South African Journal of Botany 117 (2018) 256–267

Contents lists available at ScienceDirect

South African Journal of Botany

journal homepage: www.elsevier.com/locate/sajb

Viral metagenomics reveals sweet potato virus diversity in the Eastern and Western Cape provinces of South Africa

T.F. Nhlapo [a,e,*], D.J.G. Rees [b], D.A. Odeny [c], J.M. Mulabisana [d], M.E.C. Rey [e]

[a] Biotechnology Platform, Agricultural Research Council, Private Bag X5, Onderstepoort 0110, South Africa
[b] Department of Life and Consumer Sciences, University of South Africa, Private Bag X6, Florida 1710, South Africa
[c] ICRISAT (ICRAF Campus), PO Box 39063-00623, Nairobi, Kenya
[d] Vegetable and Ornamental Plant Institute, Agricultural Research Council, Private Bag X293, Pretoria 0001, South Africa
[e] School of Molecular and Cell Biology, University of the Witwatersrand, Private Bag 3, Johannesburg PO Wits 2050, South Africa

ARTICLE INFO

Article history:
Received 7 April 2018
Accepted 29 May 2018
Available online xxxx

Edited by E Balázs

Keywords:
Viral metagenomics
Ipomoea batatas
Virus detection
Next generation sequencing (NGS)
Mixed infections

ABSTRACT

Limited studies have been undertaken with regard to virus complexes contributing to the aetiology of sweet potato virus disease (SPVD) in South Africa (SA). In this study, a metagenomic approach was adopted to reveal the genetic diversity of viruses infecting sweet potato. In order to undertake a comprehensive analysis of viral sequences, total RNA was isolated from 17 asymptomatic and symptomatic sweet potato plants that were collected from the Eastern (EC) and Western Cape (WC) provinces of SA. DNase-treated total RNA was depleted of ribosomal RNA (rRNA) and deep-sequenced using the Illumina MiSeq platform. Genomic DNA, isolated from the same plants, underwent rolling circle amplification (RCA) and deep sequencing. Sequence reads were analysed with the CLC Bio Genomics Workbench. Both *de novo* and reference-guided assemblies were performed resulting in four near full-length RNA virus genomes. BLAST searches using *de novo* assembled sequences against published virus genomes confirmed the presence of previously detected begomoviruses in the Western Cape (WC) province, namely Sweet potato mosaic virus (SPMaV) and Sweet potato leaf curl Sao Paulo virus (SPLCSPV). The begomoviruses were detected in mixed infections with two major disease-causing RNA viruses, Sweet potato feathery mottle virus (SPFMV) and Sweet potato chlorotic stunt virus (SPCSV). The sequence data further demonstrated mixed infections of RNA and DNA viruses from 11 of the 17 sequenced samples. Metagenomics is a reliable diagnostic tool for virus diversity detection, in particular virus-complexes and synergies affecting disease aetiology.

1. Introduction

Sweet potato (*Ipomoea batatas* (L.) Lam.) is a dicotyledonous perennial plant belonging to the morning glory family *Convolvulaceae*. It produces edible, highly nutritious tubers and is ranked as the 3rd most important root crop and the 7th most important staple crop in the world (Valverde et al., 2007; Clark et al., 2012). Sweet potato is an attractive crop to resource-poor farmers as it is easy to grow, high yielding, is drought and heat tolerant, and crowds out weeds (Kays, 2004). In countries such as Zambia, South Africa (SA), Uganda, Kenya, Nigeria, and Tanzania women are the primary growers of sweet potato and they use it to generate income. Sweet potatoes have a high content of carbohydrates and dietary fibre; they are rich in vitamin A, vitamin C, vitamin B6 and because of these nutritional benefits they are used for poverty alleviation (Van Jaarsveld et al., 2005). A number of DNA and RNA viruses accumulate in the crop as a result of vegetative propagation (Valverde et al., 2007), and pose a serious threat to sweet potato production (Valverde et al., 2007; Kreuze et al., 2009; Tesfaye et al., 2011). It has been recorded that viral diseases can decrease yield and quality of sweet potato storage roots by 30–100% in countries such as the United States (Valverde et al., 2007), Peru (Cuellar et al., 2008), SA (Domola et al., 2008), Kenya (Ateka et al., 2004) and Ethiopia (Tesfaye et al., 2011).

The most prominent viral disease complex known to affect sweet potato worldwide is sweet potato virus disease (SPVD), which was observed for the first time in Uganda in 1940 (Ateka et al., 2004). Although two viruses, Sweet potato chlorotic stunt virus (SPCSV) (family *Closteroviridae*) (Kreuze et al., 2002) and Sweet potato feathery mottle virus (SPFMV) (family *Potyviridae*) (Cuellar et al., 2008; Clark et al., 2012) are known to cause SPVD, there are reports in SA suggesting the involvement of several other viruses as role-players in sweet potato disease complexes (Domola et al., 2008; Tesfaye et al., 2011). The last comprehensive survey in SA documenting viral prevalence was carried out in 2003 and employed serology (ELISA assays), reverse transcription PCR (RT-PCR), electron

* Corresponding author.
E-mail address: thulilenhlapo@gmail.com (T.F. Nhlapo).

https://doi.org/10.1016/j.sajb.2018.05.024
0254-6299/

microscopy and indexing methods for virus detection (Domola et al., 2008). Using these traditional methods, a previous survey detected 9 viruses including SPFMV, SPCSV, Sweet potato mild mottle ipomovirus (SPMMV) (family *Potyviridae*), Sweet potato chlorotic fleck virus (SPCFV) (family *Flexiviridae*), Sweet potato caulimo-like virus (SPCV) (family *Caulimoviridae*), Sweet potato virus G (SPVG) (family *Potyviridae*), Sweet potato virus 2 (SPV2) (family *Potyviridae*), Sweet potato latent (SPLV) (family *Potyviridae*), and Sweet potato mild speckling virus (SPMSV) (family *Potyviridae*) (Domola et al., 2008). More recently, two begomoviruses, Sweet potato leaf curl Sao Paulo virus (SPLCSPV) and Sweet potato mosaic virus (SPMaV) were detected for the first time in the Limpopo province of SA using rolling circle amplification (RCA), cloning and sequencing (Esterhuizen et al., 2012). These sweet potato viruses, previously classified as sweepoviruses have recently been classified under begomoviruses (Brown et al., 2015). Begomoviruses belong to the family *Geminiviridae* and genus *Begomovirus* and consist of circular ssDNA viruses that can be bipartite or monopartite (Fauquet et al., 2003). The Infection of sweet potato with begomoviruses has been reported in Japan, Israel, Peru, Italy, Spain, China, Taiwan, Korea, Kenya, United States of America, Puerto Rico, Costa Rico, South Africa and Brazil (Kreuze and Fuentes, 2008; Paprotka et al., 2010; Albuquerque et al., 2011, 2012; Clark et al., 2012; Esterhuizen et al., 2012).

In the last two decades, virus detection methods have shifted from traditional techniques to metagenomic approaches coupled with high throughput sequencing (Boonham et al., 2014). Viral metagenomics has been used to identify novel viruses in plants (Kreuze and Fuentes, 2008; Idris et al., 2014). This approach is considered an unbiased one for viral detection since no prior knowledge of the virus is necessary, and neither virus-specific primers, nor antibodies are required. Consequently novel viruses, if present, can be detected, identified and quantified in a single experiment (Studholme et al., 2011). However, in the absence of reference sequences, the use of high throughput sequencing for virus detection requires *de novo* genome assembly of new virus sequences, which can be a challenge. A metagenomic approach also means that the entire microbial community within a sample can be described, even in mixed viral interactions, thus simplifying diagnostics (Idris et al., 2014). In most cases, the virus sequences generated in a metagenomic study would form a small proportion of the total nucleic acids making the removal of host sequences critical prior to, or after sequencing (Stobbe and Roossinck, 2014). For this reason, enrichment methods such as isolation of dsRNA (Clark et al., 2012) and small interfering RNA (siRNAs) (Kreuze et al., 2009) have been employed to detect DNA and RNA viruses from different hosts (Kashif et al., 2012). The availability of NGS platforms such as those supplied by Illumina (Illumina Inc., San Diego, CA, USA) has further revolutionized viral metagenomics studies. NGS technologies generate large amounts of data rapidly at reduced costs and many bioinformatic tools have been developed to handle data analysis (Massart et al., 2014). This study was carried out with the objective of establishing the current status of sweet potato viruses in two South African provinces (Western Cape and Eastern Cape). NGS of a symptomatic sweet potato field sample revealed a mixed infection of six viruses [SPMaV, SPLCSPV, SPFMV, SPCSV, SPVG and Sweet potato virus C (SPVC)], for the first time in SA.

2. Materials and methods

2.1. Sources of plant material

Sweet potato cuttings were collected from four smallholder farms in the EC province of South Africa: Alice (32°47′13.6″S 26°50′56.8″E), Zwide (33°52′12.8″S 25°34′24.0″E), Kwazakhele (33°53′11.0″S 25°36′00.7″E), and Motherwell (33°48′08.3″S 25°35′47.0″E). In the WC province, material was collected from four commercial farms, the locations included Paarl (33°40′12.0”S 18°58′08.0″E), Klawer (31°46′59.0”S 18°37′00.0″E), Franschhoek (33°55′00.1”S 19°07′59.9″E) and Lutzville (31°33′11.0”S 18°12′57.0″E). The cuttings were transplanted to potting soil in 20 cm diameter pots and grown in a greenhouse at the Agricultural Research Council – Vegetable and Ornamental Plant Institute (ARC-VOPI) in Pretoria, SA (25°40′51.67″S 28°17′10.25″E). Plants were grown at optimum temperatures of 25 °C for 16 h (day cycle) and 15 °C for 8 h (night cycle) (Domola et al., 2008). Plants were watered once a day and soluble nutrient fertilisation (Multifeed P, Plaaskem, Pty, LTD) was applied on a weekly basis. Insect pests were also monitored and controlled by spraying with recommended insecticides as required. A list of samples, abbreviations and symptoms are depicted in Table 2.

2.2. Description of symptoms

Sweet potato plants collected from the field were maintained in the glasshouse and observed for symptom development over a period of 6 months. Plants exhibiting symptoms typical of viral infection, such as upward curling of the leaves, chlorotic spots, vein clearing, and purple ring spots, were selected for analysis by sequencing. Symptom severity was scored using a 1–5 scoring scale (Mwanga et al., 2001; Domola et al., 2008) where, 1 = no virus symptoms, 2 = mild symptoms (chlorotic spots), 3 = moderate symptoms (chlorosis, chlorotic spots and vein clearing), 4 = severe symptoms (chlorotic spots, leaf curl, and leaf puckering/necrosis) and 5 = very severe symptoms (chlorotic spots, leaf curl, mottling, and stunting).

2.3. RNA library preparation and sequencing

After a period of 6 months, 10 symptomatic and 7 asymptomatic plants were randomly selected for further analysis (Table 2). Total RNA was isolated from the 17 samples using the QIAGEN RNeasy Plant Mini Kit (QIAGEN, Valencia, CA, USA) following the manufacturer's instructions. To remove DNA, total RNA underwent DNase treatment using the QIAGEN RNase-free DNase Set (QIAGEN, Valencia, CA, USA). The integrity of the extracted RNA was analysed by agarose gel electrophoresis and quantified using the Qubit™ RNA BR Assay Kit (Invitrogen, Life Technologies). The total RNA was stored at −80 °C until further use. Total RNA was treated with the Ribo-Zero™ Magnetic Kit (Plant Leaf) (Epicentre, Madison, WI, USA) to deplete ribosomal RNA (rRNA). RNA paired-end libraries were prepared using the Illumina TruSeq™ Stranded Total RNA Sample Preparation Kit (Illumina Inc., San Diego, CA, USA) following the manufacturer's instructions. The libraries were quantified using the Qubit™ dsDNA HS Assay Kit (Invitrogen, Life Technologies). The libraries were sequenced on the Illumina MiSeq (Illumina Inc., San Diego, CA, USA).

2.4. Nucleic acid isolation, rolling circle amplification (RCA), library preparation and sequencing

Genomic DNA (gDNA) was isolated from plant leaf material using the QIAGEN DNeasy Plant Mini Kit (*QIAGEN* Inc., Valencia, CA, USA) following manufacturer's instructions. The integrity of the extracted DNA was visualised by electrophoresis and quantified using the Qubit™ dsDNA BR Assay Kit (Life Technologies, Thermo Fisher Scientific Inc.). The gDNA was stored at −80 °C until further use. Prior to sequencing, genomic DNA underwent rolling circle amplification (RCA) using the Illustra™ TempliPhi™ 100 Amplification Kit (GE Healthcare, Amersham, UK) following the manufacturer's instructions. The RCA products were subjected to the Nextera DNA Sample Preparation Kit (Illumina Inc., San Diego, CA, USA) to generate 17 RCA libraries. Samples were sequenced using the Illumina MiSeq (Illumina Inc., San Diego, CA, USA) platform.