RUFORUM Working Document Series (ISSN 1607-9345) No. 14 (1): 653-660.
Available from http://repository.ruforum.org

Research Application Summary

Application of next generation sequencing in identification of cowpea viruses

Mbeyagala, K.E.[1,2,3], Mukasa, S.B.[2], Holton, T.[4], Aritua, V.[5], Manpreet, S.[6], Katari, P.[6], Tukamuhabwa, P.[2] & Bisikwa, J.[2]
[1]National Agricultural Research Organization (NARO), P. O. Box 295 Entebbe, Uganda
[2]School of Agricultural Sciences, Makerere University, P. O. Box 7062 Kampala, Uganda
[3]National Semi Arid Resources Research Institute (NaSARRI), Uganda
[4]Formerly at Biosciences eastern and central Africa (BecA-ILRI Hub) P. O. Box 30709, Nairobi 00100, Kenya
[5]CIAT- Kawanda, P. O. Box 6247 Kampala,, Uganda
[6]New York University, New York, N.Y. 10003, USA
Corresponding author: kmbeyagala@gmail.com

Abstract

Cowpea (*Vigna unguiculata* L. Walp) is an important legume particularly in the eastern and northern regions of Uganda. Its production is greatly affected by the occurrence of several diseases caused mainly by viral pathogens. Previous studies employed in the detection of cowpea viruses in the country relied largely on specific methods thus offering a partial description of the disease etiology. In this study, virus symptomatic leaf samples were assayed for the presence of viruses using a high throughput sequencing method. Leaf samples were collected from four agro-ecological zones (AEZs) in Uganda and used for total RNA extraction. From the total RNA, cDNA was synthesized and then sequenced using the Illumina Miseq platform. BLASTN analysis revealed the presence of contig sequences with >80% identity to *Cowpea aphid borne mosaic virus* (CABMV), *Peanut mottle virus* (PMV), *Crotalaria mosaic virus, South African passiflora* and *Sesame mosaic virus.* Analysis also revealed sequences that showed high similarity to *Maize dwarf mosaic virus* (MDMV), *Sugarcane mosaic virus* (SCMV) and *Maize chlorotic mottle virus* (MCMV). This is the first report of cowpea serving as a legume host for MCMV, SCMV and MDMV and given the increasing legume-cereal intensification in the country, this presents a risk of virus spread to maize.

Key words: Blastn, contig, cowpea, ELISA, maize, sequencing, Virome

Résumé

La dolique (*Vigna unguiculata* L. Walp) est une légumineuse importante en particulier dans les régions de l'est et du nord de l'Ouganda. Sa production est grandement affectée par l'apparition de plusieurs maladies causées principalement par des agents pathogènes viraux. Des études antérieures utilisées dans la détection des virus qui attaquent la dolique

dans le pays comptaient largement sur des méthodes spécifiques offrant ainsi une description partielle de l'étiologie de la maladie. Dans cette étude, le virus des échantillons de feuilles symptomatiques ont été analysés pour déterminer la présence de virus à l'aide d'un procédé de séquençage à haut débit. Des échantillons de feuilles ont été prélevés dans quatre zones agro-écologiques (ZAE) en Ouganda et utilisés pour l'extraction de l'ARN total. De l'ARN total, l'ADNc a été synthétisé et séquencé en utilisant la plate-forme Illumina Miseq. L'analyse BLASTN a révélé la présence de séquences contig avec > 80% d'identité à au *virus de la mosaïque causé par le puceron attaquant la Dolique* (VMPD), *le virus de la marbrure de l'arachide* (VMA), *le virus de la mosaïque du Crotalaria* (VMC), *passiflore sud-africain* et *le virus de la mosaïque du sésame*. L'analyse a également révélé des séquences qui ont montré une similarité élevée au *virus de la mosaïque du maïs nain* (VMMN), *le virus de la mosaïque de la canne à sucre* (VMCS) *et le virus de la marbrure chlorotique du maïs* (VMCM). Cette étude est la première à indiquer la dolique servant en tant qu'une légumineuse hôte pour VMCM, VMCS et VMMN et compte tenu l'intensification croissante de la culture mixte légumineuses-céréales dans le pays, ceci présente un risque de propagation du virus au maïs.

Mots clés: Blastn, contig, le dolique, ELISA, le maïs, le séquençage, virome

Background

Cowpea (*Vigna unguiculata* L. Walp) is an important food legume in the drier regions of the tropics especially in sub-Saharan Africa. In Uganda, it is one of the important food and cash crops especially among the small-scale farmers in drought prone areas of eastern and northern regions. It helps to bridge the "hunger period" prior to the main cereal harvest due to its short maturity period, fixes atmospheric nitrogen, and its decaying residues improve soil fertility and is a rich source of cheap dietary proteins (Casky *et al.*, 2002). Despite its importance, its production is greatly affected by occurrence of viral diseases (Edema *et al.*, 1997; Orawu *et al.*, 2005) which have negatively affected production and food security among farming households (Orawu *et al.*, 2015). Infection of cowpea by viral diseases has been reported to reduce the nutritional value (especially proteins) of the grain as well as significant reductions in yield per unit area (Raheja and Leleji, 1974; Taiwo and Akinjogunla, 2006; Kareem and Taiwo, 2007). Cowpea plants infested with viruses manifest a range of symptoms such as mosaics, mottling, interveinal chlorosis, distortion, leaf curling, necrosis and leaf vein banding (Bashir *et al.*, 2002; Aliyu *et al.*, 2012).

While several studies have been carried out to detect viruses on cowpea in the country, for instance by Edema *et al.* (1997), Orawu *et al.* (2005), Amayo *et al.* (2012), Orawu *et al.* (2015), these studies relied on very specific methods such as ELISA. Such methods however suffer from several significant drawbacks especially when trying to identify 'unknown' agents, e.g. either a pathogen infecting a new host or a previously uncharacterized pathogen. For instance, specific assays require reliable knowledge of which pathogens infect certain hosts for prediction of the infective agent. However, information is often absent, especially for many of the more obscure or novel hosts. Second, specific tests, sometimes utilize reagents such as antibodies with a finite specificity covering a particular strain, individual

species or small group of pathogens. These reagents are often incapable of detecting variants, such as new strains, that can arise unexpectedly. As a result, these can evade detection and spread rapidly (Adams *et al*., 2009, 2012).

Accurate plant virus diagnoses need novel and unbiased methods, which require no a *priori* knowledge of the host or pathogen. Next-generation sequencing (NGS) is one such method which allows pathogen sequences to be generated in a non-specific fashion and identification is based on similarity searching against GenBank sequences (Adams *et al*., 2005; Adams *et al*., 2012). The NGS, thus makes possible to obtain a more complete view of the viral population in a sample facilitating rapid identification of disease causing agents and providing sequences to which rapid diagnostics can be developed (Adams *et al*., 2012; DeBoever *et al*., 2013). Given the diverse agro-ecological zones in which cowpea is grown in the country (Wortmann and Eledu, 1999), NGS is a reliable method to accurately study the cowpea virome in the country. The objective of this study was to determine the cowpea virome in Uganda using next generation sequencing (NGS) in order to facilitate development of better strategies for management of cowpea virus diseases.

Study description

Collection of samples. Fresh leaf samples from symptomatic cowpea plants were collected from farmers' fields in eastern and northern districts of the Uganda namely; Ngora, Kumi, Pallisa, Namutumba, Serere, Dokolo,Lira, Apac, Soroti, and Amuria. These districts are located in diverse agro-ecological zones (AEZs) namely; Southern and Eastern Lake Kyoga Basin [SELKB] AEZ, Northern Moist Farmlands [NMF] AEZ and Northern Central Grass Bush Farmlands [NECGBF] AEZ (Wortmann and Eledu, 1999). Additional samples were obtained from experimental plots at Makerere University Agricultural Research Institute at Kabanyolo in Wakiso district in the Lake Victoria crescent [LVIC] AEZ. Leaf samples were immediately wrapped in paper towels and dried over silica gel (about 50g) in Ziploc bags. The Ziploc bags containing samples were later placed in a plastic container for 24 hours after which silica gel was changed twice to ensure complete drying of samples. The dried leaf samples were used for total RNA extraction.

Total RNA extraction. Total RNA was extracted from dry leaf samples using TRI Reagent method (Sigma Cat #T9424) following the manufacturer's guidelines with modifications. The quality and concentration of extracted RNA was checked using Nano Drop spectrophotometer and later visualized using TAE/Formamide agarose gel electrophoresis as described by Masek *et al*. (2005) with modifications. Samples showing intact 18S and 25S subunits were selected, diluted and pooled per agro-ecological zone giving a total of four RNA libraries (i.e. library1= LVIC, library 2= NMF, library 3= NCGBF, library 4 = SELKB).

cDNA sequencing. The cDNA libraries for sequencing were constructed from 1µg of RNA per library according to illumina Truseq RNA library preparation protocol with modifications. Quality of synthesized cDNA was checked using 2100 Bio-analyzer Desktop System (Agilent, part # G2940CA) and quantified using Qubit 2.0 Flourometer ((Invitrogen,

Carlsbad CA, USA). Paired – end sequencing (2 x 300 bp) was performed using MiSeq Desktop Sequencer platform at BecA-ILRI hub, Nairobi Kenya.

Analysis of sequence data. Raw paired-end sequences for each library were subjected to a quality control (QC) analysis, trimming and *de novo* contig assembly using CLC Genomics workbench version 7.0.4 (CLC bio, Qiagen). Based on the raw QC analysis report, length trimming of sequences was done by removing terminal nucleotides (20 nucleotides at the 5' end and 100 nucleotides at the 3' end). Only trimmed paired sequences were assembled into contiguous sequences (contigs) while the unpaired or orphan reads were discarded. Assembled contigs were subjected to BLASTN and BLASTX homology searches against a plant viral database of 57,417 sequences downloaded from National Center oor Biotechnology Information (NCBI) using NCBI-blast-2.2.29+ (Altschul *et al.*, 1997).

Results

A total of 18,318,168 raw reads with an average fragment length of 178 nucleotides (nt) was obtained from the four libraries (LVIC =library1, NMF = library 2, NECGBF = library 3, SELKB= library 4). After trimming of terminal nucleotides (at the 5' end and at the 3'end), a total of 13,781,621 reads remained with an average length of 78 nt. Assembly of trimmed paired read gave 34,851,14,580, 16,257 and 10,174 contigs for libraries 1,2,3 and 4, respectively. BLASTN and BLASTX analyses against the viral database gave similar results (except for library 3 (NECGBF) which did not give any significant hits) thus results that follow are from BLASTN analysis (Table1). The sequences obtained in this study were deposited in the NCBI genbank and assigned accession numbers. Analysis of library 1(LVIC) showed contig sequences with significant sequence similarity (>80%) over the entire length of the contigs (evalue = 0) to different isolates of *Maize dwarf mosaic virus* (MDMV) and *Sugarcane mosaic viruses* (SCMV). In library2 (NMF), contig sequences showed highly significant identity (>96%) over the entire length of the contigs (evalue = 0) to different isolates of *Maize chlorotic mottle virus*. Contig sequences in library 4 (SELKB) showed sequences with significant similarity (>80%) over the entire length of the contigs (evalue = 0) to *Cowpea aphid borne mosaic virus, Peanut mottle virus, Crotalaria mosaic virus* and *South African passiflora virus*. There were differences in virus hits detected in each of the libraries from the different locations (Table1) indicating geographical variation in distribution of cowpea viruses in Uganda.

Discussion

Cowpea production is constrained by an array of viruses wherever it is grown. In the present study, NGS showed that the main viruses infecting the crop are potyviruses; *Cowpea aphid borne mosaic virus* (CABMV) and *Peanut mottle virus* (PMV). While CABMV has been detected on cowpea in Uganda in previous studies (Orawu *et al.*, 2005, Orawu *et al.*, 2015), this is the first time that PMV has been reported on cowpea. Previous studies only show that PMV can infect groundnut and soybean (Bock, 1973) in east Africa with cowpea as a potential host (Demski *et al.*, 1983). Sequences showing identity to *Crotalaria mosaic*

Table 1. BLASTN results of selected hits for sequences from cowpea libraries (e-value=0)

Library	Accession #	Blast hit	Subject	Identity (%)	Alignment length (bp)
LVIC	KT970623	Polyprotein	[a]DQ647657.1	80.46	1341
LVIC	KT970623	Polyprotein	[a]DQ647660.1	80.24	1341
LVIC	KT970624	Coatprotein	[a]AJ310105.1	88.54	939
LVIC	KT970624	Coatprotein	[a]JX188385.1	93.62	939
LVIC	KT970624	Coatprotein	[b]D00949.1	92.91	931
LVIC	KT970624	Coatprotein	[b]S77088.1	87.27	927
LVIC	KT970624	Coatprotein	[b]AY660663.1	88.01	926
NMF	KT970620	Replicase	[c]JX286709.1	100	1009
NMF	KT970620	Replicase	[c]JQ982469.1	99.21	1009
NMF	KT970618	Replicase	[c]JX286709.1	99.58	713
NMF	KT970618	Replicase	[c]JQ982470.1	99.3	711
NMF	KT970621	Coatprotein	[c]JX286709.1	99.83	597
NMF	KT970621	Coatprotein	[c]JQ982469.1	99.33	597
NMF	KT970622	Coatprotein	[c]JX286709.1	99.59	491
NMF	KT970622	Coatprotein	[c]JQ982469.1	99.19	491
NMF	KT970627	P31	[c]JX286709.1	99.18	483
NMF	KT970627	P31	[c]JQ982470.1	98.76	483
SELKB	KT726938	Polyprotein	[d]AF348210.1	80.41	8831
SELKB	KT726938	Polyprotein	[d]HQ880243.1	79.89	8744
SELKB	KT726938	Polyprotein	[d]HQ880242.1	80.94	8734
SELKB	KT726938	Polyprotein	[e]AF368424.1	83.03	1691
SELKB	KT726938	Polyprotein	[f]EF547367.1	82.72	1678
SELKB	KT726938	Polyprotein	[g] SIU90326	83.22	1299
SELKB	KT970619	Polyprotein	[h]AF023848.1	96.03	2672
SELKB	KT970619	Polyprotein	[h]X73422.1	94.93	1203

[a]Sugarcane mosaic virus isolates, [b]Maize dwarf mosaic virus isolates, [c]Maize chlorotic mottle virus isolates, [d]Cowpea aphid-borne mosaic virus isolates, [e]Crotalaria mosaic potyvirus, [f]South African passiflora virus, [g]Sesame mosaic potyvirus polyprotein, [h]Peanut mottle virus strains

virus, Sesame mosaic virus, South African passiflora virus, Maize dwarf mosaic virus, *sugarcane mosaic viruses* and *Maize chlorotic mottle virus* were also obtained in this study. The identification of grass infecting viruses on cowpea through whole transcriptome sequencing underscores the limitations of conventional virus detection methods such as ELISA, RT-PCR, sap inoculation among others in disease diagnosis (Adams *et al.*, 2009; Adams *et al.*, 2012) since they are specific, require prior knowledge and only target a small range of potential viruses.

The identified cereal viruses are responsible for a very devastating viral disease called maize lethal necrosis disease (MLN). MLN is caused by double infection of maize plants with *Maize chlorotic mottle virus* (MCMV) and any of the cereal viruses in the *Potyviridae* group, such as *Sugarcane mosaic virus* (SCMV), *Maize dwarf mosaic virus* (MDMV), or *Wheat streak mosaic virus* (WSMV) (Goldberg and Brakke, 1987; Uyemoto *et al*,

1981). In this study, double infection of cowpea with MCMV and any of the cereal viruses that together cause MLN was not found since each of these viruses were detected in sample collected from different locations/regions. However, these viruses can be transmitted through vectors and seed and therefore present a potential for spread to other localities which may lead to an outbreak of MLN epidemic. In this study, a legume (cowpea) has for the first time been identified as potential host for MCMV, SCMV and MDMV and given the increasing legume-cereal intensification in the country increases the risk of their spread to maize. This calls for a development and routine deployment of diagnostic assays such as reverse- transcription PCR in the localities where they have been identified. The deep sequencing results need to be validated through polymerase chain reaction.

Previous studies have reported as many as eleven viruses on cowpea in Sub-Saharan Africa (SSA) (Thottappilly and Rossel, 1992; Hughes and Shoyinka, 2003; Sengooba, 2003). Of these up to eight have been reported to infect cowpea in the major growing areas of Uganda (Amayo *et al*., 2012; Orawu *et al*., 2005). However, our results reveal the presence of only *Cowpea aphid borne mosaic virus* (CABMV) among the eight viruses as opposed to those previous studies. Our results also show the absence of *Cucumber mosaic virus* (CMV), which is reported to occur ubiquitously across Sub-Saharan Africa (SSA) (Hughes and Shoyinka, 2003) as well as the absence of *Cowpea mild mottle virus* (CPMMV) and *Cowpea severe mosaic virus* (CPSMV) (Orawu *et al*., 2005; Orawu *et al*., 2015).

Acknowledgements

We thank the Regional Universities Forum for Capacity Building in Agriculture (RUFORUM) for giving us an opportunity to share these findings. Financial support by McKnight Foundation and sequencing support by Biosciences eastern and central Africa Hub at the International Livestock Research Institute (BecA-ILRI Hub) are also acknowledged. This paper is a contribution to the 2016 Fifth African Higher Education Week and RUFORUM Biennial Conference.

References

Adams, I.P., Glover, R.H., Monger, W.A., Mumford, R., Jackeviciene, E., Navalinskiene, M., Marija, M. and Boonham, N. 2009. Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology. *Molecular Plant Pathology* 10 (4):537-545.

Adams, I.P., Miano, D.W., Kinyua, Z.M., Wangai, A., Kimani, E., Phiri, N., Reeder, R., Harju, V., Glover, R., Hany, U., Souza-Richards, R., Deb Nath, P., Nixon, T., Fox, A., Barnes, A., Smith, J., Skelton, A., Thwaites, R., Mumford, R. and Boonham, N. 2012. Use of next-generation sequencing for the identification and characterization of maize chlorotic mottle virus and Sugarcane mosaic viruscausing maize lethal necrosis in Kenya. *Plant Pathology* 62 (4): 741-749.

Aliyu, T., Balogun, O.S. and Kumar, L. 2012. Survey of the symptoms and viruses associated with cowpea (*Vigna unguiculata* (L). Walp) in the agroecological zones of Kwara

State, Nigeria. *Ethiopian Journal of Environmental Studies and Management* 5 (4):613-619.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25 (17):3389-402.

Amayo, R., Arinaitwe, A., Mukasa, S.B., Tusiime, G., Kyamanywa, S., Rubaiyaho, P.R. and Edema, R. 2012. Prevalence of viruses infecting cowpea in Uganda and their molecular detection. *African Journal of Biotechnology* 11(77):14132-14139.

Bashir, M., Ahmad, Z. and Ghafoor, A. 2002. Cowpea aphid-borne mosaic potyvirus: A review. *International Journal of Pest Management* 48(2):155-168.

Bock, K.R. 1973. Peanut mottle virus in East Africa. *Annals of Applied Biology* 74 (2):171-179.

DeBoever, C., Reid, E.G., Smith, E.N., Wang, X., Dumaop, W., Harismendy, O., Dennis, C., Douglas, R., Eliezer, M. and Frazer, K.A. 2013. Whole transcriptome sequencing enables discovery and analysis of viruses in archived primary central nervous system lymphomas. *PloS One* 8(9):e73956–e73956.

Demski, J., Alexander, A., Stefani, M. and Kuhn, C.W. 1983. Natural infection, disease reactions and epidemiological implications of peanut mottle virus in cowpea. *Plant Disease* 67:267-269.

Edema, R., Adipala, E. and Florini, D. A. 1997. Influence of season and cropping system on occurrence of cowpea diseases in Uganda. *Plant Disease* 81:465-468.

Goldberg, K. and Brakke, M. K. 1987. Concentration of maize chlorotic mottle virus increased in mixed infections with maize dwarf mosaic strain B. *Phytopathology* 77:162-167.

Hughes, J. d'A, and Shoyinka, S.A. 2003. Overview of viruses of legumes other than groundnut in Africa. In: Proceedings of a conference organized by IITA/ : 4-8 June 2001.

Kareem, K.T. and Taiwo, M.A. 2007. Interactions of viruses in cowpea: Effects on growth and yield parameters. *Virology Journal* 4:15–15.

Masek, T., Vopalensky, V., Suchomelova, P. and Pospisek, M. 2005. Denaturing RNA electrophoresis in TAE agarose gels. *Analytical Biochemistry* 336 (1):46-50.

Orawu, M., Melis, R., Milliano, W.D., Laing, M. and Adipala, E. 2005. Occurrence and prevalence of cowpea virus diseases in Uganda. *Africana Crop Science Proceedings* 7: 1279-1283.

Orawu, M., Obuo, J. and Omadi, R. 2015. Distribution and detection of cowpea viruses infecting cowpea in Uganda. *American Journal of Plant Sciences* 06 (05):574-581. http://doi.org/10.4236/ajps.2015.65062

Orawu, M., Obua, J. and Omadi, R. 2015. Distribution and detection of viruses infecting cowpea in Uganda. *American Journal of Plant Sciences* 6 (5):574-581.

Sengooba, T. 2003. Synthesis of country reports on virus diseases of legume crops in sub-Saharan Africa. In *Proceedings of a conference organized by IITA/ : 4-8 June 2001.* pp. 342-347. Retrieved from http://books.google.com/books?

Taiwo, M. A. and Akinjogunla, O. J. 2006. Cowpea viruses: Quantitative and qualitative effects of single and mixed viral infections. *African Journal of Biotechnology* 5 (19):1749-1756.

Uyemoto, J.K., Claflin, L.E., Wilson, D.L. and Raney, R.J. 1981. Maize chlorotic mottle and maize dwarf mosaic viruses: Effect of single and double inoculations on symptomatology and yield. *Plant Disease* 65:39-41.

Wortmann, C.S. and Eledu, C.A. 1999. Uganda's agroecological zones: A guide to planners and policy makers. Centro Internationale de Agricultural Tropical (CIAT), Kawanda, Uganda.