

## Sequence diversity among badnavirus isolates infecting yam (*Dioscorea* spp.) in Ghana, Togo, Benin and Nigeria

A. O. Eni · J. d'A. Hughes · R. Asiedu ·  
M. E. C. Rey

Received: 16 May 2008 / Accepted: 21 October 2008 / Published online: 22 November 2008  
© Springer-Verlag 2008

**Abstract** We analysed the sequence diversity in the reverse transcriptase (RT)/ribonuclease H (RNaseH) coding region of 19 badnavirus isolates infecting yam (*Dioscorea* spp.) in Ghana, Togo, Benin, and Nigeria. Phylogenetic analysis of the deduced amino acid sequences revealed that the isolates are broadly divided into two distinct species, each clustering with *Dioscorea alata* bacilliform virus (DaBV) and *Dioscorea sansibarensis* bacilliform virus (DsBV). Fourteen isolates had 90–96% amino acid identity with DaBV, while four isolates had 83–84% amino acid identity with DsBV. One isolate from Benin, BN4Dr, was distinct and had 77 and 75% amino acid identity with DaBV and DsBV, respectively, and may be a member of a new badnavirus species infecting yam in West Africa. Viruses of the two main species were present in Ghana, Togo and Benin and were observed to infect both *D. alata* and *D. rotundata* indiscriminately. This is the first confirmed report of DsBV infection in yam in Ghana and Togo. The results of this study demonstrate that members of two distinct species of badnaviruses infect yam in the West African yam zone and suggest a putative new species, BN4Dr. We also conclude that these species are not confined to limited geographic regions or specific for yam host

species. However, the three badnavirus species are serologically related. The sequence information obtained from this study can be used to develop PCR-based diagnostics to detect members of the various species and/or strains of badnaviruses infecting yam in West Africa.

### Introduction

Yam contributes more than 200 dietary calories per person each day for millions of people in West Africa, where approximately 93% of the world production of 51.9 mt was grown in 2007 [11]. Major yam producers in West Africa are Nigeria, Ghana, Côte d'Ivoire, Benin and Togo. Average yam consumption per capita per day is highest in Benin (364 kcal), followed by Côte d'Ivoire (342 kcal), Ghana (296 kcal), and Nigeria (258 kcal) [25]. The use of infected vegetative propagules and uncontrolled introductions of infected germplasm by farmers through porous land borders have resulted in the presence of yam viruses in all yam-growing areas of West Africa [1, 9, 10, 17, 24, 32, 39, 40]. Yam virus symptoms (mainly chlorotic-mosaic symptoms on leaves) result in patchy loss of photosynthetic activity and often increased rates of chloroplast senescence, leading to reduced sugar formation and minimal starch storage, thus causing significant reduction in tuber yield and quality [40]. Yam viruses are of substantial economic importance not only because of yield losses they cause, but also due to the high cost of preventive measures [8].

Viruses in the genus *Badnavirus*, family *Caulimoviridae*, are reported to infect a wide range of economically important tropical crops such as cacao [30], banana [18], yam [5], sugarcane [3], citrus [23], and rice [3]. They are pararetroviruses, characterised by non-enveloped

A. O. Eni (✉) · J. d'A. Hughes · R. Asiedu  
International Institute of Tropical Agriculture (IITA),  
Oyo Road, Ibadan, Nigeria  
e-mail: aeni@cgiar.org

A. O. Eni · M. E. C. Rey  
School of Molecular and Cell Biology,  
University of the Witwatersrand, Johannesburg, South Africa

*Present Address:*  
J. d'A. Hughes  
AVRDC – The World Vegetable Center, Shanhua, Taiwan

bacilliform particles containing circular dsDNA genomes of 7.0–7.6 kb [12]. Badnaviruses were first reported in yam in association with a member of the genus *Potyvirus* as the causative agent of mosaic symptoms observed in *D. alata* leaves with internal brown spot disease in Barbados [22]. Badnaviruses infecting yam from several countries were partially characterised by Phillips et al. [34], and two viruses were tentatively designated *Dioscorea alata* bacilliform virus (DaBV) and *Dioscorea bulbifera* bacilliform virus (DbBV). The complete nucleotide sequences of a Nigerian isolate of DaBV and a bacilliform virus isolated from *D. sansibarensis* from Benin, DsBV, have been analysed [5, 37]. The genome sizes of DaBV and DsBV were shown to be approximately 7.4 and 7.26 kb, respectively, with 38.1% sequence variability [5, 37]. More recently, the analysis of 45 partial RT/RNaseH sequences derived from DNA extracted from yams revealed the presence of 11 new badnavirus groups in yam from the South Pacific Islands. The results also suggest that some of the badnavirus groups maybe endogenous or integrated sequences [26].

Yam-infecting badnaviruses are widespread in West Africa [9]. Analysis of 1,632 yam leaves obtained during surveys conducted in 2004 and 2005 in Ghana, Togo and Benin revealed the occurrence of yam-infecting badnavirus in 45.3% of the leaves tested. Of the four viruses detected, incidence of badnavirus was highest followed by yam mosaic virus (YMV), yam mild mosaic virus (YMMV) and cucumber mosaic virus (CMV). Badnavirus was also found to be the most widely distributed yam virus in all three countries, detected in 97.7% of 136 locations sampled [9, 10].

Knowledge of the genomic variability among badnaviruses infecting yam in West Africa is limited. Molecular studies of badnaviruses infecting yams and other crops have revealed that extremely high levels of genetic variability occur within this group of viruses [13, 26–28]. Furthermore, integrated badna-like sequences in *Musa* genomes were able to give rise to infectious episomal genomes via homologous recombination [31]. These integrated sequences and the high sequence variability among badnaviruses complicate the development of reliable molecular detection tests [4, 13, 38]. To establish the genetic variability among yam-infecting badnaviruses occurring in West Africa, more isolates need to be studied. A better understanding of the genetic diversity of yam-infecting badnaviruses in West Africa will ensure that diagnostics are robust and reliable.

In this paper, we report sequence variability in the RT/RNaseH coding region of badnavirus isolates infecting yam in Ghana, Togo, Benin and Nigeria, and molecular evidence for the presence of a putative new badnavirus species whose members infect yam in West Africa.

## Materials and methods

The BadnaFP and BadnaRP primers used in this study were designed based on the consensus sequences of the RT and RNaseH coding regions of published badnavirus sequences [41] and had been used for the amplification of the RT/RNaseH-coding region of various groups of badnaviruses [26, 37, 41]. The BadnaFP primer differs by the insertion of ITI from that reported by Yang et al. [41] for the amplification of taro bacilliform virus (TaBV). For our work, immunocapture PCR was employed because it is reported to considerably decrease non-specific PCR amplifications, favouring the amplification of DNA contained within capsids and excluding DNA derived from other genomes, thus ensuring that only episomal virus sequences are amplified. This is particularly important for the detection of badnaviruses whose sequences have been reported to be integrated into host genomes [14, 15, 19, 20]. IC-PCR also has the advantage of eliminating, in the washing step, host contaminants that may interfere with PCR.

### Leaf material

Nineteen badnavirus-infected yam leaf samples from Ghana, Benin, Togo and Nigeria (Table 1), tested previously by ELISA and IC-PCR using DaBV antibody routinely used for the certification of yam plantlets at the International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria, were selected and used for this work.

### Immunocapture-polymerase chain reaction (IC-PCR)

Rabbit polyclonal antibody against DaBV from IITA, known to detect other badnaviruses including banana streak virus (BSV) and cassava Ivorian bacilliform virus (CIBV), was used for immunocapture. Degenerate badnavirus primer pair Badna FP (5'-ATGCCITTYGGIITIAA RAAYGCICC-3'), which binds from nucleotides 5833 to 5859 of the DsBV genome, and Badna RP (5'-CCAY TTRCAIACISCICCCCAICC-3'), which binds from nucleotides 6388 to 6411 of the DsBV genome, were used for PCR amplification [37, 41]. Immunocapture followed the coating and trapping method by Clark and Adams [7]. Antibody was diluted 1:1,000 in coating buffer (0.05 M sodium carbonate buffer, pH 9.6). Leaf sap extract was prepared by grinding the test leaves in grinding buffer (phosphate-buffered saline containing 0.05% (v/v) Tween-20 (PBS-T), 0.5 mM polyvinyl pyrrolidone (PVP)-40 and 79.4 mM Na<sub>2</sub>SO<sub>3</sub>). PCR reactions comprised 50 pM each of forward and reverse primers, 1 X Taq reaction buffer, 0.25 mM of each dNTP, and 1.25 U of Taq DNA polymerase (Promega, USA). The following thermocyclic regime was used: 94°C for 4 min, followed by 40 cycles of

**Table 1** Host species, geographical origin, *EcoRI* digestion pattern and accession numbers of 19 yam-infecting badnavirus isolates from Ghana, Togo, Benin and Nigeria

Isolate number	Host species	Country	Size of released insert	EMBL accession number
NG1Da	<i>Dioscorea alata</i>	Nigeria	224, 355 <sup>a</sup>	AM944571
NG2Da	<i>Dioscorea alata</i>	Nigeria	224, 355 <sup>a</sup>	AM944572
NG3Da	<i>Dioscorea alata</i>	Nigeria	224, 355 <sup>a</sup>	AM944573
GN1Dr	<i>Dioscorea rotundata</i>	Ghana	579	AM944574
GN2Dr	<i>Dioscorea rotundata</i>	Ghana	579	AM944575
GN3Dr	<i>Dioscorea rotundata</i>	Ghana	579	AM944576
GN4Da	<i>Dioscorea alata</i>	Ghana	224, 355 <sup>a</sup>	AM944577
GN5Dr	<i>Dioscorea rotundata</i>	Ghana	224, 355 <sup>a</sup>	AM944578
TG1Dr	<i>Dioscorea rotundata</i>	Togo	579	AM944579
TG2Dr	<i>Dioscorea rotundata</i>	Togo	579	AM944580
TG3Da	<i>Dioscorea alata</i>	Togo	579	AM944581
TG4Da	<i>Dioscorea alata</i>	Togo	579	AM944582
BN1Dr	<i>Dioscorea rotundata</i>	Benin	579	AM944583
BN2Da	<i>Dioscorea alata</i>	Benin	579	AM944584
BN3Dr	<i>Dioscorea rotundata</i>	Benin	579	AM944585
BN4Dr	<i>Dioscorea rotundata</i>	Benin	579	AM944586
BN5Dr	<i>Dioscorea rotundata</i>	Benin	579	AM944587
BN6Dr	<i>Dioscorea rotundata</i>	Benin	579	AM944588
BN7Dr	<i>Dioscorea rotundata</i>	Benin	579	AM944589

<sup>a</sup> Isolates with internal *EcoRI* site

94°C for 30 s, 50°C for 30 s, 72°C for 30 s, and a final extension at 72°C for 5 min. Ten microliters of IC-PCR products were resolved on a 1.5% (w/v) agarose gel containing ethidium bromide (50 µg/ml) at 100 V for 1 h.

### Cloning and sequencing

All amplicons were purified using Wizard SV Gel and PCR clean up system (Promega, USA) and ligated into the pGEM T-Easy (Promega, USA) cloning vector following the manufacturer's protocol. The plasmids were transformed into *Escherichia coli* JM 109 high-competent cells. Plasmid DNA was purified from the cells using a QIAprep Spin Miniprep kit (Qiagen) according to manufacturer's instructions. Purified plasmids were then digested using *EcoRI* to determine if the inserts were of the expected band size. Restriction digest products were resolved on a 1.5% agarose gel using TAE (40 mM Tris-acetate pH 8.3, 1 mM EDTA) at 100 V for 1 h.

### Sequencing and sequence analysis

Purified plasmids were sequenced in both direction using SP6 and T7 primers in an automated sequencer at Inqaba Biotechnical Industries, South Africa. Sequences were edited and a consensus sequence derived for each isolate. Vector sequences were identified using the VecScreen

program [2] and removed prior to sequence analysis. Sequence similarity searches were made in the GenBank databases using the BLAST program (NCBI). Amino acid sequences were deduced using the Transeq program [35]. Multiple alignments were done using the CLUSTAL W programs [6], and phylogenetic trees were created using the neighbour-joining method [36]. The tree was viewed using the njplot program [33], and the robustness of the tree was determined by bootstrap using 1,000 replicates. Two previously sequenced yam-infecting badnaviruses, DaBV (accession number: X94575–X94582) and DsBV (accession number: DQ822073), and five members of the badnavirus genus, banana streak Obino l'Ewai virus (BSV) (accession number: AJ002234), cocoa swollen shoot virus (CSSV) (accession number: AJ781003), commelina yellow mottle virus (CoYMV) (accession number: X52938), ScBV (accession number: M89923), and TaBV (accession number: AF357836), were used for phylogenetic comparisons. Rice tungro bacilliform virus (RTBV) (accession number: X57924) was defined as the outgroup.

### Results

The expected PCR product of 579 bp was amplified from all of the 19 badnavirus isolates, and there was no amplification from negative control (badna-free yam leaf extracts). *EcoRI*

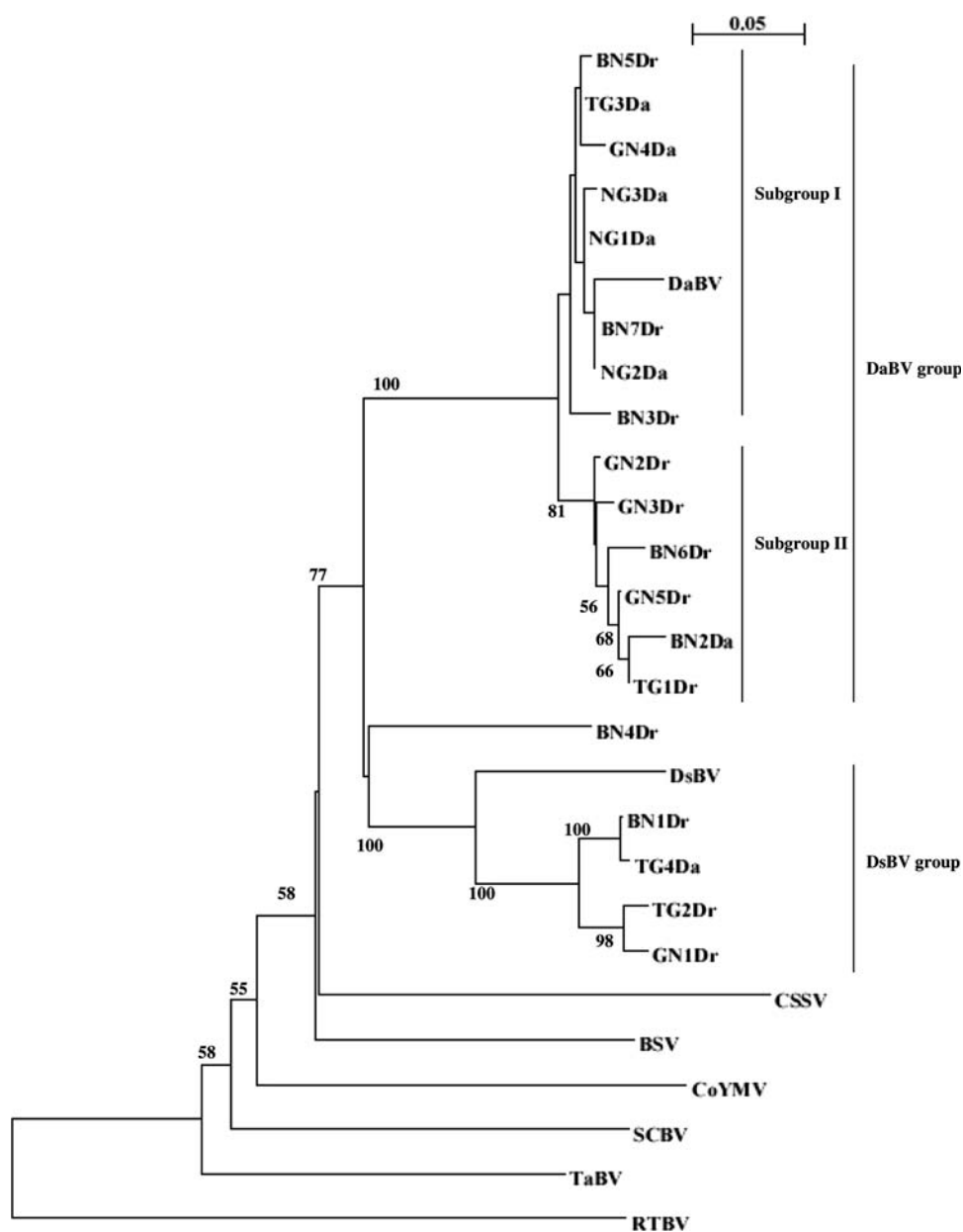
digests of purified plasmids revealed two restriction patterns. Fourteen plasmids released inserts of the expected band size of 579 bp, while the three isolates from Nigeria (NG1Da, NG2Da, NG3Da) and two isolates from Ghana (GN4Da and GN5Dr) had an internal *EcoRI* site at nucleotides 224–229 and, as a consequence, inserts were further cleaved into two fragments of 224 and 355 bp (Table 1).

Analysis of the deduced amino acid sequences of the RT/RNaseH-coding region of the badnavirus isolates in this study and published sequences of other badnaviruses showed some conserved regions typical of members of the family *Caulimoviridae* [3, 18, 29]. Phylogenetic analysis revealed that the 19 isolates formed two main groups: the DaBV group, clustering with the previously published

sequence of DaBV, and the DsBV group, clustering with the previously published sequence of DsBV (Fig. 1).

The DaBV group was further branched into two subgroups. The three Nigerian isolates (NG1Da, NG2Da, NG3Da), three Benin isolates (BN3Dr, BN5Dr and BN7Dr) and one isolate each from Ghana (GN4Da) and Togo (TG3Da) clustered with the previously published sequence of DaBV (a Nigerian isolate) in subgroup I and showed 96–100% amino acid identity to each other and 93–96% amino acid identity with DaBV. The virus isolates in subgroup II showed 95–99% amino acid identity to each other and consisted of three isolates from Ghana (GN2Dr, GN3Dr and GN5Dr), two isolates from Benin (BN6Dr and BN2Dr) and one isolate from Togo (TG1Dr).

**Fig. 1** Neighbour-joining phylogenetic tree based on the amino acid sequences of the reverse transcriptase and ribonuclease H coding regions of 19 yam-infecting badnavirus isolates from Ghana, Togo, Benin and Nigeria, DaBV, DsBV, CSSV, BSV, CoYMV, SCBV, TaBV and RTBV



Two isolates from Togo (TG2Dr and TG4 Da) and one isolate each from Ghana (GN1Dr) and Benin (BN1Dr) clustered with the previously published sequence of DsBV (from Benin) in the DsBV group. The isolates in the DsBV group had 94–99% amino acid identity to each other and had 83–84% amino acid identity with DsBV (Table 2).

One isolate from Benin (BN4Dr) did not cluster with either DaBV or DsBV and remained as a separate branch (Fig. 1). The nucleotide sequence of BN4Dr was 71–75% identical to those of the other isolates in this study, while the deduced amino acid sequence was 76–80% identical. This isolate was more closely related to DaBV (72% nucleotide identity and 77% amino acid identity) than to DsBV (70% nucleotide identity and 75% amino acid identity) (Table 2).

Nucleotide sequence identity between the two virus groups ranged from 69 to 72%, while the two groups showed 73–79% amino acid identity (Table 2). Single amino acid changes were observed at positions 14, 25, 39 and 41 of the reverse transcriptase. The major amino acid differences between the two groups were observed between positions 46 and 100 and between positions 124 and 178, with patches of conserved and semi-conserved regions for all isolates scattered between these two stretches (Fig. 2). The two subgroups of the DaBV group varied in both their nucleotide and amino acid sequences, with identities ranging from 82 to 87% and from 93 to 96%, respectively (Table 2). The major differences between the amino acid sequences of isolates in the two subgroups were found at positions 25, 46, 56, 101 and 102, where arginine, glutamic acid, lysine, and two glutamic acids in subgroup I are replaced with lysine, arginine, glutamic acid, and two aspartic acids, respectively, in subgroup II (Fig. 2).

An 18-consecutive-amino-acid identical match, LKTTKGLRSWLGLNYAR, was present at positions 104–121 of all of the yam-infecting badnavirus isolates in this study, DaBV and DsBV, but not in any of the other badnaviruses. This and the conserved region at amino acid positions 59–71 could be potential sites for specific primers for the amplification of all yam-infecting badnaviruses (Fig. 2).

The deduced amino acid sequence identities of all of the badnavirus isolates in this study were closer to DaBV (74–96% amino acid identity) than to DsBV (74–84% amino acid identity), BSV (69–74% amino acid identity), CSSV (64–68% amino acid identity), CoYMV (63–67% amino acid identity), TaBV (61–67% amino acid identity), ScBV (63–66% amino acid identity), and RTBV (44–47% amino acid identity).

In a separate analysis (details not shown), the 19 badnavirus sequences obtained from yams in West Africa in this study were found to be variously related to ten of the 45 badnavirus sequences obtained from yams in the South

Pacific Islands. Isolates in subgroup I of the DaBV group had 92–100% amino acid identity with the seven sequences in group 8 in the South Pacific Islands study, while isolates in subgroup II had 90–100% amino acid identity with the same seven sequences. One isolate each from Togo (TG3Da) and Ghana (GN5Dr) had an amino acid sequence that was identical to one sequence each from Vanuatu (Vu249\_Db) and the Philippines (PH11a\_Da), respectively.

All four isolates in the DsBV subgroup shared 82% amino acid identity with the only sequence in group 5 (Fj60b\_Dr) in the South Pacific Islands study. Interestingly, BN4Dr, the distinct isolate from this study, clustered with the two sequences in group 9 (Fj60a\_Dr and VU18\_dt) and had 94% amino acid identity with both sequences.

## Discussion

The presence of some conserved and semi-conserved regions observed in the amino acid sequences of the yam-infecting badnavirus isolates in this study and other members of the badnavirus genus, particularly within the previously described FIAVYIDDILVFS stretch of the reverse transcriptase in the C-terminal region of the open reading frame (ORF) 3 polyprotein, confirms that all 19 virus isolates in this study belong to the badnavirus genus [3, 29, 37].

The 19 badnavirus isolates from Ghana, Togo, Benin and Nigeria analysed in this study can clearly be distinguished into two distinct virus species. One group was more closely related to the published sequence of the Nigerian DaBV isolate (90–96% amino acid identity), while the second group was more closely related to the published sequence of the Benin DsBV isolate (83–84% amino acid identity). Sequence variability observed among the isolates in each of the two groups may be attributed to the inaccurate replication by reverse transcription reported for members of the family *Caulimoviridae* [18, 29].

The lower level of amino acid identity observed between one isolate from Benin, BN4Dr, and DaBV (77%) and DsBV (75%) strongly suggests that BN4Dr is a member of a distinct virus species. The 70.8% nucleotide identity reported between DaBV and DsBV in the same RT/RNaseH region [37] is similar to the nucleotide sequence identity between BN4Dr and DaBV (72%) and DsBV (70%). Considering nucleotide differences of >20% in the RT/RNaseH region as the criteria for species demarcation in the genus *Badnavirus*, BN4Dr is a member of a putative new badnavirus species that infects yam in West Africa. The presence of the previously described FIAVYIDDILVFS stretch [3, 29, 37] and the 18 consecutive identical amino acids, LKTTKGLRSWLGLNYAR, present in DaBV,

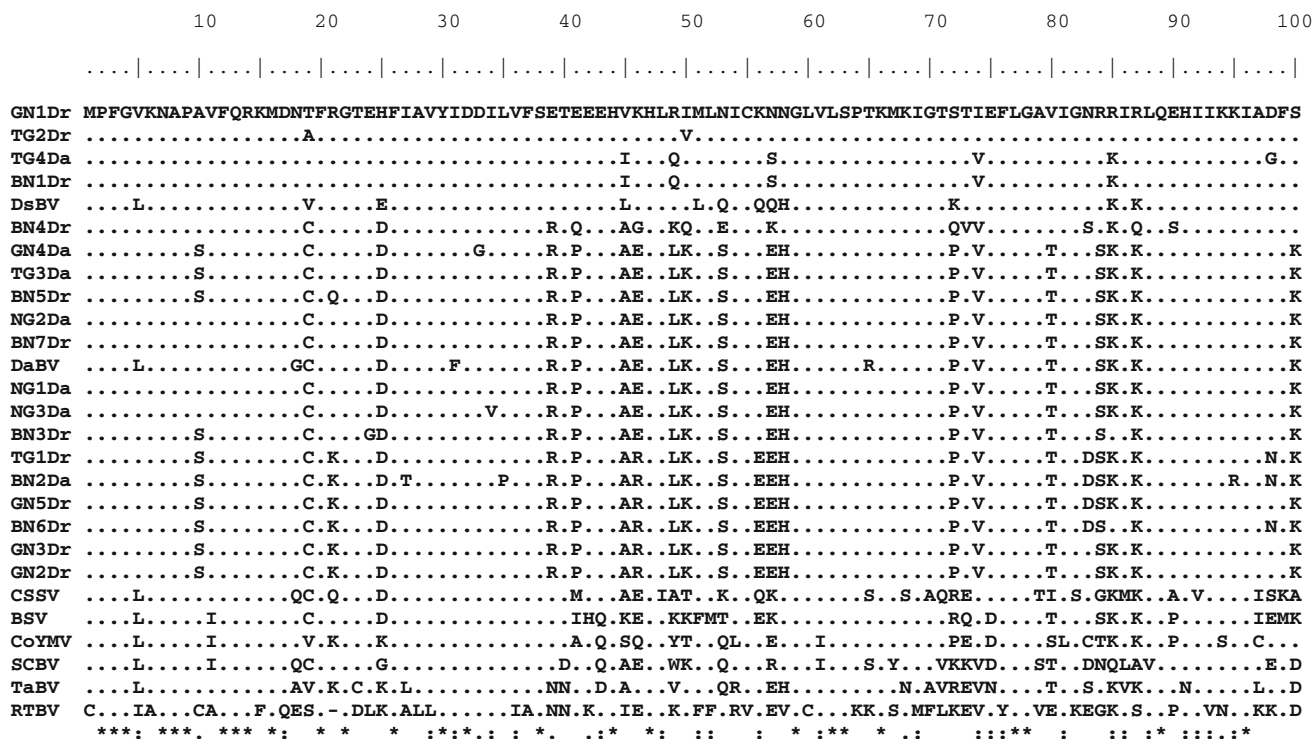
**Table 2** Nucleotide sequence and amino acid sequence identities of pair-wise combinations of the RT/RNaseH-coding region of 19 yam-infecting badnavirus isolates from Ghana, Togo, Benin and Nigeria, DaBV and DsBV

NG1	NG2	NG3	GN1	GN2	GN3	GN4	GN5	TG1	TG2	TG3	TG4	BN1	BN2	BN3	BN4	BN5	BN6	BN7	DsBV	DaBV	
100	90	90	71	87	86	90	87	87	70	87	70	69	84	86	75	85	84	89	71	95	NG1
<b>100</b>	<b>99</b>	<b>99</b>	<b>78</b>	<b>96</b>	<b>95</b>	<b>98</b>	<b>95</b>	<b>95</b>	<b>78</b>	<b>99</b>	<b>79</b>	<b>79</b>	<b>93</b>	<b>97</b>	<b>80</b>	<b>98</b>	<b>94</b>	<b>99</b>	<b>76</b>	<b>96</b>	
	100	98	71	85	86	89	84	84	71	87	70	69	83	84	74	87	83	88	72	90	NG2
	<b>100</b>	<b>98</b>	<b>77</b>	<b>96</b>	<b>96</b>	<b>97</b>	<b>95</b>	<b>94</b>	<b>77</b>	<b>98</b>	<b>78</b>	<b>79</b>	<b>93</b>	<b>96</b>	<b>80</b>	<b>98</b>	<b>94</b>	<b>100</b>	<b>76</b>	<b>96</b>	
		100	70	84	85	89	84	84	71	87	70	69	82	84	74	87	82	88	71	89	NG3
		<b>100</b>	<b>77</b>	<b>96</b>	<b>95</b>	<b>97</b>	<b>95</b>	<b>94</b>	<b>77</b>	<b>98</b>	<b>78</b>	<b>79</b>	<b>93</b>	<b>96</b>	<b>80</b>	<b>98</b>	<b>93</b>	<b>98</b>	<b>75</b>	<b>95</b>	
			100	72	72	70	72	72	93	71	89	89	72	72	71	71	72	70	77	68	GN1
			<b>100</b>	<b>76</b>	<b>76</b>	<b>76</b>	<b>75</b>	<b>75</b>	<b>97</b>	<b>77</b>	<b>94</b>	<b>95</b>	<b>73</b>	<b>76</b>	<b>79</b>	<b>77</b>	<b>76</b>	<b>77</b>	<b>84</b>	<b>74</b>	
				100	95	85	91	92	71	85	71	70	91	84	72	84	91	85	68	84	GN2
				<b>100</b>	<b>98</b>	<b>96</b>	<b>98</b>	<b>98</b>	<b>76</b>	<b>97</b>	<b>77</b>	<b>78</b>	<b>96</b>	<b>96</b>	<b>79</b>	<b>97</b>	<b>97</b>	<b>96</b>	<b>76</b>	<b>93</b>	
					100	84	91	92	72	85	71	71	91	84	73	85	91	85	70	84	GN3
					<b>100</b>	<b>95</b>	<b>97</b>	<b>97</b>	<b>76</b>	<b>96</b>	<b>77</b>	<b>77</b>	<b>95</b>	<b>95</b>	<b>79</b>	<b>96</b>	<b>97</b>	<b>96</b>	<b>76</b>	<b>93</b>	
						100	85	85	70	88	69	69	83	87	73	88	83	90	69	88	GN4
						<b>100</b>	<b>95</b>	<b>94</b>	<b>76</b>	<b>98</b>	<b>77</b>	<b>78</b>	<b>93</b>	<b>96</b>	<b>79</b>	<b>98</b>	<b>93</b>	<b>97</b>	<b>74</b>	<b>94</b>	
							100	94	72	84	70	70	94	86	73	84	94	84	70	84	GN5
							<b>100</b>	<b>99</b>	<b>75</b>	<b>96</b>	<b>76</b>	<b>77</b>	<b>97</b>	<b>96</b>	<b>78</b>	<b>96</b>	<b>97</b>	<b>95</b>	<b>76</b>	<b>92</b>	
								100	72	85	70	70	93	85	72	85	93	84	70	84	TG1
								<b>100</b>	<b>75</b>	<b>95</b>	<b>76</b>	<b>76</b>	<b>98</b>	<b>95</b>	<b>78</b>	<b>95</b>	<b>97</b>	<b>94</b>	<b>75</b>	<b>91</b>	
									100	70	88	88	70	72	72	70	72	71	77	68	TG2
									<b>100</b>	<b>77</b>	<b>94</b>	<b>94</b>	<b>73</b>	<b>76</b>	<b>79</b>	<b>77</b>	<b>75</b>	<b>77</b>	<b>84</b>	<b>74</b>	
										100	71	71	83	91	72	94	83	91	68	85	TG3
										<b>100</b>	<b>78</b>	<b>79</b>	<b>94</b>	<b>97</b>	<b>80</b>	<b>99</b>	<b>94</b>	<b>98</b>	<b>75</b>	<b>95</b>	
											100	99	70	72	72	70	70	72	76	68	TG4
											<b>100</b>	<b>99</b>	<b>75</b>	<b>76</b>	<b>79</b>	<b>78</b>	<b>76</b>	<b>78</b>	<b>83</b>	<b>75</b>	
												100	69	71	72	70	70	71	76	67	BN1
												<b>100</b>	<b>75</b>	<b>77</b>	<b>80</b>	<b>78</b>	<b>76</b>	<b>79</b>	<b>84</b>	<b>76</b>	
													100	85	72	83	97	83	67	82	BN2
													<b>100</b>	<b>94</b>	<b>76</b>	<b>94</b>	<b>96</b>	<b>93</b>	<b>74</b>	<b>90</b>	
														100	71	91	85	91	69	84	BN3
														<b>100</b>	<b>78</b>	<b>97</b>	<b>94</b>	<b>96</b>	<b>75</b>	<b>93</b>	
															100	73	72	73	70	72	BN4
															<b>100</b>	<b>79</b>	<b>77</b>	<b>80</b>	<b>75</b>	<b>77</b>	
																100	83	91	69	84	BN5
																<b>100</b>	<b>94</b>	<b>98</b>	<b>75</b>	<b>95</b>	
																	100	84	68	82	BN6
																	<b>100</b>	<b>94</b>	<b>75</b>	<b>91</b>	
																		100	70	87	BN7
																		<b>100</b>	<b>76</b>	<b>96</b>	
																			100	70	DsBV
																			<b>100</b>	<b>74</b>	
																				100	DaBV
																				<b>100</b>	

% nucleotide sequence identity (unbold values), % amino acid sequence identity (bold values)

DsBV and all of the yam-infecting badnavirus isolates in this study, confirms that BN4Dr is a yam-infecting badnavirus and not an integrated badna-like sequence. In order to further confirm the taxonomic status and to ascertain the

divergence of the badnaviruses infecting yam in West Africa, further ORFs need to be sequenced. As badnavirus strains within a single leaf can be highly diverse [16, 21, 26, 30], mixtures of viruses belonging to the three badnavirus



**Fig. 2** Amino acid sequence of the reverse transcriptase and ribonuclease H coding regions of 19 yam-infecting badnavirus isolates from Ghana, Togo, Benin and Nigeria, DaBV, DsBV, CSSV,

BSV, CoYMV, SCBV, TaBV and RTBV. An *asterisk* denotes an exact match, and *double* or *single dots* denote positions of conserved or semi-conserved amino acid changes, respectively

species described in this study may be present in some yam leaves, but this cannot be confirmed unless many more clones are screened.

Further genetic variability among the badnavirus isolates from Ghana, Togo, Benin and Nigeria, analysed in this study, was revealed by the existence of two *EcoRI* digestion patterns. Five isolates (three Nigerian isolates and two Ghana isolates) had an internal *EcoRI* digestion site and were digested into two fragments of 224 and 355 bp, while fourteen isolates were undigested. Sequence data confirmed the presence of the *EcoRI* digestion site (5' GAATTC 3') at nucleotides 224–229 in the five sequences that were digested. The *EcoRI* digestion site was also present, in the same position, in DaBV (Nigeria isolate) and the four Vanuatu sequences (Vu249\_Db, Vu252\_Db, Vu254\_Dp and Vu257\_Dp) in group 8 in the South Pacific Islands study [26]. Phylogenetic analysis revealed that four of the five isolates (NG1Da, NG2Da, NG3Da and GN4Da) which had the internal *EcoRI* site clustered with DaBV. However, the *EcoRI* digestion pattern may not be a very useful tool for differentiating isolates of yam-infecting badnavirus, since other isolates lacking the internal *EcoRI* digestion site clustered in the DaBV subgroup I, and GN5Dr, the fifth isolate which had the internal *EcoRI* digestion site, clustered in a different subgroup (Fig. 1).

The integration of badnavirus-like sequences into host genomes has been reported [15, 20, 41]. The high level of amino acid identity (>74%) observed between the badnavirus isolates in this study and previously sequenced yam-infecting badnaviruses strongly indicates that genomic virus-like sequences were likely not amplified, since Yang et al. [41] reported approximately 50–60% similarity between TaBV and TaBV-like sequences amplified from Taro plants in the RT/RNaseH region. Furthermore, the use of immunocapture in this study would favourably capture virions and therefore would increase the detection of episomal rather than integrated sequences. This was also shown to be in case with BSV, where the antiserum did not capture any *Musa* nuclear, mitochondria, or chloroplast genomes, which may have contained integrated BSV sequences [19]. Despite its advantages, particularly with regards to the presence of endogenous and integrated badnavirus sequences in the host genome, the use of immunocapture in this study may have deselected some badnaviruses that are not serologically related to the antibody used. It is therefore possible that some badnaviruses infecting yams in West Africa may have been missed due to the use of immuno-capture in this study. However, the need to amplify and analyse only episomal sequences necessitated the use of immuno-capture. Analysis of the 45 partial RT/RNaseH sequences derived from DNA extracted

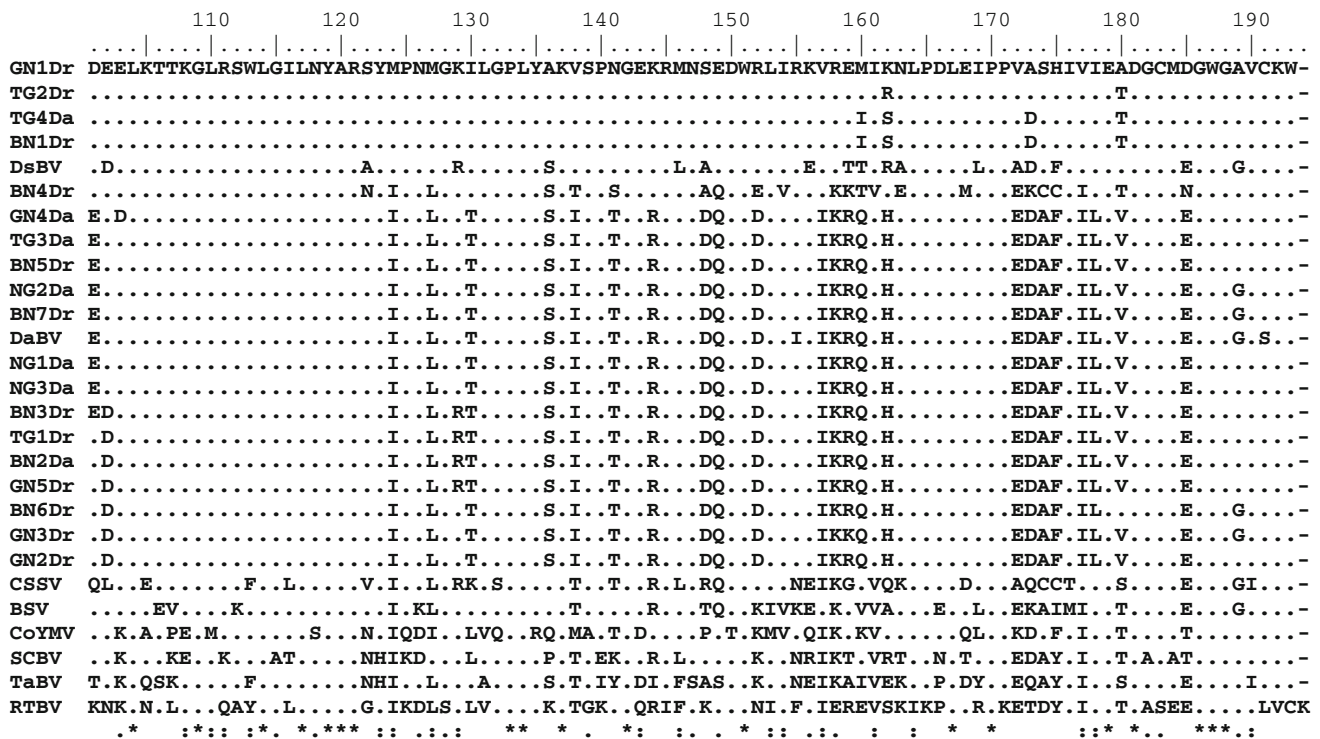


Fig. 2 continued

from yams in the South Pacific Islands using the same primer set used in this study gave strong evidence that some of the badnavirus groups may be endogenous pararetrovirus or putative integrated sequences [26]. Thus, the three virus species described in this study are inclusive but not exclusive of badnaviruses infecting yams in West Africa.

Results from current studies confirm previous reports that badnaviruses infecting yams are serologically related to each other [34, 37]. The three badnavirus species found in this study are serologically related. Phillips et al. [34] reported serological relationships among badnaviruses infecting yam and among badnaviruses in general. In their study, DaBV was trapped and decorated, in immunosorbent electron microscopy (ISEM), by antibodies raised against DbBV, ScBV and BSV. Furthermore, they found that immunoglobulin (IgG) extracted from two DbBV antisera detected both DbBV and DaBV by ELISA. Seal and Muller [37] also reported the use of a general badnavirus antibody for the detection of DsBV by ISEM and PAS-ELISA. However, unless a yam-infecting badnavirus is clearly identified by use of specific primers and/or monoclonal antibody, its actual identity remains unknown, and it is therefore generally called a yam-infecting badnavirus.

Besides the Nigerian isolates, which clustered together, there was no correlation between isolates and country of

origin or host species. Viruses of the two badnavirus species (DaBV and DsBV) were widespread and were present in Ghana, Togo and Benin and were observed to infect both *D. alata* and *D. rotundata* indiscriminately. Furthermore, BN7Dr from a *D. rotundata* plant in Benin and NG2Da from a *D. alata* plant in Nigeria both had identical amino acid sequences. Similar amino acid identities were observed between two isolates from West Africa (TG3Da and GN5Dr) and two sequences from the South Pacific Islands (Vu249\_Db and PH11a\_Da). The *D. alata* (TG3Da) and *D. rotundata* (GN5Dr) from West Africa were identical to sequences obtained from *D. bulbifera* (Vu249\_Db) and *D. alata* (PH11a\_Da) from the South Pacific Islands, respectively. This demonstrates that different badnavirus species are not restricted to single yam host species or geographical locations and confirms that badnaviruses have spread between yam-producing continents through the exchange of infected germplasm. This also suggests that DbBV, the badnavirus isolated from *D. bulbifera* [34] may belong to one of the three species found in this study.

The distribution of the two badnavirus species across Ghana, Togo and Benin and the possible infection by members of a new virus species is possibly a result of unrestricted exchange of planting materials through permeable land borders, these countries being located next to each other. Stringent quarantine laws, which are already in



place, need to be enforced in the West African yam zone, particularly with respect to the consequences of porous land borders on the containment of plant diseases. Knowledge about the actual yield losses due to yam-infecting badnaviruses will also help in convincing the farmers of the damage that could result from such unrestricted movement of planting materials. Furthermore, the wide distribution of several virus species has significant consequences in relation to diagnosis and control of infection in yam in West Africa.

**Acknowledgments** We are grateful to the Gatsby Charitable Foundation, UK for financial support to this work under Project GAT2261. We thank Dr. Bi Irie Vroh of the Biotechnology Unit, International Institute of Tropical Agriculture, and Ms. Sarah Taylor of the School of Molecular and Cell Biology, University of the Witwatersrand, for their great support during this work.

## References

- Aleman-Verdaguer ME, Goudou-Urbino C, Dubern J, Beachy RN, Fauquet C (1997) Analysis of the sequence diversity of the P1, HC, P3, Nib and CP genomic regions of several yam mosaic potyvirus isolates: implications for the intraspecies molecular diversity of potyviruses. *J Gen Virol* 78:1253–1264
- Altschul SF, Thomas LM, Alejandro AS, Jinghui Z, Zheng Z, Webb M, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402
- Bouhida M, Lockhart BEL, Olszewski NE (1993) An analysis of the complete sequence of a sugarcane bacilliform virus genome infectious to banana and rice. *J Gen Virol* 74:15–22
- Braithwaite KS, Egeskov NM, Smith GR (1995) Detection of sugarcane bacilliform virus using the polymerase chain reaction. *Plant Dis* 79:792–796
- Briddon RW, Phillips S, Brunt A, Hull R (1999) Analysis of the sequence of *Dioscorea Alata* bacilliform virus: comparison to others members of the badnavirus group. *Virus Genes* 18:277–283
- Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res* 31(13):3497–3500
- Clark MF, Adams AM (1977) Characteristics of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. *Virol Methods* 34:475–483
- Degras L (1993) *The Yam, a tropical root crop*. London, UK, 40 p
- Eni AO (2008) Epidemiology, diagnostics and molecular studies of Yam Viruses in Ghana, Togo and Benin. PhD thesis, University of the Witwatersrand, South Africa, 183 p
- Eni AO, Hughes Jd'A, Rey MEC (2008) Survey of the incidence and distribution of five viruses infecting yam in the major yam producing zones in Benin. *Ann Appl Biol* (online 27 May 2008). doi:10.1111/j.1744-7348.2008.00253.x
- FAO (2008) Food and Agricultural Organisation of the United Nations. Production Yearbook, FAO Statistics 2007, FAO, Rome
- Fauquet C, Mayo M, Maniloff J, Desselberger U, Ball L (2005) *Virus taxonomy*. Eight Report of the International committee on Taxonomy of Viruses. Elsevier, Amsterdam, 1259 p
- Geering ADW, Olszewski NE, Dahal G, Thomas JE, Lockhart BEL (2001) Analysis of the distribution and structure of integrated *Banana streak virus* DNA in a range of *Musa* cultivars. *Mol Plant Pathol* 2:207–213
- Geering ADW, McMichael LA, Dietzgen RG, Thomas JE (2000) Genetic diversity among *Banana streak virus* isolates from Australia. *Phytopathology* 90:921–927
- Geering ADW, Olszewski NE, Harper G, Lockhart BEL, Hull R, Thomas JE (2005) Banana contains a diverse array of endogenous badnaviruses. *J Gen Virol* 86:511–520
- Geijskes RJ, Braithwaite KS, Dale JL, Harding RM, Smith GR (2002) Sequence analysis of an Australian isolate of sugarcane bacilliform badnavirus. *Arch Virol* 147:2393–2404
- Goudou-Urbino C, Konate G, Quiot JB, Duber J (1996) Aetiology and ecology of a yam mosaic disease in Burkina Faso. *Trop Sci* 36:34–40
- Harper G, Hull R (1998) Cloning and sequence analysis of *Banana Streak Virus* DNA. *Virus Genes* 17(3):271–278
- Harper G, Dahal G, Thottappilly G, Hull R (1999) Detection of episomal banana streak badnavirus by IC-PCR. *J Virol Methods* 79:1–8
- Harper G, Osuji JO, Heslop-Harrison JSP, Hull R (1999) Integration of banana streak badnavirus into the *Musa* genome: molecular and cytogenetic evidence. *Virology* 255:207–213
- Harper G, Hart D, Moul S, Hull R (2004) Banana streak badnavirus is very diverse in Uganda. *Virus Res* 100:51–56
- Harrison BD, Roberts IM (1973) Association of virus-like particles with internal brown spot of yam (*Dioscorea alata*). *Trop Agric* 50:335–340
- Huang Q, Hartung JS (2001) Cloning and sequence analysis of an infectious clone of *Citrus yellow mosaic virus* that can infect sweet orange via *Agrobacterium*-mediated inoculation. *J Gen Virol* 82:2549–2558
- Hughes Jd' A, Dongo L, Atiri GI (1997) Viruses infecting cultivated yams (*Dioscorea alata* and *D. rotundata*) in Nigeria. *Phytopathology* 87:545
- IITA (2007) Yam. <http://www.iita.org/cms/yam>
- Kenyon L, Lebas BSM, Seal SE (2008) Yams (*Dioscorea* spp.) from the South Pacific Islands contain many novel badnaviruses: implications for international movement of yam germplasm. *Arch Virol* 153:877–889
- Lebas BSM (2002) Diversity of viruses infecting *Dioscorea* species in the South Pacific. PhD thesis, University of Greenwich, Greenwich, 302 p
- Lockhart BEL, Olszewski NE (1993) Serological and genomic heterogeneity of banana streak badnavirus: Implications for virus detection in *Musa* germplasm. In: Ganry J (ed) *Breeding banana and plantain for resistance to diseases and pests*. International Network for the Improvement of Banana and Plantain, Montpellier, France, pp 105–113
- Medberry SL, Lockhart BEL, Olszewski NE (1990) Properties of *Commelina yellow mottle virus*'s complete DNA sequence, genomic discontinuities and transcript suggest that it is a pararetrovirus. *Nucleic Acids Res* 18:5505–5513
- Muller E, Sackey S (2005) Molecular variability analysis of five new complete cocoa swollen shoot virus genomic sequences. *Arch Virol* 150:53–66
- Ndowora T, Dahal G, LaFleur D, Harper G, Hull R, Olszewski NE, Lockhart B (1999) Evidence that badnavirus infection in *Musa* can originate from integrated pararetroviral sequences. *Virology* 255:214–220
- Odu BO, Hughes Jd' A, Shoyinka SA, Dongo LN (1999) Isolation, characterisation and identification of a potyvirus from *Dioscorea alata* (water yam) in Nigeria. *Ann Appl Biol* 139:65–71
- Perrière G, Gouy M (1996) WWW-Query: an on-line retrieval system for biological sequence banks. *Biochimie* 78:364–369
- Phillips S, Briddon RW, Brunt A, Hull R (1999) The partial characterization of a badnavirus infecting the Greater Asiatic or Water Yam (*Dioscorea alata*). *J Phytopathol* 147:265–269

35. Rice P, Longden I, Bleasby A (2000) EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* 16(6):276–277
36. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
37. Seal S, Muller E (2007) Molecular analysis of a full-length sequence of a new yam badnavirus from *Dioscorea sansibarensis*. *Arch Virol* 152:819–825
38. Smith GR, Handley JA, Harding RM, Dale JL, Gambley CF, Braithwaite KS (1996) Variability in sugarcane bacilliform and mosaic viruses and consequences for diagnosis. In: Cock JH, Brekelbaum T (eds) *Proceeding of XXIII international society of sugarcane technologists*, Columbia, pp 390–396
39. Thouvenel JC, Fauquet C (1979) Yam mosaic, a potyvirus infecting *Dioscorea cayenensis* in the Ivory Coast. *Ann Appl Biol* 93:279–283
40. Thouvenel JC, Dumont R (1990) Yield decreases in yam infected with mosaic virus in Cote d'Ivoire. *L'Agronomie Tropicale* 45:125–129
41. Yang IC, Hafner GJ, Reville PA, Dale JL, Harding RM (2003) Sequence diversity of South Pacific isolates of Taro bacilliform virus and the development of a PCR-based diagnostic test. *Arch Virol* 148:1957–1968