Arch Virol (2005) 150: 1715–1727
DOI 10.1007/s00705-005-0567-0

Archives of Virology
Printed in Austria

Molecular characterization of banana virus X (BVX), a novel member of the *Flexiviridae* family

P.-Y. Teycheney[1], **A. Marais**[2], **L. Svanella-Dumas**[2], **M.-J. Dulucq**[2], and **T. Candresse**[2]

[1]CIRAD-FLHOR, Capesterre Belle-Eau, Guadeloupe, FWI
[2]Equipe de virologie, UMR GD2P, INRA-IBVM, Villenave d'Ornon Cedex, France

Received August 18, 2004; accepted April 28, 2005
Published online June 22, 2005

Summary. A novel virus was identified in banana (*Musa* spp). Analysis of the last 2917 nucleotides of its positive strand genomic RNA showed five open reading frames corresponding, from 5′ to 3′, to a truncated ORF coding for a replication-associated protein, three ORFs coding for a movement-associated triple gene block (TGB) and a capsid protein (CP) gene. This genome organization is similar to that of some members of the *Flexiviridae* family such as potexviruses and foveaviruses. This virus was named Banana virus X (BVX). Comparative sequence analysis showed that BVX is only distantly related to other members of the *Flexiviridae* family, in which it appears to define a new genus. BVX produces defective RNAs derived from its genomic RNA by non-homologous recombination. Three distinct pairs of donor/acceptor recombination sites involving short direct nucleotide repeats were characterized, accounting for deletions of 1268, 1358 and 1503 nucleotides. Contrary to the situation encountered for Potexviruses, these recombination sites are located within the TGB1 and CP genes and result in a truncated TGB1 protein.

Introduction

Banana and plantain (*Musa* spp.) provide hundreds of millions of people throughout the tropics and subtropics with an essential staple food and account for one of the most widely exported fruit in the world. This crop is susceptible to numerous pests and pathogens, including five reported and characterized viruses: banana bunchy top virus (BBTV, a nanovirus), banana bract mosaic virus (BBrMV, a potyvirus), banana mild mosaic virus (BanMMV, an unassigned member of the

The nucleotide sequences reported in this work have been deposited in the GenBank database under accession numbers AY710267 to AY710270

family *Flexiviridae*), banana streak virus (BSV, a badnavirus) and cucumber mosaic virus (CMV, a cucumovirus) [7]. Some of these viruses such as BBTV and, to a lesser extent, BSV, can cause major outbreaks of great economical importance. The impact of BBrMV, BanMMV and CMV seems somewhat lower but mixed infections involving them can lead to severe necrotic leaf symptoms. Nevertheless, all viruses infecting banana are severe constraints to exchanges of *Musa* germplasm while BSV is the major constraint for genetic improvement of banana, due to the ability of BSV sequences integrated into the genome of *Musa balbisiana* to be triggered by various abiotic and biotic stresses, including genetic crosses, and lead to infectious viral particles [4].

The new plant virus family *Flexiviridae* comprises eight viral genera and six viruses that are not assigned to any viral genus – including BanMMV [1]. All share flexuous viral particles as well as similarly organised and phylogenetically related positive single stranded RNA genomes. Species discrimination properties have been recently revised across the family and now include sequence based criteria so that isolates sharing less than 72% identical nucleotides between their entire coat protein (CP) or replication protein (RdRp) genes or 80% identical amino acids between the corresponding proteins are considered to represent distinct species [1]. Molecular criteria for genus discrimination are currently less precisely defined but include genome organisation and phylogenetic relatedness of the various viral proteins.

A study of the molecular variability of BanMMV isolates infecting banana in Guadeloupe (Teycheney et al., submitted) recently lead to the identification of a short nucleotide sequence belonging to a distinct viral species within one of the studied banana accessions. In the present report, the determination and analysis of the partial nucleotide sequence of this agent (2917 nucleotides from the genome 3′ end excluding the poly(A) tail) is reported and evidence is provided that it represents a new virus species in the *Flexiviridae* family for which the name Banana Virus X (BVX) is proposed. During the course of this analysis, evidence was also obtained for the presence in the BVX isolate studied of shorter viral RNA species corresponding to defective RNAs (D RNAs) generated by internal deletions spanning a region from the TGB1 to the CP genes.

Materials and methods

Virus sources and RNA extractions

Two distinct mother plants of the *Musa* accession Som (Ney mannan), numbered 5.1 and 5.3, were originally identified as hosts of BVX. These plants are part of CIRAD's *ex situ Musa* collection maintained in Guadeloupe (French West Indies). They were used as sources of BVX in all further experiments. Double stranded RNA (dsRNA) and total RNAs were extracted from banana leaf tissues as described by Morris and Dodds [13] as modified by Valverde et al. [21] and by Chang et al. [3].

cDNA synthesis and cloning

The 3′ terminal 2587 nucleotides of the BVX genome were amplified by 3′ Rapid Amplification of cDNA Ends (3′RACE), using a Smart RACE PCR kit and following the supplier's

instructions (BD Science, Palo Alto, USA). First strand cDNA was generated from purified dsRNAs using primer LD-polyT (5′-CACTGGCGGCCGCTCGAGCATGTAC $(T)_{30}$NN-3′). Long distance PCR amplification of the viral cDNA was then performed using 3′ primer LDPrim (5′-CACTGGCGGCCGCTCGAGCATGTAC-3′) and a 5′ primer named BVX (5′-CGGTACTTACTGCTACCTGAG-3′), which was designed from the short internal BVX fragment obtained in the course of a BanMMV variability study (Teycheney et al., submitted). PCR products were cloned into the pGEM-T easy vector (Promega, Madison, USA). Sequencing of recombinant cDNA clones was performed by Genome Express (Meylan, France).

Sequence analysis

Multiple sequence alignments and phylogenetic reconstructions (neighbour-joining) were performed using the CLUSTAL W program with randomized bootstrapping evaluation of branching validity [19, 20]. Phylogenetic trees were then visualised using TREEVIEW [14]. Average pair-wise distances were calculated using Mega2 [10] while the presence of potential recombination events in the data sets were evaluated using the GENECONV v. 1.81 software (www.math.wustl.edu/_sawyer/geneconv/) [16].

Characterization of D-RNA junctions

RT-PCR experiments were performed on total RNAs extracted from leaf material of infected plants or from uninfected plants used as controls. First strand cDNA was synthesized using BD Science PowerScript reverse transcription kit and oligo $(dT)_{18}$, following the conditions and protocols recommended by the supplier (BD Science, Palo Alto, USA). Subsequent PCR amplifications were performed using two distinct pairs of primers (see Fig. 1 and Table 2 for the sequence and position of the primers on the BVX genome). All PCR primers were used at 15 pmol and cycling parameters were 95 °C/1 min followed by 35 cycles at 95 °C/30 sec, annealing temperature/30 sec, 68 °C/3 min. Cloning and sequencing were performed as described above.

Attempts to mechanically transmit BVX to herbaceous hosts

Crude extract was prepared by grinding 1 g of leaf from infected banana plants 5.1 and 5.3 and from an uninfected control banana plant in 5 ml of phosphate buffer. Carborundum-dusted leaves of *Brassica napus*, *Lactuca sativa* cv Trocadéro, *Chenopodium quinoa*, *C. foetidum*, *C. album, C. murale*, *Vigna sinensis*, *Nicotiana benthamiana* and *N. tabacum* cv xanthi plants were mechanically inoculated at the 3 leaf stage. Total RNAs were extracted from inoculated and systemic leaves 15 days post inoculation using Qiagen plant Rneasy minikit and used for the amplification of part of BVX ORF5 by RT-PCR. Primer 3 (see Table 2) was used for the reverse transcription step and also for the PCR amplification, together with primer 4 (see Table 2), generating an expected 477 bp product from infected control plants. First strand cDNAs were prepared using AMV reverse transcriptase (USB, Cleveland, USA) according to the manufacturer's instructions. PCR primers were used at 10 pmol and PCR parameters were a pre-incubation at 95 °C for 3 min followed by 35 cycles at 95 °C for 30 sec, 56 °C for 1 min, 72 °C for 1 min and a final elongation step at 72 °C for 10 min.

Results

Cloning of the 3′ end of the banana virus X genome

Following screening of CIRAD's *Musa* collection for infection by BanMMV using an immunocapture reverse transcription nested PCR technique (Teycheney et al.,

submitted), two distinct mother plants of one accession (Som, Ney mannan, an ABB genotype), numbered 5.1 and 5.3, were found to host RNA-dependent RNA polymerase (RdRp) viral sequences that differed from those of BanMMV and that could not be related to any other virus reported to infect *Musa* spp. These plants displayed no particular symptoms and efforts to transmit either BanMMV or the new virus to a range of potential herbaceous hosts proved unsuccessful (result not shown). A primer was designed from the sequence of the short cDNA fragment of this unknown virus. This primer was used with an oligo(dT)-containing anchor primer in 3′RACE experiments performed on dsRNAs extracted from the 5.1 and

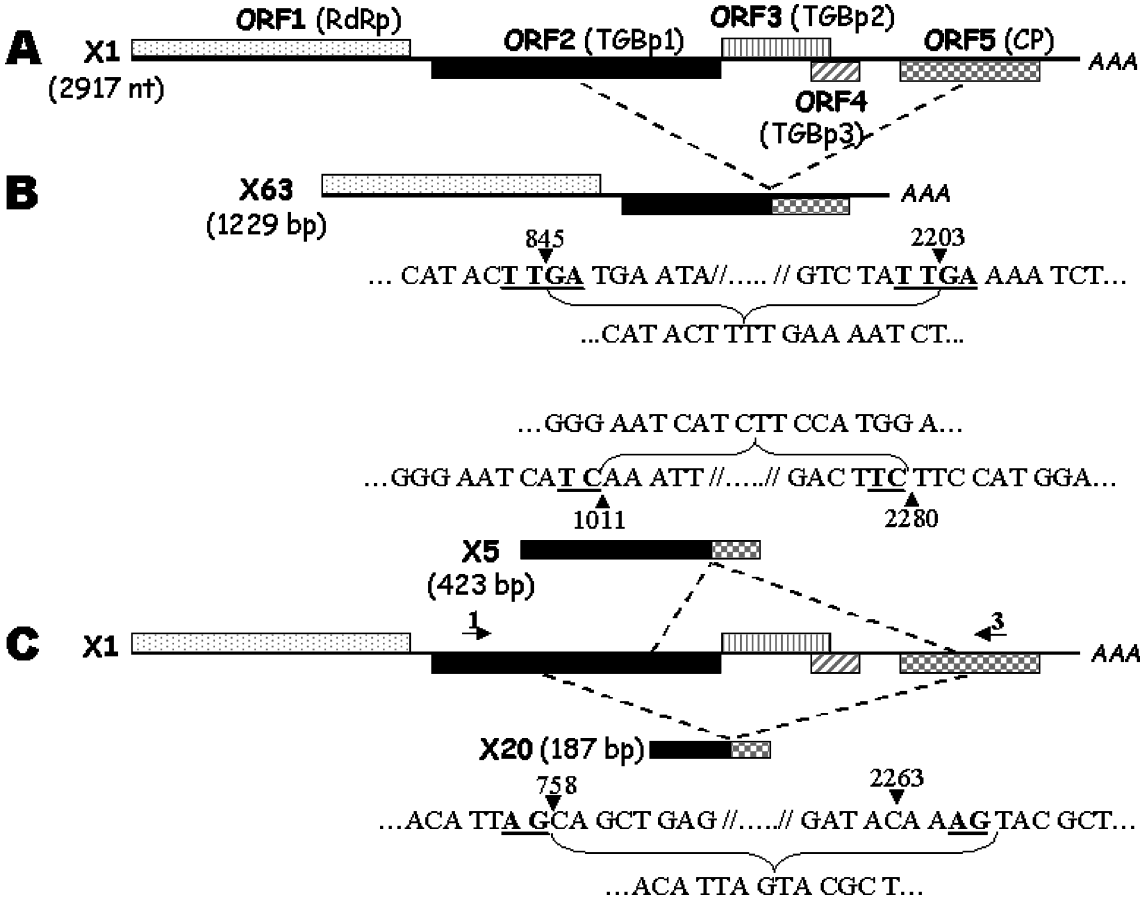

Fig. 1. Genome organization of banana virus X and structure of its defective RNAs. **A** Organization of the genome of BVX deduced from the nucleotide sequence of its 3′ terminal 2917 nucleotides. **B** Structure of defective RNA clone X63. **C** Structure of the PCR fragments generated from BVX D-RNAs using primer combinations 1 + 3. Internal deletions are illustrated as dotted lines while the name and size (between parentheses) of the cDNA clones amplified by RT-PCR are listed on the left side of the molecules. Sequences around the donor and acceptor sites are indicated (upper line), as well as the sequences of the resulting deletants (lower line). The precise breakpoints are indicated by arrowheads and the numbering of the nucleotides immediately bordering the deletions is provided. Bases duplicated between the donor and acceptor sites are indicated in bold underlined

5.3 plants as well as from healthy plants used as a control. These experiments generated a major PCR product of ca 2.5 kbp and several minor, smaller sized products when performed on dsRNAs extracted from either plant 5.1 or 5.3, while no amplification was observed from dsRNAs extracted from healthy control plants. The 2.5 kbp products amplified from plants 5.1 and 5.3 were separately cloned into pGEM-T and their ends were sequenced. All sequenced cDNA clones showed identical sequences at both ends and were therefore assumed to be identical. One of these clones, X1, which originates from plant 5.1, was selected for complete sequencing. The sequence obtained from this clone was 2587 nt long, excluding the poly-A tail. This sequence was contiged with the original internal fragment obtained during the BanMMV variability study and with a further short internal cDNA to finally yield a 2917 nt sequence.

Genome organisation of Banana virus X

As shown in Fig. 1A, the partial sequence of the BVX genome obtained encodes five open reading frames (ORFs). Based on sequence comparisons, putative functions could be attributed to the corresponding proteins. The ORF1, for which the last 906 nucleotides were sequenced, corresponds to an RNA-dependent RNA polymerase (RdRp) and comprises the characteristic RdRp core motif (S/TGx_3Tx_3NS/Tx_{22}GDD, nt 427–534) which is conserved among members of the "Sindbis-like" supergroup of RNA polymerases [9]. The next three ORFs, ORF 2 (nt 930–1613), ORF 3 (nt 1610–1963) and ORF 4 (nt 1887–2090) encode putative triple gene block proteins (TGBp1, TGBp2, TGBp3), whose sequences and function in viral movement are conserved among TGB-containing members of the family *Flexiviridae* [1]. The stop codon of ORF2 overlaps with the start codon of ORF3 (sequence **A<u>TGA</u>** where the stop codon is underlined and the start codon is in bold) while OFR3 and ORF4 show a more significant overlap of over 70 nt. The last ORF, ORF5 (nt 2181–2834), encodes the putative coat protein (CP) and contains the conserved amphipathic core CP motif (KYAAFDxFx_2Vx_3NA, nt 2588–2636) shared between many viruses belonging to the *Flexiviridae* family, including BanMMV [5]. The C terminus of this CP contains the putative catalytic motif for a threonine protease HX_3Dx_{15}TGG (nt 2709–2777), which is partly present in the coat proteins of Fovea- and Allexiviruses [5]. The BVX 3′ untranslated region (3′UTR) is 82 nt long and comprises a putative polyadenylation signal (AATAAA$x_{26}A_n$, nt 2886–2917).

With the presence of a movement-associated triple gene block and the absence of additional ORFs 3′ or 5′ of the CP gene, the genetic organisation of the genome of BVX closely resembles that of Potexviruses and Foveaviruses which are members of the *Flexiviridae* family [1, 6].

Sequence comparison between BVX and other members of the Flexiviridae family

Pair-wise percentages of amino acid sequence identity were calculated between the five proteins encoded by the BVX ORFs and the homologous proteins from

Table 1. Percentages of amino acid sequence identity between the proteins encoded by the open reading frames (ORFs) of banana virus X and the corresponding proteins from representative members of the *Flexiviridae* family

		ORF1 (RdRp)	ORF2 (TGBp1)	ORF3 (TGBp2)	ORF4 (TGBp3)	ORF5 (CP)
Allexivirus	GarV-A	30.9	32.9	38.2	na	23.8
	GarV-C	29.8	29.3	35.3	na	23.8
Potexvirus	WClMV	30.3	30.7	27.7	21.0	25.5
	ScVX	31.4	22.7	36.1	27.3	23.0
	PAMV	31.4	24.2	25.2	14.1	22.2
	SMYEV	25.0	27.5	41.3	14.5	25.1
	BaMV	29.3	26.9	34.3	17.6	20.5
	PVX	30.3	31.4	41.1	22.7	23.5
	ClYMV	27.2	30.5	42.5	23.7	22.9
Mandarivirus	ICRSV	30.9	29.8	31.1	21.7	23.6
Carlavirus	AcLV	57.6	26.9	39.8	30.6	20.5
	HpLV	57.6	30.2	35.5	26.7	22.3
	BlScV	58.1	32.9	38.1	20.3	19.8
Foveavirus	ASPV	57.6	31.5	40.9	28.4	25.7
	RSPaV	57.1	33.9	39.7	22.4	27.2
Trichovirus	ACLSV	41.4	na	na	na	10.9
	ChMLV	41.9	na	na	na	10.4
Vitivirus	GVA	43.4	na	na	na	13.9
	GVB	39.2	na	na	na	11.8
Capillovirus	ASGV	38.7	na	na	na	10.9
	CVA	44.0	na	na	na	13.0
Unassigned	SCSMaV	52.9	23.9	24.8	17.5	21.1
	CGRMV	61.3	35.8	38.2	22.7	25.7
	CNRMV	60.2	33.9	37.3	22.7	29.0
	BanMMV	52.4	25.6	38.0	13.8	24.8

na: does not apply (no homologous protein available for comparison). The abbreviations used are the following, with GenBank accession numbers shown in parentheses: GarV-A: garlic virus A (NC_003375), GarV-C: garlic virus C (NC_003376), AcLV: aconitum latent virus (NC_002795), HpLV: hop latent virus (NC_002552), BlScV: blueberry scorch virus (NC_003499), ASPV: apple stem pitting virus (NC_003462), RSPaV: rupestris stem pitting-associated virus (NC_001948), BaMV: bamboo mosaic virus (NC_001642), ClYMV: clover yellow mosaic virus (NC_001753), PVX: potato virus X (NC_001455), WClMV: white clover mosaic virus (NC_003820), SMYEV: strawberry mild yellow edge virus (NC_003794), ScVX: scallion virus X (NC_003400), PAMV: potato aucuba mosaic virus (NC_003632), ICRSV: indian citrus ringspot virus (NC_003093), GVA: grapevine virus A (NC_003604), GVB: grapevine virus B (NC_003602), ACLSV: apple chlorotic leaf spot virus (NC_001409), ChMLV: cherry mottle leaf virus (NC_002500), CVA: cherry virus A (NC_003689), ASGV: apple stem grooving virus (NC_001749), BanMMV: banana mild mosaic virus (NC_002729), CGRMV: cherry green ring mottle virus (NC_001946), CNRMV: cherry necrotic rusty mottle virus (NC_002468), ScSMV: sugarcane striate mosaic virus (NC_003870)

representative members of all genera and unassigned viruses in the *Flexiviridae* family (Table 1). Allexiviruses were not included in the TGBp3 comparisons because the size and sequence of their corresponding ORF is very different from those of any other member of the family. Likewise, members of *Trichovirus*, *Capillovirus* and *Vitivirus* genera could not be included in TGB comparisons because these genera do not have a TGB but have instead a single movement protein. Overall, the most conserved protein appears to be the partial RdRp with identity levels reaching slightly above 60% for the two closely related and unassigned cherry green ring mottle virus (CGRMV) and cherry necrotic rusty mottle virus (CNRMV). Values above 55% were also observed for members of the *Foveavirus* and *Carlavirus* genus but much lower values were observed for other genera in the family, for example, the *Potexvirus* (range 25–31%). All other ORFs were significantly less conserved with highest identity levels observed of only 35.8% (TGBp1 with CGRMV), 42.5% (TGBp2 with clover yellow mosaic virus, a member of genus *Potexvirus*), 30.6% (TGBp3 with aconitum latent virus, a member of genus *Carlavirus*) and only 29% for the coat protein (with CNRMV).

These very low values, which were confirmed by Blast analysis against databanks (data not shown), indicate that BVX is a new virus and that it is only very distantly related to any other previously sequenced virus in the *Flexiviridae* family. In fact, the amino acid identity values presented in Table 1 are in most cases in the range observed for comparison within the *Flexiviridae* between species belonging to different genera which usually show less then ca 45% nucleotide or 40% amino acid identity in these genes [1].

This analysis is further extended by the phylogenetic reconstructions for each of the BVX encoded proteins and the homologous proteins from selected representative members of the *Flexiviridae* family shown in Fig. 2. Only the RdRp sequence provided a bootstrap-supported phylogenetic association of the BVX protein (Fig. 2A), while for the TGBp1-TGBp2, TGBp3 and CP, the BVX proteins showed no phylogenetically sturdy grouping with any member of the *Flexiviridae* (Fig. 2B–E). In addition, the weak groupings observed were with different viruses/genera for each analyzed protein.

Molecular characterization of BVX D-RNAs

During the process of cloning the 3′ end of the genome of BVX, a shorter, ca. 1.2 kb cDNA was amplified during the 3′RACE experiments performed on BVX dsRNAs extracted from plant 5.1 and cloned into pGEM-T. Sequencing of the ends of this cDNA clone, named X63, showed that it had the same ends as the 2587 bp X1 clone. Complete sequencing of X63 showed that it carries a 1356 nucleotide-long internal deletion (nt 1176–2532) when compared to the sequence of clone X1 (Fig. 1B). In order to check whether this truncated cDNA resulted from an artefact during the 3′RACE or from the presence of a truncated recombinant RNA in infected plants, two pairs of primers were designed for specific amplification of putative BVX deleted RNA molecules. Table 2 and Fig. 1C show the sequence of these primers and their position on the BVX genome. RT-PCR experiments

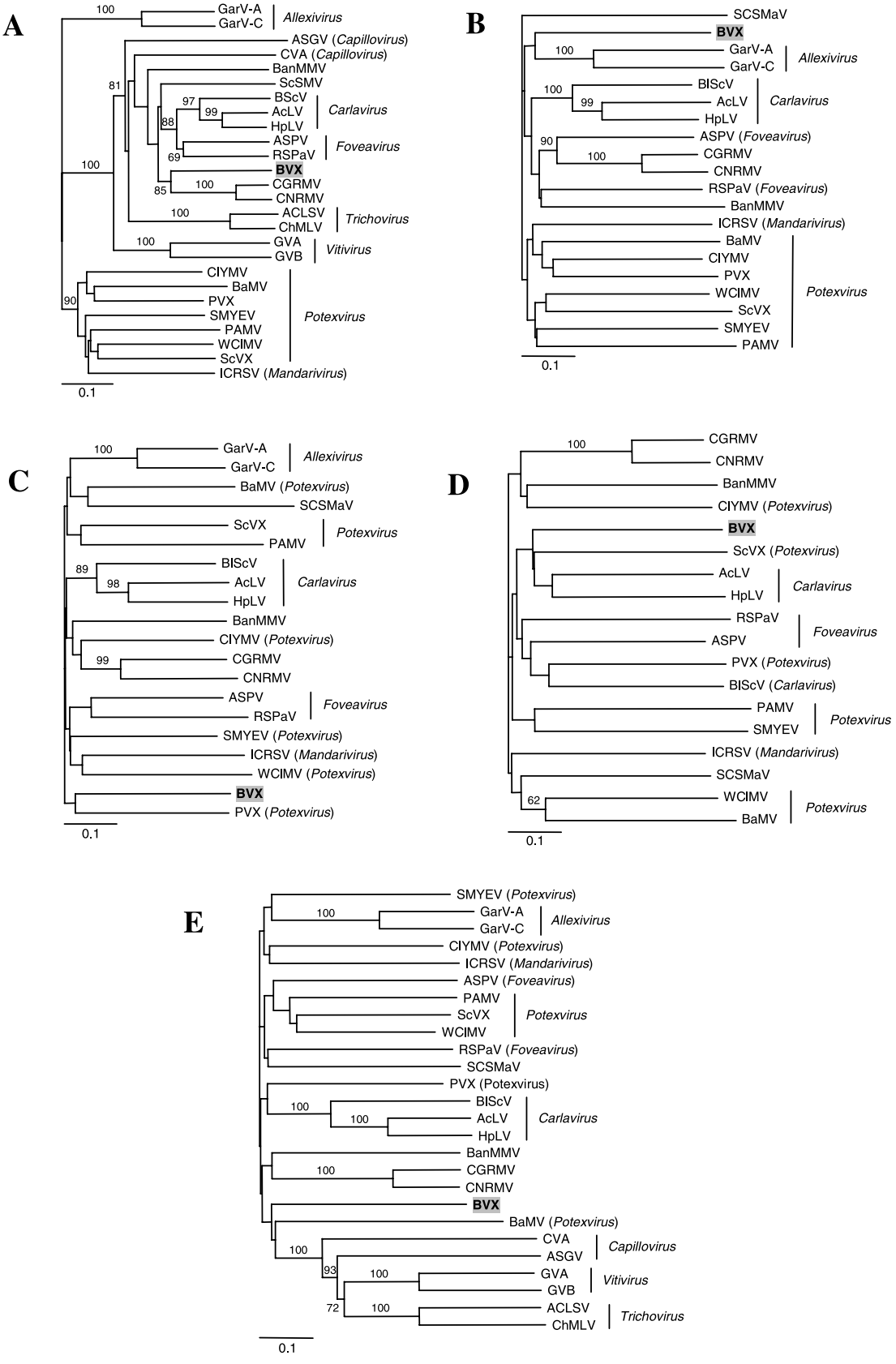
A
100
GarV-A
GarV-C
Allexivirus
ASGV (Capillovirus)
CVA (Capillovirus)
BanMMV
ScSMV
81
97
BScV
99
AcLV
HpLV
Carlavirus
88
ASPV
RSPaV
Foveavirus
69
100
BVX
85
100
CGRMV
CNRMV
100
ACLSV
ChMLV
Trichovirus
100
GVA
GVB
Vitivirus
ClYMV
BaMV
90
PVX
SMYEV
PAMV
Potexvirus
WClMV
ScVX
ICRSV (Mandarivirus)
0.1
B
SCSMaV
BVX
100
GarV-A
GarV-C
Allexivirus
100
BIScV
99
AcLV
HpLV
Carlavirus
90
ASPV (Foveavirus)
100
CGRMV
CNRMV
RSPaV (Foveavirus)
BanMMV
ICRSV (Mandarivirus)
BaMV
ClYMV
PVX
WClMV
Potexvirus
ScVX
SMYEV
PAMV
0.1
C
100
GarV-A
GarV-C
Allexivirus
BaMV (Potexvirus)
SCSMaV
ScVX
PAMV
Potexvirus
89
BIScV
98
AcLV
HpLV
Carlavirus
BanMMV
ClYMV (Potexvirus)
99
CGRMV
CNRMV
ASPV
RSPaV
Foveavirus
SMYEV (Potexvirus)
ICRSV (Mandarivirus)
WClMV (Potexvirus)
BVX
PVX (Potexvirus)
0.1
D
100
CGRMV
CNRMV
BanMMV
ClYMV (Potexvirus)
BVX
ScVX (Potexvirus)
AcLV
HpLV
Carlavirus
RSPaV
ASPV
Foveavirus
PVX (Potexvirus)
BIScV (Carlavirus)
PAMV
SMYEV
Potexvirus
ICRSV (Mandarivirus)
SCSMaV
62
WClMV
BaMV
Potexvirus
0.1
E
SMYEV (Potexvirus)
100
GarV-A
GarV-C
Allexivirus
ClYMV (Potexvirus)
ICRSV (Mandarivirus)
ASPV (Foveavirus)
PAMV
ScVX
Potexvirus
WClMV
RSPaV (Foveavirus)
SCSMaV
PVX (Potexvirus)
100
BIScV
100
AcLV
HpLV
Carlavirus
BanMMV
100
CGRMV
CNRMV
BVX
BaMV (Potexvirus)
CVA
Capillovirus
100
ASGV
93
100
GVA
Vitivirus
GVB
72
100
ACLSV
Trichovirus
ChMLV
0.1

Table 2. Primers used for the amplification of BVX D-RNAs and BVX ORF5 by RT-PCR. The position of primers on the sequence of clone X1 and their nucleotide sequences (5′ → 3′) are indicated

Primer	Position	Sequence
1	963–983 (S)	GGCTTTGTTAGAACAATAAAT
2[a]	1159–1175/2533–2536 (S)	**GGAGGGTCAACATACTT**TTGA
3	2634–2653 (AS)	CCACCTTCAGGAACTAAAGC
4	2176–2196 (S)	AGGAAATGGAAGCAGTTCAGA

[a]: primer 2 was designed so as to span the deletion found in D RNA X63. Its sequence therefore contains an upstream part corresponding to an ORF2 sequence (in bold) and a downstream part corresponding to an ORF5 sequence (underlined)
S: sense primer; AS: antisense primer

were carried out on total RNAs extracted from infected plants 5.1 and 5.3 or from uninfected control plants.

RT-PCR experiments performed using upstream primer 2 (designed so as to overlap the deletion site identified in clone X63) and downstream primer 3 (located in ORF5) generated a PCR product of the expected size (138 bp) from RNAs extracted from both infected plants, but not from uninfected control plants (data not shown). The cDNAs obtained from plant 5.1 (clones X35, X44) and plant 5.3 (clones X54, X80) were cloned and sequenced, revealing the expected nucleotide sequence (results not shown). Given that no amplification was expected from the undeleted genomic RNA because primer 2 should not anneal efficiently to it due to its hybrid nature, these results were a strong indication that deleted molecules similar to X63 pre-exist in the total RNAs extracts from infected plants.

A similar RT-PCR approach was also developed using upstream primer 1 (located in ORF2 upstream the recombination site of clone X63) and downstream primer 3. Again, amplification products shorter than the one expected from the undeleted genome were obtained and were specific of RNA extracts from infected plants. One cDNA clone (X16), obtained from plant 5.3 shows the same deletion as that of clone X63, confirming that BVX recombinant RNA molecules are present in infected plants (results not shown). Interestingly, two additional recombinant forms were cloned from plant 5.1 (clone X5) and from plant 5.3 (clone X20), showing that other deleted molecules with different breakpoints are also present

Fig. 2. Phylogenetic trees reconstructed for each of the proteins encoded by the open reading frames (ORFs) of banana virus X and the corresponding proteins from representative members of the *Flexiviridae* family. Trees were constructed by the neighbour-joining method from the amino acid sequences of the following proteins: partial RdRp **A**, TGBp1 **B**, TGBp2 **C**, TGBp3 **D** and CP **E**. Only bootstrap values above 70% are indicated. The scale bars represent 0.1 substitutions per site. The BVX proteins are indicated in a shaded box

in the infected plants (Fig. 1C). These two additional clones contained deletions of 1503 nt and 1267 nt for clone X20 and clone X5, respectively. In all three deleted molecules, the deletion spanned part of ORF2, ORFs 3 and 4 and part of ORF5 (CP) and resulted in a truncation of ORF2 while the remainder of the CP gene was joined out of frame and therefore unable to express CP-like peptides. Analysis of the region surrounding the recombination sites showed that the recombinants reported here result from non-homologous recombination events. It is likely that these deleted molecules correspond to defective RNAs (D RNAs) forms of the BVX genomic RNA.

Discussion

The work reported here describes the molecular characterization of banana virus X (BVX), a new virus infecting *Musa* spp. The genome organisation deduced from the sequence of the last 2917 nucleotides of its genomic RNA and the phylogenetic affinities of the encoded proteins unambiguously position BVX in the recently described *Flexiviridae* family. Sequence comparisons (Table 1) and phylogenetic analyses (shown in Fig. 2 for each of the BVX proteins) show that it is not possible to fit BVX in any of the existing genera of this family. The genomic organisation of BVX is similar to that of members of two genera in the family, *Potexvirus* and *Foveavirus*. The BVX RdRp, on the other hand, clearly does not belong to the *Potexvirus-Mandarivirus* cluster, which excludes BVX from the *Potexvirus* genus. Similarly, although clustering in the major group of RdRp together with the *Foveavirus* genus, the RdRp of BVX is positioned with significant bootstrap support in a different subcluster than the *Foveavirus* RdRp. Similar results together with low identity levels in the other viral proteins have recently led to the exclusion of CGRMV and CNRMV from the *Foveavirus* genus and their positioning in the *Flexiviridae* family as unassigned to any described genus. Finally, BVX is not particularly close to any of the unassigned viruses in the family. It is significantly removed from banana mild mosaic virus (BanMMV, [5]) and despite the fact that the closest RdRp affinities, supported by an 85% bootstrap value, are with CGRMV and CNRMV, there is clearly no close relationship with these two viruses in any of the other viral proteins. The most logical conclusion of this analysis is therefore that BVX is a new virus in the *Flexiviridae* family which could, in time, define a new genus in this family.

A small scale study of the molecular variability of BVX was performed by generating ten 559 nucleotides-long sequences (from within ORF1) from five *Musa* samples collected in different locations in Guadeloupe (results not shown). As previously observed for plants 5.1 and 5.3 (see above), these five infected plants were also asymptomatic, so that there is currently no evidence that infection by BVX can cause any visible symptoms in *Musa* spp. All sequences generated from these five plants were almost identical, showing a maximum of 2 nt substitutions among themselves or with the sequences derived from plants 5.1 and 5.3. These results indicate that BVX seems to be reasonably distributed in *Musa* plants in Guadeloupe because random sampling easily provided evidence

for BVX infection in 7 of the 41 plants tested. In addition, these preliminary results indicate a very low variability for BVX or at least for its Guadeloupe population, which is in marked contrast to that encountered for BanMMV, another member of the *Flexiviridae* family infecting *Musa* spp, which displays very high levels of variability in the RdRp ORF (Teycheney et al., submitted).

Although several attempts to characterize either BVX genomic RNA or D-RNAs by northern blot hybridization proved unsuccessful probably as a consequence of very low viral titers, the results of the various RT-PCR analyses reported here also demonstrate the presence of defective RNAs (D-RNAs) of BVX in the original BVX sources. Three different D-RNAs were characterized by RT-PCR in this work, which harbour deletions of 1268, 1358 and 1503 nucleotides, respectively. In each case, a single deletion event joined out of frame sequences internal to ORF2 to sequences internal to ORF5 (Fig. 1B & C). In this regard too, BVX differs from other members of the *Flexiviridae* family, in that previously characterized D-RNAs of white clover mosaic virus [22, 24], cassava common mosaic virus (CsCMV, [2]) and bamboo mosaic virus (BaMV, [25]) result from a recombination event between the 5′ part of ORF1 and the 3′ part of ORF5, and potentially encode a RdRp-CP fusion protein. In the particular case of ClYMV, White et al. [23] demonstrated that the RdRp-CP fusion protein is synthesized and that the coding capacity of the D-RNA determines its *in vivo* accumulation. Thus far, no interfering properties have been demonstrated for any of the D-RNAs found associated with members of the *Flexiviridae*, contrary to the well documented situations encountered with other plant RNA viruses. It has long been known that defective interfering RNAs (DI RNAs) observed following serial passages for many plant and animal viruses can affect viral replication and the structure of the population of the helper virus [8, 15]. On the other hand, the effect of the defective viral genomes that have been described in natural field isolates of several plant viruses such as *Citrus tristeza virus* [12], *Beet curly top virus* [18] or *Cotton leaf curl virus* [11] still needs to be precisely evaluated.

The recombination breakpoints of the BVX D-RNAs were characterised by an RT-PCR-based approach. Sequence analysis showed that non-homologous recombination mechanisms are involved in the synthesis of these D RNAs, as previously reported for other defective and defective interfering RNAs of monopartite plus-strand RNA plant viruses [24]. They also showed that in most cases short, 2 or 4 base pairs duplicated sequences are present at or near the recombination sites (see Fig. 1B & C). On the other hand, these recombination sites are not located in particularly AU-rich regions, which have been reported as promoting or facilitating RNA recombination, especially in the case of tomato bushy stunt virus (TBSV) DI RNAs [17] nor could stem-loop secondary structures, which were shown to play an important role in the selection of the recombination site within or in the vicinity of the 5′ and 3′ junction borders of *potexvirus* D-RNAs [25], be conclusively identified in the case of BVX. It is therefore possible that BVX differs also from other members of the family *Flexiviridae* in the recombination processes leading to the synthesis of its D-RNAs.

Acknowledgements

The authors thank Mark Tepfer for critical reading of the manuscript and Kathryn Mayo for help with the English of the manuscript.

References

1. Adams MJ, Antoniw JF, Bar-Joseph M, Brunt AA, Candresse T, Foster GD, Martelli GP, Milne RG, Fauquet CM (2004) The new plant virus family *Flexiviridae* and assessment of molecular criteria for species demarcation. Arch Virol 149: 1045–1060
2. Calvert LA, Cuervo MI, Ospina MD, Fauquet CM, Ramirez BC (1996) Characterization of *Cassava common mosaic virus* and a defective RNA species. J Gen Virol 77: 525–530
3. Chang S, Puryear J, Cairney J (1993) A simple and efficient method for isolating RNA from pine trees. Plant Mol Biol Rep 11: 113–116
4. Dallot S, Acuna P, Rivera C, Ramirez P, Cote F, Lockhart BE, Caruana ML (2001) Evidence that the proliferation stage of micropropagation procedure is determinant in the expression of banana streak virus integrated into the genome of the FHIA 21 hybrid (*Musa* AAAB). Arch Virol 146: 2179–2190
5. Gambley CF, Thomas JE (2001) Molecular characterisation of banana mild mosaic virus, a new filamentous virus in *Musa* spp. Arch Virol 146: 1369–1379
6. Hull R (2002) Plant virology. Academic Press, San Diego
7. Jones DR (2000) Diseases of banana, abacá and enset. CABI Publishing, Wallingford, UK, pp 241–293
8. Kirkwood TBL, Bangham CRM (1994) Cycles, chaos and evolution in virus cultures: a model of defective interfering particles. Proc Natl Acad Sci USA 91: 8685–8689
9. Koonin EV, Dolja VV (1993) Evolution and taxonomy of positive-strand RNA viruses: implication of comparative analysis of amino acid sequences. Crit Rev Biochem Mol Biol 28: 375–430
10. Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: Molecular Evolutionary Genetics Analysis (MEGA) software. Bioinformatics 17: 1244–1245
11. Liu Y, Robinson DJ, Harrison BD (1998) Defective forms of cotton leaf curl virus DNA-A that have different combinations of sequence deletion, duplication inversion and rearrangement. J Gen Virol 79: 1501–1508
12. Mawassi M, Mietiewska E, Hilf ME, Ashoulin L, Karasev AV (1995) Multiple species of defective RNAs in plants infected with citrus tristeza virus. Virology 214: 264–268
13. Morris TJ, Dodds JA (1979) Isolation and analysis of double stranded RNA from virus-infected plant and fungal tissue. Phytopathology 69: 854–858
14. Page DR (1996) TreeView: an application to display phylogenetic trees on personal computer. Comput Appl Biosci 12: 357–358
15. Roux L, Simon AE, Holland JJ (1991) Effects of defective interfering viruses on virus replication and pathogenesis *in vitro* and *in vivo*. Adv Virus Res 40: 181–211
16. Sawyer S (1989) Statistical tests for detecting gene conversion. Mol Biol Evol 6: 526–538
17. Shapka N, Nagy PD (2004) The AU-rich RNA recombination hotspot sequence of brome mosaic virus is functional in Tombusviruses: implications for the mechanism of RNA recombination. J Virol 78: 2288–2300
18. Stenger DC (1995) Genotypic variability and the occurrence of less then genome-length viral DNA forms in a field popumation of beet curly top geminivirus. Phytopathology 85: 1316–1322
19. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 11: 4673–4680

20. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acid Res 24: 4876–4882
21. Valverde RA, Dodds JA, Heick JA (1986) Double-stranded ribonucleic acid from plants infected with viruses having elongated particles and undivided genomes. Phytopathology 76: 459–465
22. White KA, Bancroft JB, Mackie GA (1991) Defective RNAs of clover yellow mosaic virus encode nonstructural/coat protein fusion products. Virology 183: 479–486
23. White KA, Bancroft JB, Mackie GA (1992) Coding capacity determines in vivo accumulation of a defective RNA of clover yellow mosaic virus. Proc Natl Acad Sci USA 66: 3069–3076
24. White KA, Morris TJ (1999) Defective and defective interfering RNAs of monopartite plus-strand RNA plant viruses. Curr Top Microbiol Immunol 239: 1–17
25. Yeh TY, Lin YC, Chang YC, Hsu YH, Lin NS (1999) A defective RNA associated with bamboo mosaic virus and the possible common mechanisms for RNA recombination in potexviruses. Virus Genes 18: 121–128

Author's address: Dr. Pierre-Yves Teycheney, CIRAD-FLHOR, Station de Neufchâteau, Sainte-Marie, F-97130 Capesterre Belle-Eau, Guadeloupe, FWI; e-mail: teycheney@cirad.fr