Nigerian Journal of Biochemistry and Molecular Biology 2018; 33(1&2): 44-59

0189-4757/96 $3.0 + 0.00
Printed in Nigeria

Available online at http://www.nsbmb.com.ng

NJBMB/009/18

Molecular and Cytogenetic Analyses of the Integration Events of Badnavirus Endogenous Pararetrovirus Sequences in Yam (Dioscorea species)

***[1,2]Turaki, A. A. and [1]Abdulrazak, U. K.**

[1]Natural Resources Institute, University of Greenwich, Central Avenue, Chatham, Kent ME4 4TB, UK.
[2]Federal University Birnin Kebbi PMB 1157 Birnin Kebbi Nigeria

Abstract: Yam (*Dioscorea* spp.) is an important staple food crop in Sub-Saharan Africa and is vegetatively-propagated. This had led to the accumulation of viruses and their integrated sequences leading to decreased yam production and hindering international movement of selected germplasm. It is unclear if badna virus sequences are integrated in the genomes of West African yam breeding lines. This study was aimed to determine that. DNAs were extracted from leaf samples of Nigerian yam breed using an optimised CTAB-extraction method and then screened by PCR using degenerate badnavirus primers targeting a 579 bp RT-RNaseH region and Rolling cycle amplification. Diversity of the RT-RNaseH region sequence was determined using Denaturing Gradient Gel Electrophoresis (DGGE). All the yam samples tested badnavirus PCR-positive (100%). Probes for the Southern Band Fluorescent ***in situ*** hybridisation (FISH) results were designed using individual DGGE partial RT-RNaseH band sequences (NGb4_Dr, NGb5_Dr and NGb6_Dr), badnavirus complete genome and partial genome cloned. The FISH result was inconclusive to suggests integration of badnavirus sequences in genomes of *D. rotundata* breeding lines. The consequences of the integrated and episomal badnavirus sequences for yam improvement programmes in West Africa are discussed.

KEYWORDS: Badnavirus, Hybridization, Integration, Dioscorea

1.0 Introduction

Integration of viral DNA into the genomes of bacteria and animals has been reported by several studies (Pistello and Antonelli 2016) as a common event because bacteriophages (Campbell, 2003) and animal retroviruses (Katzourakis *et al*., 2007) require integration for their replication (Hohn *et al*., 2008). In contrast, for plant genomes there are relatively few reported cases of viral DNA integration, and such integrations appear to involve only plant DNA viruses of the families *Caulimoviridae* (Staginnus and Richert-Pöggeler, 2006) and *Geminiviridae* (Lefeuvre *et al*., 2011). The integration of these DNA viruses into their host plant genomes is hypothesized to have happened by illegitimate recombination, as the viruses involved do not encode an integrase function (Jakowitsch *et al*., 1999; Lockhart *et al*., 2000; Kunii *et al*., 2004). The viral integrated sequences of the family *Caulimoviridae* are called endogenous pararetrovirus (EPRV) sequences (Staginnus *et al*., 2009; Geering *et al*., 2010). The integrated sequences can become fixed in the host population through vertical transmission if the EPRVs are incorporated into the host germ cells (Feschotte and Gilbert, 2012).

The family *Caulimoviridae* is composed of six genera and during the past few years various EPRVs belonging to five of these genera, namely *Badnavirus* (*Banana Streak Virus* (BSV) and *Dioscorea Bacilliform virus* (DBV)), *Cavemovirus* (*Tobacco Vain Clearing Virus* (TVCV), *Caulimovirus* (DMV), *Petuvirus* (*Petunia Vain Clearing Virus* (PVCV) and *Tungrovirus* (*Rice Tungro Bacilliform virus* (RTBV), have been described (Jakowitsch *et al*., 1999; Ndowora *et al*., 1999; Lockhart *et al*., 2000; Mette *et al*., 2002; Hansen *et al*., 2002; Richert-Pöggeler *et al*., 2003; Kunii *et al*., 2004;;Geering *et al*., 2005a, 2005b; Hansen *et*

***Corresponding Author**
Tel.: +234 8036066777
E-mail: turakialiyu@yahoo.com

al., 2005; Staginnus and Richert-Pöggeler, 2006; Staginnus *et al*., 2007; Pahalawatta *et al*., 2008;). Despite numerous reports on the detection of EPRVs in the genomes of host plants, to date EPRVs have only found to be activatable in banana, petunia and tobacco for the viruses BSV, PVCV and TVCV respectively (Harper *et al*., 1999; Jakowitsch *et al*., 1999; Ndowora *et al*., 1999; Lockhart *et al*., 2000; Richert-Pöggeler *et al*., 2003; Gayral *et al*., 2008; Chabannes *et al*., 2013). The activation of these EPRVs occur only in plant hybrids from interspecific crosses following environmental stress (Lheureux *et al*., 2003; Richert-Pöggeler *et al*., 2003; Gayral *et al*., 2008; Chabannes *et al*., 2013).

It has been previously hypothesized that DBV sequences are integrated in the genome of yam species following several reports of very high proportions of badnavirus PCR-positive samples (Eni *et al*., 2008; Kenyon *et al*., 2008; Bousalem *et al*., 2009). The integration of DBV sequences has been reported in the genomes of some yams belonging to the *D. cayenensis-rotundata* complex (Seal *et al*., 2014; Umber *et al*., 2014).

This study focused on detection of endogenous Dioscorea bacilliform virus sequences (eDBVs) in the genomes of West African yam breeding lines. The high levels of PCR-positives (100%) recorded in the collection of the IITA yam breeding lines (seal *et al.*, 2014), and the presence of several common DGGE bands in the diversity study (Turaki *et al.*, 2017a) support the existence of eDBV sequences in the genomes of selected West African yam breeding lines. Further support for this assumption came from phylogenetic analysis of the sequences that originated from DGGE bands present in all the lines, which showed these sequences to cluster with eDBV sequences in yam genomes (Seal *et al*., 2014; Umber *et al*., 2014).

The aim of this study was to use Southern blotting and *in situ* hybridisation techniques to further investigate the presence of such putative DBV integration events in West African yam breeding lines. Similar approaches were previously used in detection of EPRVs in banana, petunia and tomato (Harper *et al*., 1999; Richert-Pöggeler *et al*., 2003; Staginnus *et al*., 2007). The presence of eDBVs in yam genomes further complicates the development of reliable diagnostic tools for the indexing of yam badnaviruses, in a similar manner as highlighted by the challenges encountered in reliable detection of episomal BSV in *Musa* species (Harper *et al*., 1999; Ndowora *et al*., 1999; Le Provost *et al*., 2006). The current study will ease the complexity in developing a reliable diagnostic tool for the indexing of yam badnaviruses.

2.0 Materials and Methods

2.1 Plant Material

The yam plants (*D. alata* and *D. rotundata*) used had been grown in the NRI greenhouse from tubers that were brought from the International Institute of Tropical Agriculture (IITA) Ibadan. DNAs extracted from these plants, which were badnavirus positive (Bomer et al., 2016, Turaki et al., 2017a), were used for Southern blot hybridisation. Root tips (~10 mm long) of yams (n = 15) growing in the NRI glasshouse at the time of the study were collected for fluorescence *in situ* hybridisation (Figure 1).

Figure 1: Yam plant (*D. rotundata*) grown in the NRI glasshouse. The red circle shows the approximate root tip size that was collected for fluorescence *in situ* hybridisation (FISH).

2.2 Collection and Fixation of Root Tips

The selection of root tips for FISH was because the tips of the roots contain actively dividing cells and thus many of these cells would be at the stages of mitosis. This allows chromosomes of individual cells to be stained which makes them more readily visible under a microscope. It is equally important to select fine root tips as they generate better metaphases compared to thick root tips. The tips were placed in 1 ml Room Temperature (RT) metaphase arresting agent (2 mM 8-hydroquinoline) and kept in the dark at RT for 2 hrs. Then the root tips were transferred to 4^{o}C. After 24 hrs, the pre-treatment was replaced with fixatives using 3:1 (v:v) of ethanol:glacial acetic acids (Clarkes solution) and stored at 4^{o}C before use.

2.3 Total DNA Extraction

Total DNAs were extracted from freshly harvested yam leaf tissue (~3 g) following the optimised CTAB DNA extraction method described by Turaki *et al.* (2017b). Modifications were made by grinding leaf tissues under liquid nitrogen in a pestle and mortar that had been bleach-treated (to remove any residual contaminating DNAs), and then rinsed thoroughly in Sterile distilled water (SDW) and dried. Samples were ground to a fine powder, before mixing with 30 ml grinding buffer (Turaki *et al*., 2017b). The subsequent steps of the DNA extraction were carried out exactly as described by Turaki *et al*. (2017b). DNA pellets were re-suspended in 2 ml suspension buffer (50 mM Tris-Cl, 0.7 M NaCl, 10 mM EDTA, pH 7) and then purified through Tip-100/G columns (Qiagen, UK) according to manufacturer's instructions. The final pellets were re-suspended in 500 µl of sterile distilled deionised water and stored at –20^{o}C.

2.4 Fluorescence in situ Hybrisation

This was carried out according to the method described by Schwarzacher and Heslop-Harrison (2000). The hybridisation probes were selected from sequences that originated from DGGE bands that were common to all *D. rotundata* (band 4, 5 and 6) and hence suspected to be eDBVs (Turaki *et al*., 2017). A few DGGE bands as well as RCA-generated viral sequences in Bomer *et al.,* (2016) were also used as probes. The RCA-generated viral sequences were used as probes to determine if similar viral sequences could be detected in the chromosomes of a few *D. alata* and *D. rotundata* breeding lines, which would provide support for some eDBVs possibly being activatable EPRVs.

The double-target fluorescence *in situ* hybridisation was used with RCA-generated badnaviral clones DBALV2 (~7.1 kb) and DBSNV2 (~4.3 kb) as probes. Ribosomal DNA pTa 71 (~8.9 kb) and pTa 974 (~0.5 kb) probes were used as internal controls to locate the position of the sequences in yam chromosomes (Table 1). These probes were labeled using the DIG-random primed label method. The method incorporates a label DIG-dUTP into a newly synthesized DNA from a denatured single stranded DNA (ssDNA) primed by random sequence hexameric deoxynucleotides as primers and Klenow fragment of polymerase I. The labelled probe product is usually a collection of fragments of variable lengths because each of the random primers used has a different six-base sequence.

The badnaviral DBALV2 probe that falls into K08 group sequence (Kenyon *et al*. 2008) (Table 2.2) was from *D. alata* sample TDa85/00250, and shares 88.9% nucleotide identity with DBALV, for which particles were isolated (Briddon *et al.,* 1999; Phillips *et al.,* 1999). Equally, the DBSNV2 probe was from *D. cayenensis* sample (TDc 3709B) and clustered into Kenyon *et al*. (2008) sequence groups K04 (Table 2.2) sharing 91.8% nucleotide identity with DBSNV for which viral particles have also been isolated (Seal and Muller, 2007). The two ribosomal DNA probes pTa 71 (45S rDNA) and pTa 974 (5S rDNA) were obtained from *Triticum aestivum* and used as FISH-positive controls (Gerlach and Bedbrook, 1979; Gerlach and Dyer, 1980). These plant rDNAs 45S and 5S occur as repetitive DNAs in eukaryote genomes and are clustered at specific chromosomal positions. Therefore, pTa 71 and pTa 974 probes were expected to hybridise with yam chromosomes as they encode for the plant housekeeping genes,

and these sequences are present in high copy number.

2.5 Probe Preparation

2.5.1 PCR-synthesized Probes

The efficiency of probe labelling and concentration were determined by the band intensity of 5 µl probes electrophoresed on 1% (w/v) agarose in 0.5xTBE gels and measurement of the serial diluted probe as described in the instruction manual (Roche, UK).

2.5.2 Random Priming-synthesized Probes

Longer probes were synthesized using random priming. The efficiency of probe labelling was determined by measuring the colour intensity of 1 µl probes placed on a nylon membrane.

2.5.3 Probe Labelling by Random Primers

Double stranded DNAs were denatured to allow random primers to bind to the single-stranded DNAs. The single DNA is amplified using Ex-Klenow or Klenow fragment of *E.coli* DNA polymerase 1 according to the described Schwarzacher and Heslop-Harrison (2000) method as outlined below. The probes were labelled with Random primer labelling system (Roche, UK) for digoxigenin incorporation and Bioprime® array CGH labelling system (Bioline, UK) for biotin incorporation. Badnavirus RCA insert plasmids were used as the DNA templates.

Fifty microliters (50 µl) reactions were set up for both Digoxigenin and Biotin using two step mixtures. DNA (2500 ng/25 µl) was mixed with 20 µl Primers (2.5x). The mixed DNA was denatured in boiling water (95^{o}C for 5 min) in a water bath, and immediately chilled on ice for 5 min. The outlined components below were added on ice. The second mix was 3 µl dNTPs mix (10x) and 1 µl Exo-Klenow fragment (40 U) for Digoxigenin and Biotin labelling. However, the dNTPs mixture includes 5 µl (10x) (biotin-dUTP) and 1.8 µl DIG fluoro dUTP. The reaction was mixed by votexing and then incubated at 37^{0}C for 2 hrs. The reactions were stopped at the end of incubation by the addition of 5 µl of stop buffer (0.5 M EDTA pH 8.0). The labelled probes were purified using NucleoSpin® extract II kit (MACHERY-NAGEL, UK) according to manufacturer's instructions (http://www.mn-net.com/tabid/1452/default.aspx), and the probes were stored at -20^{0}C before use.

2.5.4 Dot Blot for Checking DNA Probe Labelling

Colorimetric dot blot test was performed to check the efficiency of labelled nucleotide incorporation in labelled probes. A piece (3x3 cm) of positively charged nylon membrane (Hybond-N^{+}, Amersham Biosciences, UK) was soaked in buffer 1 (containing 100 mM Tris-HCl, pH 7.5, 15 mM NaCl) at RT for 5 min after marking on the membrane with pencil depending on the number of probes. The membrane was dried between two filter papers. Labelled probes (1 ul) along with a positive control were placed on the membrane closed to the marked points, air-dried (allowed to absorb) for 7 min and then re-soaked in buffer 1 for 2 min. The membrane was incubated at RT for 30 min in buffer 2 (0.5% (w/v) blocking reagent (Roche, UK) in buffer 1). The membrane was drained from excess buffer 2, and then incubated under a plastic coverslip at 37^{o}C for 30 min with gentle agitation, with diluted 1:500 of 0.75 U/ml antibody solutions (containing anti-biotin-alkaline phosphatase and anti-digoxigenin-alkaline phosphatase, Roche, UK) in buffer 1. The membrane was equilibrated in buffer 3 (100 mM Tris-HCl, 100 mM NaCl, 50 mM $MgCl_2$ pH 9.5), and then followed by adding 5 µl detection solution (containing diluted (1:500) stock of INT/BCIP (33 mg/ml 2-(4-iodophenyl)-5-(4-nitrophenyl)-3-phenyltetrazolium chloride and 33 mg/ml 5-bromo-4-chloro-3-indolyl-phosphate, toluidine-salt in DMSO, Roche, UK) in 1500 µl buffer 3. The membrane was then incubated at RT for 7 min in the dark. The labelled probes appeared as dark brown dot on the membrane due to colorimetric reaction of the detection reagents. The labelling efficiency was scored by colour intensity of the dots.

2.6 Chromosome Preparation

Preparation of yam chromosomes was performed as previously described by Schwarzacher and Heslop-Harrison (2000). The fixed root tips were incubated twice for 10 min at RT in 1x enzyme buffer (40 mM Citric acid, 60 mM Tri-sodium citrate, pH 4.6) to remove the fixative and then digested at 37^{o}C for 2 h with enzyme solution (3% (w/v) pectinase 450 U/ml (Sigma, UK), 1.8% (w/v) cellulase 4000 U/g (Calbiochem, UK), 0.2% (w/v) cellulose 5000 U/g (Onozuka, RS) in 1x enzyme buffer). After the digestion the root tips were washed in 1x enzyme buffer for 10 min. The chromosomes were prepared on a clean glass slide (SuperFrost® Menzel-Glasër, Thermo Scientific, UK) under stereo microscope (Carl Zeiss, Germany). A single root tip was placed in a drop (10-30 µl) of 60% (v/v) acetic acid. The meristematic tissue was dissected, spread and then squashed under 18x18 mm coverslip by applying thumb pressure after removing other permanent tissue from the root tip. The coverslips were removed with a razor blade after freezing the slide on dry ice for 5 min. The slides were air-dried, labelled, scanned and kept overnight at RT before FISH subsequent analysis.

2.7 Pre-hybridisation and hybridisation for Fluorescent in situ Hybridization (FISH)

Air-dried slides were marked on the edges of the spread area, and then re-fixed in fresh absolute ethanol: acetic acid (3:1 v/v) for 25 min at RT. The slides were dehydrated twice with 100% (v/v) ethanol for 5 min and allowed to air-dry for 30 min, and then treated with 250 µl RNase A (10 unit/ml, Bioline, UK) diluted 1:100 in 2x SSC under a plastic coverslip in a humid chamber at 37^{o}C for 1 h. The slides were washed twice in 2x SSC at RT for 2 min and 10 min, then equilibrated in 0.01 M HCl for 2 min. Equilibrated slides were treated with 200 µl of 1:100 diluted pepsin (5 unit/ml, Bioline, UK) in 0.01 M HCl under a plastic coverslip in a humid chamber at 37^{o}C for 10 min. Subsequently, the slides were washed in 2x SSC at RT for 5 min and re-fixed in freshly prepared 4% (w/v) paraformaldehyde pH 7.0 (Agar Scientific, UK) for 10 min at RT, then followed by washing twice in 2x SSC for 5 min. Slides were then dehydrated in a series of 70% (v/v), 85% (v/v) and 100% (v/v) ethanol for 2 min each and air-dried. The slides were re-scanned for possible loss of cells that might have occurred during storage or pre hybridisation wash before probing.

A total of 40 µl probe hybridisation mixture (containing 50% (v/v) formamide, 20% (w/v) dextran sulphate, 2x SSC, 25-100 ng probe, 0.0025 µg of salmon sperm DNA, 0.125% (w/v) SDS and 0.125 mM EDTA) was applied per slide. The probe hybridisation mixture was first denatured at 80^{o}C for 10 min and immediately chilled on ice for 10 min. Probe hybridisation mixture was then placed on the slide and denatured together with chromosomal DNA on a Hybaid OmniSlide (Thermo Scientific, UK) at 75^{o}C for 7 min under plastic coverslip and slowly cooled to the hybridisation temperature of 37^{o}C for 18 h with vibration set up to 1.

2.7.1 Post Hybridisation Washes

Post hybridisation washes were carried out to remove the hybridisation mixture and any unbound probe. Coverslips were floated off by incubating the slides in 2x SSC at 42^{o}C for 2 min. Two low stringent washes were carried out with 0.1x SSC at 42^{o}C for 5 min each, and one low stringent wash was carried out with 0.1x SSC at 42^{o}C for 10 min. Slides were then washed in 2x SSC at 42^{o}C for 5 min and followed by cooling down to RT in 2x SSC.

2.7.2 Slides Detection

Slides were incubated in detection buffer (containing, 4x SSC and 0.2% (v/v) Tween-20) at RT for 5 min and then blocked at 37^{o}C for 30 min with 200 µl of 5% (w/v) BSA made in detection buffer. Hybridisation sites were detected with 40 µl of 2 µg/ml streptavidin conjugated to Alexa594 (Molecular Probes, UK) and 4 µg/ml antidigoxigenin conjugated to fluorescein isothiocyanate (FITC) (, Roche, UK) made up in 5% (v/v) BSA solution. Slides were incubated at 37^{o}C for 1 h in a humid chamber,

and then followed by two washes in detection at 40°C for 8 min each.

2.7.3 Mounting of Slides

Chromosomes were counterstained with 100 µl of 4 µg/ml 4', 6-diamidino-2-phenylindole (DAPI) diluted in McIlvaine's buffer (0.1 M citric acid, 0.2 M di-sodium hydrogen phosphate, pH 7.0) at RT for 30 min in dark, then followed by rinsing in a detection buffer before mounting in 2 drops (~60 µl) of antifade solution (Citiflour, Agar Scientific, UK) under a coverslip (24x40 mm). Slides were incubated at 4°C overnight to allow binding of the antifade solution to the fluorophores that stabilize the fluorescence when viewed under a microscope.

2.7.4 Photography and Image Processing

Slides were analysed on a Zeiss epifluorescence microscope (Carl Zeiss, UK) with single band filters equipped with a CCD[defined earlier?] camera (ProgRes™ C12, Optronics, model S97790, UK). Signals were analysed using Filter Set 10 (excitation = BP450-490, bean splitter = FT510 and emission = BP515-565) for degoxigenin-labelled probes and Filter Set 15 (excitation = BP546/12, bean splitter = FT580 and emission = LP590) for biotin-labelled probes. Whereas the DAPI-stained chromosomes were analysed using UV band pass filter (Filter Set 01; excitation = BP365/12, bean splitter = FT395 and emission = LP397). Each metaphase was captured in three different filter sets and then overlaid and analysed using adobe Photoshop CS3 (Sofonic International, Barcelona, Spain). Only those functions that treat all pixels of the image equally were used for colour balance and processing.

2.8 Southern Blotting

Restriction enzyme digestions were performed directly on column-purified DNAs in relatively large volumes (150 µl). It was then followed by DNA precipitation at the final stage to concentrate samples to allow 10-20 µg of restricted DNA in a single gel well. DNA samples were separated by gel electrophoresis in an undigested and restriction enzyme-digested form, and then transferred to a solid membrane support to allow hybridisation of the viral DNA probes.

Southern blot hybridisation was used with badnavirus probes to differentiate badnavirus sequences that hybridise to yam genomic DNA samples (evidenced by >8 kb bands in undigested and additional bands in digested material), from active replicating virus (evidenced by ~7–8 kb bands in the undigested and several smaller fragments in digested). Similar approaches were employed for the detection of EPRVs sequences in potato (Hansen *et al*., 2005), tomato (Staginnus *et al*., 2007), banana (Geering *et al*., 2005a, 2005b; Gayral *et al*., 2008) and petunia (Richert-Pöggeler *et al*., 2003). Most Southern hybridisations utilize radioactive ^{32}P-labelled probes (Geering *et al*., 2001; Harper *et al*., 1999; Seal *et al*., 2014; Umber *et al*., 2014). In this study, non-radioactive digoxigenin (DIG) labelled probes were used like those reported for the detection of activatable ePVCV in petunia (Richert-Pöggeler *et al*., 2003). The probes used were from one representative each of DGGE bands 4, 5 and 6 of *D. rotundata* samples (Turaki *et al*., 2017a) as well as a ~ 3.5 kb RCA clone from a *D. rotundata* sample (TDr 89/02574). The name of the DGGE clone sequences are NGb4_Dr, NGb5_Dr and NGb6_Dr, and fall into Kenyon *et al*. (2008) and Umber *et al*. (2014) sequence groups K09, K08 and U12 respectively. The RCA clone used (GV1) represented the putative badnavirus new species group T13.

3.0 Results

Despite optimisation of the CTAB method for its ability to extract high quality DNAs from 100 mg leaf tissue, this modified method failed to yield DNA of sufficient quality on the 'maxi'-preps (~3 g of leaf tissue) needed for Southern blot hybridisation. Therefore, further optimisation of the CTAB method had to be carried out. Yields from two additional methods were compared to the modified CTAB method. The two additional methods tried were the inclusion of a polysaccharide pre-treatment step using HEPES, and a Qiagen Tip100/G column

purification of the DNA generated with the modified CTAB method (Turaki *et al.,* 2017b). Figure 2 shows that Qiagen Tip100/G column purification of extracted DNA (QC1-QC4) yielded DNA of better quality compared to the modified CTAB method (CT1-CT4) developed in Turaki *et al*., 2017b or the modified CTAB method on HEPES treated grounded leaf (HP1-HP4). From the result in Figure 2, the low molecular weight DNA/RNAs were successfully removed from the DNAs of *D. rotundata* samples (lane QC1-2) using Qiagen Tip100/G column purification. In contrast, the low molecular weight DNA/RNA was not completely removed from *D. alata* sample (lane QC3-4). This could be due to large quantities of mucilaginous substances that were observed to drop through the columns of *D. alata* compared to that of *D. rotundata* samples during purification. This led to slow passage of the buffers through the columns in *D. alata* samples until a gentle force was applied to enable the flow of the solutions as suggested by the manufacturers of the columns (Qiagen, UK). This additional force might have led to the passage of unwanted products through the columnsfxc.

The PCR method used in this study for labelling was found not to be efficient for long DNA template as shown in lane PG1, (Figure 3).

Southern hybridisation with probe NGb5_Dr (group K08) revealed high molecular weight (>8 kb) plant genomic bands on *D. rotundata* (Figure 4B) as well as viral particle-sized bands (~7-8 kb) on *D. alata* (Figure 4A). The plant genomic bands hybridising were detected in undigested (Figure 4B, lanes 1 and 3) as well as *Nde*I-digested B26 and *Eco*RV/*Hin*dIII-digested G2 (Figure 4B, lanes 2 and 4). The viral particle bands hybridising were detected in undigested (Figure 4A, lanes 1, 3 and 5) and in *Hin*dIII-digested G5, *Pst*I-digested G11 and *Hin*dIII-digested G16 (Figure 4A, lanes 2, 4 and 6).

Metaphase chromosomes of *D. rotundata* and *D. alata* samples were stained with DAPI as shown in Figure 5. Figures 5B and 5C shows the probe signals labelled with digoxigenin and biotin, while while Figure 5D showed the overlay images of Figures 5A, 5B and 5C.

Fluorescence hybridisation with probe

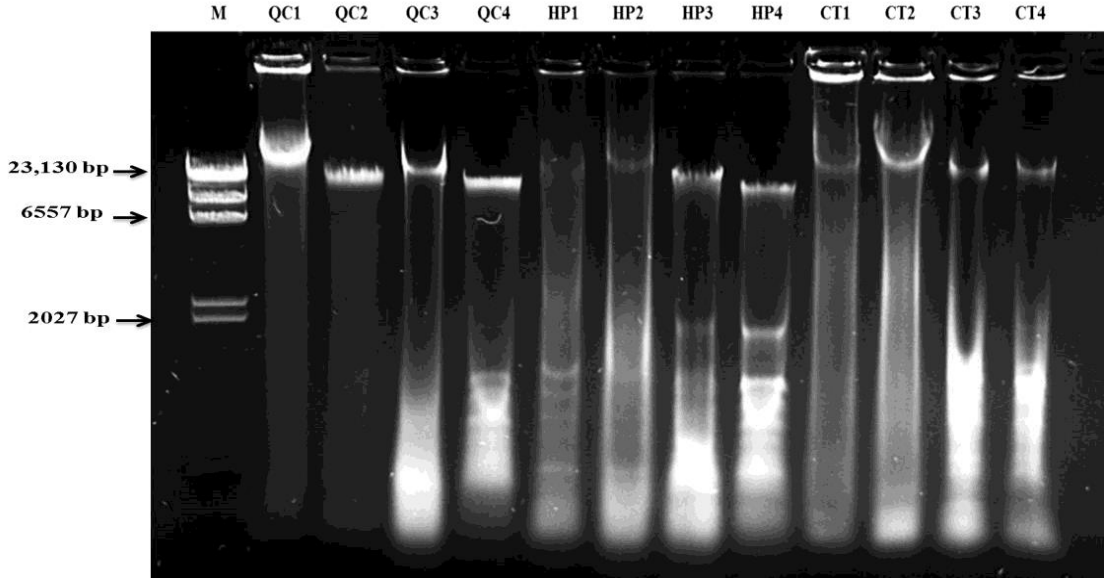

Figure 2: Total genomic DNA extracted from *D. alata* and *D. rotundata* using three methods, namely; Qiagen Tip100/G column purification of extracted DNA (QC), modified CTAB method on HEPES-treated leaf (HP) and modified CTAB method (CT). All PCR products were electrophoresed through a 0.8% (w/v) agarose in 0.5x TBE gel. Lane M = lambda *Hin*dIII ladder (NEB, UK), lane QC1, HP1 and CT1 = *D. rotundata* sample (G9), lane QC2, HP2 and CT2 = *D. rotundata* sample (G12), lane QC3, HP3 and CT3 = *D. alata* sample (G5) and lane QC4, HP4 and CT4 = *D. alata* sample (G16).

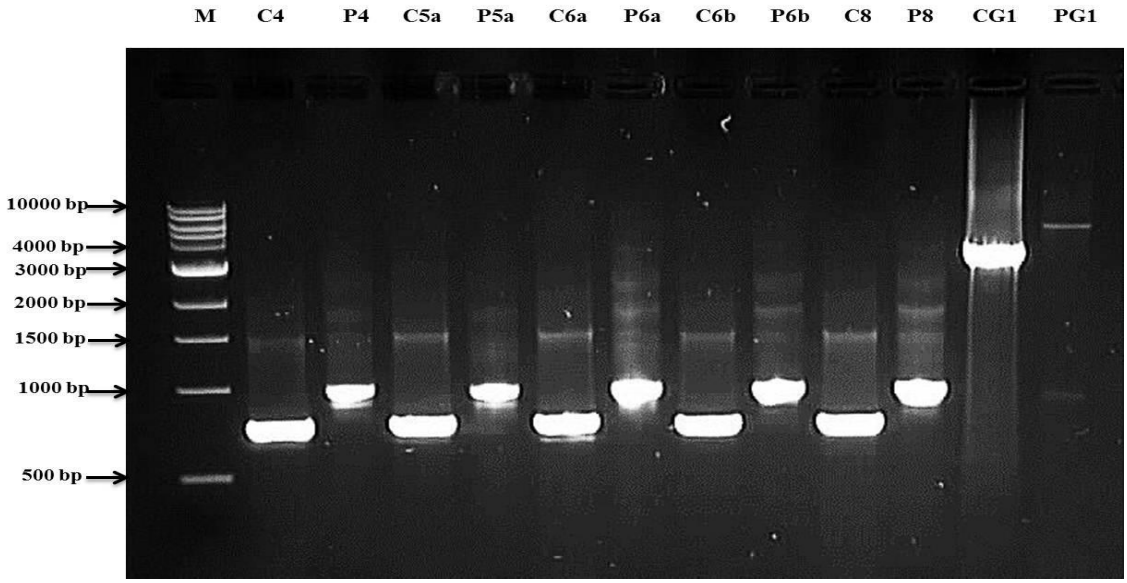

Figure 3: Determination of PCR-DIG labelling of probes used on Southern blots by 1% (w/v) agarose gel electrophoresis. Lane M = 1 kb ladder (NEB, UK); lane C4 and P4 = control and probe from DGGE clone 4; lane C5a and P5a = control and probe from DGGE clone 5a; lane C6a and P6a = control and probe from DGGE clone 6a; lane C6b and P6b = control and probe from DGGE clone 6b; lane C8 and P8 = control and probe from DGGE clone 8; lane CG1 and PG1 = control and probe from RCA clone GV1.

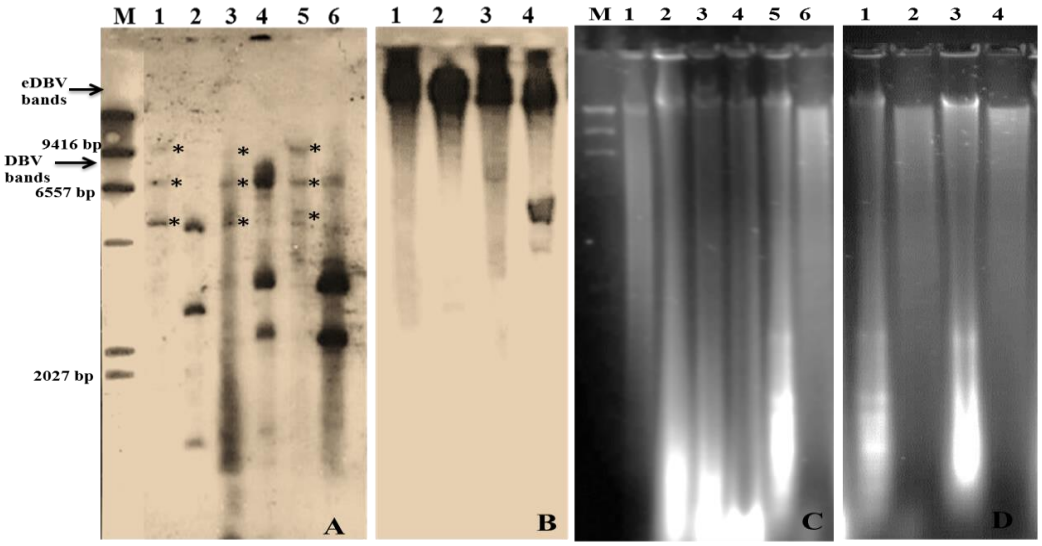

Figure 4: Southern blot hybridisation of total DNA from *D. alata* and *D. rotundata* breeding lines (undigested and restriction enzyme-digested), revealing hybridisation with DIG-labelled DGGE clone sequence (NGb5_Dr, group K08) as probe. **Plate A)** represents probe hybridisation on episomal virus DNA in *D. Alata*. **Plate B**) represents hybridisation of plant DNA (integrated) in *D. rotundata*. **Plates C** and **D** represent total uncut and restriction enzyme-digested *Dioscorea* species DNA samples were electrophoresed through 0.8% (w/v) agarose gel in 0.5x TBE buffer. DNA samples were loaded in the following general order: undigested DNA (odd lane numbers) and restriction enzyme-digested DNA (even lane numbers). The asterisk (*) in plate A indicates present of viral band in undigested *D. alata* DNA. Lanes represent **Plate A)** 1 = G5 undigested, 2 = G5 *Hin*dIII-digested, 3 = G11 undigested, 4 = G11 *Pst*I-digested, 5 = G16 undigested, 6 = G16 *Hin*dIII –digested. **Plate B)** 1 = B26 undigested, 2 = B26 *Nde*I-digested, 3 = G2 undigested, 4 = G2 *Eco*RV/*Hin*dIII-digested.

DBALV2 revealed weak hybridising signals on *D. rotundata* (Figure 5 D) over randomly placed dots representing background or non-specific hybridisation with signal strength of several magnitudes lower than that observed with the FISH-positive control (pTa 71).

4.0 Discussion

Separation of polysaccharides from DNA is one of the major challenges of DNA extraction from plant tissues containing high levels of polysaccharides (Sharma *et al*., 2002; Ghaffari *et al*., 2011). Separation of these two polymers (polysaccharides and polyphenols) is usually achieved using high salt concentration in the extraction buffers and the DNA precipitation step (Fang *et al*., 1992). Evidence arising from this study indicates that the modified CTAB extraction yielded good quality DNA for a 'mini prep' but the method could not yield enough DNA of high quality for a 'maxi prep'. To obtain higher quality and sufficient quantity of DNA needed for the Southern hybridisation analyses, additional purification step was carried out. The purification step which was performed using a Qiagen Tip100\G column yielded better qualitative/purified DNA compared to adding HEPES-treatment to leaves (Figure 2), a suggestion offered by IITA to overcome this problem (Bhattacharjee and Gezagegh, personal communication). Two-thirds of the DNA was lost during purification when Qiagen Tip100/G columns purification step was incorporated. This posed a problem for Southern blot hybridisation as large quantities of DNA (10-20 µg/lane) are required. Therefore, for the hybridisation purposes there is need to employ better method that will produce high quantities of DNA. However, in this study several extractions were carried out and then the DNA were pooled together.

The results obtained in this study have further provided supporting evidence that some badnavirus sequences are integrated into the genomes of *D. rotundata* breeding lines that originated from West Africa. The results agreed with those findings reported by Seal *et al.* (2014) and Umber *et al*. (2014). It is important to differentiate between endogenous and episomal viral sequences since endogenous viral sequences may or may not be capable of causing infection (Umber *et al*. 2014). This will equally assist in developing diagnostic tools that can help in avoiding false positives due to 'inert' endogenous sequences.

The probe sequence NGb5_Dr that hybridised to eDBV was shown to be common to all *D. rotundata* by DGGE (Turaki *et al*., 2017a), and falls into species group K08 (Kenyon *et al*., 2008), which clusters with yam badnavirus DBALV for which particles have been isolated (Briddon *et al*., 1999; Phillips *et al*., 1999) and nine RCA-characterised sequences (Bomer *et al.,* 2016). Therefore, hybridisation of NGb5_Dr, a DBALV-like group K08 sequence, to badnaviral particles in *D. alata* and to the plant genome in *D. rotundata* samples (Table 1) suggested that the group K08 eDBV also exists in viral particles in *D. rotundata*. This was already confirmed in this study by the RCA detection of DBALV2 isolate in *D. rotundata* (TDr 95/18544) sample. Furthermore, the probe sequence NGb5_Dr has shown 100% nucleotide identity with episomal sequence NGb0477_Dr and known eDBV8 sequence S2h9Dr (KF829997, Umber *et al*., 2014), and this, therefore, supports the hybridisation result.

A similar hybridisation in the plant genome and episomal virus genomes was recorded with species group K08 neighbouring DGGE probe sequence (NGb6_Dr) that falls into species group U12 [Endeavour not to make too much reference to the tables in the Discussion Section] (Umber *et al.,* 2014). In contrast, the probe sequence NGb6_Dr neither clustered nor showed high nucleotide identity with any RCA-generated episomal sequence. However, the sequence NGb6_Dr showed 100% nucleotide identity with known eDBV12 sequence S1a4Dr (KF829956). In contrast, the DGGE probe sequence NGb4_Dr (100% nucleotide identity with S1h2Dr (KF829975) that was also common to all *D. rotundata* (Turaki *et al*., 2017a), and which falls into species group K09 (Kenyon *et al.,* 2008), only hybridised to eDBV (plant genomic) bands (Table 1), with no hybridising bands in the viral particle range. These suggest that the eDBVs belonging to the K09 group might have originated from more ancient events. This may not lead to episomal viral infection

Table 1 Properties of probe sequences and notes on the nature of hybridisation studies (FISH and Southern blot)

Probe	Sizes (kb)	Amplification	Species group	Description	Sources	Hybridisation notes
pTa71	8.9	-	-	45S rDNA	*Triticum aestivum*	Fish = *D. rotundata* (+)
pTa794	0.41	-	-	5S rDNA	*Triticum aestivum*	Fish = *D. rotundata* (+) and *D. alata* (+/?)
DBALV2	7.1	RCA	K08	ORF 1, 2, 3	*D. alata*	Fish = *D. rotundata* (+/?) and *D. alata* (-)
DBSNV2	4.3	RCA	K04	ORF 2 and 3	*D. cayenensis*	Fish = *D. rotundata* (+/?) and *D. alata* (-)
NGb4_Dr	0.6	DGGE	K08	RT-RNaseH	*D. rotundata*	Southern = *D. rotundata* (plant genome) and *D. alata* (no hybridisation)
NGb5_Dr	0.6	DGGE	K09	RT-RNaseH	*D. rotundata*	Southern = *D. rotundata* (plant genome) and *D. alata* (viral genome)
NGb6_Dr	0.6	DGGE	U12	RT-RNaseH	*D. rotundata*	Southern = *D. rotundata* (plant/viral genome) and *D. alata* (viral genome)
GV1	3.5	RCA	T13	ORF 3	*D. rotundata*	Southern = *D. rotundata* (viral genome) and *D. alata* (viral genome)

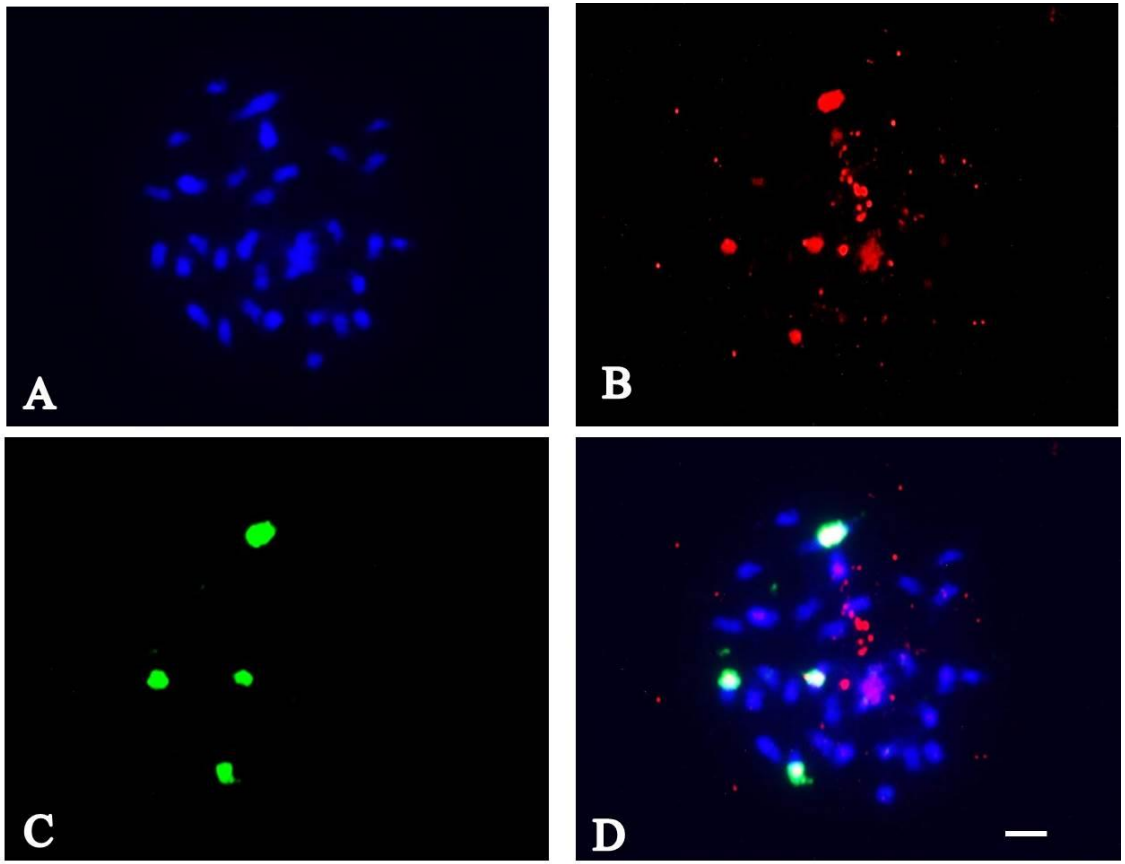

Figure 5: Double target fluorescent *in situ* hybridisation performed on yam (*D. rotundata* TDr 89/02475) root tip metaphase chromosome (2n = 40). **Plate A)** yam (*D. rotundata* TDr 89/02475) chromosomes appeared blue with DAPI fluorescence. **Plate B)** *in situ* hybridisation of the yam badnavirus DNA clone (DBALV2) from *D. alata* labelled with biotin 16-dUTP (detected in red), the fluorescence *in situ* hybridisation allows the detection of physical location of badnavirus sequence on yam chromosome section. **Plate C)** Hybridisation pattern of the pTa71 clone labelled with digoxigenin 11-dUTP (detected in green) showing the physical location of major 45S rDNA sites of yam. **Plate D)** Overlay of A, B and C images. Bar represents 10 µm.

due to having suffered many nucleotide changes from its initial homologous viral form. Further support to this suggestion was the reported partial RT-RNaseH sequences belonging to the K09 group contained stop codons which cannot translate into a functional protein (Kenyon *et al*., 2008). In addition, further support for these sequences representing only 'dead' eDBVs was the inability to characterise full-length genomic sequence belonging to K09 through RCA (Turaki *et al*., 2017a) despite the suggestion that episomal sequences of this group may exist (Umber *et al*., 2014). Southern results using a probe from this group sequence (K09) was found to be repetitive in *D. rotundata* sample (Seal *et al.*, 2014).

The GV1 probe from DBRT1V (new putative badnavirus species) was found to hybridise to only viral genomes in total DNA of *D. rotundata* (TDr 89/02475) and *D. alata* (TDa 95/00310) samples (Table 1). These samples were confirmed to contain episomal sequences of this virus by RCA (Bomer *et al*., 2016). The non-hybridisation of this probe with the plant genome sequences of *D. rotundata* or *D. alata* samples suggest that there are no 'DBRT1V-like' eDBV sequences in the *D. rotundata* or *D. alata* genomes analysed in this study. A similar situation has been reported in *D. sansibarensis* (Seal *et al*., 2014) where a probe from DBSNV sequence could only hybridise to *D. sansibarensis.* The DBSNV sequence was also found to exist as an episomal sequence in *D. cayenensis* and *D. rotundata* (Bomer *et al.*, 2016) indicating that DBSNV virus sequence is not restricted to only a single host (*D. sansibarensis*) as suggested (Seal and Muller, 2007).

The hybridisation signals to plant genome eDBV sequences are much stronger in some plants than others, presumably due to the quality/quantity of genomic DNA per lane. It could also be due to variation in the copy number of the integration event between cultivars, as observed in BSV (Geering *et al*., 2001).

Based on the findings reported by Bomer *et al*. (2016) and Turaki *et al.* (20017a), it can be hypothesized that several independent integration events have occurred in *D. rotundata* genomes, resulting in the presence of eDBVs from at least two distinct badnavirus species groups (K08 and K09). This supports the data recently published for other yam samples (Seal *et al.* 2014; Umber *et al*., 2014). It was further supported by the Blast analyses of yam genome DNA scaffolds[refs needed], and this confirmed the existence of several rearranged eDBVs in the partial RT-RNaseH domain with 83-88% nucleotide identity to K08 and K09 species groups (Tamiru *et al.,* 2017).

Results presented in this study were from short amplicon sequences of the partial RT-RNaseH domain. Further research on yam genome sequences outside the partial RT-RNaseH sequence region is required as this will assist in understanding the nature of the integration events. Such studies on *Musa, Petunia* and *Oryza* genomes have resulted in the identification of the junctions of virus sequence integration (Ndowora *et al.,* 1999; Richert-Pöggeler *et al*., 2003; Kunii *et al.,* 2004; Geering *et al*., 2005a; Chabannes *et al*., 2013). These studies have further assisted in illustrating the importance of integration information, as for BSV integration where the integrated viral DNA detected in *Musa* cultivars was traced back to the parent plant (Geering *et al*., 2001; Harper *et al*., 1999, 2002; Iskra-Caruana *et al*., 2014a, 2014b). The studies were successful due to the availability of genomic libraries for these plant hosts. Full genome sequence of yam has just been recently released (Tamiru et al., 2017). Therefore, further cytological studies can also be utilized to determine the yam chromosome(s) that host eDBV sequence using badnavirus-specific fluorescent probes. Unfortunately, the FISH experiments undertaken in this study were inconclusive due to limited availability of plant material at the time of the experiment.

For eBSVs, it has been highlighted that the integrated sequences could be either detrimental by being infectious or advantageous by protecting their host from infection by closely related badnaviruses (Iskra-Caruana *et al*., 2010). For the activatable EPRVs, it has been proposed that the interaction of genetic factors after interspecies genome hybridisation supported their activation in the hybrid, and the degree of symptom severity is correlated with environmental conditions such as temperature and day length (Harper *et al.,* 2002). Similar

environmental factors may play a role in the expression of infectious endogenous yam badnaviruses as reported in petunia, tobacco and banana by PVCV (Richert-Pöggeler and Shepherd, 1997; Richert-Pöggeler *et al*., 2003), TVCV (Lockhart *et al*., 2000), and BSV (Harper *et al*., 1999; Ndowora *et al.,* 1999; Dallot *et al*., 2001) respectively. Therefore, the presence of activatable EPRV sequences has an important impact on yam breeding strategies. The finding of this study and other related studies (Seal *et al*., 2014; Umber *et al*., 2014) have shown that eDBV sequences are widely spread in West African native *D. rotundata*. This strong instance of DBVs existing as EPRVs implies serious consideration for both yam breeding and exchange of genetic materials because it is currently impossible to eliminate the eDBVs from propagative stock (Alangar *et al.,* 2016).

The use of disease-free yam planting materials is believed to be critical in increasing yam production, particularly in SSA where yam is a significant stable crop (Aighewi et al., 2014). Results from Southern blot hybridisation confirm that improvements of current PCR-based diagnostic tools are required to prevent false positives. Likewise, rapid diagnostic techniques that can allow the differentiation of non-functional eDBVs from those in which viral particles occur are required. This situation is similar to the diagnosis of BSV in *Musa* species, whereby optimisation had to be carried out to prevent false positive results arising from the presence of eBSVs (Le Provost *et al*., 2006). It is equally important for the vegetative propagation of yam materials or developing virus management tactics to distinguish clearly whether viral or host plant sequences are being detected. However, in a situation where DBVs are actively integrated into yam genome(s), this will necessitate studying the factors involved in the expression of EPRVs resulting in pathogenicity.

This study further analysed the location of such badnavirus sequences in the chromosome of West African yam breeding lines. Metaphase chromosome hybridisation was performed using FISH with badnavirus cloned sequences as probes to locate the physical position of the eDBV sequence on *Dioscorea* chromosome. This was similar approach employed by previous studies in demonstrating the location of eBSV, ePVCV and *Solanum lycopersicum* endogenous pararetrovirus (*Lyc*EPRV) sequences on their host chromosomes *Musa*, *Petunia* and *Solanum lycopersicum* (tomato) respectively (Harper *et al*., 1999; Richert-Pöggeler *et al*., 2003; Staginnus *et al*., 2007). This is also illustrated in BSV and PVCV, in which the integrated viral DNA was detected and linked to a specific chromosome, then traced back to a parent during species separation/breeding (Geering *et al*., 2001; Harper *et al*., 1999, 2002; Richert-Pöggeler *et al*., 2003). However, the results presented in this study demonstrated unconvincing FISH signal over a diffusing background because of the degeneracy of the primers, thus making the results inconclusive.

In this study, the probes used were obtained from the episomal DNAs of yam badnaviruses by RCA. The probe DBALV2 was chosen because it clusters in the K08 group, which contains both episomal and integrated sequences. Similarly, the probe DBSNV2 clusters in K04 group, a neighbour to K05 group, also reported to contain both episomal and integrated sequences (Umber *et al*., 2014). Further support for the choice of probes was due to high nucleotide identity observed with some yam DNA scaffold (Tamiru *et al.,* 2017). Furthermore, the choice was due to poor availability of plant materials (root tips) at the time of FISH analyses. High hybridisation stringency was used in this study to make the probes not hybridise to other reverse transcriptase-like sequences in the yam genome. The genome of several plant species has been shown to contain dispersed distribution of reverse transcriptase gene of *Ty1-copia*-like retrotransposons on their chromosomes (Heslop-Harrison *et al*., 1997).

Conclusion

DGGE, Southern hybridisation and RCA have confirmed the existences of both episomal and endogenous sequences in *D. rotundata* accessions. However, some of these *Dioscorea* accessions could not give convincing signals on the chromosome metaphases by FISH that could assist in drawing a clear conclusion like BSV

(Harper *et al*., 1999). These could be attributed to either due to low copy number of integration sites or the inability of the probes to hybridise with integrated sequence considering the extensive viral heterogeneity. Therefore, further optimisations of the technique on many yam breeding lines accessions is required to conclude on the integration sites of badnavirus sequences in yam chromosome. Similarly, FISH probes need to be improved since poor probe labelling will affect the resolution.

Acknowledgements

We wish to thank Professor Susan E Seal for mentoring and all the support she has given, We also wish to thank Dr Trude Schwarzacher and Professor Heslop-Harrison university of Leicester UK for their support and allowing us to carry out the research in their laboratory.

References

Aighewi B .A., Maroya, N. G. and Asiedu, R. (2014). Seed yam production from minisetts: A training manual. YIIFSWA International Institute of Tropical Agriculture (IITA).

Alangar, I. B., Thomas, H. and Ramasamy, S. (2016). Badnaviruses: The Current Global Scenario. Viruses 8(6):177; doi:10.3390/v8060177

Bömer, M., Rathnayake, A. I., Visendi, P., Silva, G. and Seal, S. E. (2017). Complete genome sequence of a new member of the genus *Badnavirus*, *Dioscorea bacilliform* RT virus 3, reveals the first evidence of recombination in yam badnaviruses. Archive of Virology 163(2):533–538.

Bömer, M., Turaki, A. A., Silva, G., Kumar, P. L. and Seal, S. E. (2016). A sequence-independent strategy for amplification and characterisation of episomal badnavirus sequences reveals three previously uncharacterised yam badnaviruses. Viruses 8(7):188-209.

Bousalem, M., Durand, O., Scarcelli, N., Lebas, B.S.M., Kenyon, L., Marchand, J.L., Lefort, F. and Seal, E.S. (2009). Dilemmas caused by endogenous pararetrovirus regarding the taxonomy and diagnosis of yam (*Dioscorea* spp.) badnaviruses: analyses to support safe movement. Archives of Virology 154 (2):297-314.

Briddon, R. W., Phillips, S., Brunt A. and Hull R. (1999). Analysis of the sequence of *Dioscorea alata* bacilliform virus; comparison to other members of the badnavirus group. Virus Genes 183: 277-283.

Campbell, A (2003). The future of bacteriophage biology. Nature Review Genetics 4:471-477.

Chabannes, M. and Iskra-Caruana M. L. (2013). Endogenous pararetroviruses-a reservoir of virus infection in plants. Current Opinion Virology 3: 615-620.

Chabannes, M., Baurens, F. C., Duroy, P. O., Bocs, S., Vernerey, M. S., Rodier-Goud, M., Barbe, V., Gayral, P. and Iskra-Caruana, M. L. (2013) Three infectious viral species lying in wait in the banana genome. Journal of Virology 87: 8624-8637.

Dallot S., Acuña P., Rivera C., Ramírez P., Côte F., Lockhart, B. E. L. and Iskra-Caruana, M. L. (2001). Evidence that the proliferation stage of micropropagation procedure is determinant in the expression of *Banana streak virus* integrated in to genome of the FHIA21 hybrid (*Musa* AAAB). Archives of Virology 146: 2179-2190.

Eni, A. O., Hughes, J. d'A., Asiedu, R. and Rey M. E. C. (2008). Sequence diversity among badnavirus isolates infecting yam (*Dioscorea* spp.) in Ghana, Togo, Benin and Nigeria. Archives of Virology 153: 2263-2272.

Fang, G., Hammar, S. and Grumet, R. (1992). A quick and inexpensive method for removing polysaccharides from plant genomic DNA. Biotechniques 13:52.

Feschotte, C. and Gilbert, C. (2012). Endogenous viruses: insights into viral evolution and impact on host biology. Nature Review Genetics 13:283-296.

Gayral, P., Noa-Carrazana, J. C., Lescot, M., Lheureux, F., Lockhart, B. E. L., Matsumoto, T., Piffanelli, P. and Iskra-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.

Geering, A. D. W., Olszewski, N. E., Dahal, G., Thomas, J. E. 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.

Geering, A. D. W., Olszewski, N. E., Harper, G., Lockhart, B. E. L., Hull, R. and Thomas, J. E. (2005a). Banana contains a diverse array of endogenous badnaviruses. Journal of General Virology 86:511-520.

Geering A. D. W., Pooggin, M. M., Olszewski, N. E., Lockhart, B. E. L. and Thomas, J. E. (2005b). Characterisation of *Banana streak Mysore virus* and evidence that its DNA is integrated in the B

genome of the cultivated *Musa* sp. Archives of Virology 150:787-796.

Geering, A. D. W., Tanya, S. and Teycheney, P. Y. (2010). The classification and nomenclature of endogenous viruses of the family *Caulimoviridae.* Archives of Virology 155:123-131.

Gerlach, W. L. and Bedbrook, J. R. (1979). Cloning and characterization of ribosomal RNA genes from wheat and barley. Nucleic Acids Research 7:1869-1885.

Gerlach, W. L. and Dyer, T. A. (1980). Sequence organization of the repeating units in the nucleus of wheat which contain 5S rRNA genes. Nucleic Acids Research 8:4851-4865.

Ghaffari, S., Hasnaoui, N. and Ferchichi, A. (2011). Rapid, high quality DNA isolation from Tunisian Grapevine (*Vitis vinifera* L.) cultivars and optimization of the RAPD Marker Technique. Romanian Biotechnological Letters 16:2881-2890.

Hansen, C. N., Harper, G. and Heslop-Harrison, J. S. (2005). Characterisation of pararetrovirus-like sequences in the genome of potato (*Solanum tuberosum*). Cytogenetic and Genome Research 110:559-565.

Harper, G., Hull, R., Lockhart, B. E. L. and Olszewski, N. E. (2002). Viral sequences integrated into plant genomes. Annual Review of Phytopathology 140:119-136.

Harper, G., Osuji, O. J., Heslop-Harrison, J. S. and Hull, R. (1999). Integration of banana streak badnavirus into the *Musa* genome: molecular and cytogenetic evidence. Virology 255:207-213.

Heslop-Harrison, J. S. P, Brandes, A., Taketa, S., Schmidt, T., Vershinin, A. V., Alkhimova, E. G., Kamm, A., Doudrick, R. L., Schwarzacher, T., Katsiotis, A., Kubis, S., Kumar, A., Pearce, S. R., Flavell, A. J., Harrison, G. E. and Capy, P. (1997). The chromosomal distributions of Ty1-*copia* group retrotransposable elements in higher plants and their implications for genome evolution. Genetica 100:197-204.

Hohn, T., Richert-Pöggeler, K. R., Harper, G., Schwarzacher, T., Teo, C. H., Techeney, P. Y., Iskra-Caruana, M. L. and Hull, R. (2008). Evolution of integrated plant viruses. In*:* M. Roossinck (ed.), Plant virus evolution. Springer, Heidelberg, Germany, pp. 58–81.

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. The American Phytopathological Society 23:1394-1402.

Iskra-Caruana, M. L., Chabannes, M. and Duroy, P. O. (2014b). A possible scenario for the evolution of Banana streak virus in banana Virus Research http://dx.doi.org/10.1016/j.virusres.2014.01.005.

Iskra-Caruana, M. L., Duroy, P. O., Chabannes, M. and Muller, E. (2014a). The common evolutionary history of badnaviruses and banana. Infectious Genetic Evolution 21:83-89.

Jakowitsch J., Mette, M. F., Van der Winden, G., Matzke, M. A. and Matzke, A. J. M. (1999). Integrated pararetroviral sequences define a unique class of dispersed repetitive DNA in plants. Evolution 96:13241-13246.

Katzourakis, A., Vini Pereira, V. and Tristem, M. (2007). Effects of recombination rate on human endogenous retrovirus fixation and persistence. Journal of Virology 81: 10712-10717.

Kenyon, L., Lebas, B. S. M. and Seal, S. E. (2008). Yams (*Dioscorea* spp.) from the South Pacific Islands contain many novel badnaviruses: implications for international movement of yam germplasm. Archives of Virology 153:877-889.

Kunii, M., Kanda, M., Nagano, H., Uyeda, I., Kishima, Y. and Sano, Y. (2004). Reconstruction of putative DNA virus from endogenous *Rice tungro bacilliform* virus-like sequences in the rice genome: implications for integration and evolution. BMC Genomics 5:80.

Le Provost, G., Iskra-Caruana, M. L., Acina, I. and Teycheney, P. Y. (2006). Improved detection of episomal banana streak viruses by multiplex immunocapture PCR. Journal of Virological Methods 137:7-13.

Lefeuvre, P., Harkins, G. W., Lett, J. M., Briddon, R. W., Chase, M. W., Benoit, M. and Darren, P. M. (2011). Evolutionary time-scale of the Begomoviruses: evidence from integrated sequences in the *Nicotiana* genome. PLoS ONE, 6:e19193.

Lheureux F., Carreel, F., Jenny, C., Lockhart, B. E. L. and Iskra-Caruana, M. L. (2003). Identification of genetic makers linked to banana streak disease expression in inter-specific *Musa* hybrids. Theoretical Applied Genetics 106: 594-598.

Lockhart. B. E. L., Menke, J., Dahal, G. and Olszewski, N. E. (2000). Characterisation and genome analysis of *Tobacco vein clearing virus*: a plant pararetrovirus that is transmitted vertically and related to sequences integrated in the host genome. Journal of General Virology 81:1579-1585.

Mette, M. F., Kanno, T., Aufsatz, W., Jakowitsch, J., Van der Winden, J., Matzke, M. A. and Matzke, A. J. M. (2002). Endogenous viral sequences and their potential contribution to heritable virus resistance in plants. The EMBO Journal 21:461-469.

Tamiru, M., Satoshi, N., Hiroki, T., Benjamen, W., Hiroki, Y., Motoki, S., Kentaro, Y., Aiko, U., Kaori, O., Akira, A., Naoya, U., Hideo, M., Pachakkil, B., Shinsuke, Y., Ryo, M., Satoru, M., Gezahegn, G., Antonio, L. M., Melaku, G., Ranjana, B., Michael, A., Lava, K. P., Ismail, R., Mai, T., Toru, T., Wilfried, H., Manuel, C., Sophien, K., Günter, K., Hiroko, T., Robert, A. and Ryohei, T. (2017). Genome sequencing of the staple food crop white Guinea yam enables the development of a molecular marker for sex determination. BMC Biology 15:86 DOI 10.1186/s12915-017-0419-x

Ndowora, T. C., Dahal, G., LaFleur, D., Harper, G., Hull, R., Olszewski, N. E. and Lockhart, B. E. L. (1999). Evidence that badnavirus infection in *Musa* can originate from integrated pararetroviral sequences. Virology 225:214-220.

Pahalawatta, V., Druffel, K. and Pappu, H. (2008). A new and distinct species in the genus *Caulimovirus* exists as an endogenous plant pararetroviral sequence in its host, *Dahlia variabilis*. Virology 376:253-257.

Phillips, S., Briddon, R. W., Brunt, A. A. and Hull, R. (1999). The partial characterisation of a badnavirus infecting the greater Asiatic or water yam (*Dioscorea alata*). Journal of Phytopathology 147:265-26.

Pistello, M. and Antonelli, G. (2016). Integration of the viral genome into the host cell genome: a double-edged sword. Clinical Microbiology and Infection 22(4):296-298.

Richert-Pöggeler, K. R., Noreen, F., Schwarzacher, T., Harper, G. and Hohn, T. (2003) Induction of infectious petunia vein clearing (pararetro) virus from endogenous provirus in petunia. The EMBO Journal 22:4836-4845.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular cloning: a laboratory manual*, 2rd edition Cold Spring Harbor Laboratory Press, Cold Spring Harbour, New York, USA. pp. 931-957.

Seal, S. E., Turaki, A., Muller, E., Kumar, P. L., Kenyon, L., Filloux, D., Galzi, S., Lopez-Montes, A. and Iskra-Caruana, M. L. (2014). The prevalence of badnaviruses in West African yams (*Dioscorea cayenensis-rotundata*) and evidence of endogenous pararetrovirus sequences in their genomes. Virus Research 186:144-154.

Seal, S. E. and Muller, E. (2007). Molecular analysis of a full-length sequence of a new yam badnavirus from *Dioscorea sansibarensis*. Archives of Virology 7:1-7.

Sharma, A. D., Gill, P. K. and Singh, P. (2002). DNA isolation from dry and fresh samples of polysaccharide-rich plants. Plant Molecular Biology Reporter 20:415-419.

Staginnus, C. and Richert-Pöggeler, K. R. (2006). Endogenous pararetroviruses: two-faced travellers in the plant genome. Trends in Plant Science 11:485-491.

Staginnus, C., Gregor, W., Mette, M. F., Teo, C. H., Borroto-fernàndez, E. G., Machado, M. L., Matzke, M. and Schwarzacher, T. (2007). Endogenous pararetroviral sequences in tomato (*Solanum lycopersicum*) and related species. BMC Plant Biology 7:24-32.

Staginnus, C., Iskra-Caruana, M. L., Lockhart, B. E. L., Hohn, T. and Richert-Poggeler, K. R. (2009). Suggestions for nomenclature of endogenous pararetroviral (EPRV) sequences in plants. Archives of Virology 154:1189-1193.

Tamiru, M., Natsume, S., Takagi, H., Babil, P., Yamanaka, S., Lopez-Montes, A., Gedil, M., Bhattacharjee, R., Takagi, H., Asiedu, R. and Terauchi, R. (2013). Whole genome sequencing of Guinea yam (*Dioscorea rotundata*). In: Book of abstracts for Yams-2013, First global conference on Yam, 3–6 October 2013, Accra, Ghana, pp. 20.

Turaki,, A. A., Bömer, M., Silva, G., Kumar, P. L. and Seal, S. E. (2017a). PCR-DGGE analysis: Unravelling complex mixtures of *Badnavirus* sequences present in yam germplasm. Viruses 9(7);181; doi:10.3390/v9070181

Turaki A. A., Ahmad, B., Magaji, U. F., Abdulrazak U. K., Yusuf, B. A. and Hamza, A. B. (2017b). Optimised cetyltrimethylammonium bromide (CTAB) DNA extraction method of plant leaf with high polysaccharide and polyphenolic compounds for downstream reliable molecular analyses African Journal of Biotechnology 16(24):1354-1365.

Umber, M., Filloux, D., Muller, E., Laboureau, N., Galzi, S., Roumagnac, P., Iskra-Caruana, M. L., Pavis, C., Teycheney, P. Y. and Seal, S. E. (2014). The genome of African yam (*Dioscorea cayenensis-rotundata* complex) hosts endogenous sequences from four distinct badnavirus species. Molecular Plant Pathology 1-12 DOI: 10.1111/mpp.12137.