1.57 | 0.96 | 0.89 | 1.29 | 0.67 |
| | CUC | 0.07 | 1.03 | 0.52 | 0.43 | 0.43 | 0.06 |
| | CUA | 0.88 | 0.54 | 1.52 | 1.08 | 2.14 | 0.49 |
| | CUG | 1.42 | 1.42 | 1.28 | 0.86 | 0.86 | 0.31 |
| Ile | AUU | 1.74 | 1.16 | 0.91 | 2.08 | 1.17 | 1.34 |
| | AUC | 0.96 | 0.93 | 1.17 | 0.42 | 0.5 | 0.24 |
| | AUA | 0.3 | 0.91 | 0.91 | 0.49 | 1.33 | 1.42 |
| Val | GUU | 0.44 | 1.72 | 0.13 | 0.66 | 0.67 | 1.81 |
| | GUC | 1.78 | 0.34 | 1.25 | 1.52 | 2 | 0.19 |
| | GUA | 1.33 | 0.76 | 1.75 | 1.62 | 0 | 1.48 |
| | GUG | 0.44 | 1.17 | 0.88 | 0.2 | 1.33 | 0.52 |
| Ser | UCU | 1.48 | 1.12 | 0.7 | 0.65 | 0.92 | 0.75 |
| | UCC | 0.62 | 1.12 | 0.17 | 2.29 | 0.46 | 0.6 |
| | UCA | 2.73 | 1.56 | 2 | 1.08 | 1.38 | 0.75 |
| | UCG | 0.08 | 0.1 | 0.43 | 0.65 | 0 | 0.15 |
| | AGU | 0.39 | 1.53 | 0.35 | 0.99 | 0.46 | 3.3 |
| | AGC | 0.7 | 0.58 | 2.35 | 0.35 | 2.77 | 0.45 |
| Pro | CCU | 0.43 | 1.14 | 0.26 | 0.14 | 1.33 | 0.8 |
| | CCC | 1.79 | 1.77 | 1.9 | 1.49 | 0 | 1.2 |
| | CCA | 1.57 | 0.83 | 1.84 | 2.02 | 2.67 | 0.8 |
| | CCG | 0.21 | 0.26 | 0 | 0.34 | 0 | 1.2 |
| Thr | ACU | 0.62 | 1.53 | 1.42 | 1.16 | 0 | 1.78 |
| | ACC | 1.62 | 0.44 | 0.08 | 0.22 | 2.29 | 0.44 |
| | ACA | 1.38 | 1.59 | 1.5 | 1.96 | 1.71 | 1.56 |
| | ACG | 0.38 | 0.44 | 1 | 0.67 | 0 | 0.22 |
| Ala | GCU | 1.33 | 0.94 | 0.58 | 2.24 | 0.8 | 1.67 |
| | GCC | 1.08 | 1.45 | 1.25 | 0.55 | 2.4 | 0.33 |
| | GCA | 1 | 1.12 | 1.67 | 1.09 | 0.8 | 1.33 |
| | GCG | 0.58 | 0.48 | 0.5 | 0.12 | 0 | 0.67 |

**Continued**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Tyr | UAU | 1.91 | 1.51 | 1.44 | 1.56 | 1.2 | 1.87 |
| | UAC | 0.09 | 0.49 | 0.56 | 0.44 | 0.8 | 0.13 |
| His | CAU | 1.29 | 1.33 | 1.63 | 0.31 | 0 | 1.6 |
| | CAC | 0.71 | 0.67 | 0.38 | 1.69 | 0 | 0.4 |
| Gln | CAA | 1.12 | 1.51 | 1.6 | 0.97 | 1.14 | 1 |
| | CAG | 0.88 | 0.49 | 0.4 | 1.03 | 0.86 | 1 |
| Asn | AAU | 1 | 1.4 | 1.21 | 1.35 | 1.45 | 1.94 |
| | AAC | 1 | 0.6 | 0.79 | 0.65 | 0.55 | 0.06 |
| Lys | AAA | 1.44 | 1.53 | 1.42 | 1.35 | 1 | 1.29 |
| | AAG | 0.56 | 0.47 | 0.58 | 0.65 | 1 | 0.71 |
| Asp | GAU | 1.56 | 1.55 | 1.38 | 0.59 | 1.11 | 1.83 |
| | GAC | 0.44 | 0.45 | 0.62 | 1.41 | 0.89 | 0.17 |
| Glu | GAA | 1.32 | 1.29 | 1.2 | 1.59 | 1 | 0.8 |
| | GAG | 0.68 | 0.71 | 0.8 | 0.41 | 1 | 1.2 |
| Cys | UGU | 0 | 0.37 | 1.64 | 0.83 | 0 | 1.82 |
| | UGC | 2 | 1.63 | 0.36 | 1.17 | 2 | 0.18 |
| Trp | UGG | 1 | 1 | 1 | 1 | 1 | 1 |
| Arg | CGU | 0.15 | 0.46 | 0.1 | 0.08 | 0 | 0.32 |
| | CGC | 0.15 | 0.35 | 0.39 | 0.12 | 0.43 | 0.16 |
| | CGA | 0.31 | 0.71 | 0 | 1.29 | 0 | 0.32 |
| | CGG | 0.62 | 0.57 | 0.34 | 0.12 | 0.43 | 0.16 |
| | AGA | 2.62 | 3.39 | 3.97 | 2.62 | 4.29 | 2.37 |
| | AGG | 2.15 | 0.52 | 1.21 | 1.77 | 0.86 | 2.68 |
| Phe | UUU | 0.67 | 1.29 | 0.95 | 1.04 | 0.29 | 1.6 |
| | UUC | 1.33 | 0.71 | 1.05 | 0.96 | 1.71 | 0.4 |
| Gly | GGU | 0 | 0.59 | 0.52 | 0.12 | 1.43 | 1.12 |
| | GGC | 0.91 | 0.45 | 0.71 | 1.65 | 1.14 | 0.16 |
| | GGA | 2.11 | 1.61 | 1.79 | 1.57 | 0.86 | 1.44 |
| | GGG | 0.98 | 1.35 | 0.97 | 0.67 | 0.57 | 1.28 |
| Stop | UAA | 1.11 | 0.87 | 1 | 1.7 | 0.75 | 1.05 |
| | UAG | 0.6 | 1.03 | 0.25 | 0.36 | 1.13 | 1.2 |
| | UGA | 1.29 | 1.11 | 1.75 | 0.94 | 1.13 | 0.75 |

AGA (Arginine) and GGA (Glycine). However, in some cases, additional over-expressed codons discriminated the three groups of BSV thus highlighting relationship between some groups to the detriment of others. This is the case for the

overexpressed codons CUG (Leucine), GCC (Ala), CAU (His), CAA (Gln), GAU (Asp) and UGA (Stop) common to Groups 1, 2 and 3; AUU (Ile), UGC (Cys) common to Groups 1, 2, 4; GUA (Val), CCA (Pro), AGG (Arg); and ACU (Thr), GCA (Ala), AAU (Asn) common to Groups 2, 3 and 4. Codons overexpressed individually in one or all groups were also observed. The four arginine codons were very weakly used by four BSV groups. Only the CGA codon was overexpressed by Group 4.

Table 3 summarises the codon usage bias for the two SCBV isolates. For many amino acids, these two isolates show the same overexpressed codons. This is the case for codons UUA (Leu), AUU and AUA (Ile), ACA (Thr), UAU (Tyr), AAU (Asn), GAU (Asp), AGA (Arg) GGU (Gly) and UAG (Stop). However, differences in overexpressed codons were recorded between the two isolates. In particular, codons CUU and CUA (Leu), GUC and GUG (Val), UCA and AGC (Ser), CCU and CCA (Pro), ACC (Thr), GCC and GCA (Ala), CAA (Gln), UGC (Cys), UUC (Phe), GGC (Gly) and UGA (Stop) were overexpressed by the SCBV_Cd isolate. The codons GUU (Val), AGU (Ser), CCC and CCG (Pro), ACU (Thr), GCU (Ala), CAU (His), AAA (Lys), GAG (Glu), UGU (Cys), UUU (Phe), AGG (Arg), GGA and GGG (Gly) were overexpressed by the isolate SCBV_CE.

In particular, the single codon UGG of the amino acid tryptophan is not overexpressed by any of the BSV groups and SCBV isolates. It shows no usage bias (RSCU = 1) as do the codons ACG (Thr), UAA (Stop) and the codons GCA (Ala), AAU and AAC (Asn) for BSV Groups 1 and 3, respectively. The same observation has been made for codons AAA and CAG (Lys), GAA and GAG (Glu), and codons CAA and CAG (Gln) for isolates SCBV_Cd and SCBV_CE.

### 3.4. Phylogenetic Relationships

As shown in Figure 3, phylogenetic tree was constructed by the RT/RNase H regions of the 38 BSV isolates collected in this study. The results indicated the existence of four well defined groups, named 1, 2, 3 and 4. Groups 3 and 4 are more heterogenous, while groups 1 and 2 are fairly homogenous. Group 1 included twelve Hauts Bassins isolates and one Boucle of Mouhoun isolates; and Group 2 contained five Boucle of Mouhoun isolates. Group 3 included three North isolates, two Boucle of Mouhoun isolates, two Centre West isolates and one Centre isolates whereas Group 4 was comprised four Boucle of Mouhoun isolates, four Hauts Bassins isolates, three Cascades isolates and one East Centre isolate.

To further determine the phylogenetic relationship, another phylogenetic tree was constructed based on the 38 sequences of RT/RNase H regions from this study and 35 additional sequences from different countries from GenBank (Figure 4(a)). Phylogenetic analysis separated the sequences into four major group (Groups A, B, C and D). The four previous BSV group (Groups 1, 2, 3 and 4) from Burkina Faso felled into three different groups (Groups A, B and C).

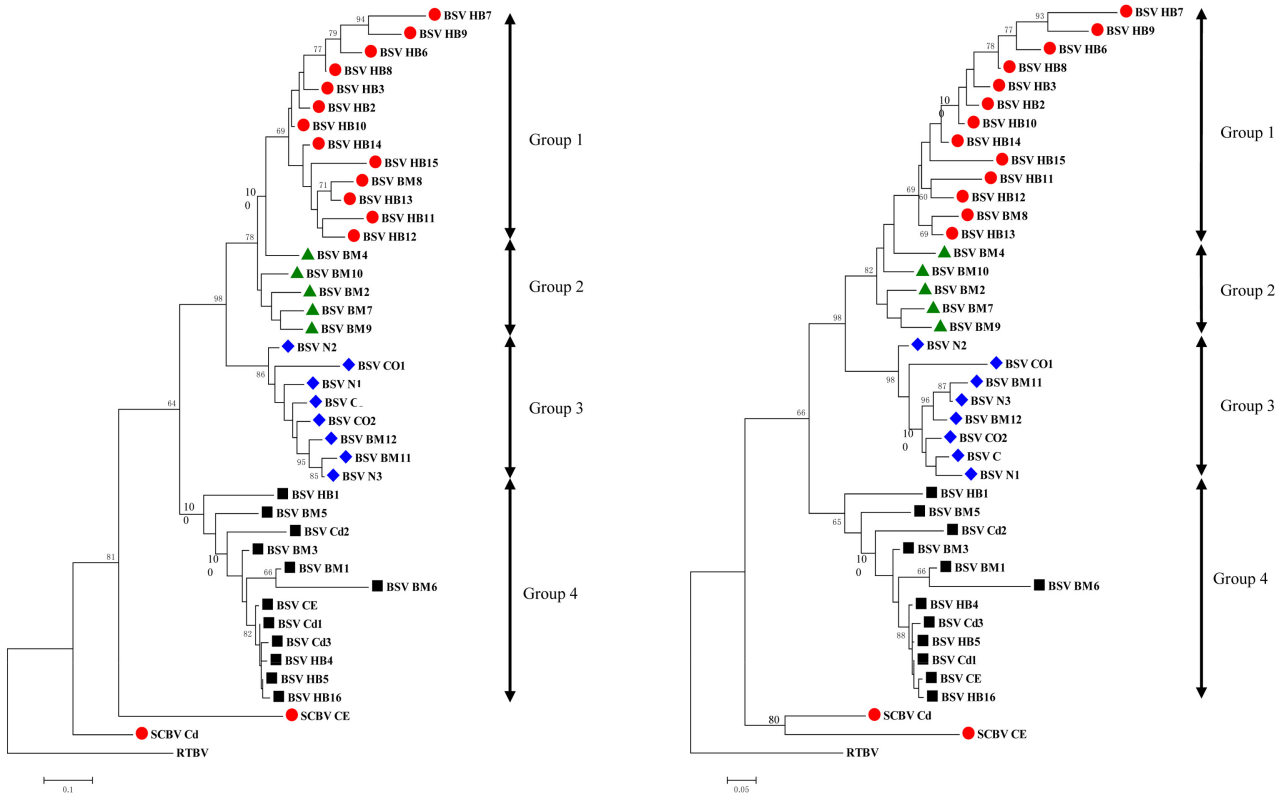


**Figure 3.** Phylogeny tree was established with (a) nucleotide sequences and (b) amino acid of RT/RNase H region of 38 BSV isolates from Burkina Faso. The trees were constructed using the Maximum Likelihood method. The numbers at the nodes indicate bootstrap support (1000 replicates). RTBV: *Rice Tungro Bacilliform Virus* (D10774) was used as the outgroup.

Group A consisted of Group 3 with BSV isolates from Australia, China, Cuba, Egypt, France, Equator, Guadeloupe and France. Two Burkina Faso BSV Groups 1 and 2 were grouped together with China isolates on Group B. Group 4 clustered with other BSV isolates from eight countries (Brasilia, China, Kenya, Nigeria, Uganda, India and France) on Group C. Group D consisted of other BSV originating from seven countries.

To determine the relationship of two SCBV sequences determined in this study with other SCBV species, phylogenetic analyses were conducted using RT/RNase H region (**Figure 5(a)**). Phylogenetic analyses separated the SCBV sequences into three major phylogenetic groups (Groups 1, 2 and 3). Our two SCBV isolates were grouped in Group 3 along with the SCBV isolates from Guadeloupe and China. Groups 1 and 2 were constituted of SCBV isolates from other countries.

The dataset of SCBV sequences consisted to 25 previously-published SCBV genotypes (from SCBV-A to Y) were used to assess our SCBV isolates through phylogenetic analysis (**Figure 5(b)**). The maximum-Likelihood phylogenetic tree revealed that SCBV_Cd isolates clustered into the SCBV-C genotype, while the SCBV_CE isolate was segregated into a new SCBV genotype separated under Genotype Z.

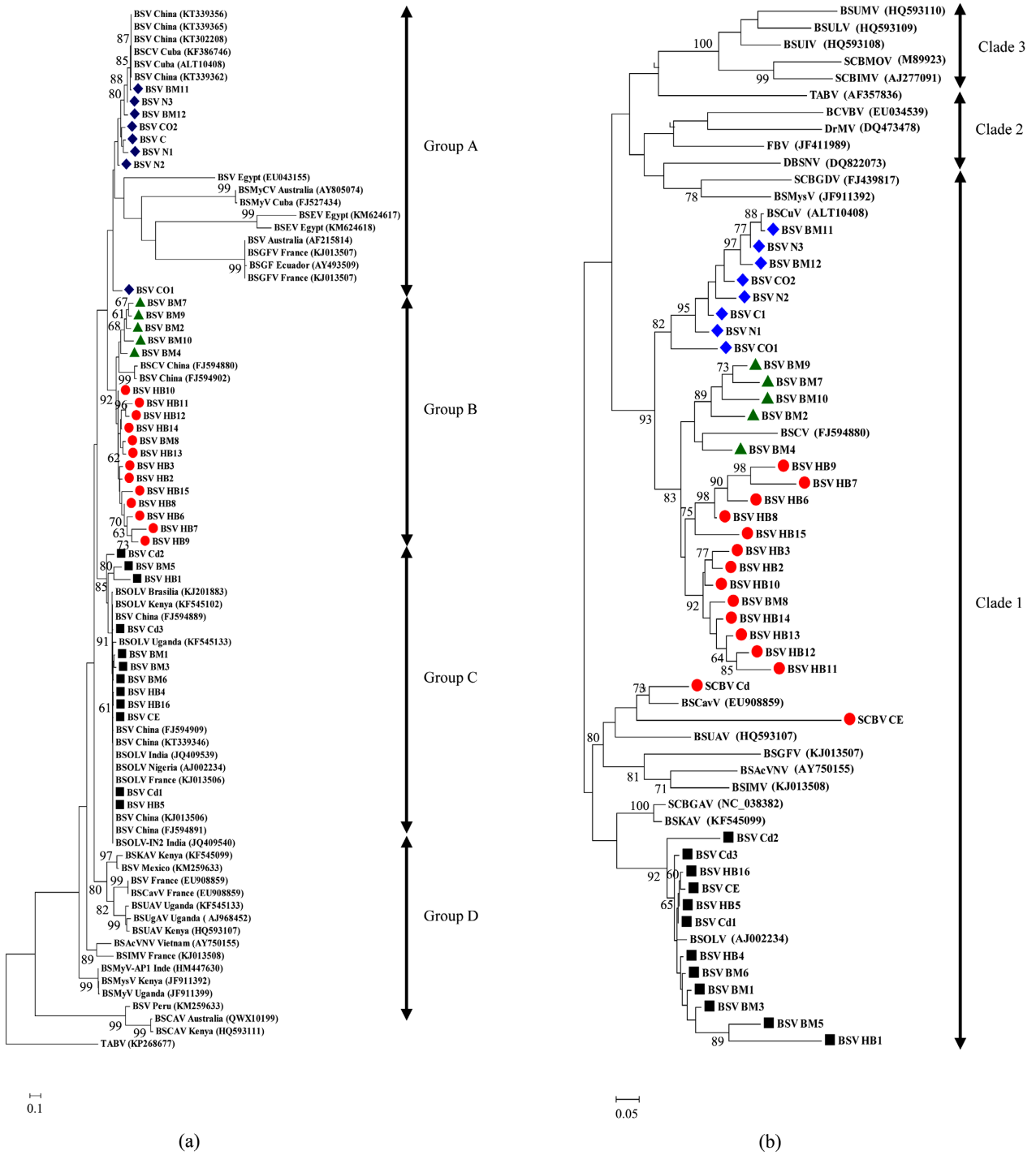


**Figure 4.** (a) Phylogenetic tree based on the multiple sequence alignment of the nucleotide sequences of RT/RNase H region from 38 BSV isolates from this study and 44 from the GenBank database. The tree was constructed using the Maximum Likelihood method and bootstrap resampling of 1000 replications. TABV: *Taro Bacilliform Virus* (KP268677) was used as the outgroup. (b) Phylogenetic tree depicting the relationships of 38 BSV and two SCBV tested isolates with other badnaviruses based on nucleotides sequences of RT/RNase H using the Maximum Likelihood method, Bootstrap values (1000 replicates) above 60% are indicated for each node. TABV: *Taro Bacilliform Virus*, BSUAV*: Banana Streak Virus*, BSUIV, BSULV, BSUMV, BCVBV: *Bougainvillae spectabilis chlorotic vein-banding virus*, DBSNV: *Dioscorea Bacilliform SN Virus*, DrMV: Dracaena Mottle Virus, FBV: *Figbadnavirus*, SBMOV: *Sugarcane Bacilliform MO Virus*, SCBIMV: *Sugarcane Bacilliform IM Virus*, SCBGAV: *Sugarcane Bacilliform Guadeloupe A Virus*, SCBGDV: *Sugarcane Bacilliform Guadeloupe D Virus*.

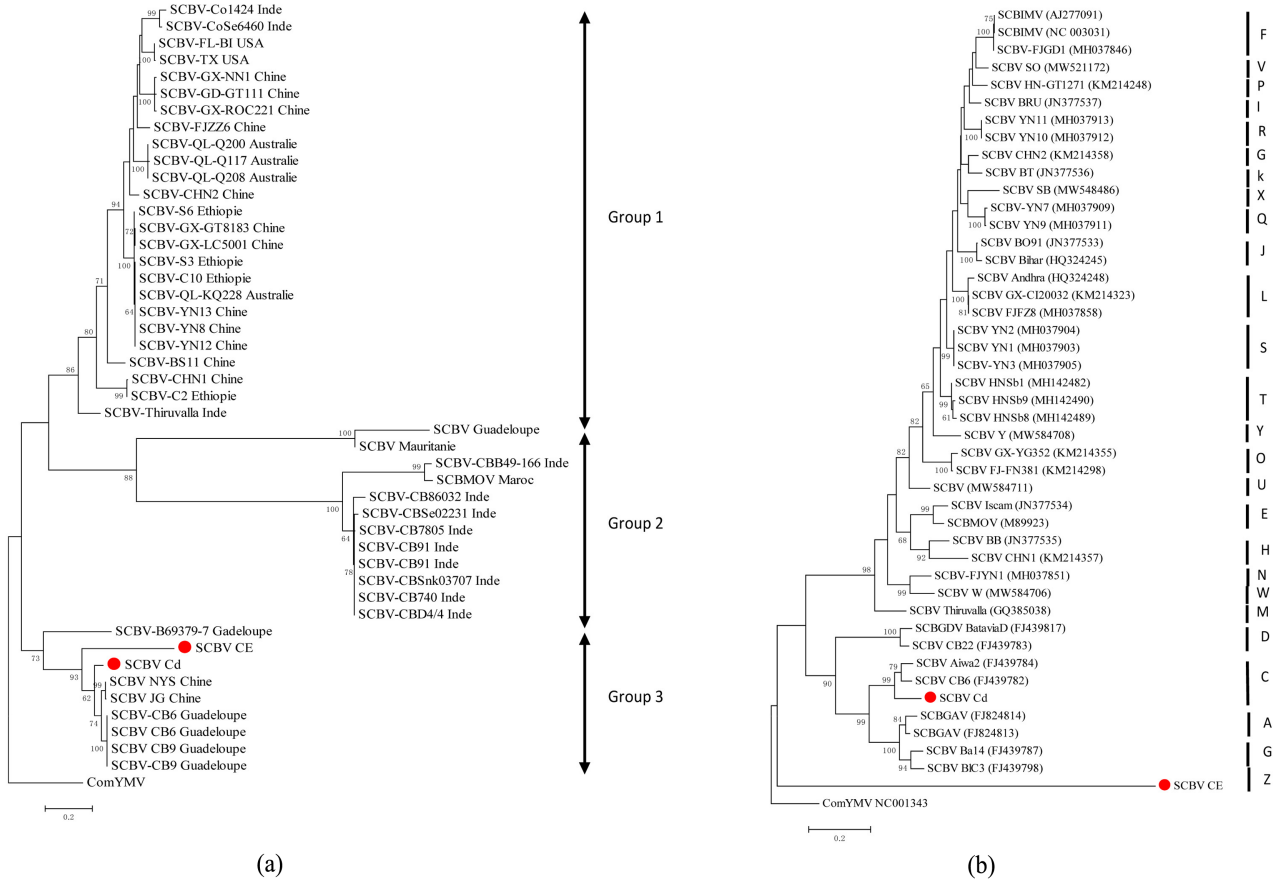


**Figure 5.** Phylogenetic relationship based on nucleotide sequences of RT/RNase H region of two Burkina Faso SCBV isolates with (a) SCBV originated from other countries and (b) 25 SCBV genotypes rooted with *Commelina* yellow mottle virus (NC001343). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (shown only when >60%).

In order to analyse the evolutionary relationships between our BSV and SCBV, and other badnaviruses, a phylogenetic tree based on their RT/RNase H sequence was constructed (**Figure 4(b)**). The results showed that these badnaviruses can be clustered into three distinct badnaviruse clades (Clades I, II and III). All Burkina Faso BSV and SCBV isolates were classified into Clade I with nine BSV species and two SCBV species. Clade II contained other badnaviruses species (DBSNV, FBV, DrMV, BCVBV and TABV) and Clade III included three Ugandan BSV species (BSUIV, BSULV and BSUMV) and two SCBV species (SCBMOV and SCBIMV).

## 3.5. Recombination Analysis

Recombination analysis of the RT/RNase H sequences using the seven methods of the RDP software revealed potential recombination event on one Sugarcane baciliform virus (SCBV_CE) and thirteen Banana Streak Virus (BSV) genome sequence (**Table 4**). Only SCBV_CE recombination was detected with all recombinant method detection while major parent was unknown and minor parent was SCBV_Cd. However, six BSV recombination isolates BSV_BM6, BSV_CO1,

**Table 4.** Recombination events detected in the RT/RNase H region of BSV and SCBV isolates using RDP methods showing the parental and recombinant sequences.

| Recombinant isolate | Minor parent | Major parent | Recombination break point | Positive method* | P-Value |
|---|---|---|---|---|---|
| SCBV_CE | Unknown | SCBV_Cd | 23 - 281 | RGBM**C**ST | $8.447 \times 10^{-11}$ |
| BSV_BM6 | BSV_HB16 | Unknown | 593 - 66 | R**G**MT | $1.195 \times 10^{-8}$ |
| | BSV_HB5 | Unknown | 561 - 591 | **R**GT | $1.729 \times 10^{-3}$ |
| BSV_Cd2 | BSV_BM7 | BSV_HB16 | 550 - 174 | MC**S** | $2.937 \times 10^{-15}$ |
| BSV_CO1 | BSV_HB2 | Unknown | 580 - 182 | RGM**C**T | $3.220 \times 10^{-4}$ |
| BSV_HB4 | BSV_CE | BSV_HB6 | 582 - 26 | R**M**T | $4.934 \times 10^{-4}$ |
| BSV_HB11 | Unknown | BSV_HB10 | 60 - 574 | MC**S**T | $4.168 \times 10^{-19}$ |
| BSV_HB1 | BSV_HB7 | BSV_CE | 488 - 38 | R**M**C | $2.116 \times 10^{-6}$ |
| BSV_BM5 | BSV_HB2 | BSV_Cd3 | 66 - 346 | GMC**S** | $1.222 \times 10^{-19}$ |
| BSV_Cd3 | BSV_CE | BSV_BM10 | 580 - 36 | RG**M**CT | $3.776 \times 10^{-5}$ |
| BSV_N3 | BSV_CO2 | BSV_BM9 | 1 - 29 | **G**T | $6.231 \times 10^{-3}$ |
| BSV_HB12 | BSV_BM7 | BSV_HB7 | 331 - 582 | M**C**S | $3.790 \times 10^{-5}$ |
| BSV_HB15 | Unknown | BSV_HB7 | 14 - 489 | MC**S**T | $1.967 \times 10^{-3}$ |
| BSV_BM2 | BSV_HB5 | Unknown | 347 - 32 | MC**S** | $7.427 \times 10^{-5}$ |
| BSV_C1 | Unknown | BSV_HB7 | 28 - 464 | MC**S** | $1.378 \times 10^{-3}$ |

*The methods used to infer recombination break points were R: RDP, G: GENECONV, B: BOOTSCAN, M: MAXCHI, C: CHIMAERA, S: SISCAN, L: LARD, 3S: 3SEQ. The methods whose P-values are shown are indicated in boldface.

BSV_HB11, BSV_BM5, BSV_Cd3 and BSV_HB15 was detected at least four methods. Two recombination break points were detected in the isolate BSV_BM6 while minor parent was unknown and major parent were BSV_HB16 (nts 593 - 66) and BSV HB5 (nts 561 - 591).

## 4. Discussion

A previous study showed that Banana Streak Virus (BSV) was the main virus of banana (*Musa* spp.) in Burkina Faso. And that Sugarcane Baciliform Virus (SCBV) was only distributed in a few banana growing areas in the Cascades and East Centre regions [17]. In this study, we performed the molecular characterization of BSV and SCBV in samples collected from major banana growing areas in the country. The RT/RNase H region in the viral genome of Badnavirus commonly was used for diversity and phylogenetic analysis [4] [40]. The results showed viruses of the Badnavirus genus were clearly detected by PCR from field samples of banana using a universal primer widely used by several authors for each virus. In addition, sequencing of the RT/RNase gene and sequence analysis confirmed that virus isolates belonged to the genus Badnavirus, especially BSV and SCBV.

Understanding the genetic diversity and population structure of viruses can

help us to develop disease management strategies. In this study, we analysed the genetic diversity and population structure of BSV isolates from Burkina Faso using RT/RNase H sequences. The RT/RNase H sequences of BSV isolates obtained showed a large degree of genetic identity ranging from 58.9% to 98.1%. In fact, this was not surprising since several investigators had considered BSV as a generic name of several species showing up to 30% nucleotide divergence but provoking the same disease in banana plants [41] [42] [43]. According to [44], this divergence could reach 40% in some cases.

Analysis of the nucleotide sequence identities of 38 Burkina Faso isolates based on RT/RNase H regions resulted in assigning virus isolates to four groups. The higher genetic diversity observed between groups suggests the existence of different species of BSV in Burkina Faso. In the genus Badnavirus, nt sequence identity of less than 80% in the RT/RNase H nucleotide sequence can be used to describe new species [45]. Comparison of group sequences with 15 BSV species published sequences indicated lower identity. BSV Groups 1 and 3 recorded higher nucleotide identity ranging from 76.8% to 79.2% and 68.8% to 79.7%, with BSCV; and Group 2 recorded 72.8% to 79.0% with BSOLV. In addition, the topology of the phylogenetic trees obtained in this study with Burkina Faso isolates and with BSV sequences from GenBank are in accordance with nucleotide similarities. All Burkina Faso BSV groups were distributed into Groups A, B and C with BSV isolates from other countries. Similar results were obtained by Rao *et al.* [16] suggesting introduction from the exchange of vegetative materials or regrouping according to BSV species. According to Ermini *et al.* [46], the introduction of banana into the subtropical zone was associated with a broader genetic variation in order to increase the genetic homeostasis necessary for adapting the crop to the suboptimal environment. Yet, it is possible to observe the coevolution between badnaviral sequences and bananas [47]. However, we noted a weak link between our isolates and those of other countries, with the segregation of isolates into Groups A and B. These results suggest that isolates of Groups 1, 2 and 3 of new BSV species typical of Burkina Faso such as those obtained by [48] in Australia, [49] in China, [26] in Uganda and [43] in Vietnam. These authors in molecular characterization of BSV isolates collected from different countries identified a new BSV species Banana Streak CA Virus (BSCAV) in Australia, Banana Streak Acuminata Yunnan Virus (BSAcYuV) in China, Banana streak UA, UI, UL, UJ, UM and UK viruses in Uganda and Banana streak VN virus (BSVNV) in Vietnam in banana. However, the high nucleotide identity between Group 4 and BSOLV species as well as the heterogeneity of isolates in Group C suggests that Group 4 isolates from Burkina Faso belong to BSOLV species. The higher genetic diversity observed between groups than within groups of BSV populations suggests that there is more frequent gene flow within groups than between groups. Further studies should be undertaken to better understand the diversity of BSV isolates from Burkina Faso.

Partial RT/RNase H sequences have been used as taxonomic markers for addressing the sequence diversity among SCBV isolates. A comparison of nucleo-

tide sequences of two SCBV isolates collected from banana growing regions Cascades and Centre East of Burkina Faso has confirmed the existence of high SCBV genetic diversity with 68% identity between two isolates. Similar to other badnaviruses, SCBV is genetically diverse, and the large pool of SCBV variants present in sugarcane is probably due to the vegetative nature of propagation of the host and its long history of movement and cultivation [50]. However, SCBV complexes pose high genetic variation and were clustered into at least 25 genotypes [23] [24]. According to genetic differentiation and phylogenetic analysis of RT/RNase H sequences, SCBV_Cd clustered into SCBV genotype C. Sequence analysis of SCBV_Cd isolates infecting banana revealed a very high nucleotide homology 91.3% with SCBV genotype C (SCBV_C). The phylogenetic analysis indicates that the SCBV_Cd isolates are very close to genotype C with SCBV_Aiwa and SCBV_CB6. As for SCBV_CE isolate, genetic differentiation together with phylogenetic analysis revealed that it would be a new SCBV species according to ICTV demarcation criteria of species in the Badnavirus genus [45] [51]. Overall, this analysis revealed a lower level of nucleotide identity range to 50.4% - 68.7% with 25 SCBV genotypes (A-Y). A such, phylogenetic tree was constructed based on RT/RNase H region revealed that SCBV_CE isolates were segregated into an apart cluster namely Z. Similar results have been obtained by many authors [22] [23] [24], leading to the identification of new SCBV genotypes. These results could be related to agro-climatic conditions or recombination processes within our isolates.

Highly divergent BSV and SCBV variants have been reported worldwide. Genetic recombination, as one of the evolutionary forces, may have contributed to the emergence of genetic diversity in plant viruses such as BSV and SCBV. Stronger evidence of recombination was detected within and between BSV and SCBV in the RT/RNase H gene coding region. One recombination signal was detected in SCBV_CE isolate some two SCBV sequences analysed in this study. This recombinant detected for all methods would be used to explain the particularity of SCBV_CE isolate. Similarly, [24] was observed potent recombination events in Indian SCBV isolates, especially from the proposed new genotypes SCBV-U, -W and -X. However, 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 [22]. According to [52], recombination is considered to be a major source of variation and the driving force for the evolution of several viral families. Thirteen recombination events were identified among the BSV sequences analysed in this study. Six of these were detected by at least four different methods. These recombinant isolates belong to the different BSV groups, thus highlighting the role of recombination in virus diversification.

The results of the RT/RNase H region sequence analysis allowed the classification of BSV isolates into four groups and SCBV isolates into two distinct genotypes which were confirmed by phylogenetic analyses. The codon usage bias between these groups/genotypes is in agreement with the clustering of the BSV

and SCBV isolates studied. Major differences were found for the RSCU with codons overexpressed to a specific group/genotype. However, some overexpressed codons were sometimes common to two groups/genotypes or more groups in the case of BSV. Some codons can undergo more changes and therefore result in codon equilibrium frequencies [24]. According to [53], the RSCU is considered a species-specific statistic. As such, it can be useful for species description. In addition to recombination, codon usage bias analysis has pointed out possible reasons for the genetic diversity found in BSV and SCBV groups/genotypes through this study [24]. Similar conclusions were deduced by [24] in 2021 and [54] in 2019 in their studies on SCBV and voandzou Potyviruses, respectively.

The topology of the phylogenetic tree constructed with sequences of BSV and SCBV isolates from Burkina Faso and the sequences of other viruses belonging to the Badnavirus genus showed a distribution of viruses in the three Clades, 1, 2 and 3. [2] [4] [26] [55] showed a similar distribution of badnaviruses in the three clades. All BSV and SCBV isolates from Burkina Faso clustered in Clade 1 with other BSV and SCBV sequences published in GenBank. Indeed, most of the BSV species are known to belong to Clade 1 of badnaviruses. This clade contains the major BSV species (BSOLV, BSGFV, BSIMV and BSMYV) distributed worldwide with integrated eBSV counterparts [55] [56] and the other species isolated during geographically distinct BSV epidemics BSPEV, BSCAV, BSUAV, BSVNV, BSAcYuV [29] [44]. Three BSV species originating from Uganda clustered in Clade 3 in accordance with the results of studies conducted by [55] on the possible evolutionary scenario of the Banana Streak Virus. As for SCBV also belonging to Clades 1 and 3, it could be considered with BSV as descended from the same common ancestor virus [13] [14]. It is interesting to note that SCBV_Cd was closer to the Mauritius BSCavV than the SCBV_CE, which are from other geographical locations. Although they infect two different crops, these two viruses are very similar. According to [27], banana could have been the ancestral host plant of SCBV as well as for other badnaviruses like Kalanchoe Top-Spotting Virus (KTSV) and Turmeric Mild Chlorosis Virus (TMCV). In many countries, sugarcane and banana are grown close to each other and this could have played a part in the coevolution of the two viruses, BSV and SCBV [57]. It is possible that infection of a different crop, *i.e.* sugarcane by BSV happened during its evolution, when a form infective to sugarcane was generated or vice versa. The polyphylogenetic structure of SCBV and BSV phylogeny suggests that sugarcane and banana could be the original source of BSV [21]. Such host changes can occur when two host-plant species are colonized by the same viral insect vector [58]. A nucleotide similarity of up to 90% was recorded between SCBV and BSV isolates when PCR products from the RNase H and intergenic regions of four SCBV isolates (IJ76-468, SCBV-Morocco, IJ76-465 and Ireng Maleng) and two BSV isolates (both Australian) were compared [18]. 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 [59]. According to the same authors, SCBV experimen-

tally causes symptoms similar to those of BSV on banana, rice and sorghum. All other (partial) Badnavirus sequences described in banana are in Clade 2 [27]. But according to some authors, this clade also contains sequences of endogenous Badnaviruses present in the *Musa* genome without the episomal form among current BSV species [27] [29] [60].

Our findings further elucidate the variability of Badnavirus and several putative recombinant isolates infecting banana in Burkina Faso, and should be useful for improving the classification of Badnavirus, and more broadly, for understanding the evolution of BSV and SCBV. The results of our study indicate that both isolates belong to the SCBV species. Although SCBV has already been identified in banana in China [16], it has a strong serological relationship and nucleotide identity with BSV. Therefore, it would be interesting to sequence both isolates with specific primers and then clone them to obtain the complete genome to ensure the identity of the virus. This study will lay the foundation for the implementation of appropriate measures to control the spread of BSV and SCBV.

## Acknowledgements

We are grateful to G. Konaté for helpful discussions and constructive criticisms of the manuscript. This work was supported by the National Fund for Research and Innovation for Development (FONRID) and partially by the International Joint Laboratory (LMI) Patho-Bios. B. Ouattara was granted a fellowship from Burkina Faso's Ministry of research. We would also like to thank our reviewers for their helpful comments in improving the manuscript.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

[1] Vo, J.N., Campbell, P.R. and Mahfuzc, N.N. (2016) Characterization of the Banana Streak Virus Capsid Protein and Mapping of the Immunodominant Continuous B-Cell Epitopes to the Surface-Exposed N-Terminus. *Journal of General Virology*, **97**, 3446-3457. https://doi.org/10.1099/jgv.0.000643

[2] 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 Research*, **186**, 155-162. https://doi.org/10.1016/j.virusres.2014.01.005

[3] Alangar, B., Thomas, H. and Ramasamy, S. (2016) Badnaviruses: The Current Global Scenario. *Viruses*, **8**, Article 177. https://doi.org/10.3390/v8060177

[4] Harper, G., Hart, D., Moult, S. and Hull, R. (2002) Detection of Banana Streak Virus in Field Samples of Bananas from Uganda. *Annals of Applied Biology*, **141**, 247-257. https://doi.org/10.1111/j.1744-7348.2002.tb00216.x

[5] Yot-Dauthy, D. and Bové, J.-M. (1966) Mosaïque du bananier. Identification et purification de diverses souches du virus. *Fruits*, **21**, 449-466.

[6] Dallot, S., Acuna, P., Rivera, C., Ramirez, P., Cote, F., Lockhart, B.E.L. and Is-

kra-Caruana, M.-L. (2001) Evidence That the Proliferation Stage of Micro Propagation Procedure Is Determinant in the Expression of Banana Streak Virus Integrated into the Genome of the FHIA21 Hybrid (*Musa* AAAB). *Archive of Virology*, **146**, 2179-2190. https://doi.org/10.1007/s007050170028

[7] Geering, A.D., Maumus, F., Copetti, D., Choisne, N., Zwickl, D.J., Zytnicki, M., McTaggart, A.R., Scalabrin, S., Vezzulli, S., Wing, R.A., Quesneville, H. and Teycheney, P.Y. (2014) Endogenous Florendoviruses Are Major Components of Plant Genomes and Hallmarks of Virus Evolution. *Nature Communications*, **5**, Article No. 5269. https://doi.org/10.1038/ncomms6269

[8] Gayral, P., Carlos, J., Noa-Carrazana, N., Lescot, M., Lheureux, F., Lockhart, E.L.B., Mutsumoto, Piffanelli, P. and Iscra-Caruana, M.-L. (2008) A Single Banana Streak Virus Integration Event in the Banana Genome as the Origin of Infectious Endogenous Pararetrovirus. *Journal of Virology*, **82**, 6697-6710. https://doi.org/10.1128/JVI.00212-08

[9] Cote, F.-X., Galzi, S., Folliot, M., Lamagnere, Y., Teycheney, P.-Y. and Iskra-Caruana, M.-L. (2010) Micropropagation by Tissue Culture Triggers Differential Expression of Endogenous Banana Streak Virus (eBSV) in the B Genome of Natural and Synthetic Interspecific Banana Plantain. *Molecular Plant Pathology*, **11**, 137-144. https://doi.org/10.1111/j.1364-3703.2009.00583.x

[10] Harper, G., Dahal, G., Thottapilly, G. and Hull, R. (1999) Detection of Episomal Banana Streak Badnavirus by IC-PCR. *Journal of Virological Methods*, **79**, 1-8. https://doi.org/10.1016/S0166-0934(98)00158-X

[11] Geering, A.D.W., Olszewski, N.E., Harper, G., Lockhart, B.E.L., Hull, R. and Thomas, J.E. (2005) Banana Contains a Diverse Array of Endogenous Badnaviruses. *Journal of General Virology*, **86**, 511-520. https://doi.org/10.1099/vir.0.80261-0

[12] Hohn, T. and Rothnie, H. (2013) Plant Pararetroviruses: Replication and Expression. *Current Opinion in Virology*, **3**, Article ID: 621628. https://doi.org/10.1016/j.coviro.2013.08.013

[13] 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. *Journal of General Virology*, **74**, 15-22. https://doi.org/10.1099/0022-1317-74-1-15

[14] Frison, E.A. and Putter, C.A. (1993) FAO/IBPGR Technical Guidelines for the Safe Movement of Sugarcane Germplasm. FAO/IBPGR.

[15] Abdel-Salam, A.M., Dawoud, R.A., Aly, A.M.E. and El Saghir, S.M. (2014) The Detection and Diversity of Banana Streak Virus Isolates in Egypt. *Egyptian Journal of Virology*, **11**, 136-149.

[16] Rao, X.-Q., Wu, Z.-L., Wang, W., Zhou, L., Sun, J. and Li, H.-P. (2020) Genetic Diversity Analysis Reveals New Badnaviruses Infecting Banana in South China. *Journal of Plant Pathology*, **102**, 1065-1075. https://doi.org/10.1007/s42161-020-00646-8

[17] Ouattara, B., Sérémé, D., Koala, M., Nitiema, L., Koïta, K., Kaboré, E. and Wonni, I. (2023) Prevalence and Spatial Distribution of Badnavirus in the Banana (*Musa* spp.) Major Growing Areas in Burkina Faso. In Press.

[18] 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. *Archives of Virology*, **147**, 2393-2404. https://doi.org/10.1007/s00705-002-0879-2

[19] Adams, M.J. and Carstens, E.B. (2012) Ratification Vote on Taxonomic Proposals to the International Committee on Taxonomy of Viruses. *Archives of Virology*, **157**, 1411-1422. https://doi.org/10.1007/s00705-012-1299-6

[20] Karuppaiah, 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. https://doi.org/10.1007/s11262-013-0890-6

[21] Muller, E., Dupuy, V., Blondin, L., Bauffe, F., Daugrois, J., Nathalie, L. and Iskra-Caruana, M. (2011) High Molecular Variability of Sugarcane Bacilliform Viruses in Guadeloupe Implying the Existence of at Least Three New Species. *Virus Research*, **160**, 414-419. https://doi.org/10.1016/j.virusres.2011.06.016

[22] 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. *Journal of Phytopathology*, **164**, 595-607. https://doi.org/10.1111/jph.12483

[23] Ahmad, K., Sun, S.-R., Chen, J.-L., Huang, M.-T., Fu, H.-Y. and Gao, S.-J. (2019) Presence of Diverse Sugarcane Bacilliform Viruses Infecting Sugarcane in China Revealed by Pairwise Sequence Comparisons and Phylogenetic Analysis. *The Plant Pathology Journal*, **35**, 41-50. https://doi.org/10.5423/PPJ.OA.08.2018.0156

[24] Janiga, P.K., Nithya, K. and Viswanathan, R. (2021) Dynamics of Genetic Diversity among Indian Sugarcane Bacilliform Virus Species and Implications of Associated Recombination Events in the Virus. Research Square. https://doi.org/10.21203/rs.3.rs-713719/v2

[25] Lockhart, B.E.L. and Autrey, L.J.C. (1988) Occurrence in Sugarcane of a Bacilliform Virus Related Serologically to Banana Streak Virus. *Plant Disease*, **72**, 230-233. https://doi.org/10.1094/PD-72-0230

[26] Harper, G., Hart, D., Moult, S., Hull, R., Geering, A. and Thomas, J. (2005) The Diversity of Banana Streak Virus in Uganda. *Archives of Virology*, **150**, 2407-2420. https://doi.org/10.1007/s00705-005-0610-1

[27] Gayral, P. and Iskra-Caruana, M.L. (2009) Phylogeny of Banana Streak Virus Reveals Recent and Repetitive Endogenization in the Genome of Its Banana Host (*Musa* sp.). *Journal of Molecular Evolution*, **69**, 65-80. https://doi.org/10.1007/s00239-009-9253-2

[28] Mukwa, F.T.L. (2016) Les virus du bananier et plantain (*Musa* spp.) en République démocratique du Congo: Occurrence, identification de nouveaux virus et diversité génétique. Thèse de Doctorat, Université Catholique de Louvain, Belgique, 257 p.

[29] James, A.P., Geijskes, R.J., Dale, J.L. and Harding, R.M. (2011) Development of a Novel Rolling Circle Amplification Technique to Detect Banana Streak Virus That Also Discriminates between Integrated and Episomal Virus Sequences. *Plant Disease*, **95**, 57-62. https://doi.org/10.1094/PDIS-07-10-0519

[30] Permingeat, H.R., Romagnoli, M.V. and Vallejos, R.H. (1998) A Simple Method for Isolating High Yield and Quality DNA from (*Gossypium hirsutum* L.) Leaves. *Plant Molecular Biology Reporter*, **16**, 1-6.

[31] 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. *Archives of Virology*, **148**, 1957-1968. https://doi.org/10.1007/s00705-003-0163-0

[32] Muhire, B., Varsani, A. and Martin, D.P. (2014) SDT: A Virus Classification Tool Based on Pairwise Sequence Alignment and Identity Calculation. *PLOS ONE*, **9**, e108277. https://doi.org/10.1371/journal.pone.0108277

[33] Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, **30**, 2725-2729. https://doi.org/10.1093/molbev/mst197

[34] Martin, D.P., Williamson, C. and Posada, D. (2005) RDP2: Recombination Detection and Analysis from Sequence Alignments. *Bioinformatics*, **21**, 260-262. https://doi.org/10.1093/bioinformatics/bth490

[35] Martin, D.P., Lemey, P. and Posada, D. (2011) Analysing Recombination in Nucleotide Sequences. *Molecular Ecology Resources*, **11**, 943-955. https://doi.org/10.1111/j.1755-0998.2011.03026.x

[36] Padidam, M., Beachy, R. and Fauquet, C.M. (1995) Classification and Identification of Geminiviruses Using Sequence Comparisons. *Journal of Virology*, **76**, 249-263. https://doi.org/10.1099/0022-1317-76-2-249

[37] Posada, D. and Crandall, K.A. (2001) Evaluation of Methods for Detecting Recombination from DNA Sequences: Computer Simulations. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 13757-13762. https://doi.org/10.1073/pnas.241370698

[38] Gibbs, M.J., Armstrong, J.S. and Gibbs, A.J. (2000) Sister-Scanning: a Monte Carlo Procedure for Assessing Signals in Recombinant Sequences. *Bioinformatics*, **16**, 573-582. https://doi.org/10.1093/bioinformatics/16.7.573

[39] Boni, M.F., Posada, D. and Feldman, M.W. (2007) An Exact Nonparametric Method for Inferring Mosaic Structure in Sequence Triplets. *Genetics*, **176**, 1035-1047. https://doi.org/10.1534/genetics.106.068874

[40] Bousalem, M., Douzery, E.J. and Seal, S.E. (2008) Taxonomy, Molecular Phylogeny and Evolution of Plant Reverse Transcribing Viruses (Family Caulimoviridae) Inferred from Full-Length Genome and Reverse Transcriptase Sequences. *Archives of Virology*, **153**, 1085-1102. https://doi.org/10.1007/s00705-008-0095-9

[41] Gayral, P., Blondin, L., Guidolin, O., Carreel, F., Hippolyte, I., Perrier, X., *et al.* (2010) Evolution of Endogenous Sequences of Banana Streak Virus: What Can We Learn from Banana (*Musa* sp.) Evolution? *Journal of Virology*, **84**, 7346-7359. https://doi.org/10.1128/JVI.00401-10

[42] Iskra-Caruana, M.-L., Baurens, F.-C., Gayral, P. and Chabannes, M. (2010) A Four-Partner Plant-Virus Interaction: Enemies Can Also Come from Within. *Molecular Plant-Microbe Interaction*, **23**, 1394-1402. https://doi.org/10.1094/MPMI-05-10-0107

[43] Ferreira, C.H.L.H., Jordäo, L.J., Ramos-Sobrinho, R., Ferro, M.M.M., Cavalcanti da Silva, S.J. and Assunção Lima, G.S.A. (2019) Diversification into Genus Badnavirus: Phylogeny and Population Genetic Variability. *Ciência Agricola*, **17**, 59-72. https://doi.org/10.28998/rca.v17i2.6286

[44] Lheureux, F., Laboureau, N., Muller, E., Lochart, E.L. and Iskra-Caruana, M.L. (2007) Molecular Characterization of Banana Streak Acuminata Vietnam Virus Isolated from *Musa acuminata* Siamea (Banana Cultivar). *Archive of Virology*, **152**, 1409-1416. https://doi.org/10.1007/s00705-007-0946-9

[45] King, M.Q., Adams, M.J., Carstens, E.B. and Lefkowitz, E.J. (2012) Virus Taxonomy: Classification and Nomenclature of Viruses. 9th Report of International Committee on Taxonomy of Viruses. Elsevier, San Diego.

[46] Ermini, J.L., Tenaglia, G.C. and Pratta, G.R. (2018) Molecular Diversity in Selected Banana Clones (*Musa* AAA "Cavendish") Adapted to the Subtropical Environment of Formosa Province (Argentina). *American Journal of Plant Sciences*, **9**, 2504-2513. https://doi.org/10.4236/ajps.2018.912181

[47] Chabannes, M., Gabriel, M., Aksa, A., Galzi, S., Dufayard, Iskra-Caruana, M.-L. and Muller, E. (2020) Badnaviruses and Banana Genomes: A Long Association Sheds Light on *Musa* Phylogeny and Origin. *Molecular Plant Pathology*, **22**, 216-230.

https://doi.org/10.1111/mpp.13019

[48] Geering, A.D.W., McMichael, L.A., Dietzgen, R.G. and Thomas, J.E. (2000) Genetic Diversity among Banana Streak Virus Isolates from Australia. *Phytopathology*, **90**, 921-927. https://doi.org/10.1094/PHYTO.2000.90.8.921

[49] Zhuang, J., Wang, J.-H., Zhang, X. and Liu, Z.-X. (2011) Molecular Characterization of Banana Streak Virus Isolate from *Musa acuminata* in China. *Virologica Sinica*, **26**, 393-402. https://doi.org/10.1007/s12250-011-3212-y

[50] Geering, A.D.W. and Hull, R. (2012) Family Caulimoviridae. In: King, A.M.Q., Adams, M.J., Carestens, E.B. and Lefkowitz, E.J., Eds., *Virus Taxonomy Classification and Nomenclature of Viruses Ninth Report of the International Committee on Taxonomy of Viruses*, Elsevier, San Diego, 424-443.

[51] Banerjee, A., Roy, S., Behere, G.T., Roy, S.S., Duttad, S.K. and Ngachan, S.V. (2014) Identification and Characterization of a Distinct Banana Bunchy Top Virus Isolate of Pacific-Indian Oceans Group from North-East India. *Virus Research*, **183**, 41-49. https://doi.org/10.1016/j.virusres.2014.01.017

[52] Fox, M.J. and Erill, I. (2010) Relative Codon Adaptation: A Generic Codon Bias Index for Prediction of Gene Expression. *DNA Research*, **17**, 185-196. https://doi.org/10.1093/dnares/dsq012

[53] Zongo, A.E., Neya, B.J., Palanga, E., Sereme, D., Koala, M., Longue, R.D.S., Zabre, J., Bouda, Z., Barro, N. and Traore, O. (2019) Molecular Characterization of a Large Diversity of Bambara Groundnut Potyviruses (*Vigna subterranea* (L.) Verdc.) in Burkina Faso. *Advances in Bioscience and Biotechnology*, **10**, 270-281. https://doi.org/10.4236/abb.2019.109021

[54] Harper, G., Hart, D., Moult, S. and Hull, R. (2004) Banana Streak Virus Is Very Diverse in Uganda. *Virus Research*, **100**, 51-56. https://doi.org/10.1016/j.virusres.2003.12.024

[55] Iskra-Caruana, M.-L., Duroy, P.-O., Chabannes, M. and Muller, E. (2013) The Common Evolutionary History of Badnaviruses and Banana. *Infection, Genetics and Evolution*, **21**, 83-89. https://doi.org/10.1016/j.meegid.2013.10.013

[56] James, A.P., Geijskes, R.J., Dale, J.L. and Harding, R.M. (2011) Molecular Characterisation of Six Badnavirus Species Associated with Leaf Streak Disease of Banana in East Africa. *Annals of Applied Biology*, **158**, 346-353. https://doi.org/10.1111/j.1744-7348.2011.00466.x

[57] Jaufeerally-Fakim, Y., Khorugdharry, A. and Harper, G. (2006) Genetic Variants of Banana Streak Virus in Mauritius. *Virus Research*, **115**, 91-98. https://doi.org/10.1016/j.virusres.2005.06.015

[58] Lockhart, B.E.L. and Jones, D.R. (2000) Banana Streak. In: Jones, D.R., Ed., *Diseases of Banana, Abaca and Enset*, CAB International, Wallingford, 263-274.

[59] 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. https://doi.org/10.1007/s11262-015-1179-8

[60] Geering, A.D.W., McMichael, L.A., Dietzgen, R.G. and Lockhart, B.E.L. (2001) Analysis of the Distribution and Structure of Integrated Banana Streak Virus DNA in a Range of *Musa* Cultivars. *Molecular Plant Pathology*, **2**, 207-213. https://doi.org/10.1046/j.1464-6722.2001.00071.x