

Characterization of Sirnas Derived from Two Species of Cassava-Infesting Geminiviruses in *Nicotiana Benthamiana*

Paul Kuria

Kenya Agricultural and Livestock Research Organization, Biotechnology Research Institute, Kabete

Postal address: PO Box 14733 00800 Nairobi, Kenya

Tel: +254791081756, Email: kuriapk@gmail.com

Abstract

The infection dynamics of cassava mosaic geminiviruses was determined in *Nicotiana benthamiana*. Following inoculation, symptoms were observed at 5 and 10 dpi for *African cassava mosaic virus* (ACMV-CM) and the Kenyan strain K201 of *East African cassava mosaic virus* (EACMV KE2 [K201]). Northern and Southern blots indicated higher accumulation of ACMV-CM at 10 dpi compared with EACMV KE2 (K201) correlating with more severe symptoms at this time point. However, unlike for EACMV KE2 (K201), symptom remission characterized by appearance of non-asymptomatic leaves was observed after 20 dpi in plants infected with ACMV-CM. Deep sequencing identified between 969,015 and 1,307,689 total reads of which >80% mapped to the host genome. For both virus species, 22 nt virus derived sRNAs (vsRNAs) reads were the most abundant (34% - 53%) followed by 21 nt (27% - 27%) and 24 nt (9 - 23%) vsRNA reads. EACMV KE2 (K201) vsRNAs were proportionally higher (12.31% and 4.51%) than those derived from ACMV-CM (5.70% and 0.45%) for DNA A and DNA B respectively. vsRNAs were covered the entire virus genome with dense accumulation at regions where the viral ORFs overlapped.

Key words; RNA interference, Cassava mosaic disease, Geminiviruses

DOI: 10.7176/JNSR/10-8-06

Publication date: April 30th 2020

1. Introduction

Eukaryotes utilizes a conserved RNA silencing as a potent defense against invading foreign nucleic acids among other regulatory mechanisms (Yang *et al.*, 2014; Martínez de Alba *et al.*, 2013; Csorba *et al.*, 2009). RNA silencing is induced by recognition and processing of double stranded RNAs (dsRNAs) molecules or hairpin-like RNA secondary structures into 21 - 24 nucleotides (nt) primary small interfering RNAs duplexes (siRNAs) (Deleris *et al.*, 2006). Biogenesis of siRNAs is coordinated by ribonuclease III-like enzymes, termed DICER-like proteins (DCLs) in plants (Pooggin, 2016; Deleris *et al.*, 2006). The siRNAs associate with effector molecules to guide specific degradation or translation inhibition of cognate messenger RNA (mRNA) and DNA/histone modification (Carbonell and Carrington, 2015). RNA silencing is amplified by host RNA-dependent RNA-polymerases (RDRs) through generation of secondary siRNA that mediate systemic silencing throughout the plant (Molnar *et al.*, 2010; Wang *et al.*, 2010).

The core protein components of RNA silencing have been described and shown to be diverse in different plants species. Four DCLs, ten ARGONAUTE (AGO) family of proteins and six RDRs have been reported in *Arabidopsis thaliana* (Margis *et al.*, 2006), rice genome encodes eight DCLs, five RDRs, and nineteen AGO proteins (Kapoor *et al.*, 2008), in tomato seven DCL, 15 AGO and six RDR genes have been reported (Bai *et al.*, 2012) whereas maize genome encodes five DCLs, five RDRs, and 18 AGOs (Qian *et al.*, 2011). In *A. thaliana* antiviral response is mediated through biogenesis of 21 nt, 22 nt and 24 nt siRNAs conditioned by DCL4, DCL2, and DCL3 in a hierarchical manner (Garcia-Ruiz *et al.*, 2010). Redundancy of functions in processing of 21 and 22 nt siRNAs has been reported for DCL4 and DCL2 (Axtell, 2013; Cao *et al.*, 2014). Plants infected with RNA viruses have been shown to accumulate abundant amounts of 21 nt and 22 nt virus-derived small RNAs (Ogwok *et al.*, 2016; Deleris *et al.*, 2006) whereas plants infected with DNA viruses mostly accumulate 24 nt siRNAs derived from virus genome that direct antiviral RNA silencing (Pooggin *et al.*, 2016; Incarbone and Dunoyer, 2013). However, diverse plant species systemically infected with geminiviruses accumulate different profiles of virus derived sRNAs (Kuria *et al.*, 2017; Rogans *et al.*, 2016; Aregger *et al.*, 2012) therefore, germplasm displaying different genetic backgrounds respond differently to virus infection.

Recruitment of siRNAs by different AGO protein complexes have been reported to be mostly determined by the 5' nucleotide and the structure of siRNAs duplex (Carbonell and Carrington, 2015; Poulsen *et al.*, 2013). Antiviral

defense in *Arabidopsis* have been reported to involve diverse AGO families including, AGO1, AGO2, AGO5, and AGO10 that act in PTGS targeting RNA viruses (Brosseau and Moffett, 2015; Carbonell and Carrington, 2015; Garcia-Ruiz *et al.*, 2015; Ma *et al.*, 2015). On the other hand AGO4, AGO6, and AGO9 preferentially associate with 24 nt siRNAs to mediated methylation viral DNA genome (Havecker *et al.*, 2010). RDR1, RDR2 and RDR6 have been implicated in amplification of virus derived siRNAs to induce antiviral defense (Wang *et al.*, 2010).

Cassava-infecting geminiviruses (Family Geminiviridae, Genus Begomovirus) are majorly transmitted by whiteflies (Legg *et al.*, 2015) and are the causal agent of cassava mosaic disease (CMD). The disease is endemic to cassava production areas of sub-Saharan Africa (Patil and Fauquet, 2009) and has been ranked as the seventh most important viral disease worldwide (Rybicki, 2015). CMD is further exacerbated by recombination and pseudo-recombination of different species leading to evolution of more virulent strains (Legg, and Thresh, 2000; Olsen *et al.*, 1999). Emergence of recombinant strains such as East African cassava mosaic virus-Uganda Variant (EACMV-UG) have been associated with severe CMD pandemic in the 1990s and 2000s (Legg, and Thresh, 2000; Olsen *et al.*, 1999; Zhou *et al.*, 1997). However, the interaction between different cassava-infecting geminiviruses and host plant silencing machinery has not been precisely characterized. Cassava-infecting geminiviruses possess two similar sized but independent single stranded DNA (ssDNA) genomic components designated DNA A and DNA B (Brown *et al.*, 2015; Hanley-Bowdoin *et al.*, 2013). Both viral genomic components are essential for full disease establishment. DNA A encodes four genes in complementary sense orientation involved in replication and transcription and two genes in virion sense orientation involved in virus encapsidation and pathogenicity determinant. DNA B encodes two genes each in complementary and virion sense orientation that are involved in virus trafficking (Hanley-Bowdoin *et al.*, 2013). The two viral genomic components share homologous sequence of 200 nucleotide in length known as common region (CR) that contain promoters involved in initiation of replication.

In order to develop effective management strategies of cassava-infecting geminiviruses, better understanding host-virus interaction is essential. Therefore, this study reports response of *Nicotiana benthamiana* to infection by two species of cassava-infecting geminiviruses.

Material and methods

Plant material and Agro-inoculation

Ten *Nicotiana benthamiana* plants were inoculated with *Agrobacterium tumefaciens* (strain GV3103) transformed with AKK1420 binary vector carrying infectious clones of ACMV-CM (AF112352 and AF112353) and EACMV KE2 (K201) (AJ717541 and AJ704953), respectively, cloned as head to tail partial dimers (Patil and Fauquet, 2015). *Agrobacterium* cultures were established at 28°C for 12h with shaking in Luria-Bertani (LB) broth supplemented with appropriate antibiotics. At an optical density of 1.0, the culture was spun at 10,000 g, for 2 min. After discarding the supernatant, the resulting pellet was re-suspended in 25 ml of infiltration buffer (0.5 strength Murashige and Skoog (MS) media, supplemented with 5 mM MES, 5 mM MgSO₄, pH 5.7) and the culture incubated without shaking for three hours at room temperature. The culture was then diluted to an OD of 0.5 using the infiltration solution supplemented with 200µM acetosyringone. The *Agrobacterium* culture mixtures were infiltrated on the abaxial side of three week old *N. benthamiana* leaves as described by Leuzinger *et al.* (2013).

Nucleic acid extraction

Young leaves of *N. benthamiana* plants showing typical CMD symptoms were collected at 20 days post inoculation (dpi). The leaf samples were derived from four independent biological replicates for ACMV-CM and five biological replicates for EACMV KE2 (K201) infected plants. Total nucleic acids were extracted as previously described (Patil *et al.*, 2015) and apportion into two. One of the samples aliquots was subjected to TURBO™ DNase (Ambion, Carlsbad, CA, USA) treatment as per kit instructions. The other aliquot was treated 4 µl DNase free ribonuclease (Roche, Indianapolis, IN, USA) for 2h at 37°C to remove RNA. The RNA and DNA samples were quantified on a NanoDrop 2000c spectrophotometer (Thermo Scientific, DE, USA).

Northern blot analysis

Virus derived small RNAs were detected using Northern blotting. Forty micrograms total RNA were fractionated on 15% Criterion TBE-Urea polyacrylamide gel (Bio-Rad, Hercules, CA, USA) at 100V for 2 h. RNA was electro-transferred onto positively charged Hybond nylon membrane (Amersham, UK) using trans-Blot Turbo Transfer System (Bio-Rad, Hercules, CA, USA) at 25V for 30 min and immobilized by crosslinking to the membrane twice

at 120,000 microjoules/cm² using a Stratalinker UV crosslinker 1800 (Stratagene, La Jolla, CA, USA). RNAs probes for virus derived siRNAs detection were PCR amplified from AC2/AC3 region of each virus genome using primers listed on Table 1. In vitro transcription was performed using DIG RNA labelling kit SP6/T7 (Roche, Indianapolis, IN, USA) following manufacturer's instructions. The ensuing steps were performed as described by Kuria *et al.* (2017). Blots were exposed to Amersham high-performance chemiluminescence film (GE Healthcare, Pittsburgh, PA, USA) for 15 min and processed on an automated developer (Konica Minolta-SRX-101A). The autoradiographs were scanned on Epson Perfection V700 photo scanned (Epson, CA, USA) and the signal quantified on image J.

Southern blot analysis

Southern blot analysis was performed to detect viral DNA titer using 5 µg genomic DNA derived from systemically infected *N benthamiana* plants at 20 dpi. Samples were fractionated on 1 % (w/v) agarose gel at 30V for 12 h. The DNA was deproteinated in a solution of 0.2 N HCl for 15 min followed by denaturation in 0.5 N NaOH and 1.5 M NaCl for 30 min. Prior to transfer the pH of the gel was neutralized with 0.5 M Tris-HCl and 1.5 M NaCl for 25 min. The DNA was transferred overnight (12 h) onto a positively charged nylon membrane (Amersham, NJ, USA) using 20X SSC. The DNA was immobilized on the membrane through exposure to UV at 120,000 microjoules/cm² using a Stratalinker UV crosslinker 1800 (Stratagene, La Jolla, CA, USA). Probes for virus detection were PCR amplified using primers listed on Table 1 targeting replicase and movement protein gene. Probes were PCR labelled using digoxigenin (DIG) DNA labelling kit (Roche, Indianapolis, IN, USA) as recommended by the manufacturer. Membranes were hybridized overnight at 55°C. Subsequent steps are previously described (Patil *et al.*, 2015).

Small RNA library preparation

Small RNA libraries were prepared using NEBNext® Multiplex Small RNA Library Prep Set for Illumina. The libraries were submitted to the Genome Technology Access Center (GTAC), Washington University in St. Louis, Missouri, USA followed by quality control on Agilent 2100 Bioanalyzer (Agilent Technologies, Inc., Santa Clara, CA). The small RNA libraries were subjected to Illumina HiSeq 2,500 using 1 × 50 single-end read protocol. After sequencing raw reads were downloaded from GTAC website and de-multiplexed by QIIME (Caporaso *et al.*, 2010) and reads with quality score below 19 were discarded. Adapter sequences were trimmed using Cutadapt (Martin, 2011). Small RNA sequences in the size range of 21–24 nt were selected and mapped to the virus reference genome with zero mismatches. Statistical analysis was performed using BEDTools (Quinlan and Hall, 2010) and all outputs were graphically presented by Shell scripts (Fahlgren *et al.*, 2009).

Results

Response of *N benthamiana* to inoculation with ACMV-CM and EACMV KE2 (K201)

The infection dynamics of two species of cassava-infecting geminiviruses were studied in *N benthamiana* through agro-inoculation. The onset of CMD symptoms on *N benthamiana* leaves was observed at 5 and 10 dpi for ACMV-CM and EACMV KE2 (K201), respectively (Fig. 1). Co-inoculation with both ACMV-CM and EACMV KE2 (K201) produced the most severe symptoms starting from 5 dpi resulting into leaf curling, yellow mosaic and plant stunting (Fig. 1). Based on the intensity of the Northern and Southern blots ACMV-CM seem to have accumulated to higher amounts at 10 dpi compared with EACMV KE2 (K201) correlating with more severe symptoms at this time point. However, after 20 dpi plants infected with ACMV-CM displayed symptoms remission characterized by appearance of non-asymptomatic leaves (Fig. 1). Contrastingly EACMV KE2 (K201) induced symptoms persisted throughout the experimental period (Fig. 1).

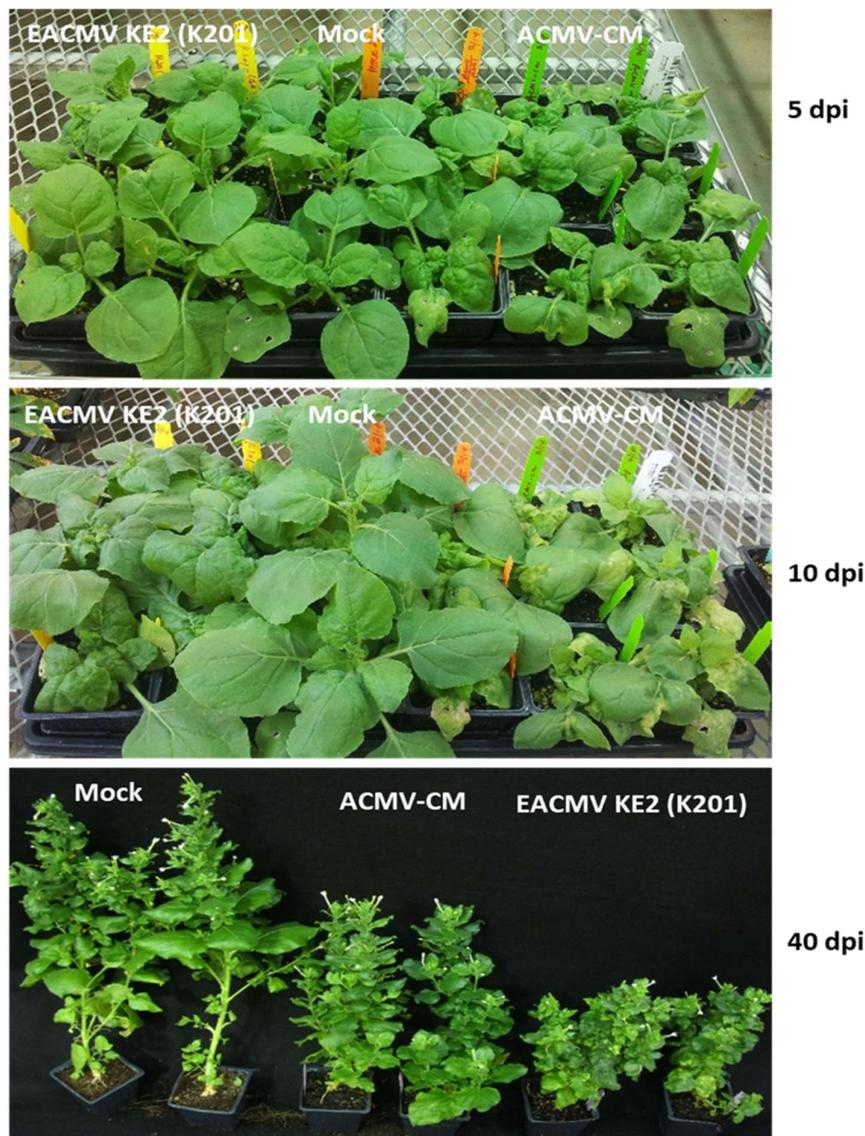


Figure 1 *Nicotiana benthamiana* plants inoculated with ACMV CM and EACMV KE2 (K201). The onset of CMD symptoms were displayed at 5 dpi for ACMV-CM and 10 dpi for EACMV KE2 (K201). Co-inoculation with ACMV CM and EACMV KE2 (K201) produced more severe symptoms compared with ACMV CM and EACMV KE2 (K201) individually. Symptoms recovery was observed after 20 dpi in plants infected with ACMV-CM.

Accumulation of virus mRNAs, siRNAs and DNA was detected in plants showing systemic CMD symptoms at 10 dpi (Fig. 2). At 10 dpi, there was direct correlation between virus titer, virus derived siRNAs, mRNAs and symptom expression (Figs. 1, and 2). The DNA conformations of the two virus species differed whereby plants systemically infected with ACMV-CM accumulated more single stranded DNA (ssDNA) forms than those infected with EACMV KE2 (K201) (Fig. 2C) indicating active replication and encapsidation of ACMV-CM virus. However, there was no obvious differences in the accumulation of double stranded DNA (dsDNA) forms for the two virus species (Fig. 2C). The two viral species were seen to accumulate differently whereby ACMV-CM accumulated more viral mRNAs, siRNAs and DNA compared with EACMV KE2 (K201) producing severe symptoms (Fig. 2).

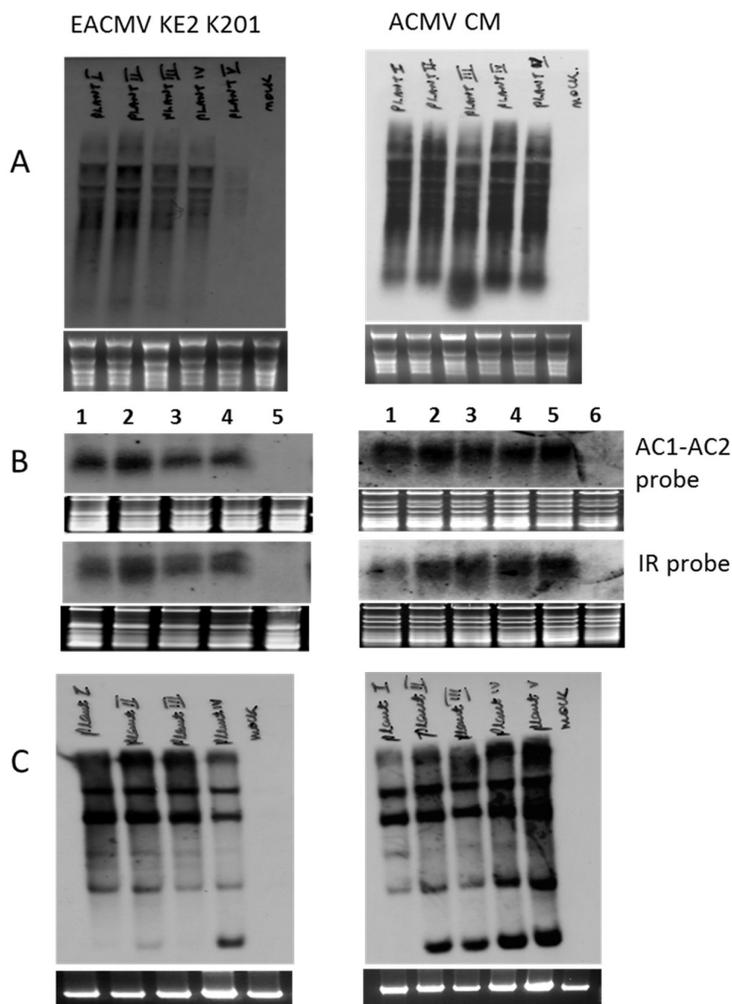


Figure 2 Detection of ACMV CM and EACMV KE2 (K201) systemically infected leaves of in individual *Nicotiana benthamiana* plants at 20 dpi. (A) Accumulation of virus mRNAs as detected on RNA denaturation gel. (B) Virus derived siRNAs were detected in all plants showing systemic symptoms. (C) Virus DNA titer detected on systemically infected leaves using Southern blot.

Deep sequencing of small RNAs in geminiviruses infected *Nicotiana benthamiana*

Sequence data showing the average total small RNAs reads obtained from *Nicotiana benthamiana* leaves systemically infected with ACMV-CM and EACMV KE2 (K201) respectively is presented in Table 1. After adaptor removal the average total of 21 - 24 nt reads with Phred Quality score of above 20 were between 969,015 and 1,307,689 for libraries constructed from EACMV KE2 (K201) and ACMV-CM infected leaf tissues. Of the total small RNAs reads >80% mapped to the host genome. In *Nicotiana benthamiana* systemically infected with EACMV KE2 (K201), 12.31% and 4.51% reads were mapped to the viral DNA A and DNA B (Table 1). On the other hand *Nicotiana benthamiana* systemically infected with ACMV-CM DNA accumulated 5.70% and 0.45% vsRNAs reads mapping to ACMV-CM DNA A and DNA B components (Table 1). For both virus species, majority of vsRNAs were mapping to DNA A component. However, EACMV KE2 (K201) derived small RNAs were proportionally higher than those derived from ACMV-CM (Table 1).

Table 1 Total number of small RNAs reads from EACMV KE2 (K201) and ACMV-CM mapped to viral

Host	Tissue	Virus species	Symptom severity	Total raw reads	Total clean reads	Reads mapping viral genome A	Reads mapping viral genome B	Reads mapping host genome
<i>N. benthamiana</i>	leaf	EACMV KE2 (K201)	4	4408572	969015	119382 (12.31%)	44319 (4.51%)	805314 (83.11%)
						130768	74567 (5.70%)	1226671 (93.80%)
<i>N. benthamiana</i>	leaf	ACMV CM	3	5341422	9			

Characterization of small RNAs derived from ACMV-CM and EACMV KE2 (K201) respectively

The total sRNAs reads of 21 - 24 nt in length from libraries prepared from EACMV KE2 (K201) and ACMV-CM infected leaf samples were mapped to the respective viral genome with zero mismatches. Different classes of vsRNAs showed a differential accumulation for both virus species with 22 nt being the most abundant (Fig. 3). The total mapped reads of 22 nt vsRNAs varied between the two viruses, DNA components and polarity. EACMV K201 derived sRNAs ranged between 9888 and 42501 reads compared with 2828 and 34045 reads for ACMV (Fig. 3). Overall for both viruses DNA A accumulated the most vsRNAs compared with DNA B but for both components there was no apparent differences in polarity for all populations of vsRNAs (Fig. 3).

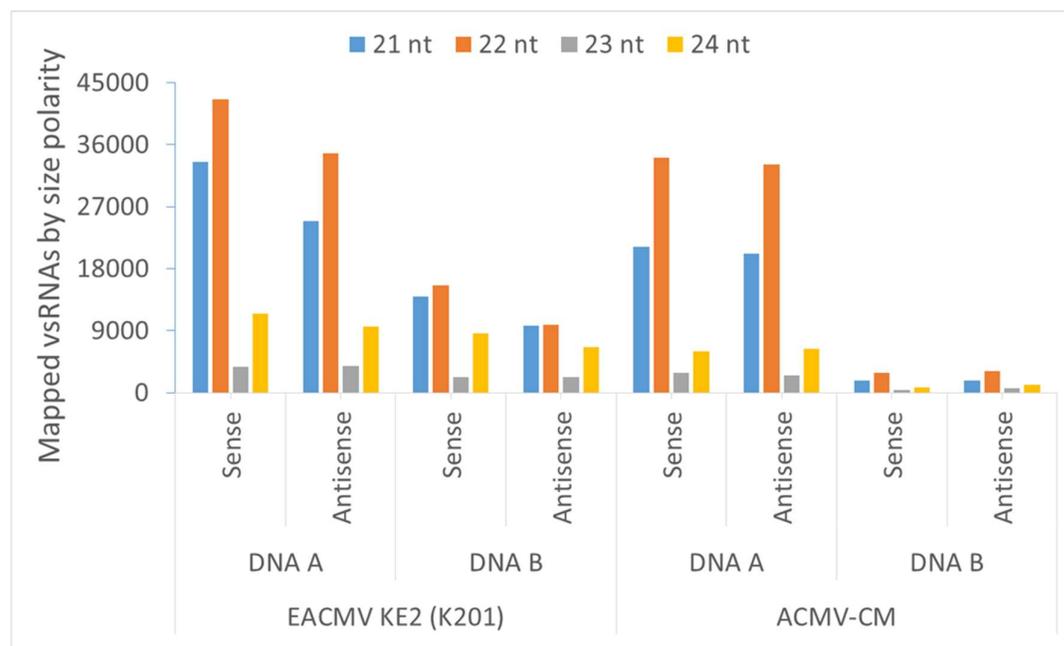


Figure 3 Illumina deep sequencing of small RNAs in *N. benthamiana* infected with EACMV KE2 (K201) and ACMV-CM respectively. The graphs shows total vsRNAs of size-classes 21-24 nt vsRNAs mapped to the virus genome with zero mismatches.

For both virus species, 22 nt vsRNAs reads were the most abundant (34% - 53%) followed by 21 nt (27% - 27%) and 24 nt (9 - 23%) vsRNAs reads (Fig. 4). The proportion of vsRNAs derived from sense and antisense polarities was almost equal for the two genomic components of the two virus species (Fig. 4).

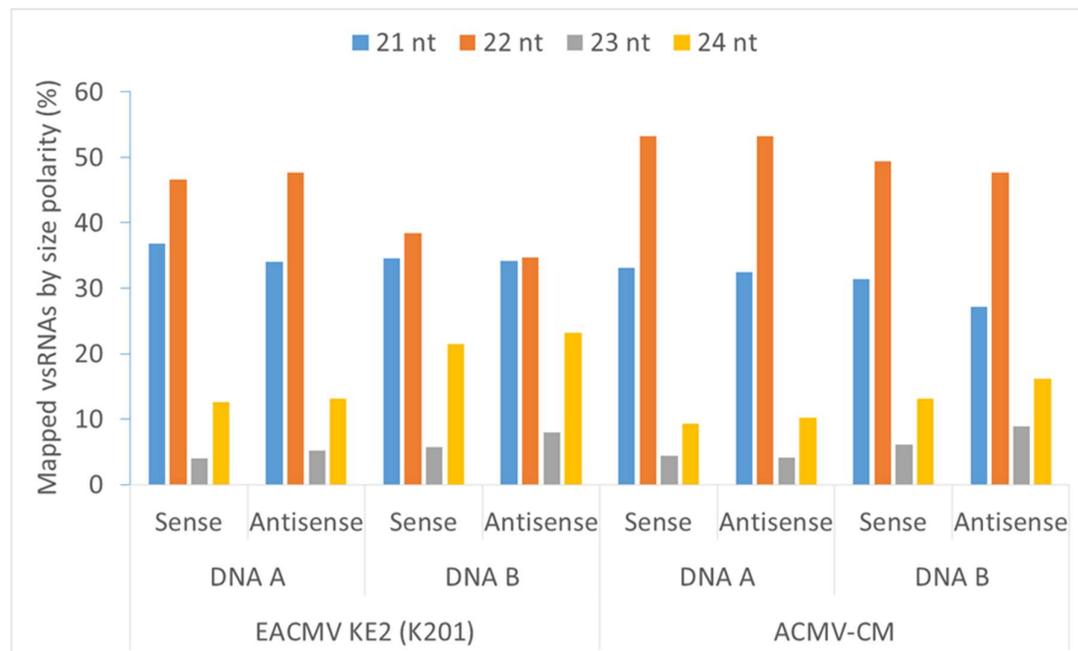


Figure 4 Abundance of EACMV KE2 (K201) and ACMV CM vsRNAs in inoculated *Nicotiana benthamiana*. Bars represent reads matched to sense or antisense strands of viral genomic DNA, respectively.

Distribution and frequency of vsRNAs along ACMV-CM and EACMV KE2 (K201) genomes

Single base resolution maps of all ACMV-CM and EACMV KE2 (K201) redundant vsRNAs reads revealed that vsRNAs densely covered the entire virus genome in sense and antisense polarities. However, several hotspots of 21 22 and 24 nt vsRNAs were identified within the overlapping regions of AC3/AC2/AC1 and AC1/AC4 of EACMV KE2 (K201) DNA A (Fig. 5A). The most prominent vsRNAs peaks of EACMV KE2 (K201) DNA B were mostly occupied by 24 nt vsRNAs and were identified within BC1 and IR regions (Figs 5B). On the other hand two prominent peaks of 22 nt ACMV-CM vsRNAs were found in overlapping regions of AV1/AV2 and AC1/AC2/AC3 (Fig. 5C). On ACMV-CM DNA B 22 and 24 nt vsRNAs formed nine prominent peaks on BV and BC1 respectively (Fig. 5C). The promoter regions of ACMV-CM DNA A and DNA B and certain transcriptional units were poorly covered with vsRNAs (Fig. 5). Gaps were identified in the following nucleotide positions; 1-134, 659-760, 850-882, 1158-1262, 2581-2606, 1964-2099, 2659-2777 of ACMV-CM DNA A (Fig. 5C) and 1-604, 683-736, 786-925, 975-1010, 1199-1246, 1289-1338, 1430-1459, 1542-1566, 2199-2473, 2630-2704 of ACMV-CM DNA B (Fig. 5D).

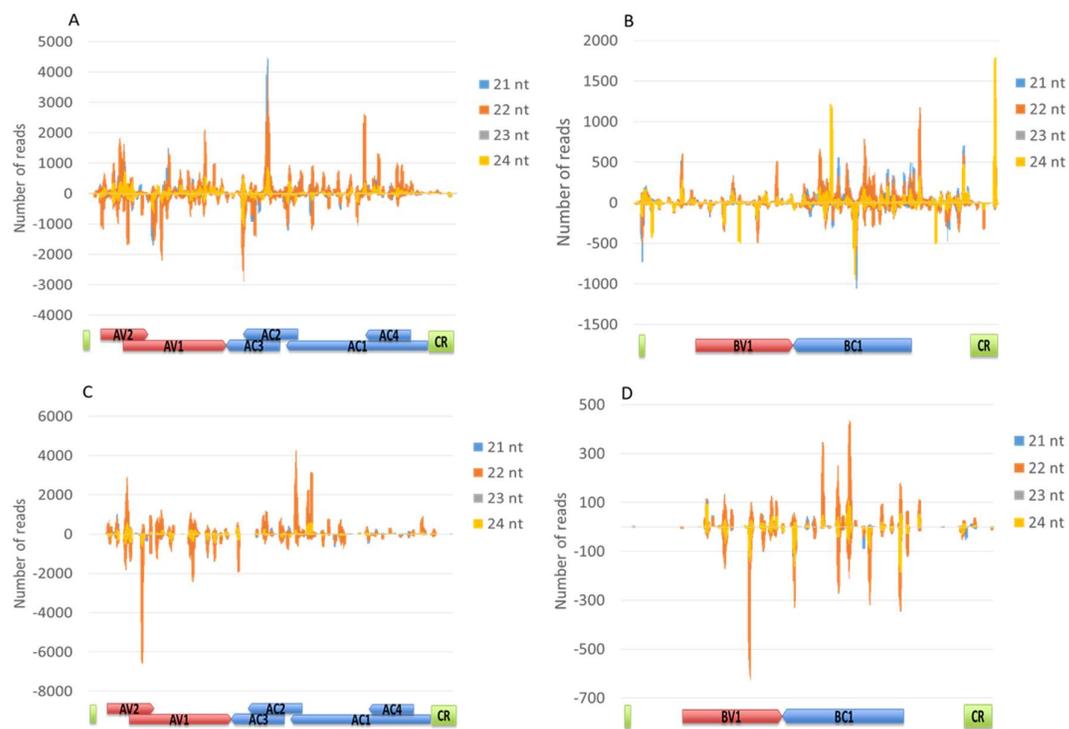


Figure 5 Profiles of 21–24 nucleotide virus-derived small RNAs on each of the genomic segments of EACMV KE2 (K201) and ACMV-CM. The graphs represents the number of 21-24 nt virus-derived small RNAs at each position along the (A) EACMV KE2 (K201) DNA A (B) EACMV KE2 (K201) DNA B (C) ACMV-CM DNA A and (D) ACMV-CM DNA B in positive and negative polarities.

The density of vsRNAs relative to each virus genome transcription units was evaluated. High abundance of EACMV KE2 (K201) vsRNAs was observed in AV1, AV2, AC3, AC1 and BC1 (Fig. 6A), with both polarities equally represented for AV1, AV2 and AC3 while reads from AC1 and BC1 were mostly derived from sense genomic region. For ACMV-CM the most abundant vsRNAs reads were mapped to antisense region of AV1 and AV2 and sense region of AC1 (Fig. 6B).

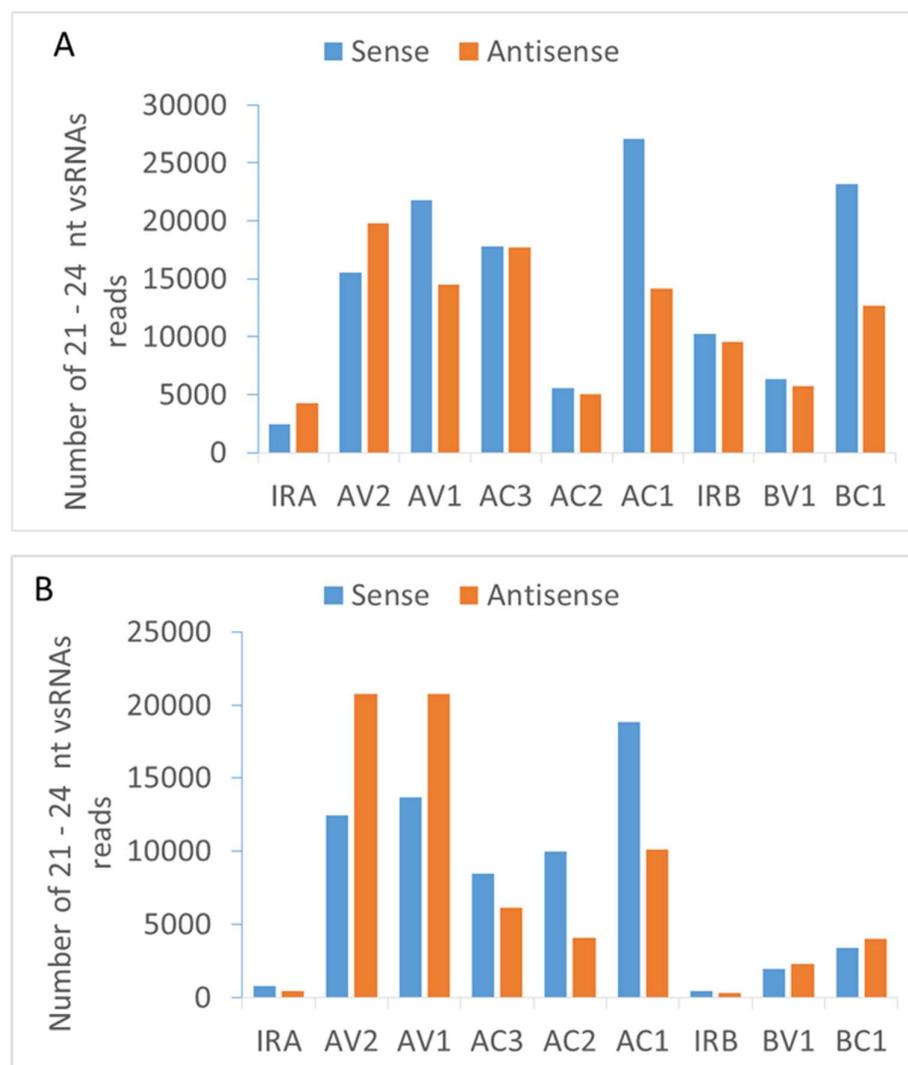


Figure 6 Distribution of (A) EACMV KE2 (K201) and (B) ACMV-CM vsRNAs within the virus transcriptional units

The relative abundance of the four different 5'-terminal nucleotides of cassava-infecting geminiviruses

The AGO protein complex incorporate one strand of sRNAs to guide in silencing target mRNAs. The association between AGO and sRNAs is mostly determined by first 5' nucleotide (Carbonell and Carrington 2015; Garcia-Ruiz *et al.*, 2015). The interaction of vsRNAs with various AGOs encoded by *N benthamiana* genome was predicted based on relative abundance of the vsRNAs 5'terminal nucleotide (Fig. 7). Overall, vsRNAs with preference to start with G as their 5'terminal nucleotide were the lowest (6 - 12%) for the two virus species (Fig. 6A and B). The majority of EACMV KE2 (K201) vsRNAs had U (36 - 57%) followed by A (22 - 37%) as the most abundance 5' end nucleotide (Fig. 7A). The ACMV-CM derived small RNAs containing U (35 - 58%) at the 5' end were the most abundant followed by A (17 - 44%) and C (13 - 18%) (Fig. 7B). Analysis of 21, 22 and 24 nt vsRNAs showed tendency to start with U (>35%) and A (>17%) for the two viruses an indication of sorting of vsRNAs into AGO4 AGO3 and AGO2 homologues.

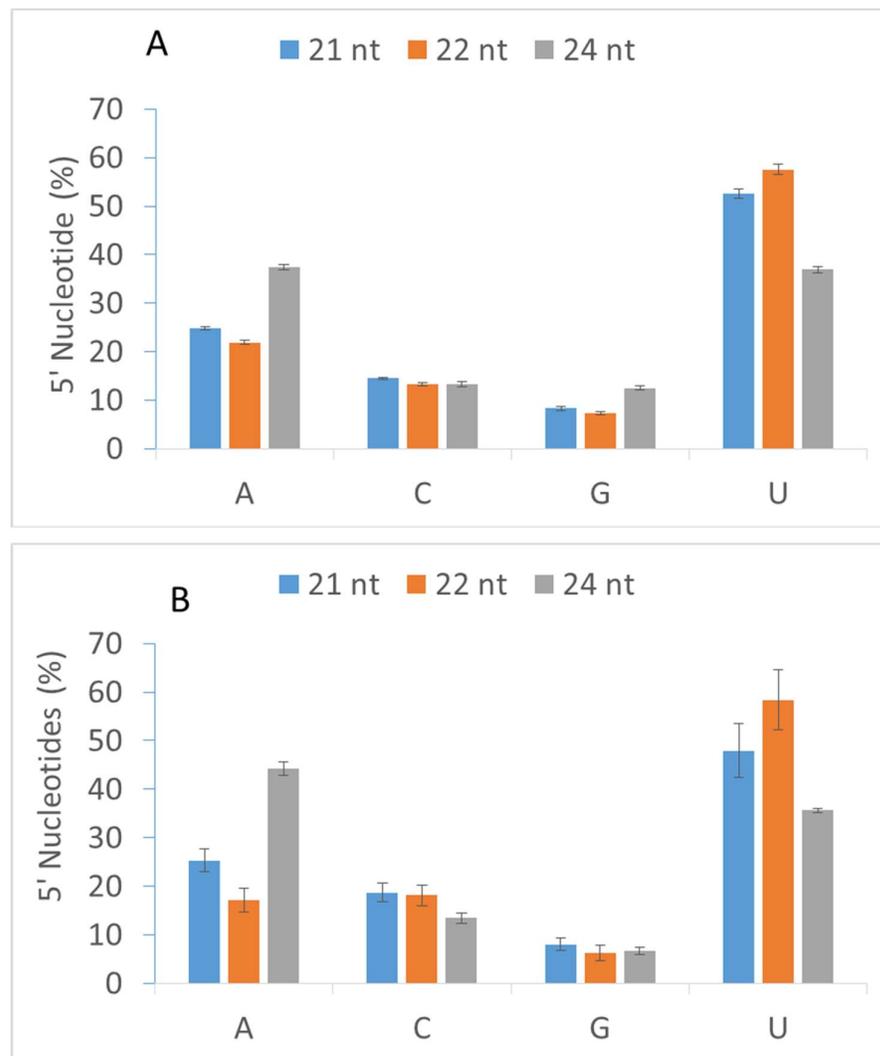


Figure 7 Relative frequencies of (A) EACMV KE2 (K201) and (B) ACMV-CM 5' terminal nucleotide

DISCUSSION

The interaction between cassava-infecting geminiviruses and host plant *N benthamiana* was studied through agro-inoculation. Differential response was observed between the two viral species. *N benthamiana* plants infected with ACMV-CM expressed symptoms after 5 dpi while those infected with EACMV KE2 (K201) developed symptoms starting at 10 dpi. Previous studies have demonstrated different symptom dynamics in *N benthamiana* plants with ACMV like viruses inducing symptoms at 4 dpi and EACMV like viruses symptoms appearing after 8 dpi (Patil and Fauquet, 2015). Studies by Shen *et al.* (2014) demonstrated that phosphorylation of AC2 proteins from *Cabbage leaf curl virus* (CaLCuV) and *Tomato mottle virus* (ToMoV) by SnRK, a signaling kinase involved in regulating sugar metabolism in plants resulting in delayed symptom development in *Arabidopsis thaliana* plants.

Data presented here showed recovery from ACMV-CM infection beginning 20 dpi whereas EACMV KE2 (K201) symptoms were persistent throughout the experimental period. A range of symptom phenotypes have been previously reported for various cassava-infecting geminiviruses in different host plants (Patil and Fauquet, 2015). Past studies have associated viral load and symptoms severity with virus pathogenicity (Sun *et al.*, 2015). The results described here showed that EACMV KE2 (K201) induced more severe persistent symptoms compared with ACMV-CM. EACMV-like virus have been shown to encode suppressors of gene silencing that target both post transcriptional gene silencing (PTGS) and transcriptional gene silencing (TGS) pathways whereas ACMV-CM encoded suppressors only subvert PTGS (Vanitharani *et al.*, 2004).

Sequence analysis of libraries derived from *N benthamiana* infected with cassava-infecting geminiviruses

identified 4-12 % sRNAs reads mapping to EACMV KE2 and 0.5% - 5% sRNAs reads mapping to the ACMV-CM genome. Similar results have been reported in tomato plants infected with *Tomato yellow leaf curl virus* (Bai *et al.*, 2016). However, other authors have reported wide variation (3 to 64%) of geminiviruses vsRNAs and this mostly depend with the host plant (Akbergenov *et al.* 2006; Donaire *et al.* 2009; Qi *et al.* 2009; Wang *et al.* 2010).

In *N. benthamiana* plants infected with ACMV-CM and EACMV KE2 respectively, the most frequent class of vsRNAs were 22 nt, followed by 21 nt and 24. Similar results have been reported by Akbergenov *et al.*, (2006) using blot based assays. Deep sequencing results previously reported by Yang *et al.*, (2011) on *Tomato yellow leaf curl China virus* (TYLCCNV) -infected *S. lycopersicum* and *N. benthamiana* plants identified 22 nt vsRNAs as the most abundant reads. On the contrary, 24 nt and 21 nt vsRNAs were reported to dominate in CaLCuV and CaMV infected *Arabidopsis* plants (Blevins *et al.* 2011; Aregger *et al.* 2012). Therefore, the processing of vsRNAs differ within the species of virus and host plants. The various DCLs process dsRNAs into distinct sizes (Blevins *et al.*, 2006). Data presented here showed accumulation of 22 nt, 21 nt and 24 nt vsRNAs indicating involvement of several *N. benthamiana* DCL homologues in biogenesis of vsRNAs.

Genome wide resolution maps revealed heterogeneous distribution of vsRNAs throughout the virus genome in sense and antisense polarities. This collaborate well with formation of long dsRNA covering the entire virus genome in sense and antisense direction as a result of POI II-mediated bidirectional transcription (Poogin *et al.*, 2016; 2013). Readthrough transcripts and/or their degradation products are processed by DCLs into distinct classes of vsRNAs covering the entire virus genome (Seguin *et al.*, 2014). Several hotspot for were identified within the coding regions and were more pronounced within the overlapping regions of DNA A components for both virus species as previously described (Bai *et al.*, 2016; Miozzi *et al.* 2013; Yang *et al.* 2011). However, for DNA B components hotspots were identified in non-overlapping BC1 and BV1 regions (Fig. 5B and D). One prominent 24 nt vsRNAs peak was identified within the promoter region of EACMV KE2 (K201) sense strand. The 24 nt vsRNAs potentially direct transcriptional silencing of virus BC1 as previously demonstrated (Poogin *et al.*, 2016). The distinct size classes of vsRNAs were localized in the same hotspots indicating targeting of same regions along viral genome by different DCLs. Comparison of ACMV-CM and EACMV KE2 (K201) vsRNAs distribution along the viral genome revealed that certain regions in ACMV-CM genome such as AC1, AC4 and non-coding regions of DNA B were poorly saturated with vsRNAs. Interaction between sRNAs and AGO proteins is mostly determined by 5' terminal nucleotide (Fang and Qi 2016; Carbonell and Carrington, 2015). For both virus species, vsRNAs exhibited a strong bias to uridine 21 nt (52.5%), 22 nt (57.6%) and 24 nt (36.9%) for EACMV KE2 (K201) and 21 nt (47.9%), 22 nt (58.4%) and adenosine 24 nt (44.3%) for ACMV-CM. This result support the participation of AGO1, AGO2 and AGO4 for geminiviruses defense in *N. benthamiana* and is consistent with previous studies (Piedra-Aguilera *et al.*, 2019; Wu *et al.*, 2019; Margaria *et al.*, 2016). Loading of vsRNAs into diverse AGOs mediated by the 5'-terminal U and A may play a role in targeting expression of multiple host genes involved in different processes and pathways that enhance virus infection. Recently, Wu *et al.* (2019) identified 19 vsRNAs from *Tobacco curly shoot virus* (TbCSV) that inhibit expression of complementary cellular transcripts involved in molecular functions and biological processes and enhances expression of RNA-dependent RNA polymerase 1 (RDR1) resulting in ideal conditions for viral infection. Furthermore, vsRNAs derived potato spindle tuber viroid (PSTVd)-derived vsRNAs degrade two callose synthase genes exacerbating disease severity and accumulation of viroids (Adkar-Purushothama *et al.*, 2015). Similarly, peach latent mosaic viroid (PLMVd)-derived siRNAs target the chloroplastic heat shock protein 90 gene in peach to induce albinism and potentially create a favorable host environment for viroid infection (Navarro *et al.*, 2012). To further understand geminiviruses and host interaction we suggest AGO vsRNAs immunoprecipitation studies.

Conclusion

This study demonstrate infection dynamics of important viruses that inhibit potential yields in cassava. The profiles of EACMV KE2 (K201) and ACMV-CM derived small RNAs were different in *N. benthamiana* whereby more vsRNAs were derived from EACMV KE2 (K201) indicating differences in virus pathogenicity. The outcome of this study highlight possible role of RNA silencing pathway in geminiviruses pathogenicity.

References

- Adkar-Purushothama, C.R., Brosseau, C., Giguere, T., Sano, T., Moffett, P., Perreault, J.P., 2015. Small RNA derived from the virulence modulating region of the potato spindle tuber viroid silences callose synthase genes of tomato plants. *Plant Cell* 27 (8), 2178–2194.
- Axtell, M. J. (2013). Classification and comparison of small RNAs from plants. *Annual review of plant biology*, 64, 137-159.
- Bai, M., Yang, G. S., Chen, W. T., Lin, R. M., Ling, J., Mao, Z. C., & Xie, B. Y. (2016). Characterization and function of Tomato yellow leaf curl virus-derived small RNAs generated in tolerant and susceptible tomato varieties.
- Bai, M., Yang, G. S., Chen, W. T., Mao, Z. C., Kang, H. X., Chen, G. H., Yang, Y. H., & Xie, B. Y. (2012). Genome-wide identification of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families and their expression analyses in response to viral infection and abiotic stresses in *Solanum lycopersicum*. *Gene*, 501(1), 52-62.
- Brosseau, C., & Moffett, P. (2015). Functional and genetic analysis identify a role for *Arabidopsis* ARGONAUTES5 in antiviral RNA silencing. *The Plant Cell*, 27(6), 1742-1754.
- Brown, J. K., Zerbini, F. M., Navas-Castillo, J., Moriones, E., Ramos-Sobrinho, R., Silva, J. C., Fiallo-Olivé, E., Briddon, R. W., Hernández-Zepeda, C., Idris, A., Malathi, V. G., Martin, D. P., Rivera-Bustamante, R., Ueda, S., & Varsani, A. (2015). Revision of Begomovirus taxonomy based on pairwise sequence comparisons. *Archives of virology*, 160(6), 1593-1619.
- Cao, M., Du, P., Wang, X., Yu, Y. Q., Qiu, Y. H., Li, W., ... & Ding, S. W. (2014). Virus infection triggers widespread silencing of host genes by a distinct class of endogenous siRNAs in *Arabidopsis*. *Proceedings of the National Academy of Sciences*, 111(40), 14613-14618.
- Carbonell, A., & Carrington, J. C. (2015). Antiviral roles of plant ARGONAUTES. *Current opinion in plant biology*, 27, 111-117.
- Rybicki, E.P., 2015. A Top Ten list for economically important plant viruses. *Archives of virology*, 160(1), 17-20.
- Csorba T, Pantaleo V, Burgyan J (2009) RNA silencing: an antiviral mechanism. *Adv Virus Res* 75:35–71.
- Deleris A, Gallego-Bartolome J, Bao J, Kasschau KD, Carrington JC, Voinnet O (2006) Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. *Science* 313:68–71
- Deleris, A., Gallego-Bartolome, J., Bao, J., Kasschau, K. D., Carrington, J. C., & Voinnet, O. (2006). Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. *Science*, 313(5783), 68-71.
- Ding SW, Voinnet O (2007) Antiviral immunity directed by small RNAs. *Cell* 130:413–426
- Fang, X., & Qi, Y. (2016). RNAi in Plants: An Argonaute-Centered View. *The Plant Cell*, T
- Garcia-Ruiz, H., Carbonell, A., Hoyer, J. S., Fahlgren, N., Gilbert, K. B., Takeda, A., ... & Baladejo, M. T. M. (2015). Roles and programming of *Arabidopsis* ARGONAUTE proteins during Turnip mosaic virus infection. *PLoS Pathog*, 11(3), e1004755.
- Garcia-Ruiz, H., Takeda, A., Chapman, E. J., Sullivan, C. M., Fahlgren, N., Brempelis, K. J., & Carrington, J. C. (2010). *Arabidopsis* RNA-dependent RNA polymerases and dicer-like proteins in antiviral defense and small interfering RNA biogenesis during Turnip Mosaic Virus infection. *The Plant Cell*, 22(2), 481-496.
- Hanley-Bowdoin, L., Bejarano, E. R., Robertson, D., & Mansoor, S. (2013). Geminiviruses: masters at redirecting and reprogramming plant processes. *Nature Reviews Microbiology*, 11(11), 777-788.
- Havecker E.R., Wallbridge L.M., Hardcastle T.J., Bush M.S., Kelly K.A., Dunn R.M., Schwach F., Doonan J.H., Baulcombe D.C. (2010). The *Arabidopsis* RNA-directed DNA methylation argonautes functionally diverge based on their expression and interaction with target loci. *Plant Cell* 22: 321–334.
- Incarbone, M., & Dunoyer, P. (2013). RNA silencing and its suppression: novel insights from in planta analyses. *Trends in plant science*, 18(7), 382-392.
- Kapoor, M., Arora, R., Lama, T., Nijhawan, A., Khurana, J. P., Tyagi, A. K., & Kapoor, S. (2008). Genome-wide identification, organization and phylogenetic analysis of Dicer-like, Argonaute and RNA-dependent RNA Polymerase gene families and their expression analysis during reproductive development and stress in rice. *BMC genomics*, 9(1), 451.

- Legg, J. P., & Thresh, J. M. (2000). Cassava mosaic virus disease in East Africa: a dynamic disease in a changing environment. *Virus research*, 71(1), 135-149.
- Leuzinger, K., Dent, M., Hurtado, J., Stahnke, J., Lai, H., Zhou, X., & Chen, Q. (2013). Efficient agroinfiltration of plants for high-level transient expression of recombinant proteins. *JoVE (Journal of Visualized Experiments)*, (77), e50521-e50521.
- Ma, X., Nicole, M. C., Meteignier, L. V., Hong, N., Wang, G., & Moffett, P. (2014). Different roles for RNA silencing and RNA processing components in virus recovery and virus-induced gene silencing in plants. *Journal of experimental botany*, eru447.
- Margaria, P., Miozzi, L., Ciuffò, M., Rosa, C., Axtell, M. J., Pappu, H. R., & Turina, M. (2016). Comparison of small RNA profiles in *Nicotiana benthamiana* and *Solanum lycopersicum* infected by *Polygonum ringspot tospovirus* reveals host-specific responses to viral infection. *Virus research*, 211, 38-45.
- Margis R, Fusaro AF, Smith NA, Curtin SJ, Watson JM, Finnegan EJ, Waterhouse PM (2006) The evolution and diversification of Dicers in plants. *FEBS Lett* 580:2442–2450
- Martínez de Alba AE, Elvira-Matelot E, Vaucheret H (2013) Gene silencing in plants: A diversity of pathways. *Biochim Biophys Acta* 1829(12):1300–1308.
- Navarro, B., Gisel, A., Rodio, M.E., Delgado, S., Flores, R. and Di Serio, F., 2012. Small RNAs containing the pathogenic determinant of a chloroplast-replicating viroid guide the degradation of a host mRNA as predicted by RNA silencing. *The Plant Journal*, 70(6), pp.991-1003.
- Ndunguru, J., De León, L., Doyle, C. D., Sseruwagi, P., Plata, G., Legg, J. P., Thompson, G., Tohme, J., Aveling, T., Ascencio-Ibáñez, J. T., & Hanley-Bowdoin, L. (2016). Two Novel DNAs That Enhance Symptoms and Overcome CMD2 Resistance to Cassava Mosaic Disease. *Journal of virology*, 90(8), 4160-4173.
- Olsen, K., Schaal, M., Barbara, A., 1999. Evidence on the origin of cassava: phylogeography of *Manihot esculenta*. *Proc. Natl. Acad. Sci. U. S. A. (PNAS)* 96 (10), 5587e5590.
- Patil, B. L., & Fauquet, C. M. (2015). Studies on differential behavior of cassava mosaic geminivirus DNA components, symptom recovery patterns, and their siRNA profiles. *Virus genes*, 50(3), 474-486.
- Poulsen, C., Vaucheret, H., & Brodersen, P. (2013). Lessons on RNA silencing mechanisms in plants from eukaryotic argonaute structures. *The Plant Cell*, 25(1), 22-37.
- Qian, Y., Cheng, Y., Cheng, X., Jiang, H., Zhu, S., & Cheng, B. (2011). Identification and characterization of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in maize. *Plant cell reports*, 30(7), 1347-1363.
- Seguin J, Rajeswaran R, Malpica-López N, Martin RR, Kasschau K, Dolja VV, Otten P, Farinelli L, Pooggin MM (2014) De novo reconstruction of consensus master genomes of plant RNA and DNA viruses from siRNAs. *PLoS One* 9:e88513
- Shen, W., Dallas, M. B., Goshe, M. B., & Hanley-Bowdoin, L. (2014). SnRK1 phosphorylation of AL2 delays cabbage leaf curl virus infection in *Arabidopsis*. *Journal of virology*, 88(18), 10598-10612.
- Sun, Y. W., Tee, C. S., Ma, Y. H., Wang, G., Yao, X. M., & Ye, J. (2015). Attenuation of Histone Methyltransferase KRYPTONITE-mediated transcriptional gene silencing by Geminivirus. *Scientific reports*, 5.
- Wang XB, Wu QF, Ito T, Cillo F, Li WX, Chen X, Yu JL, Ding SW (2010) RNAi-mediated viral immunity requires amplification of virus-derived siRNAs in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* 107:484–489
- Wu, G., Hu, Q., Du, J., Li, K., Sun, M., Jing, C., Li, M., Li, J. and Qing, L., 2019. Molecular characterization of virus-derived small RNAs in *Nicotiana benthamiana* plants infected with tobacco curly shoot virus and its β satellite. *Virus research*, 265, pp.10-19.
- Yang J, Zheng SL, Zhang HM, Liu XY, Li J, Li JM, Chen JP (2014) Analysis of small RNAs derived from Chinese wheat mosaic virus. *Arch Virol* 159:3077–3082
- Zhou, X., Liu, Y., Calvert, L., Munoz, C., Otim-Nape, G. W., Robinson, D. J., & Harrison, B. D. (1997). Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. *Journal of General Virology*, 78(8), 2101-2111.
- Piedra-Aguilera, Á., Jiao, C., Luna, A.P., Villanueva, F., Dabad, M., Esteve-Codina, A., Díaz-Pendón, J.A., Fei, Z., Bejarano, E.R. and Castillo, A.G., 2019. Integrated single-base resolution maps of transcriptome, sRNAome and methylome of Tomato yellow leaf curl virus (TYLCV) in tomato. *Scientific reports*, 9(1), pp.1-16.