Virus Genes (2019) 55:429–432
https://doi.org/10.1007/s11262-019-01650-6

Diversity and distribution of *Maize-associated totivirus* strains from Tanzania

David Alan Read[1] · Jonathan Featherston[1] · David Jasper Gilbert Rees[1] · Genevieve Dawn Thompson[1] · Ronel Roberts[2] · Bradley Charles Flett[3] · Kingstone Mashingaidze[3] · Gerhard Pietersen[4] · Barnabas Kiula[5] · Alois Kullaya[6] · Ernest R. Mbega[7]

Received: 30 January 2019 / Accepted: 13 February 2019 / Published online: 21 February 2019

Abstract

Typically associated with fungal species, members of the viral family *Totiviridae* have recently been shown to be associated with plants, including important crop species, such as *Carica papaya* (papaya) and *Zea mays* (maize). *Maize-associated totivirus* (MATV) was first described in China and more recently in Ecuador, where it has been found to co-occur with other viruses known to elicit maize lethal necrosis disease (MLND). In a survey for maize-associated viruses, 35 samples were selected for Illumina HiSeq sequencing, from the Tanzanian maize producing regions of Mara, Arusha, Manyara, Kilimanjaro, Morogoro and Pwani. Libraries were prepared using an RNA-tag-seq methodology. Taxonomic classification of the resulting datasets showed that 6 of the 35 samples from the regions of Arusha, Kilimanjaro, Morogoro and Mara, contained reads that were assigned to MATV reference sequences. This was confirmed with PCR and Sanger sequencing. Read assembly of the six MATV-associated datasets yielded partial MATV genomes, two of which were selected for further characterization, using RACE. This yielded two full-length MATV genomes, one of which is divergent from other available MATV genomes.

Keywords *Totivirus* · Maize · Tanzania · Maize lethal necrosis disease · Illumina sequencing

Introduction

Members of the viral family *Totiviridae* have been shown to infect a wide diversity of fungal genera and have been implicated in the "killer" phenotype of certain yeast strains [1]. They have also been hypothesized to modulate virulence in parasites, such as *Leishmania, Trichomonas* and *Giardia* [2] and shown to infect arthropod hosts, including mosquitos [3], ants [4] and shrimp [5]. More recently, members of the family *Totiviridae* have been detected in a number of plant species including *Carica papaya* (papaya) [6], *Panax* sp [7]. and *Zea mays* (maize) [8, 9]. Given the opportunities for transfer during fungal colonization, it is possible that *Totiviruses* were initially transmitted to maize via a fungal host species [10]. *Totivirus* genomes are double-stranded RNA molecules, typically non-segmented containing two open reading frames (ORF), with the 5′ proximal ORF encoding for a coat protein (CP) and the 3′ proximal ORF encoding for an RNA-dependant-RNA-polymerase (RdRp). Genome sizes vary in length from − 4000 to 8500 bp, based on publicly available sequences in GenBank. *Maize-associated totivirus*

Edited by Dr. Karel Petrzik.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11262-019-01650-6) contains supplementary material, which is available to authorized users.

✉ David Alan Read
readd@arc.agric.za

[1] Agricultural Research Council - Biotechnology Platform, Private Bag X05, Onderstepoort 0110, South Africa

[2] Agricultural Research Council - Plant Protection Research, Private Bag X134, Queenswood 0001, South Africa

[3] Agricultural Research Council - Grain Crops, Private Bag X1251, Potchefstroom 2520, South Africa

[4] Department of Genetics, Stellenbosch University, Stellenbosch 7600, South Africa

[5] Tanzania Agricultural Research Institute (TARI, Ilonga), Kilosa, Morogoro, Tanzania

[6] Tanzania Agricultural Research Institute (TARI, Mikocheni), Dar es Salaam, Tanzania

[7] Nelson Mandela African Institution of Science and Technology (NM-AIST), Arusha, Tanzania

(MATV) was first detected in China [8] and subsequently in Ecuador [9].

Maize leaf material was collected from 35 plants as part of a survey for maize-associated viruses in Tanzania [11]. Total RNA isolations were rRNA depleted, using a Ribo-Zero rRNA Removal Kit—Plants (Illumina, San Diego, CA, USA). Tagged libraries were prepared according to the RNA-tag-seq protocol of Shishkin et al. [12] and were sequenced using an Illumina HiSeq2500 instrument (paired-end; 2 × 125 nt) (Ilumina, San Diego, CA, USA) by the Agricultural Research Council—Biotechnology Platform (ARC-BTP).

Quality and adapter trimming was done using CLC Genomics Workbench 9 (Qiagen Bioinformatics, Aarhus, Denmark). Initial taxonomic identification of trimmed unassembled reads was performed using the browser-based version of Kaiju software package [13]. Six of the 35 NGS datasets suggested the presence of MATV, with between 0.001 and 20.8% of virus-associated reads mapping to *Totivirus* references. CLC Genomics Workbench 9 was used to assemble contigs, with the de novo assembly function, using the following adjustable parameters: mismatch cost: 2, insertion cost: 3, deletion cost: 3, length fraction: 0.5, similarity fraction: 0.7, global alignment: off, update contigs: on, minimum contig length: 1000 base pairs (bp). The identities of the contigs were determined using BLASTx [14] and contigs showing homology to reference sequences within the *Totiviridae* family were analysed further.

Seven partial MATV genomes were generated and were assigned GenBank Accession Numbers MK037419 – MK037425. Two complete/near complete genomes were generated from a single sample (16–0130). These genomes represent two distinct MATV strains, one of which shares a 99% nucleotide sequence homology to MATV-Ec from Ecuador (KT722800) [9] and is 5006 bp in length. The genome of this isolate has been named Maize-associated totivirus-1-Tanz (MATV-1-Tanz), with GenBank accession number MK066242. The other strain represents a divergent MATV strain, with a sequence length of 5583 bp and has been named Maize-associated totivirus-4-Tanz (MATV-4-Tanz) and was assigned GenBank Accession Number MK066243.

The terminal nucleotides of the two near complete genomes derived from sample 16–0130 were confirmed using RACE-PCR. RACE-PCR products were cleaned-up using a Nucleospin® Gel and PCR clean-up kit (Macherey–Nagel, Düren, Germany) and submitted to the ARC-BTP Core facility (Agricultural Research Council—Biotechnology Platform, Pretoria, South Africa) for Illumina MiSeq sequencing. RACE contigs were generated using the de novo assembly tool within the CLC Genomics Workbench 9 suite. Whole-genome contigs of MATV-1-Tanz and MATV-4-Tanz were generated, using the Contig Assembly Program (CAP) assembly program" within BioEdit 7.1.3.0 [15], using the partial genome contigs and RACE-PCR contigs as input sequences.

The complete genomes generated from this study represent two divergent MATV strains that were obtained from the same plant. MATV-1-Tanz (MK066242) has a total genome length of 5006 bp with two non-overlapping ORFs (separated by 27 bp). ORF1 (115–2409 bp) encodes for a putative capsid protein of 764 amino acids (aa). ORF2 encodes for a putative RdRp of 834aa. The amino acid sequences of the putative coat-protein and RdRp gene products share a > 99.9% aa homology with the MATV 2 isolate EC_Portoviejo strain (KT722800) from Ecuador. MATV-4-Tanz (MK066243) represents a divergent MATV strain with a total genome length of 5583 bp with ORF1 and ORF2, sharing an overlap of 36 bp. The aa sequence of the putatively expressed coat protein of ORF1 shares a 35% aa identity with the homologous protein product of Panax notoginseng virus B (MF614102). The RdRp amino acid sequence shares a > 99% aa similarity with the cognate gene product of MATV 1 from South Sudan (MF372915).

Coat protein sequences derived from the contigs generated in this study, as well as references, were aligned using the "create alignment" function in CLC Genomics Workbench 9. MEGA 6.06 [16] was used for the determination of a best-fit model and the subsequent generation of a maximum-likelihood phylogeny (Fig. 1), using the LG matrix model and 1000 bootstrap replicates.

Separate RT-PCR assays were designed to target the two MATV strains from this study. MATV1 was targeted using the MATV1F/R primer pair (5′CTACCTCCGATGCACAATGAGTTC3′/5′ GGATAGAGTGCGCTTGACGATG3′) and MATV4 was targeting using the MATV4F/R primer pair (5′ATCGTAGTGTGTCGTTCACAGG3′/5′CAACATTAGATCGTCTGCCGACG3′). All primers have a Tm of 60 °C. Positive reactions for each primer pair/sample combination are listed in Supplementary Table 1. PCR products derived from samples 16–0111 and 16–0130 were submitted for bidirectional Sanger sequencing (Inqaba Biotechnical Industries, Pretoria, South Africa). The identity of each sequence was confirmed through alignment against the MK066242 and MK066243 genome sequences.

This is the first report of MATV in Tanzania and the first formal report of the presence of the virus in Africa. The occurrence of MATV sequences on GenBank originating from Kenya, Ethiopia and South Sudan, however, suggests that the virus is widespread in East Africa. Prior to this report, the virus was detected in geographically disparate countries of China [8] and Ecuador [9], suggesting that the virus may have a worldwide distribution. The contribution of MATV to symptom expression in plants still requires further research. Redinbaugh [17] recently suggested that MATV is an MLND-associated virus, along with Maize yellow mosaic

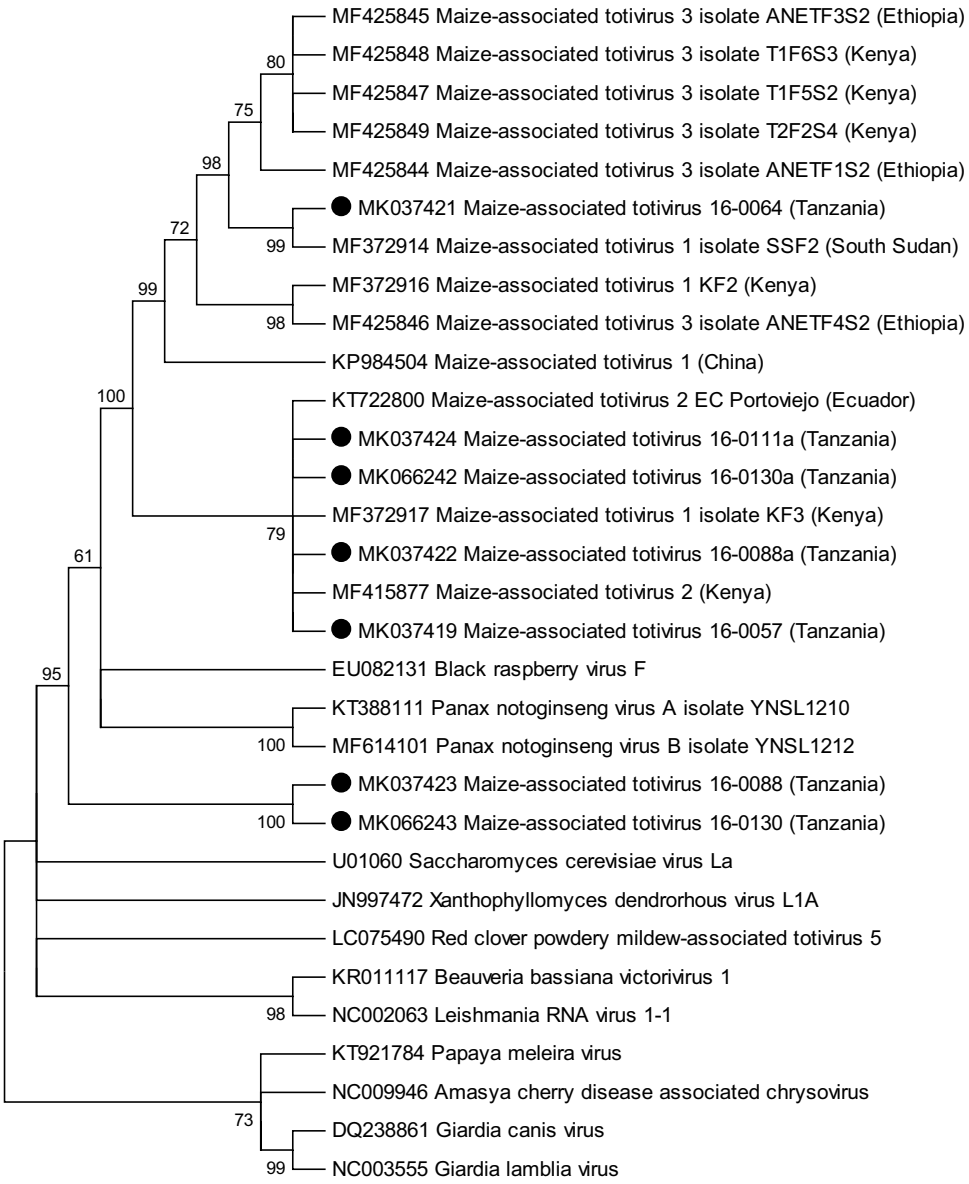

Fig. 1 Maximum likelihood condensed tree with a ≥ 60% cutoff value. The tree was generated using the complete coat protein amino acid sequences derived either whole or partial MATV genomes from Tanzania (this study), as well as MATV and other members of the Totiviridae family, available on GenBank. Solid circles represent the sequences associated with this study. The phylogram was produced in MEGA 6.06 using the best-fit LG matrix model, with 1000 bootstrap replicates

virus (MaYMV). This trend was observed among the samples from this study, where Maize chlorotic mottle virus (MCMV), potyvirus and MaYMV associated-reads formed part of each viral population. Future research will require determining the ability of MATV strains to elicit symptoms as single viral isolates, its potential to act synergistically with other MLND-associated viruses and the mechanisms involved in its dispersal.

Author Contributions DJGR, BCF and KM contributed to the study design. DAR, JF, GDT, RR, GP, BK, AK and ERM contributed to various aspects of sample collection. DAR, GDT, RR and GP wrote the manuscript.

Funding This work received no specific Grant from any funding agency.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical standards The manuscript constitutes original research and has not been previously published. All authors have agreed to the submission of the manuscript, which complies with the ethical standards of the journal.

References

1. Ghabrial SA, Castón JR, Jiang D, Nibert ML, Suzuki N (2015) 50-plus years of fungal viruses. Virology 479:356–368
2. Gomez-Arreaza A, Haenni A-L, Dunia I, Avilán L (2017) Viruses of parasites as actors in the parasite-host relationship: a "ménage à trois". Acta Trop 166:123–132

3. de Lara Pinto AZ, de Carvalho MS, de Melo FL, Ribeiro ALM, Ribeiro BM, Slhessarenko RD (2017) Novel viruses in salivary glands of mosquitoes from sylvatic Cerrado, Midwestern Brazil. PLoS ONE 12(11):e0187429
4. Koyama S, Urayama S, Ohmatsu T, Sassa Y, Sakai C, Takata M et al (2015) Identification, characterization, and full-length sequence analysis of a novel double-stranded RNA virus isolated from an arboreal ant, *Camponotus yamaokai*. J Gen Virol 96:1930–1937
5. Poulos BT, Tang KFJ, Pantoja CR, Bonami JR, Lightner DV (2006) Purification and characterization of infectious myonecrosis virus of penaeid shrimp. J Gen Virol 87:987–996
6. Antunes TFS, Amarel RJV, Ventura JA, Godinho MT, Amaral JG, Souza FO et al (2016) The dsRNA virus papaya meleira virus and an ssRNA virus are associated with papaya sticky disease. PLoS ONE 11(5):e0155240
7. Guo L, Yang X, Wu W, Tan G, Fang S, Zhang S (2016) Identification and molecular characterization of *Panax notoginseng virus* A, which may represent an undescribed novel species of the genus *Totivirus*, family *Totiviridae*. Arch Virol 161:731–734
8. Chen S, Cao L, Huang Q, Qian Y, Zhou X (2016) The complete genome sequence of a novel maize-associated totivirus. Arch Virol 161:487–490
9. Alvarez-Quinto RA, Espinosa-Lozano RF, Mora-Pinargote CA, Quito-Avila DF (2017) Complete genome sequence of a variant of maize-associated totivirus from Ecuador. Arch Virol 162:1083–1087
10. Roosinck MJ (2018) Evolutionary and ecological links between plant and fungal viruses. New Phytol. https://doi.org/10.1111/nph.15364
11. Read DA, Featherston J, Rees DJG, Thompson G, Roberts R, Flett BC et al (2018) First report of maize yellow mosaic virus (MaYMV) on maize (*Zea mays*) in Tanzania. J Plant Pathol. https://doi.org/10.1007/s42161-018-0152-5
12. Shishkin AA, Giannoukos G, Kucukural A, Ciulla D, Busby M, Surka C et al (2015) Simultaneous generation of many RNA-seq libraries in a single reaction. Nat Methods 12:323–325
13. Menzel P, Ng KL, Krogh A (2016) Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat Commun 7:11257
14. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410
15. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98
16. Tamura K, Stecher G, Petersen D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30(12):2725–2729
17. Redinbaugh MG, Stewart LR (2018) Maize lethal necrosis: an emerging, synergistic viral disease. Annu Rev Virol 5:301–322