High-Throughput Sequencing of Virus-infected Cucurbita pepo Samples Revealed The Presence of Zucchini Shoestring Virus in Zimbabwe

Charles Karavina
Bindura University of Science Education Faculty of Agriculture and Environmental Science
Jacques Davy Ibaba (janick.ibaba@gmail.com)
University of KwaZulu-Natal https://orcid.org/0000-0002-6935-112X
Augustine Gubba
University of KwaZulu-Natal School of Agricultural Earth and Environmental Sciences

Research note

Keywords: Keywords Next generation sequencing, Potyvirus, Plant virus, Cucurbit, Zimbabwe, Zucchini shoestring virus, high-throughput sequencing

Posted Date: October 8th, 2019

DOI: https://doi.org/10.21203/rs.2.15833/v1

Version of Record: A version of this preprint was published on February 3rd, 2020. See the published version at https://doi.org/10.1186/s13104-020-4927-3.

Abstract

Objectives: Plant-infecting viruses remain a serious challenge towards achieving food security worldwide. Cucurbits, in Zimbabwe, like in the other parts of the world, are used in various ways. A small-scaled cucurbit virus survey was conducted in Zimbabwe during the 2014 and 2015 growing seasons. Cucurbit leaf samples displaying virus-like symptoms were collected and stored until analysis. The samples were then subjected to next-generation sequencing (NGS). The data generated from NGS were analysed using genomics technologies. Zucchini shoestring virus (ZSSV), a cucurbit-infecting potyvirus previously described in South Africa was one of the viruses identified. The genomes of three ZSSV isolates from Zimbabwe are described in this note. Results: The three ZSSV isolates had the same genome size of 10297 bp excluding the polyA tail with a 43% GC content. The large open reading frame (ORF) was found at positions 69 to 10106 on the genome and encodes a 3345 amino acids long polyprotein which had the same cleavage site sequences as those described on the South African isolates except for the P1-pro site. The smaller ORF, also called the pretty interesting Potyviridae ORF, was located at positions 3611 to 3793 on the genomes for all three ZSSV isolates.

Introduction

Cucurbits are among the most economically important vegetables worldwide. In Zimbabwe, they are widely grown by both commercial and smallholder farmers as food and cash crops. Virus diseases on cucurbits produce diverse symptoms that results in yield reduction and in severe instances compromised fruit quality. The negative effects of plant-infecting viruses on agricultural crops are more prominent especially in countries where their studies are limited.

Zucchini shoestring virus (ZSSV) is a member of the genus Potyvirus in the Family Potyviridae [1, 2]. Viruses that constitute the genus Potyvirus are non-enveloped, flexuous and filamentous virions with a length varying between 680–900 nm long and 11–20 nm in diameter. The potyvirus genome, normally enclosed in the capsid, consists of a monopartite, single-stranded, positive-sense RNA. The 5’ terminus of the genomic RNA forms covalent bonds with the viral protein genome-linked (VPg) and a polyadenylic acid (PolyA) tail makes up the 3’ terminus. Eleven multifunctional proteins are produced from the potyvirus genome through polyprotein expression and polymerase slippage mechanisms [3].

The first confirmed report of ZSSV dates back to 2015 from the studies on viruses infecting cucurbits in South Africa in which ZSSV was identified on *Cucurbita pepo* plants, locally known as baby marrow, displaying severe foliar symptoms and malformed fruits [4]. Molecular studies of the ZSSV capsid have shown that ZSSV is of the “Papaya ringspot virus cluster” of cucurbit-infecting potyviruses [2].

Methods

Sample sources and high-throughput sequencing

Three baby marrow (C*ucurbita pepo)* leaf samples displaying symptoms of viral aetiology were collected between 2014 and 2015 at three different cucurbit-growing farms, all located in Harare, Zimbabwe. Total RNA was extracted from each sample using the Quick-RNA Miniprep Kit (Zymo Research, USA) as per the manufacturer's instructions and shipped on dry ice to the Agricultural Research Council Biotechnology Platform (ARC-BTP) in Pretoria, South Africa, to be used for high-throughput sequencing on the Illumina HiSeq platform using the 2 x 125 bp paired-end reads. For each sample, the data generated from the sequencing was analysed as follows. Read quality control checks were assessed using FastQC version 0.11.5 (Babraham Bioinformatics). Trimmomatic version 0.36 [5] was used to trim the reads when necessary. *De novo* assembly was performed with SPAdes version 3.10.1[6] according to the developer's instructions. Nucleotide blast was performed on all contigs using BLAST+ [7].

Results And Discussion

A contig from each sample matched the full-length genome sequence of the ZSSV isolates with 30x, 66x and 80x coverage (Table 1). The genome size was the same for the three isolates under study and consisted of 10297 bp excluding the polyA tail with GC contents varying between 42.92 and 42.96%. The large open reading frame (ORF) was found at positions 69 to 10106 on the genome using the ORFfinder web version (https://www.ncbi.nlm.nih.gov/orffinder/). The 3345 amino acids long polyprotein resulting from the direct translation of the large ORF had the same cleavage site sequences as those described on the South African isolates [2] except for the protein 1 protease (P1-pro) site. The motifs, RITC and PTR, variants of the highly conserved KITC and PTK were also part of the polyprotein. The smaller ORF, also called the pretty interesting *Potyviridae* ORF, was located at positions 3611 to 3793 on the genomes for all three ZSSV isolates. Sequence identity was evaluated between the ZSSV isolates from Zimbabwe and South Africa using SIAS, an online tool, (http://imed.med.ucm.es/Tools/sias.html). The nucleotide sequence identity was 91–92% across the entire genome and the polyprotein ORF. Sequence comparisons of the polyprotein at the amino acid level showed a 96% sequence identity over the entire polyprotein. Looking at the mature peptide products of the polyprotein, the sequence identities were 87–90% for the P1-pro, 94% for the six kilodalton peptide (6k) 1, 95% for the protein 3, 97% for the nuclear inclusion A protease and the capsid protein, 98% for the VPg, the nuclear inclusion B RNA-dependent RNA polymerase and the 6k2, 99% for the helper-component proteinase and the cytoplasmic inclusion. The detection of ZSSV in cultivated baby marrow plants from the surveyed farms may indicate either a broader geographical distribution of the virus or its spreading across borders. The occurrence of ZZSV in Zimbabwe highlights the need to conduct further studies on its epidemiology and to develop effective management strategies.

Table 1: Overview of the ZSSV genome sequences of the isolates found in the samples

Isolate names

Data repository and identifier (accession number)

Zucchini shoestring virus isolate F7-Art

DDBJ/ENA/GenBank (MK204479.1)

Zucchini shoestring virus isolate S6-Prime

DDBJ/ENA/GenBank (MK204480.1)

Zucchini shoestring virus isolate S7-Prime

DDBJ/ENA/GenBank (MK204481.1)

Limitations

The small sample size in that survey was the main limitation.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data material

The data described in this note can be freely and openly accessed on DDBJ/ENA/GenBank. Please see Table 1 for details and links to the data.

Competing interests

The authors declare that they have no competing interests.

Funding

The research was funded by the W. K. Kellogg Foundation Southern African Scholarship as part of Charles Karavina's PhD studies at the University of KwaZulu-Natal. The funding body had no role in the experimental design, collection, analysis and interpretation of data or in writing the manuscript.

Authors' contributions

Charles Karavina collected samples and performed the RNA extractions. Jacques Davy Ibaba did the NGS data analysis and submission into the appropriate repository. Augustine Gubba advised on the study design. All the authors read and approved the final manuscript.

Acknowledgements

The authors will like to acknowledge the NGS technical team from the ARC-BTP for providing good services. Charles Karavina expresses his heartfelt gratitude to the farmers who willingly let him collect samples from their farms.

References

1. Adams MJ, Lefkowitz EJ, King AMQ, Harrach B, Harrison RL, Knowles NJ, et al. Changes to taxonomy and the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2017). Arch Virol. 2017;162:2505–38. doi:10.1007/s00705-017-3358-5. 2. Ibaba JD, Laing MD, Gubba A. Zucchini shoestring virus: a distinct potyvirus in the papaya ringspot virus cluster. Arch Virol. 2016;161:2321–3. doi:10.1007/s00705-016-2899-3. 3. Wylie SJ, Adams M, Chalam C, Kreuze J, López-Moya JJ, Ohshima K, et al. ICTV Virus Taxonomy Profile: Potyviridae. J Gen Virol. 2017;98:352–4. doi:10.1099/jgv.0.000740. 4. Ibaba JD, Laing MD, Gubba A. First Report of a Novel Potyvirus from the Papaya ringspot virus Cluster Infecting Zucchini (Cucurbita pepo) in KwaZulu-Natal, Republic of South Africa. Plant Dis. 2015;99:1289. doi:10.1094/PDIS-02-15-0143-PDN. 5. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20. doi:10.1093/bioinformatics/btu170. 6. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20. doi:10.1093/bioinformatics/btu170. 7. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST: architecture and applications. BMC Bioinformatics. 2009;10:421. doi:10.1186/1471-2105-10-421.

Table 1

Table 1: Overview of the ZSSV genome sequences of the isolates found in the samples

Isolate names	Data repository and identifier (accession number)
Zucchini shoestring virus isolate F7-Art	DDBJ/ENA/GenBank (MK204479.1)
Zucchini shoestring virus isolate S6-Prime	DDBJ/ENA/GenBank (MK204480.1)
Zucchini shoestring virus isolate S7-Prime	DDBJ/ENA/GenBank (MK204481.1)