Full Length Research Paper

The rolling circle amplification and next generation sequencing approaches reveal genome wide diversity of Kenyan cassava mosaic geminivirus

T. M. Kathurima[1, 2]*, E. M. Ateka[2], A. B. Nyende[2] and T. A. Holton[1]

[1]Biosciences East and Central Africa Hub, C/O ILRI, P. O. Box 30709-00100, Nairobi, Kenya.
[2]Jomo Kenyatta University of Agriculture and Technology, P. O. Box 62000-00200, Nairobi, Kenya.

Rolling circle amplification is a simple approach of enriching populations of single-stranded DNA plant begomovirus genomes (genus, *Begomovirus*; family, Geminiviridae). This is an innovative approach that utilizes the robustness of the bacteriophage phi29 DNA polymerase used in circle amplification, together with deep sequencing using Illumina Miseq and bioinformatics to assess population diversity of begomoviruses in naturally infected cassava. The approach is suitable for detecting rare members in a population in begomoviral populations in situation where mixtures of isolates, strains, and multiple species occur. The main objectives were to increase the sensitivity of detection of next generation sequencing by enriching it using rolling circle amplification then determination of the diversity of the cassava mosaic geminivirus. This was done by total nucleic acids isolated from symptomatic, field cassava infected plants, then using rolling circle amplification to multiply the less abundant viral sequences. Enriched and non-enriched virus-libraries were subjected to deep sequencing using Illumina Miseq. Using bioinformatic CLC Genomics 5.5.1 software programs the quality assessment of reads and contig assembly of viral sequences. This was done through de novo and reference-guided assembly. The identity and diversities of the begomoviral sequences were compared with sequences in Sanger sequencing of viral components deposited in the NCBI Gene Bank. In this study we have demonstrated that RCA increases the chances of detecting the virus by approximately 10 to 1000 fold and wide genome diversity of cassava mosaic geminivirus in various cassava growing zones in Kenya were detected. In conclusion, this approach described herein is simple and will enhance the exploration of begomovirus diversities from cassava infected plants, irrespective of their viral abundance. This will make it possible for routine screening of field samples as the cost of deep sequencing NGS is decreasing and the advances of bioinformatic software development become enhanced. This is the first report of the RCA-Illumina-NGS approach to explore cassava infected with begomoviruses under field conditions and their diversities.

Key words: Illumina Miseq sequencing, geminivirus, ssDNA viruses, viral sequence enrichment, de novo genome assembly, rolling cycle amplification (RCA).

INTRODUCTION

Cassava is cultivated as a subsistence crop in developing countries across the world where its roots provide a source of dietary carbohydrates for over 700 million people. In East and Central Africa, cassava mosaic disease (CMD) is the most damaging plant virus disease as well as the world with an epidemic causing annual crop losses valued at between US$ 1.9 and 2.7 billion (Patil, 2009). Cassava mosaic disease is caused by a

complex of cassava mosaic geminiviruses (CMG's), family Geminiviridae, and genus *Begomovirus*. The viruses are spread by the whitefly *Bemisia tabaci* (Gennadius) (Brown, 2007). CMG's are of two types monopartite and bipartite genomes. The bipartite has genome components, called DNA-A and DNA-B, single-stranded DNA comprising 2.7 kb circular molecules while the monopartite lacks the DNA-B component and it is approximately 2.9 kb. DNA-A encodes proteins and regulatory elements that play the role of replication, encapsidation and control of gene expression while DNA-B encodes proteins enabling viral movement (Jeske, 2009). The rolling-circle mechanism just like the ssDNA-containing bacteriophages is the mechanism utilized by the genomic ssDNA to replicate utilizing double-stranded DNA (dsDNA) intermediates, in the nucleus of the host cell.

The DNA-A and DNA-B components share no much sequence similarity except for a short sequence regarded to as the 'common region' (CR) of approximately 200 nucleotides. The common region is the initiation replication site for rolling circle replication, and it is conserved among members of the family Geminiviridae (Harrison and Robinson, 2002). Both DNA components contain protein-coding regions in the viral strand and in the complementary strand. Six such genes seem to be universally present. On the complementary strand for the New World (NW) bipartite begomoviruses, contains one gene (*AV1*) on the viral sense strand and three genes (*AC1*, *AC2*, *AC3*) for the DNA-A component (Harrison and Robinson, 1999) and for the Old World (OW) bipartite begomoviruses an additional gene *AV2* in the viral sense strand and *C4* on the complementary strand (Hanley et al., 1999). The DNA-B component contains one gene each, *BV1* and *BC1*, on the sense and complementary strands of respectively (Sanderfoot et al., 1996).

In Africa, nine species of Begomovirus have so far been associated with cassava mosaic disease (Fargette et al., 1996; Bull et al., 2006; Tiendrébéogo et al., 2012). In Kenya, only four begomovirus species have previously been reported, namely ACMV, EACMV, and EACMZV (Stanley and Gay, 1983; Were et al., 2004a, 2004b).

The current experimental procedures used for detecting and determining the   of genetic variability within begomoviral genomes utilize DNA by polymerase chain reaction (PCR) (Saiki et al., 1988) using specific or degenerate sequence primers, or use of random primers in rolling circle amplification (RCA)   (phi29 DNA polymerase) (Dean et al., 2001), followed by cloning and sanger sequencing. Normally, the monopartite or bipartite DNA complete genome components of begomoviruses

and their associated circular, ssDNA satellites are cloned from the products of RCA (Inoue-nagata et al., 2004) or virus-specific PCR (Briddon et al., 2000). The amplification of begomoviral genomic and associated components by PCR, followed by cloning and capillary DNA sequencing, are limited by the variant numbers produced at early steps of amplification specificity of the primers, and then by selection during the molecular cloning step. In another approach, begomoviral genome characterization has been achieved using a combination of restriction fragment length polymorphism (RFLP) and pyro-sequencing (Wyant et al., 2012).

Normally RCA produces high molecular weight products as dsDNA that are digested into unit-length components and cloned, with the inserts then verified by Sanger sequencing (Dean et al., 2001; Johne et al., 2009). This procedure is a time-consuming, and in addition, a few numbers of variants are represented among the resultant clones, based on the expectation that one or a few most frequent genotypes are the once represented in the starting material. In addition technical limitations can result in the inability to detect very low abundant begomoviral and/or associated DNA satellite molecules.

An innovative approach described here utilizes the robustness of the bacteriophage phi29 DNA polymerase used in RCA, together with deep sequencing using Illumina (Mardis, 2008; Bentley et al., 2008) and bioinformatics to assess population diversity of begomoviruses in naturally infected cassava. For virus discovery from field samples at the population level, traditional procedures such as RCA or PCR thereafter followed by cloning and Sanger DNA sequencing, are ineffective. The approach is suitable for detecting rare members in a population in begomoviral populations in situations where mixtures of isolates, strains, and multiple species occur. In the current study we demonstrated that circular single-stranded DNA-containing begomoviruses were enriched by RCA from total DNA extracts of naturally infected symptomatic cassava field plants. The enriched begomoviral genomes were subjected to Illumina Miseq sequencing. The short sequence reads obtained were assembled using bioinformatics tools, and were compared with genome sequences deposited in the genebank

MATERIALS AND METHODS

Cassava leaves identified with cassava mosaic disease were collected from Coast, Eastern, Western and Nyanza regions of Kenya where cassava growing is important. Twenty four samples

*Corresponding author. E-mail: kathurima07@gmail.com.

that detected positive for cassava mosaic geminiviruses with specific isolates primers on PCR were selected. Universal primers were used for detection of African cassava mosaic virus (ACMV) were JSP001 (5'-ATGTCGAAGCGACCAGGAGAT-3') ACMV for the forward primer and JSP002 (5'-TGTTTATTAATTGCCAATACT-3') ACMV for the reverse primer with an expected amplicon of 774bp. The detection of EACMV was done using EAB555 F/R primers whose sequences were EAB555/F (5'TACATCGGCCTTTGAGTCGCATGG-3') EACMV DNA-B and EAB555/R (5'CTTATTAACGCCTATATAAACACC-3') EACMV DNA-B with an expected amplicon of 556 bp fragment of EACMV DNA B component (Fondong et al., 1998). Of these four samples total DNA extract containing begomoviruses were enriched by RCA from and another four of the same samples were not enriched with RCA. The remaining 16 samples were not enriched but were chosen to represent major growing regions in Kenya.

The rolling circle amplification (TempliPhi)

The cassava geminivirus virus genome was amplified and isolated using the TempliPhi Kit (GE Healthcare, Buckinghamshire, United Kingdom) according to James et al. (2011). Two mixes were prepared. For master mix 1 (MM1), 5 µl of TempliPhi sample buffer was mixed with 1 µl of the isolated DNA and 1 µl of a 50 µM stock solution (4.16 pmol/ul of each primer) of TempliPhi degenerate primers. The mix was then heated at 95⁰C for 3 min to denature the DNA followed by cooling to room temperature or 4°C. Master Mix2 was prepared by mixing 5 µl of TempliPhi reaction buffer and 0.2 µl of TempliPhi Enzyme Mix. Five micro liter of the TempliPhi premix (mix 2) was transferred to the cooled, denatured sample (MM1) then incubated at 30⁰C for 18 h. After the incubation period, the enzyme (Phi29DNA polymerase) was heat-inactivated at 65⁰C for 10 minutes. The samples were then cooled and stored at 4°C. Cleaning of the samples was done using Qiagen kit following the manufacturers' procedure. The amplified clean DNA was used directly in Illumina next generation sequencer.

Preparation of nextera libraries for Illumina MiSeq sequencing

A good quality DNA presented by optical density of 260/280 and 260/230 purity indices equal to or greater than 1.8 to 2.0 values from spectrophotometry were selected from 24 samples for next generation sequencing. The DNA libraries were prepared from total DNA extract concentration from ranging from 0.05 to 3.3 ng/µl using the Illumina nextera DNA Sample Preparation kit[TM] according to the manufacturer's instructions (Illumina, San Diego, California). The first step involved DNA fragmentation with addition of adapter sequences to the ends to allow for amplification by PCR. Addition of indexes and enrichment was done. The final size and concentration of each library was estimated using a Bioanalyzer (Agilent, Santa Clara, CA, USA) and the Qubit (Invitrogen, Carlsbad, CA, USA), respectively. Library pools of 2 nM were prepared by mixing the libraries from each sample to achieve an equal molar concentration of each. Libraries were normalized, pooled and sequenced using a 2×300 -cycle PE V3 Illumina kit. Paired end reads were generated using the Illumina MiSeq System at the Biosciences Eastern and Central Africa - International Livestock research Institute (BecA-ILRI) Hub in Nairobi, Kenya.

Comparison of sensitivity between RCA enriched and non RCA enriched samples

Serial dilutions of 10^{-1} to 10^{-7} of the same positive batch of enriched and non-enriched total DNA extracted were made in duplicates and a negative control included. Polymerase chain reaction was carried out in a thermocycler in a reaction of 20 µl in AccuPower®Bioneer premix, 0.1 µM forward and reverse primer and 2.0 µl DNA template. The reaction profile was followed by 35 cycles of 94°C (30s), 52°C (30s), 72°C for 1 min and 72°C for final extension. PCR products were analyzed by electrophoresis in 1X TAE buffer on 2% agarose gel stained with gel red and image captured by a camera under UV light. The PCR assays were performed in technical triplicates.

Illumina Miseq Sequence short sequence assembly and analysis

Sequences from the 24 libraries were screened for quality where sequences with a value less than 25 were trimmed. The short reads were then subjected to reference/de novo assembly using CLC Genomics 5.5.1 software using default settings. The number of sequence reads from CMGs contigs extracted from non RCA and RCA were compared. The genome sequences assembled were also subjected to phylogenetic analysis together with those from genebank sequences by the neighbor-joining method, using MEGA6 software.

RESULTS

Of the 24 libraries sequenced 20 had concentrations ranging from 1.0 to 3.3 ng/ul produced good number of pair ended reads while 4 libraries had less than 0.06 ng/ul and had very low number of reads and consequently could not be assembled. The de novo assembly of the high-throughput Illumina 6 million (6M) reads was carried out using CLC Genomics 5.5.1 software default settings resulted in a large number of contigs that ranged from 1,456 to 42,181 with an average length of 254 to 388 nucleotides. The contigs were used to search in the NCBI-gene bank database in order to identify the most closely related begomovirus(es).

The search result showed that the contigs assembled from RCA containing begomovirus libraries consistently comprised the highest number of assembled reads, indicating that the RCA successfully enriched the begomoviruses DNA components looking at the number that assembled to arrive at a consensus genome (Figure 1). The ACMV_A was found in 4 libraries in samples collected from Western and Eastern regions, ACMV_B was found in 3 libraries in samples collected from western and eastern, EACMV_A in 10 libraries in Eastern, Western, Coast and Nyanza and EACMV_B was found in 6 libraries in Eastern, western, coast and Nyanza.

The complete DNA-A genome of CMG from *de novo* assembly resulted in 2754 to 2801 nucleotides, DNA-B genome 2737-2787nucleotides, ACMV_A genome 1835-2781 and DNA_B 2672-2718 nucleotides long.

The limits of detection for RCA enriched DNA derived from symptomatic cassava leaves by the PCR amplifications was observed to be up to the dilution factor of 10^{-7} for enriched RCA compared to 10^{-5} and 10^{-6} of non-enriched DNA contents (Table 1).

Table 1. Comparison of sensitivity between RCA enriched and non RCA enriched samples.

	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
Isolate Eastern_14 RCA enriched	+	+	+	+	+	+	+
Isolate Eastern _14 Non RCA enriched	+	+	+	+	+	-	-
Isolate Western_2 RCA enriched	+	+	+	+	+	+	+
Isolate Western_2 Non RCA enriched	+	+	+	+	-	-	-

+, Positive amplification; -, no amplification.

Table 2. Percent nucleotide sequence similarities of complete Kenyan EACMV DNA_A sequences.

Sequences	western_6	western_2	Eastern-14	Coast_42	Eastern_14	Coast_17	Nyanza_9	Nyanza_66	Western_40	Nyanza_11	Western_2	Coast_4	Eastern_16	Eastern_37	coast_4
K02029.TGMV	61	61	61	65	65	65	63	63	63	63	63	63	63	63	63
ACMV-UG_Nam	**99**	**96**	**96**	72	72	72	72	73	73	73	73	70	70	70	70
ACMV/Ug//Mld	**95**	**97**	**98**	72	72	72	73	73	74	74	73	70	70	70	70
EACMZV-K10	72	72	72	**98**	**98**	**98**	82	83	83	83	83	87	87	87	87
EACMZV-K212	72	72	72	**98**	**98**	**98**	82	83	83	83	83	87	87	87	87
Nam_CMD-MI54_12	73	73	73	83	83	83	**97**	**98**	**98**	**98**	**99**	92	92	92	92
gi|61191723_TZ10	73	74	74	82	82	82	**97**	**98**	**98**	**98**	**99**	92	92	92	92
EACMV-UGK66	73	74	74	83	83	83	**98**	**99**	**99**	**99**	**99**	93	93	93	93
EACMV-UGK127	73	74	74	82	83	83	**98**	**98**	**99**	**99**	**99**	92	92	92	92
EACMV-UGK115	73	74	74	82	83	83	**98**	**98**	**99**	**99**	**99**	92	92	92	92
EACMV-K6	70	70	70	86	86	86	92	92	92	92	93	**97**	**97**	**97**	**97**
EACMV-K312	70	70	70	86	86	86	91	92	92	92	92	**97**	**97**	**97**	**97**
EACMV-K313	70	70	70	86	86	86	91	92	92	92	92	**98**	**98**	**98**	**98**

Phylogenetic analysis of full length DNA_A of Kenyan CMGs

The thirteen complete genome DNA_A sequences were aligned and compared with NCBI sequences which grouped into two major clusters. Cluster one was EACMV containing sub cluster A, B and C which grouped according to three geographical regions. The sub cluster A contained isolates coast_42, eastern_14, coast_17, with 98% nucleotide sequence identity clustered with Zanzibar isolates (EACMVZV), sub cluster B isolates nyanza_66, nyanza_11, western_40 and western_2 with 97%-99% nucleotide identity clustered with Ugandan isolates (EACMV_UG) and sub cluster C isolates coast_4, eastern_16 and eastern_37 with 97 to 98% nucleotide identity clustered with Kenyan isolates (EACMVK) and cluster D was ACMV isolates eastern_6, western_2 and eastern_14 (Table 2) (Figure 2).

Phylogeny analysis of full length DNA_B of Kenyan CMGs

The nine complete DNA_B genome sequences obtained

from *de novo* assembly were aligned and compared with NCBI sequences. They clustered into two major clusters which were ACMV and EACMV. Just like the EACMV_A, they clustered into four sub clusters with isolate cmd16; western_52 and nyanza_66 clustering with UG isolate eastern_8 clustered with KE isolate and isolate coast_42, eastern_17 and eastern_14 clustering with ZV isolates. Isolate cmd14 and cmd4 clustered with ACMV.

DISCUSSION

In this study we enriched circular single-stranded DNA-containing begomoviruses by RCA from total DNA extracts of naturally infected symptomatic cassava field plants. The enriched begomoviral genomes were subjected to Illumina Miseq sequencing. The short sequence reads obtained were assembled using bioinformatics tools, and were compared with genome sequences deposited in the genebank. We have demonstrated from the serial dilution assay that the viral components that were enriched using RCA procedure increased its detection limit by approximately 10 to 1,000 fold by RCA (Table 1), consequently observation from the de novo assembled
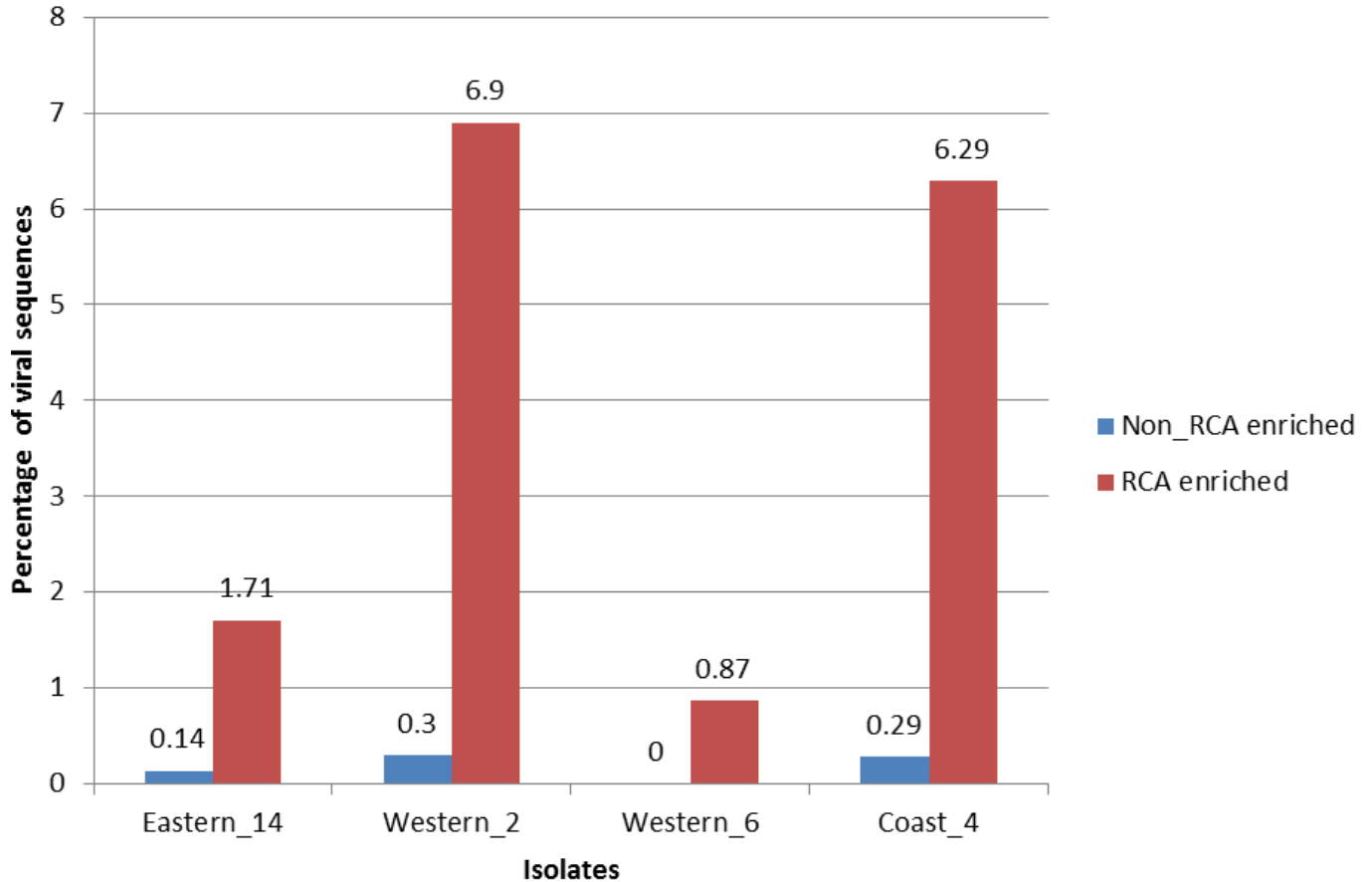
Figure 1. A graph showing percentage of viral de novo assembled sequences of various cassava geminivirus DNA components sequences obtained from total DNA extract enriched RCA and non-enriched with RCA.

reads also indicates generally a high number of sequences of CMGs were assembled from the total sequenced reads from the enriched isolates libraries than non-enriched viral components (Figure 1). Interestingly there was a library that did not record any geminivirus without enrichment except on enriched libraries. Then it is right to infer that RCA increases the chances of detecting ssDNA viruses in samples with very low virus titer. This procedure therefore steps up the sensitivity of detection by increasing the copy numbers of circular DNA. Contrary to the conventional sequencing that relies on primers to amplify a section of the genome resulting in limited information on similarities. It is possible to obtain a complete genome from this RCA and next generation sequencing approach because you need not to have prior knowledge of the viral sequences. Out of 24 libraries, we were able to assemble 25 complete genomes of Geminivirus from both DNA_A and DNA_B component infecting Kenyan cassava. To achieve such results with conventional Sanger method it will be tedious. The utility of metagenomics in detecting and determining the genome sequence and nucleotide similarities of cassava infecting begomoviruses has been used successfully demonstrated in this study (Table 3) which can be utilized in diagnostics and diversities studies of begomovirus in cassava.

The phylogenetic analysis of sequences obtained from the assembled reads from the isolates of this study and compared with genebank sequences grouped in distinct geographical distribution and others have overlapping patterns for both DNA components (Figures 2 and 3). The EACMV-UG isolates were detected in regions neighboring Uganda such Nyanza and Western regions and there was no evidence of EACMV-UG in eastern and coastal region. Previous study reports have been unable to identify EACMV-UG in coastal regions of either Kenya or Tanzania (Were et al., 2004a; Ndunguru et al., 2005). Both EACMV and EACMZV were identified in coastal and eastern region. Detection proportions result here showed that EACMV is more widespread than ACMV in the country it agrees well with (Mwatuni et al., 2015) report. Synergistic interaction between ACMV and EACMV were observed and this result could, lead to severe symptoms as reported by farmers and as observed in the CMD pandemic in Uganda (Bull et al., 2006).

The sequencing technologies cost is on downward
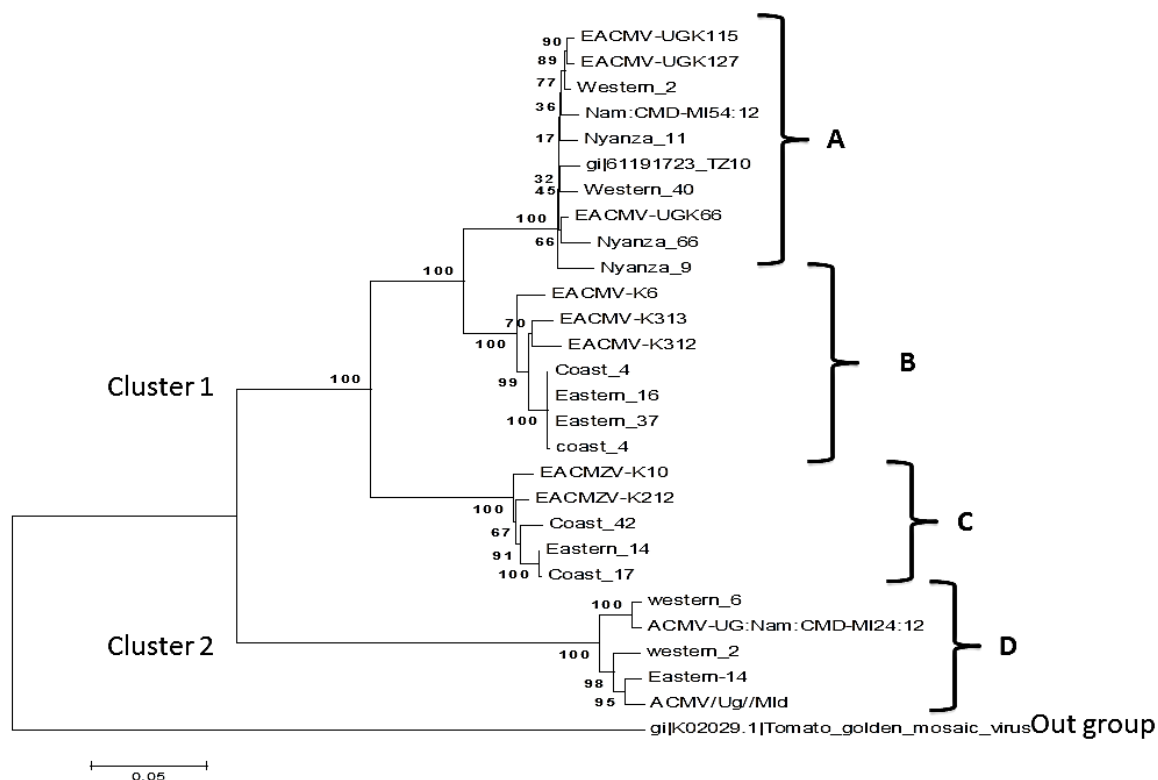
Figure 2. Neighbor-joining tree of the complete genome of DNA_A sequences obtained from de novo assembly of the resultant illumina Miseq deep sequencing. Numbers at nodes indicate percentage bootstrap values of 1000 replicates. Cluster 1 (EACMV) East Africa Cassava Mosaic Virus; sub_cluster (A) Ug isolate, (B) Ke isolate and (D) Zv isolate: Cluster 2 (D) (ACMV) Africa Cassava Mosaic Virus. The alignment was arbitrarily rooted with an out-group of (K02029) Tomato golden mosaic virus DNA_A component.

Table 3. Percentage nucleotide sequence similarities of EACMV DNA_B component and selected other begomoviruses from NCBI.

Virus isolate	Eastern-14	Western_6	Nyanza_66	Eastern_16	Western_52	coast_42	Eastern_17	Eastern_14	Eastern_8
gi|K02030.1_TGMV	46	44	44	44	45	45	44	44	44
ACMV-UG_Nam	**93**	**99**	50	50	50	50	50	50	50
EACMV/Ug3//Svr	51	50	**97**	**97**	**98**	89	89	89	90
EACMZV-K19_2_	51	50	88	89	89	**96**	**96**	**96**	90
EACMV-KE2K27	51	50	89	90	90	91	91	91	**96**
EACMV-UGK72	51	50	89	90	90	91	91	91	**96**

trend and with possibilities of multiplexing strategies. The exploration of begomovirus composition in populations can be done for as many as 96 samples per flow cell lane, infact decreasing the cost of deep sequencing. This allows large numbers of virus samples to be processed simultaneously. MiSeq libraries can be constructed using a high multiplexing barcode to lower the sequence throughput and reduce the period dedicated forcomputation and assembly (Smith et al., 2010) in order to considerably increase the number of samples per lane to make reasonable coverage for this short genome sizes

and reduce the cost per sample.

Conclusion

From the results reported herein, we have demonstrated that enriching of the circular ssDNA begomoviral first by RCA, from total DNA extracts of symptomatic cassava by selectively boosting the viral concentration of the starting DNA template thereby increasing the sensitivity of detection. These results demonstrate a reliable approach
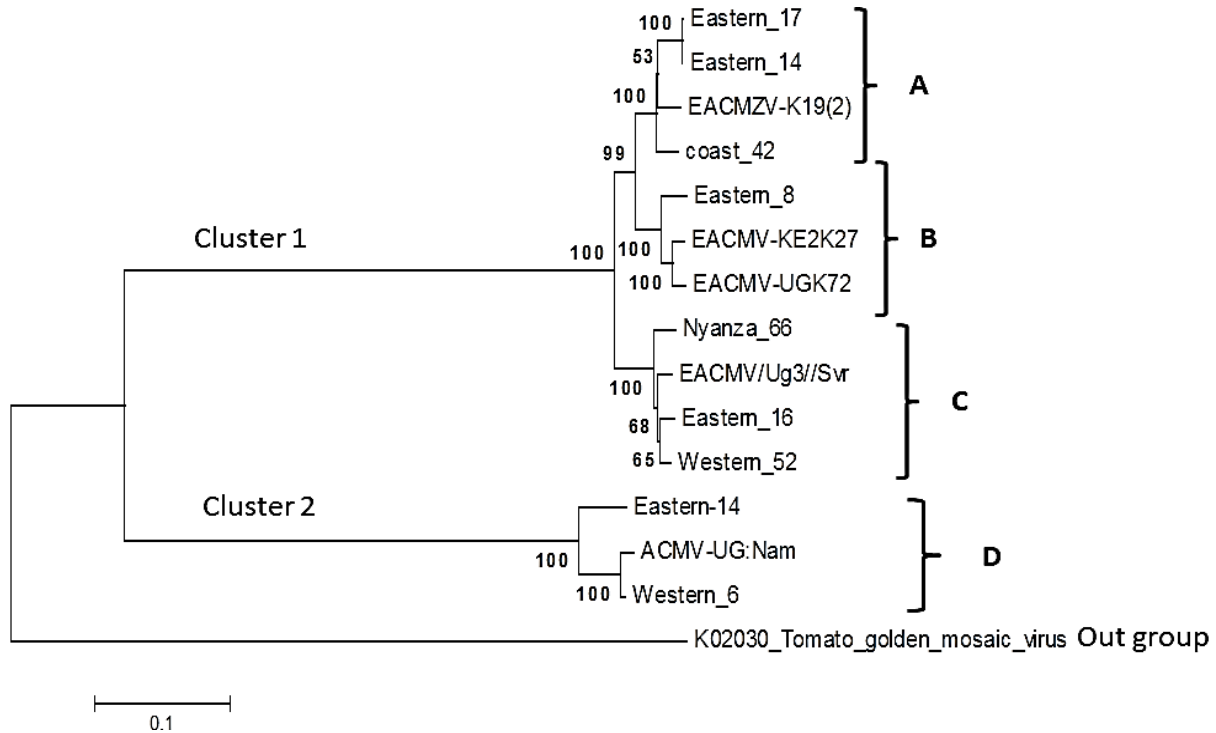
Figure 3. Neighbor-joining tree analysis of the complete genome of DNA_B component sequences obtained from de novo assembly of the resultant illumina deep sequencing. Numbers at nodes indicate percentage bootstrap values of 1000 replicates. Cluster 1(EACMV strain), sub cluster A, Uganda isolates, sub cluster B, Kenya isolate, sub cluster C, Zanzibar isolate and cluster 2 (ACMV strain) sub cluster D. The alignment was arbitrarily rooted with an out group of (K02030) Tomato golden mosaic virus DNA_B component.

for implementation of deep sequencing of begomovirus associated with of ssDNA that infections cassava plants and for exploring their diversities among viral genomes.

Conflict of Interests

The authors have not declared any conflict of interests.

REFERENCES

Bentley DR., Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall, KP, Evers DJ, Barnes CL, Bignell HR. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456(7218):53-59.

Briddon RW, Mansoor S, Bedford ID, Pinner MS, Markham PG (2000). Clones of cotton leaf curl geminivirus induce symptoms atypical of cotton leaf curl disease. Virus Genes 20(1):19-26.

Brown JK (2007). The Bemisia tabaci complex: Genetic and phenotypic variability drives begomovirus spread and virus diversification. Plant Dis. 1:25-56.

Bull SE, Briddon RW, Ssserubombwe WS, Ngugi K, Markham PG, Stanely J (2006). Genetic diversity and phylogeography of cassava mosaic viruses in Kenya. J. Gen. Virol. 87:3053-3065.

Dean FB, Nelson JR, Giesler TL, Lasken RS (2001). Rapid Amplification of Plasmid and Phage DNA Using Phi29 DNA Polymerase and Multiply-Primed Rolling Circle Amplification. Genome Res. 11(6):1095-1099.

Fargette D, Colon LT, Bouveau R, Fauquet C (1996). Components of resistance to African cassava mosaic virus. Euro. J. Plant Pathol. 102:645-654.

Fondong VN, Pita JS, Rey C, Beachy RN, Fauquet CM (1998). First report of the presence of East African cassava mosaic virus in Cameroon. Plant Dis. 82:1172.

Hanley-Bowdoin L, Settlage SB, Orozco BM, Nagar S, Robertson D (1999). Geminiviruses: models of plant DNA replication, transcription, and cell cycle regulation. Crit. Rev. Plant Sci. 18:71-106.

Harrison BD, Robinson D (1999). Natural Genomic and Antigenic Variation in Whitefly-Transmitted Geminiviruses (Begomoviruses). Annu Rev Phytopathol. 37(11):369-98.

Harrison BD, Robinson D (2002). Green shoots of geminivirology. Physiol. Mol. Plant Path. 60:215-218.

Inoue-nagata AK, Albuquerque LC, Rocha WB, Nagata T (2004). Short communication a simple method for cloning the complete

begomovirus genome using the bacteriophage phi29 DNA polymerase. J. Virol. Methods 116(2):209-211.

James AP, Geijskes RJ, Dale JL, Harding R (2011). Development of a novel Rolling Circle Amplification technique to detect Banana Streak Virus that also discriminate between integrated and episomal virus sequences. Plant Dis. 95:57-62.

Jeske H (2009). Geminiviruses. Curr. Top Microbiol. Immunol. 331:185-226.

Johne R, Mu H, Rector A, Ranst M, Van Stevens H (2009). Rolling-circle ampli cation of viral DNA genomes using phi29 polymerase. Trends Microbiol. (April):205-211.

Mardis E (2008). Next-generation DNA sequencing methods. In Annu Rev Genomics Hum Genet. Annual Reviews: Palo Alto, CA, USA, 9:387-402.

Mwatuni F, Ateka E, Karanja L, Mwaura S, Obare I (2015). Distribution of Cassava Mosaic Geminiviruses and their Associated DNA Satellites in Kenya. Am. J. Exp. Agric. 9(3):1-12.

Ndunguru J, Legg JP, Aveling TA, Thompson G, Fauquet CM (2005). Molecular biodiversity of cassava begomoviruses in Tanzania: evolution of cassava geminiviruses in Africa and evidence for East Africa being a center of diversity of cassava geminiviruses. Virol. J. 2:21.

Patil BL, Fauquet CM (2009). mosaic geminiviruses: actual knowledge and perspectives. Mol. Plant Pathol. 10(5):685-701.

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi RHorn GT, Erlich HA (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239(4839):487-491.

Sanderfoot AA, Ingham DJ, Lazarowitz SG (1996). A viral movement protein as a nuclear shuttle. Plant Physiol. (110):23-33.

Smith AM, Heisler LE, St Onge, RP, Farias-Hesson E, Wallace IM, Bodeau J, Harris A, Perry KM, Giaever G, Pourmand N, Nislow C (2010). Highly-multiplexed barcode sequencing: An efficient method for parallel analysis of pooled samples. Nucleic Acids Res. 38:13

Stanley J Gay, MR, Stanley J, Gray MG, Stanley J, Gay MR (1983). Nucleotide sequence of cassava latent virus DNA. Nature. 301:260-262.

Tiendrébéogo F, Lefeuvre P, Hoareau M, Harimalala MA, De Bruyn A, Villemot J, Traoré VSE, Konaté G, Traoré AS, Barro N, Reynaud B, Traoré OLJ (2012). Evolution of African cassava mosaic virus by recombination between bipartite and monopartite begomoviruses. Virol. J. 9:67.

Were HK, Winter S, Maiss E (2004a). Occurrence and distribution of cassava begomoviruses in Kenya. Ann. Appl. Biol. 145:175-184.

Were HK, Winter S Maiss E (2004b). Viruses infecting cassava in Kenya. Plant Dis. 88:17-22.

Wyant PS, Strohmeier S, Schafer B, Krenz B, Assuncao IP, Lima GSD, Jeske H (2012). Circular DNA genomics (circomics) exemplified for geminiviruses in bean crops and weeds of northeastern Brazil. Virol. 427(2):151-157.