SEGS-1 episomes generated during cassava mosaic disease enhance disease severity

Evangelista Chiunga[1,2,3‡], Catherine D. Aimone[2†‡],
Cyprian Rajabu[2], Mary M. Dallas[2], Joseph Ndunguru[4],
José T. Ascencio-Ibáñez[5]*, Elijah M. Ateka[1]
and Linda Hanley-Bowdoin[2]*

[1]Department of Horticulture, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya,
[2]Department of Plant and Microbial Biology, North Carolina State University, Raleigh, NC, United
States, [3]Tanzania Agricultural Research Institute (TARI)-Mikocheni, Dar es Salaam, Tanzania, [4]Tanzania
Plant Health and Pesticide Authority, Arusha, Tanzania, [5]Department of Molecular and Structural
Biochemistry, North Carolina State University, Raleigh, NC, United States

Cassava is an important root crop that is produced by smallholder farmers across
Sub-Saharan Africa. Cassava mosaic disease (CMD), which is caused by a group
of cassava mosaic begomoviruses (CMBs), is one of the most devastating
diseases of cassava. A previous study showed that SEGS-1 (sequences
enhancing geminivirus symptoms), which occur both in the cassava genome
and as episomes during CMD, can increase CMD disease severity and overcome
host resistance. In this report, we examined the effects of exogenously applied
SEGS-1 on the incidence of CMB infection, symptom severity, and viral DNA copy
number in five cassava cultivars that ranged from highly susceptible to highly
resistant to CMD. These studies revealed that the effect of SEGS-1 is cultivar
dependent. Susceptible cultivars developed severe CMD with or without
exogenous SEGS-1, while exogenous SEGS-1 increased disease severity in
cultivars carrying CMD2 resistance, which is conferred by a single locus, but
not CMD1 resistance, which is polygenic. Analysis of infected plants in the
absence of exogenous SEGS-1 revealed that some, but not all, cultivars form
SEGS-1 episomes during CMD. The presence of endogenous SEGS-1 episomes
in TME14, a CMD2 resistant cultivar, correlated with CMD severity. In contrast,
TME3, a closely related CMD2 cultivar, did not produce endogenous SEGS-1
episomes and was more resistance than TME14. The different capacities of TME3
and TME14 to form SEGS-1 episomes is unlikely due to sequence differences in
and around their genomic SEGS-1 loci. The functional regions of SEGS-1 were
mapped using TME3 to sequences flanking the episome junction, but the
junction itself was not required for activity. All cassava cultivars have SEGS-1
sequences in their genomes that have the potential to negatively impact the
development of stable CMD resistance by cassava breeding programs.

1 Introduction

Cassava (*Manihot esculenta* Crantz) is an important staple crop in many tropical and subtropical regions of the world (Bayata, 2019). More than half of cassava produced worldwide is grown in Sub-Saharan Africa region (FAOSTAT, 2021), where it is cultivated primarily by smallholder farmers. Cassava is a drought tolerant crop and can grow at high temperatures, but its production is severely limited by viral diseases (Okogbenin et al., 2013). Cassava mosaic disease (CMD) is endemic across the African continent and causes substantial crop losses on smallholder and commercial farms (Legg et al., 2011). CMD is caused by 11 DNA viruses collectively known as cassava mosaic begomoviruses (CMBs) (Crespo-Bellido et al., 2021). Nine CMB species occur in Africa, where they are transmitted by whiteflies (*Bemisia tabaci* Genn.) and by propagation of infected stem cuttings (Legg et al., 2015).

Begomoviruses comprise the largest genus in the *Geminiviridae*, a family of plant viruses with single-stranded DNA (ssDNA) genomes that are encapsidated into double icosahedral particles (Fiallo-Olivé and Navas-Castillo, 2020). CMBs have bipartite genomes that consist of two small, circular DNA molecules designated as DNA-A and DNA-B, both of which are required for systemic infection (Stanley and Gay, 1983). Together, the two genome components encode 8-9 canonical proteins that are involved in viral replication, transcription, encapsidation, movement, and countering host defenses (Yang et al., 2016). The viral proteins interact with diverse host proteins to perform their functions (Hanley-Bowdoin et al., 2013). Recent studies have suggested that the coding capacity of begomoviruses is not limited to the canonical open reading frames (ORFs), and that the number of viral proteins may be greater due to the presence of previously unidentified ORFs and alternative splicing (Gong et al., 2021).

Begomovirus genomes undergo mutation (nucleotide substitution and small indels), recombination, and reassortment (Duffy and Holmes, 2009; Crespo-Bellido et al., 2021; Aimone et al., 2021b). CMBs often occur in coinfections, in which different viruses can act synergistically and cause more severe disease (Pita et al., 2001). It is likely that the different CMB species arose via recombination between coinfecting viruses (Crespo-Bellido et al., 2021). In the 1990s and 2000s, synergy between African cassava mosaic virus (ACMV) and a recombinant CMB contributed to a severe CMD pandemic that spread from Uganda to other Sub-Saharan countries (Deng et al., 1997; Zhou et al., 1997; Pita et al., 2001). In response to the pandemic, many African farmers adopted cassava cultivars with the CMD2 locus, which confers resistance to CMBs (Akano et al., 2002; Rabbi et al., 2014). Recent studies have indicated that CMD2 resistance is due to mutations in the cassava gene encoding the catalytic subunit of the host DNA polymerase delta (Lim et al., 2022).

Two novel DNAs, designated SEGS-1 (sequences enhancing geminivirus symptoms; DNA-II; NCBI accession number: AY836366) and SEGS-2 (DNA-III; NCBI accession number: AY836367), were amplified from cassava plants showing severe CMD symptoms in Tanzania (Ndunguru et al., 2016). Laboratory studies showed that the presence of either SEGS-1 or SEGS-2 increases CMD symptom severity and that SEGS-1 can overcome CMD2 resistance (Ndunguru et al., 2016). SEGS-1 and SEGS-2 only show 23% overall sequence identity but display 99% and 84-87% identity, respectively, to sequences in the cassava genome. All cassava genomes examined to date encode a full-length copy of SEGS-1, but no full-length copy of SEGS-2 has been found. SEGS-1 and SEGS-2 also occur as small, circular DNA episomes in infected cassava plants. SEGS-2, but not SEGS-1, has been detected in virions in infected cassava and viruliferous whiteflies, suggesting that SEGS-1 and SEGS-2 are functionally distinct.

Begomovirus infection in Arabidopsis is also enhanced by SEGS-1 or SEGS-2 when either is provided as exogenous DNA or as a transgene (Aimone et al., 2021a; Rajabu et al., 2023). Experiments in Arabidopsis established that SEGS-2 activity is dependent on a small open reading frame. SEGS-2 occurs as both double and single-stranded episomes in ACMV-infected plants, and the single-stranded form is packaged into virions in both cassava and Arabidopsis. These results, in combination with results in tobacco cells showing that SEGS-2 replicates in the presence of ACMV DNA-A, established that SEGS-2 is a novel satellite (Ndunguru et al., 2016; Aimone et al., 2021a). In contrast, there is no evidence that SEGS-1 forms single or double-stranded episomes in Arabidopsis or replicates in tobacco cells co-transfected with ACMV DNA-A. Instead, a linear copy of SEGS-1 integrated into the Arabidopsis genome enhances ACMV infection (Rajabu et al., 2023), raising the possibility that the full-copy of SEGS-1 in the cassava genome also enhances CMD severity.

In the study reported here, we examined the effects of exogenous and endogenous SEGS-1 sequences on CMB infection in five cassava cultivars that ranged from highly susceptible to highly resistant to CMD. We asked if there is a correlation between the presence of SEGS-1 episomes and disease severity, determined the regions of SEGS-1 that mediate disease enhancement, and assessed the role of episome junction sequences on SEG-1 activity. Together, these experiments provided insight into the functional form of SEGS-1 in cassava and the effect of cassava genotype on SEGS-1 activity.

2 Materials and methods

2.1 SEGS-1 plasmids

Plasmids corresponding to the SEGS-1 monomer (S1-1.0; pNSB2000), and a SEGS-1 partial dimer with two copies of the GC-rich region (S1-1.5a; pNSB1829) were described previously (Rajabu et al., 2023). SEGS-1 regions G (positions 1–277), I (positions 275–646), J (positions 644–756), and F (positions 757–1007) (Supplementary Figure 1) were amplified in reactions containing S1-1.0 plasmid DNA as template. All the primer sets added *Not*I restriction sites at the 5' and 3'ends of the PCR products. G and F also included short sequences (G: 3 nt at 5' end; F: 13 nt at 3' end) that flank the cloned SEGS-1 sequence and are not present in the cassava genomic sequence. PCR was performed for 30 cycles (denaturation: 10

sec at 98°C; annealing: 30 sec at the temperatures in Supplementary Figure 1; extension: 2 min at 72°C), and the products were gel purified using a QIAquick Gel Extraction kit (Qiagen, Hilden, Germany). The purified products were digested with *Not*I and ligated into pMON721 linearized with *Not*I to generate the SEGS-1 G, I, J, F and N clones (Supplementary Figure 1).

To generate the SEGS-1 region N, the *Kpn*I site introduced during cloning between the F and G regions in S1-1.5a (pNSB1829) was deleted using the Q5-site directed mutagenesis protocol (New England Biolabs, Ipswich, MA) and the divergent primer pair, SEGS-1Nxfwd and SEGS-1Nxrev (Supplementary Table 1). The N region with an authentic junction sequence (SEGS-1 positions 757–1007 + 1–277) (Supplementary Figure 1) was amplified in reactions containing the modified S1-1.5a plasmid DNA as a template, and the PCR product was cloned as a *Not*I fragment into pMON721 as described above. All clones were confirmed by Sanger sequencing.

2.2 Cassava infection

Cassava plants (*Manihot esculenta* cv. Namikonga, Kibaha, TME14, TME3, and TMS30572) were propagated from stem cuttings and grown at 28°C under a 12-h light/dark cycle. Plants with 8-10 nodes and stems 1.5 cm in diameter (ca. 2 months after propagation) were inoculated near the apical meristem using a hand-held micro sprayer (40 psi) to deliver gold particles coated with four plasmid DNAs (1.7 μg/plasmid/plant) (Aimone et al., 2021b, 2022; Ascencio-Ibáñez et al., 2023). The clones contained partial tandem dimers corresponding to DNA-A or DNA-B of African cassava mosaic virus (ACMV; GenBank accessions MT858793.1 and MT858794.1) or East African cassava mosaic Cameroon virus (EACMCV; AF112354.1 and FJ826890.1) (Hoyer et al., 2020). ACMV and EACMCV were co-inoculated to generate coinfections with or without a SEGS-1 plasmid DNA. Mock controls were inoculated with the ACMV and EACMCV DNA-B plasmids using the protocol described above.

Experiments characterizing the effect of SEGS-1 on viral infection in different cultivars included ca. 30 plants for each treatment (ACMV+EACMCV or ACMV+EACMCV+S1-1.5a) and 3 mock-inoculated plants as negative controls. Experiments characterizing the impact of different SEGS-1 fragments on TME3 infection included 10 plants for each treatment and 3 mock-inoculated plants. All experiments were repeated a minimum of 2 times with similar results. Disease symptoms were monitored visually at 28 days post inoculation (dpi). Symptoms were scored using a disease severity scale (scale: 1 = no symptoms to 5 = very severe) in new growth (Supplementary Figure 2). The percentage of plants showing symptoms was calculated from the number of plants with symptom scores ≥ 2 relative to the total number of inoculated plants within a treatment. Average symptom scores were determined using all scores ≥ 2 within a treatment. The percent symptomatic plants and average symptom scores were compared between treatments using Fisher's exact test and Wilcoxon rank sum test, respectively, and a p value < 0.05 cutoff.

2.3 DNA extraction, viral DNA quantification, and SEGS-1 episome analysis

Six leaf punches from the base of the second visible leaf (L2) relative to the top of a cassava plant were sampled at 28 days post inoculation (dpi), flash-frozen in liquid nitrogen, and stored at -80˚C. The frozen tissue was pulverized using a homogenizer (MM 301, RETSCH-Laboratory Mills, Clifton, NJ), and total DNA was extracted using the MagMax™ Plant DNA Isolation Kit (Thermo Fisher Scientific, Waltham, MA). The copy numbers of ACMV DNA-A and EACMCV DNA-A were measured by quantitative PCR (qPCR) in total DNA samples as described previously (Aimone et al., 2022). The ACMV-A primer pair (P3P-AA2F and P3P-AA2R+4R) (Supplementary Table 1) was used to generate the copy number results in Figure 1. The ACMV-A primer pair (P3P-AA2F and P3P-AA2R+4R; Supplementary Table 1) was used to generate the ACMV DNA-A copy number results for Figures 2–5. EACMCV DNA-A was quantified using the primer pair, EACMVQ1 and EACMVQ2 (Supplementary Table 1). The qPCR reactions (25 μL) contained 5 μL (10 μM) of forward + reverse primers, 2.5 μL ultrapure water, 10 μL of Power SYBR Green PCR Master Mix (Applied Biosystems, Foster City CA), and 2.5 μL of total DNA (10 ng) from infected cassava leaves. After an initial denaturation step at 95°C for 2 min, 30 amplification cycles were performed with each consisting of 15 s at 94°C, 1 min at 60°C and 30 s at 72°C. Three technical replicates were generated for each cassava DNA sample. Viral DNA was quantified using a qPCR standard curve generated from a 10-fold dilution series (10^{-8} to 10^{-14} ng/μL) of plasmid DNA containing either a single copy of ACMV DNA-A or EACMCV DNA-A. A standard curve, which was generated in triplicate on the same 98-well plate as the test samples, was used to determine the $\log_{10}$ copy number of DNA-A/μg total DNA (Aimone et al., 2022). Viral DNA copy number was compared between treatments using a two-tailed Students' T test and a p-value < 0.05 cutoff.

Analysis of variance (ANOVA) was performed using the Proc GLM procedure in SAS version 9.4. to test for effects of SEGS-1 and CMB treatments on DNA-A copy number in TME14, TME3, and TMS30572 (Supplementary Table 2). Data were log transformed prior to analysis to stabilize variance. Separate analyses were conducted for each of the cultivars because they were tested in separate experiments. The analyses included replicate to account for variation between experiments with the same variety. An initial analysis showed that SEGS-1*virus interaction effects were not significant for any variety, and the interaction term was not included in subsequent analyses. The final model tested was: $\log_{10}$ copy number = replicate + SEGS-1 + virus + error.

The presence of SEGS-1 episomes was assessed in total DNA samples from cassava plants. Total DNA (2 μL) was amplified by rolling circle amplification (RCA) at 40°C for 2 h using the EquiPhi29 DNA polymerase (ThermoFisher Scientific, USA), as described previously (Rajabu et al., 2023). The RCA product was diluted 10-fold with DNase-free water and 1 μL was used as template in a 50-μL PCR reaction containing the primer pair, S1-
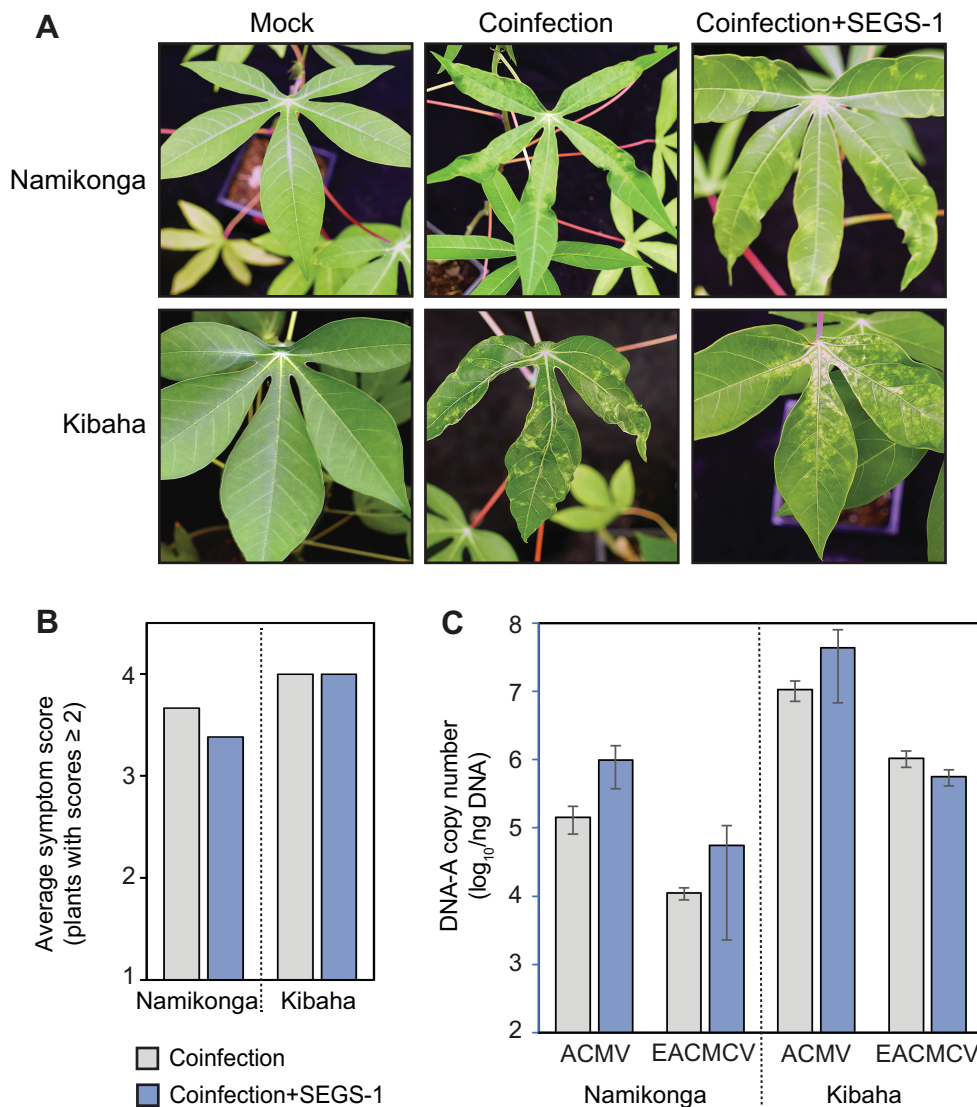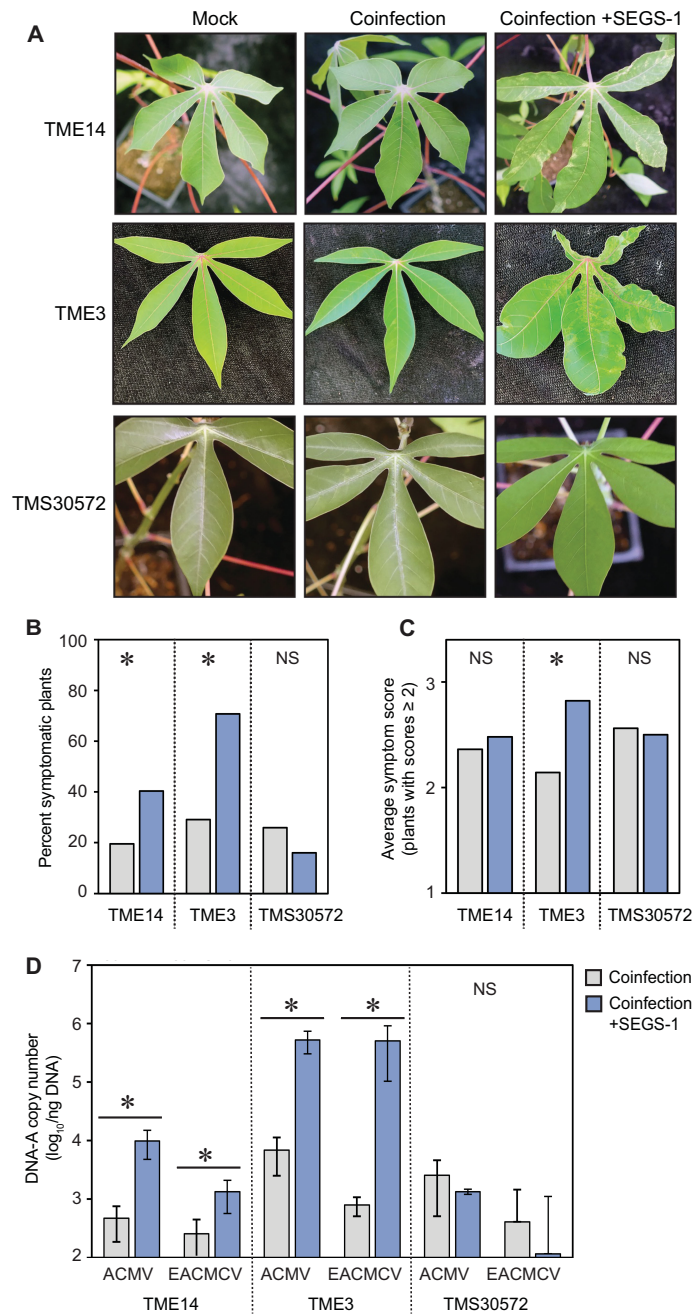
FIGURE 1
Exogenous SEGS-1 does not enhance symptoms during CMB coinfections of susceptible cassava. Cassava cv. Namikonga and Kibaha were co-inoculated with ACMV+EACMCV alone (coinfection) or in combination with exogenous SEGS-1 DNA (coinfection+S1-1.0) under controlled conditions and monitored for symptom severity, and viral copy number at 28 dpi. **(A)** Images of coinfected leaves. Mock-inoculated controls are shown on the left. **(B)** Average symptom scores of plants with scores ≥ 2. **(C)** Average copy number of viral DNA-A/ng total DNA on a $\log_{10}$ scale as determined by qPCR analysis. The bars represent 2 standard errors. No significant differences were detected with (blue) or without (gray) exogenous SEGS-1 DNA.

4F and S1-2R (Supplementary Table 1), using previously established conditions (Ndunguru et al., 2016; Rajabu et al., 2023).

2.4 Sequencing CMD2 mutations and SEGS-1 genomic copies

The published CMD2 mutations in *MePOLD1* (Manes. 12G077400) in the TME3, TME14, and Kibaha plants (Lim et al., 2022) in our studies were confirmed by Sanger sequencing. Total DNA was extracted from healthy, young leaves corresponding to the three cassava varieties using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany). The *M. esculenta* reference genome v8.1 on Phytozyme 13

(https://phytozome-next.jgi.doe.gov) was used to design primer pairs that amplified across the CMD2 SNPs associated with the G680V and A684G amino acid changes (primers: G680V-F and GV680R) and with the V528L replacement (primers: CMD2snEF and CMD2snER) (Supplementary Table 3). The primers were used for end-point PCR to amplify the target regions from plant genomic DNA in 50-μL reactions containing 100 ng of total DNA, 1.25 U of Hotstart Taq Polymerase (New England Biolabs), 10 μM of each primer, and 1× PCR buffer. PCR was performed for 30 cycles (denaturation: 10 sec at 98°C; annealing: 30 sec at 55°C; extension: 2 min at 72°C). The products were gel purified and subjected to Sanger sequencing using the G680V-F primer for the adjacent G680V and A684G SNPs and the V528L Seq2 primer for the

FIGURE 2

Exogenous SEGS-1 can overcome CMD2 resistance in cassava. The CMD resistant cultivars, TME-14 (CMD2), TME3 (CMD2), and TMS300572 (CMD1) were co-inoculated with ACMV+EACMCV alone (coinfection) or in combination with exogenous SEGS-1 DNA (coinfection+S1-1.0) under controlled conditions and monitored for the of percent symptomatic plants, symptom severity, and viral copy number at 28 dpi for TME14 and TME3 and at 46 dpi for TMS300572. **(A)** Images of coinfected leaves from the resistant cassava cultivars. Mock-inoculated controls are shown on the left. **(B)** Percent of plants with symptom scores ≥2. **(C)** Average symptom scores of plants with symptom scores ≥2. **(D)** Average copy number of viral DNA-A/ng total DNA on a log_{10} scale as determined by qPCR analysis. The bars represent 2 standard errors. Asterisks indicate significant differences (p-value < 0.05) between coinfection (gray) and coinfection+SEGS-1 (blue) treatments within a cultivar. NS, No statistically significant.

V528L SNP. The target mutations were decoded by superimposing the reference sequence with the sequencing chromatograms, followed by manual examination for a heterozygous (double) peaks representing the SNP-specific bases. Validation of the Sanger sequencing results was done by examining mismatches with the reference sequence and the extent of baseline noise in the sequencing chromatograms.

The full-length copies of SEGS-1 and flanking genomic sequences were amplified, cloned, and sequenced from healthy TME3 and TME14 plants. The *M. esculenta* v8.1 reference
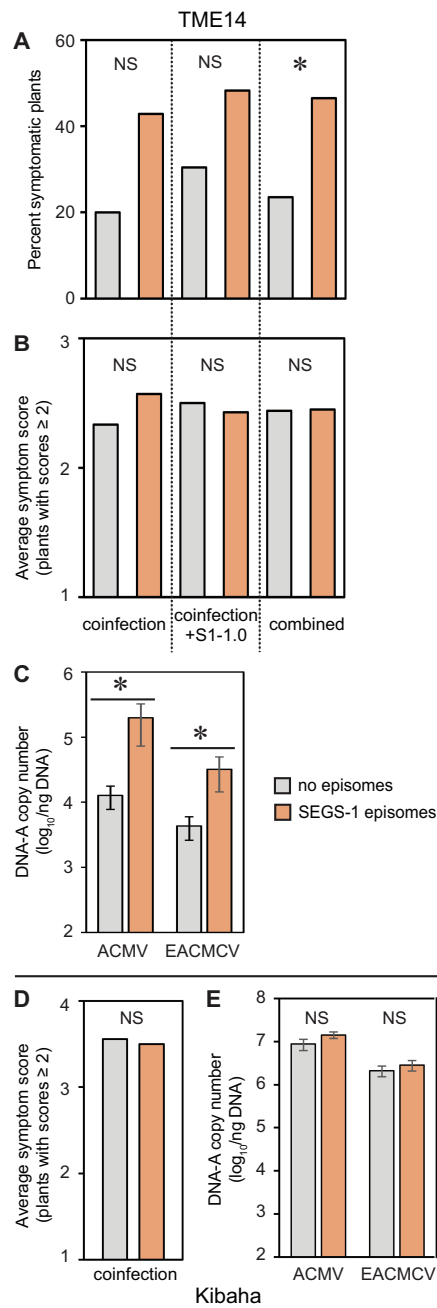
FIGURE 3
Endogenous SEGS-1 episomes increase CMD symptoms and viral copy number in a CMD2 resistant cultivar. Cassava plants were co-inoculated with ACMV+EACMCV alone (coinfection) or in combination with exogenous SEGS-1 DNA (coinfection+S1-1.0) under controlled conditions and assessed for the presence of SEGS-1 episomes by RCA followed by endpoint PCR at 28 dpi. Infection parameters at 28 dpi were compared for plants with SEGS-1 episomes (orange) versus plants without SEGS-1 episomes (gray). **(A)** Percent of plants with symptom scores $\geq$2. **(B, D)** Average symptom scores of plants with symptom scores $\geq$2. **(C, E)** Average copy number of viral DNA-A/ng total DNA on a $\log_{10}$ scale as determined by qPCR analysis. The bars represent 2 standard errors. Results for TME14 (CMD2 resistant) are shown panels **(A–C)**. Results for Kibaha are shown in panels **(D, E)**. All Kibaha plants were infected in all treatments (not shown). Asterisks indicate significant differences (p-value < 0.05) between plants with (orange) or without (gray) SEGS-1 episomes during coinfection or coinfection+S1.0 treatments within a cultivar. NS, No statistically significant.

genome (https://phytozome-next.jgi.doe.gov/) was used to design a primer pair (S1_PyR3 and S1_PyF9; Supplementary Table 3) that amplified a 1639-bp product containing 89 bp of upstream sequence and 543 bp of downstream sequence flanking the 1007-bp SEGS-1 sequence. The presence of a 1.5-kb AT-rich region precluded the design of a more distal upstream PCR primer.

LongAmp® Hot Start Taq DNA Polymerase kit (New England Biolabs) was used to amplify the SEGS-1 region from DNA samples collected from 4 plants of each genotype. The PCR products were gel purified and cloned into the pMiniT 2.0 vector using the New England Biolabs PCR Cloning Kit. Colony PCR using the LongAmp® Hot Start Taq DNA Polymerase and the forward and
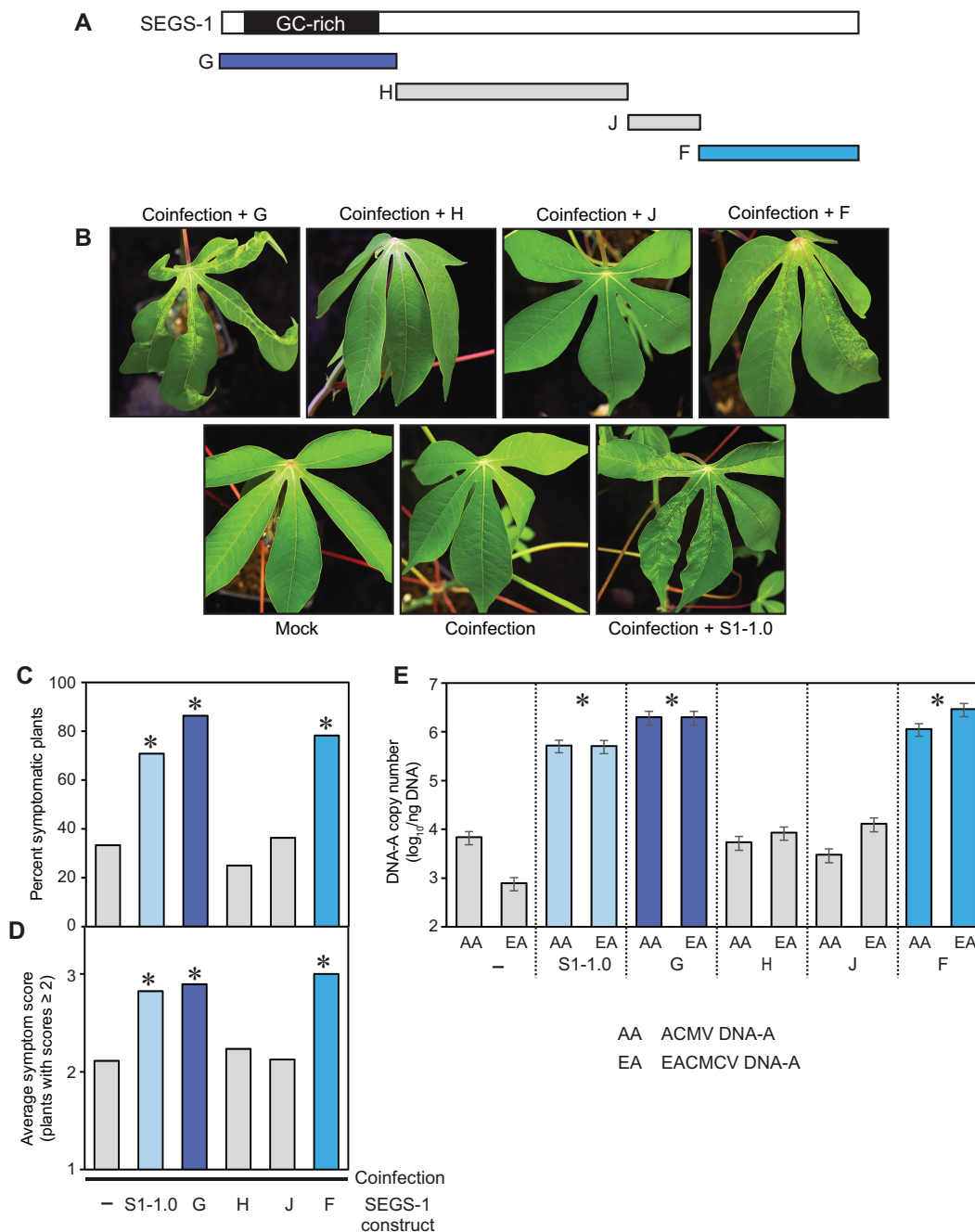
FIGURE 4
The regions of SEGS-1 that enhance CMB infection in a CMD2 cultivar. **(A)** SEGS-1 was cloned into four regions G, H, J and F (see Supplementary Figure 1). **(B)** Images of leaves from TME3 plants co-inoculated with ACMV+EACMCV alone (coinfection) or in combination with exogenous SEGS-1 DNA regions (coinfection + G, H, J or F) at 28 dpi. The coinfection + S1-1.0 treatment was a positive control **(C)** Percent of plants with symptom scores ≥2. **(D)** Average symptom severity of plants with symptom scores ≥2. **(E)** Average copy number of viral DNA-A/ng total DNA on a log₁₀ scale as determined by qPCR analysis. The bars represent 2 standard errors. Asterisks indicate significant differences (p-value<0.05) between coinfection and coinfection+SEGS-1 segments treatments. The blue bars designate treatments showing SEGS-1 activity, while the gray bars denote treatments with no SEGS-1 activity. NS, No statistically significant.

reverse primers provided in the PCR cloning kit were used to identify positive colonies. Plasmids from positive colonies were purified using QIAprep Spin Miniprep Kit (Qiagen, Hilden, Germany) and subjected to Sanger sequencing using overlapping primers designed based on the *M. esculenta* v8.1 reference genome

to obtain coverage of the entire cloned SEGS-1 region. A list of the sequencing primers is found in Supplementary Table 3. Sequences from each plant were aligned to the reference sequence using the SnapGene desktop software (Version 7.2.1, Dotmatics) to assemble a consensus sequence for each full length insert. The resulting
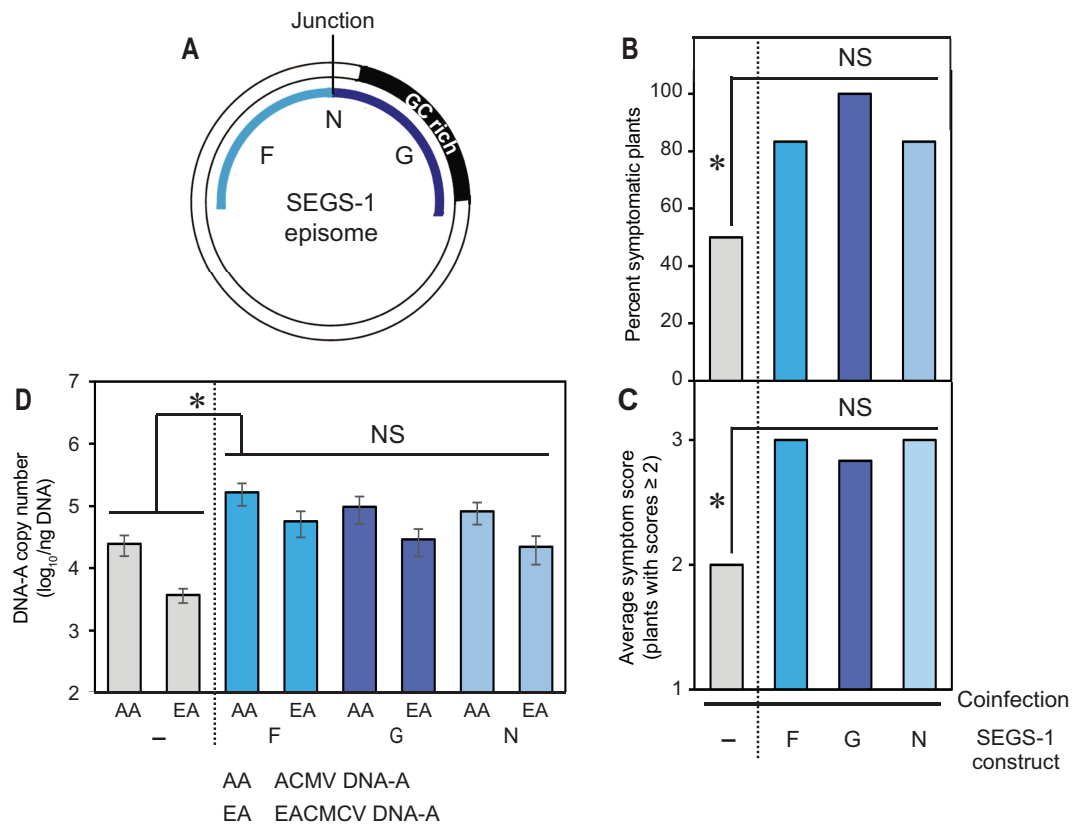
FIGURE 5

The SEGS-1 episome junction does not increase activity. **(A)** Diagram of the SEGS-1 episome showing the N region that combines the F and G regions to reconstitute episome junction sequences. **(B)** Percent of plants with symptom scores ≥ 2. **(C)** Average symptom severity of plants with symptom scores ≥ 2. **(D)** Average copy number of viral DNA-A per ng total DNA on a $\log_{10}$ scale as determined by qPCR analysis. The bars represent 2 standard errors. Asterisks indicate significant differences (p-value<0.05) between coinfection and coinfection+SEGS-1 region treatments. No differences were detected between the SEGS-1 regions. The blue bars designate treatments showing SEGS-1 activity, while the gray bars denote treatments with no SEGS-1 activity.

consensus sequences were compared using the alignment tool in SnapGene to the cloned SEGS-1 sequence reported by Ndunguru et al., 2016, and the *M. esculenta* v8.1 reference genome.

3 Results

3.1 SEGS-1 does not increase disease severity in susceptible cassava varieties

An earlier study showed that exogenously applied SEGS-1 DNA increases the symptoms caused by ACMV in the hypersusceptible cassava cv. 60444 (Ndunguru et al., 2016). We asked if exogenous SEGS-1 also impacts CMD symptoms and viral DNA copy number in two African cultivars – Namikonga and Kibaha. Both cultivars are highly susceptible to coinfection by ACMV and EACMCV, which have been reported to act synergistically during infection (Fondong et al., 2000). Unlike the earlier studies, we co-inoculated plants with ACMV + EACMCV because of less plant-to-plant variability in coinfections than single infections.

Namikonga and Kibaha plants were co-infected with ACMV + EACMCV in the presence or absence of exogenous SEGS-1 DNA. These experiments used a partial tandem copy of SEGS-1 with duplicated GC-rich regions (S1-1.5a). Our microspray inoculation protocol resulted in 100% infection rates in both cultivars, with all inoculated plants showing leaf yellowing and deformation at 28 dpi (Figure 1A). Namikonga and Kibaha had symptom scores ranging from 3 to 4 (on a scale of 1-5) (Supplementary Figure 2). There were no differences in the average symptom scores for either variety whether the plants were co-inoculated with ACMV + EACMCV with or without exogenous SEGS-1 (Figure 1B). There were also no differences in the DNA copy numbers of ACMV-A and EACMCV-A between the inoculation treatments for both cultivars. These conclusions were supported by within-variety statistical comparisons of the two inoculation treatments, all of which gave p-values > 0.05 in Wilcoxon ranked sum tests for symptom scores and Student's t tests for viral DNA copy number. Thus, exogenous SEGS-1 did not enhance CMD in Namikonga or Kibaha co-infected with ACMV + EACMCV.

3.2 SEGS-1 alters the CMD response of a CMD2 resistant cultivar

An earlier study (Ndunguru et al., 2016) showed that exogenously applied SEGS-1 DNA can overcome host resistance in the cassava cultivar TME3. We compared the impact of exogenous SEGS-1 on CMD resistance using three resistant varieties, TMS30572, TME14 and TME3. TMS30572 carries the polyploid CMD1 resistance, while TME14 and TME3 are closely related varieties with the CMD2 locus. CMD2 resistance has been correlated with nonsynonymous, single nucleotide polymorphisms in the *DNA polymerase δ subunit* 1 gene (*MePOLD1*) (Lim et al., 2022). TME14 and TME3 have the same chimeric mutation that results in a G680V amino acid substitution in POLD1. We confirmed that the TME14 and TME3 plants used in our studies carry the mutation using PCR to amplify the exon 18 of the *MePOLD1* gene from genomic DNA followed by Sanger sequencing (Supplementary Figure 3).

TME14, TME3, and TMS30572 plants were co-inoculated with ACMV + EACMCV alone or in combination with SEGS-1 DNA (S1-1.5a). Mild symptoms were observed on TME14 and TME3 plants (Figure 2A), with the percent of symptomatic plants approximately 2-fold higher in the presence of SEGS-1 DNA (Figure 2B). Interestingly, the average symptom scores only showed a significant increase for TME3 and not for TME14 (Figure 2C). Both CMD2 cultivars showed significant increases in the DNA copy numbers for ACMV-A and EACMCV-A in the presence of SEGS-1, with TME3 showing a larger increase than TME14 (Figure 2D). The presence of SEGS-1 had no detectable effects on the percent of symptomatic plants, average symptom scores, or viral DNA copy numbers in TMS30572 at 46 dpi

(Figure 2). No viral DNA was detected by qPCR in TMS30572 plants at 28 dpi and only in some TMS30572 plants at 46 dpi, as indicated by the large error bars.

ANOVA (Supplementary Table 2) that assessed the effect of SEGS-1 on viral DNA copy number across experiments showed similar trends for TME14, TME3, and TMS30572 as seen in Figure 2D. The presence of SEGS-1 was associated with significant increases in the $\log_{10}$ copy number of ACMV-A and EACMV-A in TME14 and TME3 but not TMS30572 (Supplementary Table 2A). Interestingly, TME14 also showed a significant difference between the viral treatments (Supplementary Table 2B) that is also apparent in the mean separations of ACMV-A and EACMCV-A $\log_{10}$ copy numbers (Supplementary Table 2C).

3.3 Episomal copies of SEGS-1 in cassava varieties

SEGS-1 episomes have been detected in leaves collected from infected plants in Tanzania and Cameroon (Ndunguru et al., 2016), but it was not possible to assess whether the SEGS-1 episomes in the field-grown plants were from endogenous or exogenous sources. Hence, we asked if SEGS-1 episomes could be detected in infected cassava plants grown under controlled laboratory conditions that precluded potential exogenous sources. For this experiment, cassava plants were co-inoculated with ACMV + EACMCV only. None of the plants were treated with exogenous SEGS-1 plasmid DNA. Total DNA was isolated at 28 dpi and used as template for RCA, which efficiently amplifies small, circular DNA molecules like SEGS-1 episomes. The RCA products were then amplified using a divergent primer pair, (S1-4F/S1-2R) to give a 552-bp product that includes
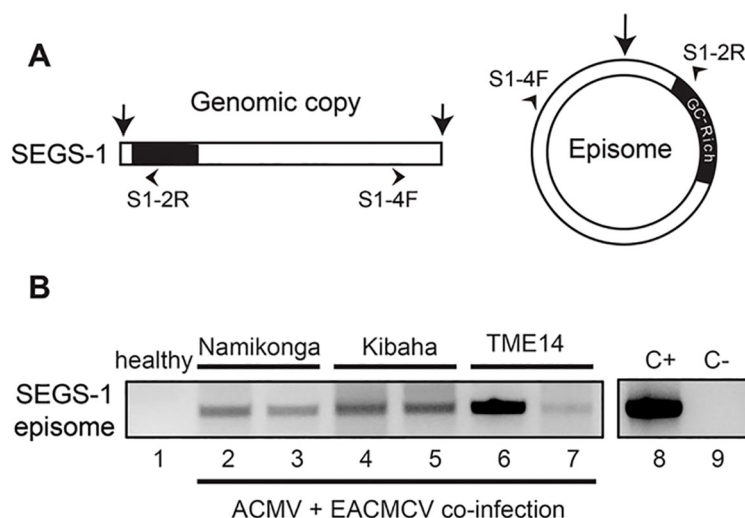
FIGURE 6
Production of endogenous SEGS-1 episomes during CMB coinfection. **(A)** Diagram of divergent primers that amplify across the junction of a SEGS-1 episome but not the linear SEGS-1 sequence in the cassava genome. **(B)** SEGS-1 episomes were detected by RCA followed by endpoint PCR in Namikonga (lanes 2 and 3), Kibaha (lanes 4 and 5), and TME-14 (CMD2 resistant, lanes 6 and 7) plants co-inoculated with ACMV+EACMCV alone at 28 dpi. No episomes were detected in the healthy cassava control (lane 1), demonstrating that the divergent primers do not amplify the SEGS-1 genomic sequence. Lane 8 is the positive control (+C) using SEGS-1 plasmid S1-1.5a that contains the junction sequences as template. Lane 9 is the no template control (-C).

the SEGS-1 episomal junction (Figure 6A) (Ndunguru et al., 2016). The divergent primers do not amplify the SEGS-1 sequence in the cassava genome. Figure 6B shows representative examples of SEGS-1 episomes that were detected in infected Namikonga (lanes 2 and 3), Kibaha (lanes 4 and 5), and TME14 (lanes 6 and 7) plants. No episomes were observed in healthy plants (an example is shown in Figure 6B, lane 1). These results indicated that SEGS-1 episomes are generated during viral infection by a process that most likely involves the full copy of the SEGS-1 sequence in the cassava genome. These results also unequivocally established that SEGS-1 episomes are not from exogenous sources like satellite molecules.

We then surveyed the Namikonga, Kibaha, TME14, TME3, and TMS300572 cultivars to gain insight into how often SEGS-1 episomes form when plants are inoculated with ACMV + EACMCV alone or in combination with exogenous SEGS-1 DNA (S1-1.5a). SEGS-1 episomes were detected in ≥ 50% of Namikonga and Kibaha plants in both inoculation treatments (Table 1). SEGS-1 episomes were also detected in 12% and 58% TME14 plants not treated or treated, respectively, with exogenous SEGS-1 DNA at the time of viral inoculation. The higher frequency of SEGS-1 episomes in TME14 plants in the virus plus exogenous SEGS-1 treatment may be due to residual SEGS-1 plasmid DNA, but we think this is unlikely because systemically infected leaves that emerged after inoculation were sampled at 28 dpi. No SEGS-1 episomes were detected in TME3 or TMS300572 plants co-inoculated with ACMV + EACMCV alone, and only a few plants had episomes in the virus plus exogenous SEGS-1 treatment. These results indicated that the endogenous release of SEGS-1 episomes from the cassava genome is cultivar dependent.

The difference in the capacities of TME3 and TME14 to form SEGS-1 episomes could be due to sequence differences between the full-length SEGS-1 sequences in their genomes. To address this possibility, we amplified and sequenced the full-length SEGS-1 genomic copy and 89 bp of upstream sequence and 543 bp of downstream sequence from 4 independent plants for each genotype. Comparison of the 8 sequences uncovered heterogeneity between the sequences from individual plants within a genotype as well as differences between the genotypes (Figures 7A, B). Eight SNPs occurred in a subset of TME3 and TME14 sequences (Figure 7A, white lines). The SNPs observed in one genotype only occurred in one of the four sequences from that genotype (Figure 3A, green lines – TME3; pink lines – TME14). No conserved genotype-specific SNPs were observed. Instead, the 4 SNPs (2 in SEGS-1 and 2 in the distal

downstream region) that were conserved across all TME3 sequences were also conserved in all TME14 sequences (Figure 7A, yellow lines). No SNPs occurred around the junction sequences of either genotype (Figure 7C). Thus, it is unlikely that sequence differences in and around the SEGS-1 locus account for the different capacities of TME3 and TME14 to form SEGS-1 episomes.

3.4 Correlation of SEGS-1 episomes with CMD severity

We asked if the presence of SEGS-1 episomes in TME14 correlates with increased frequency of infection, symptom severity and/or viral copy number. When TME14 plants were co-inoculated with ACMV + EACMCV alone or in combination with SEGS-1 DNA (S1-1.5a), the frequency of symptomatic plants trended higher (ca. 2-fold) when SEGS-1 episomes were present (Figure 3A). The copy numbers of ACMV-A and EACMCV-A were significantly higher (ca. 10-fold) at 28 dpi in plants with SEGS-1 episomes versus plants without episomes (Figure 3C). In contrast, there was no difference in the average symptom scores of plants with or without SEGS-1 episomes (Figure 3B). However, it is important to note that the average is determined using only symptomatic plants (score ≥ 2), and more plants were symptomatic when SEGS-1 episomes were present (Figure 3A). A similar analysis of Kibaha did not detect any differences in symptom severity (Figure 3D) or viral DNA copy numbers (Figure 3E) between plants with or without SEGS-1 episomes.

3.5 Characterization of sequences required for SEGS-1 activity

TME3 plants do not form SEGS-1 episomes (Table 1) but showed increases in the percent of symptomatic plants, symptom scores, and DNA copy numbers of ACMV-A and EACMCV-A when treated with exogenous SEGS-1 DNA (Figure 2). These properties allowed us to characterize the sequence requirements for SEGS-1 activity in the absence of endogenous background activity that could mask the effects of exogenously applied SEGS-1 sequences. SEGS-1 was subcloned into 4 fragments, designated as SEGS-1 regions G (positions 1–277), I (positions 275–646), J (positions 644–756), and F (positions 757–1007) (Figure 4A; Supplementary Figure 1). Each cloned fragment was co-inoculated with ACMV + EACMCV onto TME3 plants. The positive control was co-inoculation of both viruses with full-length SEGS-1 (S1-1.0), while the negative control was plants only inoculated with both viruses. The percent symptomatic plants, average symptom scores, and viral DNA-A copy numbers were assessed at 28 dpi for all inoculation treatments. Virus-inoculated plants treated with fragments G or F developed clear symptoms that were similar to the full-length SEGS-1 positive control (S1-1.0), while plants treated with fragments J and H showed few symptoms like the virus only negative control (Figure 4B). The percent of infected plants were similar for plants treated with S1-1.0, G or F and ca. 2-fold higher than the virus alone

TABLE 1 Percent of infected plants with episomes.

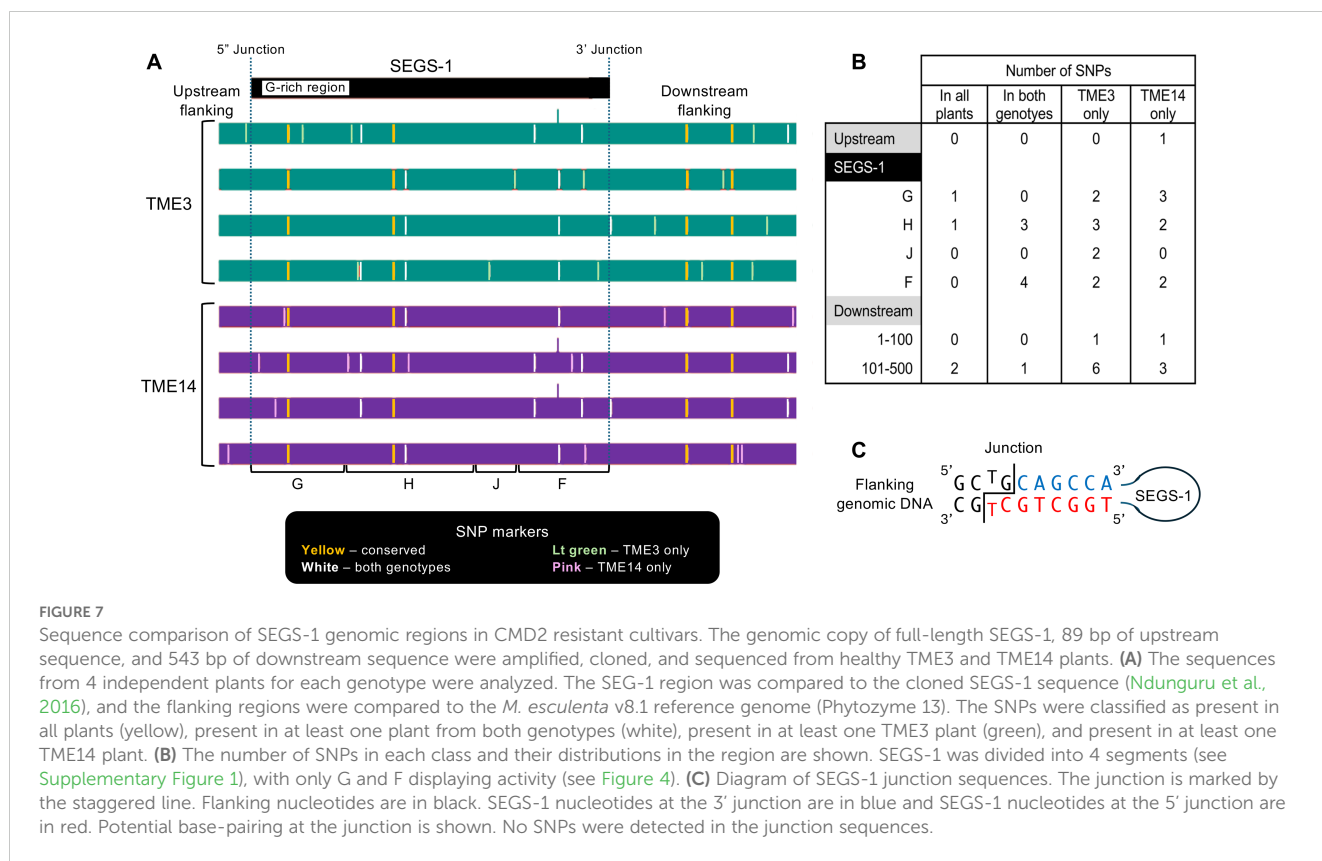| Cultivar	Coinfected	Coinfected +SEGS1	# Plants Coinfected/ Coinfected +SEGS-1
Namikonga	66.7%	50%	6/6
Kibaha	57%	50%	28/10
TME14	12%	58%	50/50
TME3	0	1.19%	54/54
TMS300572	0	9.68%	31/31

FIGURE 7
Sequence comparison of SEGS-1 genomic regions in CMD2 resistant cultivars. The genomic copy of full-length SEGS-1, 89 bp of upstream sequence, and 543 bp of downstream sequence were amplified, cloned, and sequenced from healthy TME3 and TME14 plants. **(A)** The sequences from 4 independent plants for each genotype were analyzed. The SEG-1 region was compared to the cloned SEGS-1 sequence (Ndunguru et al., 2016), and the flanking regions were compared to the *M. esculenta* v8.1 reference genome (Phytozyme 13). The SNPs were classified as present in all plants (yellow), present in at least one plant from both genotypes (white), present in at least one TME3 plant (green), and present in at least one TME14 plant. **(B)** The number of SNPs in each class and their distributions in the region are shown. SEGS-1 was divided into 4 segments (see Supplementary Figure 1), with only G and F displaying activity (see Figure 4). **(C)** Diagram of SEGS-1 junction sequences. The junction is marked by the staggered line. Flanking nucleotides are in black. SEGS-1 nucleotides at the 3' junction are in blue and SEGS-1 nucleotides at the 5' junction are in red. Potential base-pairing at the junction is shown. No SNPs were detected in the junction sequences.

control or plants treated with J or H (Figure 4C). The S1-1.0, G, and F treatments also resulted in more severe symptoms (Figure 4D) and >100-fold increases in viral DNA-A copy numbers (Figure 4E) than the virus alone, H and J treatments. These differences were statistically significant (p-values < 0.05). These results established that the G and F regions are separately sufficient for SEGS-1 activity, and that the intervening J and H regions are not required for activity.

Fragments G and F are at the opposite ends of the full-length copy of SEGS-1 in the cassava genome but are adjacent in the episome and flank the junction sequence (Figure 5A). We fused G and F to reconstitute the junction region in fragment N (Supplementary Figure 1). Co-inoculation of the cloned fragment N with ACMV + EACMCV also resulted in statistically significant increases in the percent of symptomatic plants, symptom scores, and viral DNA-A copy numbers when compared to virus alone plants (Figures 5B–D). However, no differences were detected when the frequency of infected plants, symptom severity, and viral DNA copy numbers were compared between the F, G and N treatments, indicating that joining the F and G regions and reconstituting the junction sequence did not result in increased SEGS-1 activity.

4 Discussion

SEGS-1 was first identified as a cassava genomic sequence that also occurs as an episomal DNA in cassava plants showing severe CMD symptoms in the field (Ndunguru et al., 2016). A previous study showed that SEGS-1 can increase disease severity in cassava plants co-inoculated with a CMB and exogenous SEGS-1 DNA (Ndunguru et al., 2016). We report here that the effects of exogenous SEGS-1 on CMD varies between cassava cultivars, with CMD2 cultivars most affected. We also show that cultivars differ in their capacities to produce endogenous SEGS-1 episomes from the full-length genomic copy of SEGS-1, and that there is a correlation between the generation of endogenous SEGS-1 episomes and disease severity in a CMD2 cultivar. Taken together, these results established that the host genetic background is an important determinant of SEGS-1 activity during CMB infection in cassava.

In the studies reported here, the susceptible cultivars (Namikonga and Kibaha) did not develop more severe disease when ACMV and EACMCV were co-inoculated with exogenous SEGS-1 DNA. We reported previously that exogenous SEGS-1 DNA increases symptom severity in a susceptible cultivar infected with one CMB species (Ndunguru et al., 2016). Infection by a single virus is often milder than a coinfection with two CMBs due to synergistic interactions (Fondong et al., 2000; Vanitharani et al., 2004). However, it is unlikely that the synergistic interactions between ACMV+EACMCV during coinfection of susceptible cultivars suppress SEGS-1 activity because exogenous SEGS-1 can increase disease severity during coinfection of CMD2 cultivars. The CMD2 cultivars have lower disease parameter values in the absence of SEGS-1 that are comparable to single-virus infections of susceptible cultivars. Thus, we think it is likely the high disease parameter values of the susceptible cultivar coinfections precluded the detection of SEGS-1 activity.

Exogenous SEGS-1 DNA significantly increased the percentage of infected plants and viral DNA copy number in two CMD2 resistant cultivars (TME14 and TME3) but not in a CMD1 resistant cultivar (TMS300572). The resistance phenotypes of CMD1 and CMD2 plants have different genetic origins that may determine how SEGS-1 affects CMD. CMD1 resistance is polygenic and recessive (Okogbenin et al., 2013), and the genes involved in the resistance are not known. CMD2 resistance is mediated by a single, dominant gene encoding DNA polymerase delta subunit 1 (*MePOLD1*) (Akano et al., 2002; Lim et al., 2022). TME3 and TME14 carry the same *MePOLD1* amino acid mutation (G680V in exon 18), and they have similar responses to exogenous SEGS-1. Interestingly, Kibaha is also a *MePOLD1* variant but has a different amino acid polymorphism (A684G in exon 18) that does not confer resistance in our CMB coinfection studies.

SEGS-1 episomes were first detected in field grown plants with CMD (Ndunguru et al., 2016). We report here the presence of SEGS-1 episomes in cassava plants coinfected with ACMV and EACMCV under laboratory conditions. Importantly, the detection of SEGS-1 episomes did not depend on treatment of the infected plants with exogenous SEGS-1 DNA. The only source of the episomes in untreated plants is the full-length copy of SEGS-1 that has been identified in all characterized cassava genomes to date. Although the genomic sequence of SEGS-1 is highly conserved, cultivars differ in their capacities to form SEGS-1 episomes during CMD. More than 50% of infected Namikonga and Kibaha plants formed episomes. Kibaha plants with endogenous SEGS-1 episomes did not have higher symptom scores or viral titers than infected plants lacking episomes, supporting the idea that SEGS-1 effects are not readily detected in highly susceptible cultivars. Endogenous SEGS-1 episomes were also detected in the CMD2 resistant cultivar TME14 during coinfection, with 14% of the plants containing episomes. Unlike the susceptible cultivars, the presence of SEGS-1 episomes in TME14 correlated with higher incidence of infection and viral DNA copy number, suggesting that that the formation of endogenous SEGS-1 episomes causes more severe disease in a CMD2 resistant cultivar. This effect is cultivar dependent because no endogenous SEGS-1 episomes were found in TME3 or TMS30572.

The failure of TME3 to generate endogenous episomes was unexpected because of its close relationship to TME14 (Bredeson et al., 2016). Hierarchical clustering of residual genetic distances has indicated that TME3 and TME14 genomes are not identical (Rabbi et al., 2014), supporting the idea that there is something different between their genomes that impacts the generation of SEGS-1 episomes. Sequence comparison of the genomic copies of SEGS-1 and flanking regions did not uncover any differences in TME3 and TME14 that are likely to reflect this genetic difference. Thus, we hypothesize that the different capacities of TME3 and TME14 to form endogenous SEGS-1 episomes is due to a SNP elsewhere in the cassava genome. Accordingly, TME14 and other episome-forming cultivars may encode a protein that acts in trans to generate SEGS-1 episomes, and the TME3 genome contains a mutation that interferes with its production or activity. The observation that the longest open reading frame in SEGS-1 only specifies a 48 amino acid peptide (ORF positions 164–18), further supports our premise

that episome release depends on a protein encoded at another location in the genome. A similar scenario occurs for transposable elements, which are often defective and depend on transposase genes encoded distally in a plant genome for their mobilization (Pulido and Casacuberta, 2023).

A SEGS-1 transgene increases ACMV symptom severity in Arabidopsis, establishing that SEGS-1 by itself is sufficient to enhance begomovirus infection (Rajabu et al., 2023). This observation provides an explanation for why exogenous SEGS-1 increases disease severity in TME3. It also raises the question as to whether SEGS-1 episomes are necessary for disease enhancement in cassava or if the SEGS-1 genomic copy is sufficient. The strongest evidence supporting the involvement of episomes comes from TME14, which showed a positive correlation between episome presence and disease severity. Moreover, the SEGS-1 genomic sequences of TME3 and TME14 are essentially identical and should impact CMD similarly if the genomic copy is the active form. Given that SEGS-1 belongs to a repetitive sequence family, it is likely located in a heterochromatic region of the genome that is functionally inactive. SEGS-1 episomes are only found in CMB-infected plants, consistent with infection promoting episome formation. Begomovirus infection induces host DNA replication (Nagar et al., 2002; Ascencio-Ibáñez et al., 2008) and interferes with DNA methylation (Gui et al., 2022). The opening of host chromatin during replication and/or the decrease in DNA methylation could lead to activation of the SEGS-1 genomic copy or the generation of SEGS-1 episomes.

The strong effect of exogenous SEGS-1 DNA on CMB infection in TME3 argues that its genomic copy does not enhance CMD. Instead, SEGS-1 activity in cassava is likely mediated by extrachromosomal SEGS-1 DNA through the formation of endogenous episomes or the application of exogenous SEGS-1 plasmid DNA. Endogenous SEGS-1 episomes have some features in common with but are distinct from extrachromosomal circular DNAs (eccDNAs) that have been characterized in plants (Peng et al., 2022; Zhuang et al., 2024). Like SEGS-1, the generation of some eccDNAs has been associated with stress and epigenetic changes. Conversely, SEGS-1 episomes are larger than most eccDNAs and correspond to a specific cassava sequence with a conserved junction, while eccDNAs appear to be generated randomly from the plant genome and show little reproducibility or sequence conservation with each other.

We mapped SEGS-1 activity to two regions, F and G that flank the episomal junction. When F and G were combined to restore the episomal junction, no additional increase in disease enhancement was observed. The lack of synergy between F and G is consistent with both regions of SEGS-1 targeting the same host defense pathway. Moreover, the episomal junction sequences do not contribute directly to SEGS-1 disease enhancement. The SEGS-1 episomal junction is imbedded in a 10-bp repeat motif with a single mismatch at the junction site that may be involved in SEGS-1 episome formation. Thus, we hypothesize that SEGS-1 disease enhancement involves two separable steps – [1] episome formation involving the junction sequences and a transacting protein encoded elsewhere in the genome and [2] disease enhancement mediated by the F and/or G regions. The G region

overlaps 15 SEGS-1 related sequences that are distributed across 10 chromosomes of the cassava genome, while the F region overlaps 2 SEGS-1 related sequences that occur on separate chromosomes (Ndunguru et al., 2016). It is not known if any of the SEGS-1 related sequences occur in episomes or can enhance CMD severity, but several have low E values and high identities in BLAST comparisons to the full-length genomic copy of SEGS-1, raising the possibility that they are also active during CMD. Future studies that identify the SEGS-1 products and how they function will provide insight into whether other members of the SEGS-1 sequence family affect susceptibility and resistance to CMB infection.

The studies reported here describe a new type of plant virus/host interaction in which viral infection leads to the generation of DNA episomes from SEGS-1 genomic DNA and an increase in CMD severity. The SEGS-1 genomic sequence is ubiquitous across cassava cultivars, and 3 of the 5 cultivars in this report generate SEGS-1 episomes during CMB infection. Thus, the endogenous SEGS-1/virus interaction is likely to be a common feature of the CMB infection process that is reflected by severe CMD observed in some African fields planted with CMD-resistant cassava. This is most concerning for CMD2-resistant plants, which typically become infected and then undergo recovery. CMD2 cultivars that produce SEGS-1 episomes during the infection phase are likely to be less able to undergo recovery and, instead, develop severe CMD. Hence, it is important to identify cultivars like TME14 that produce SEGS-1 episomes and remove them from cassava breeding programs, if alternative cultivars like TME3 that do not generate episomes are available. As more information becomes available about cultivar-specific generation of SEGS-1 episomes, its inclusion into cassava genetic databases could be used to identify molecular markers associated with episome formation. Such markers could be used to identify new cultivars with reduced risk of episome formation and potential gene editing targets to prevent episome formation, as well as provide insight into the mechanisms underlying SEGS-1 episomes formation and mode(s) of action. An integrated approach involving virology, genetics, and breeding to minimize the effects of SEGS-1 episomes will be essential to develop effective and sustainable CMD resistant cassava.

Data availability statement

The datasets presented in this article are not readily available because No data was generated. Requests to access the datasets should be directed to Trino Ascencio-Ibáñez, jtascenc@ncsu.edu.

Author contributions

EC: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing, Validation. CA: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. CR: Conceptualization, Investigation, Methodology, Writing – review & editing. MD: Conceptualization, Investigation, Methodology, Resources, Supervision, Writing – review & editing. JN: Conceptualization, Funding acquisition, Resources, Supervision, Writing – review & editing. JA: Conceptualization, Funding acquisition, Investigation, Methodology, Resources, Writing – original draft, Writing – review & editing, Data curation, Formal analysis. EA: Supervision, Writing – review & editing. LH: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1469045/full#supplementary-material

SUPPLEMENTARY FIGURE 1
SEGS-1 fragments. (Top) Properties of SEGS-1 fragments. In the primer sequences, black marks SEGS-1 sequences, red marks the *Not*I linker, and blue marks added residues not in the linker or SEGS-1. (Bottom) The sequences of the SEGS-1 fragments. The green shading indicates the GC-

rich region. The blue line marks the junction position. The underlined nucleotides overlap the adjacent fragment.

SUPPLEMENTARY FIGURE 2

CMD symptom scoring. Images of leaves illustrating the symptom scores. A description of the extent of symptoms is below each image.

SUPPLEMENTARY FIGURE 3

Confirmation of CMD2 mutations. Sanger sequencing profiles for the TME3, TME14, Kibaha, and TMS30752 plants used in our studies. The codons in exon 18 of the *MePOLD1* gene (Manes. 12G077400) containing the nonsynonymous SNP haplotypes, G680V and A684G, are shown at the top. The G680V and A684G SNPs are highlighted in red in each profile and indicated by asterisks (*). The TME3 and TME14 chromatograms show double peaks in codon 680 indicating that they are chimeric for wild-type GGT (glycine) and mutant GTT

(valine) at this position. The Kibaha chromatogram shows a double peak in codon 684 indicating that it is chimeric for wild-type GCC (alanine) and mutant GGC (glycine) at this position. TMS30752, which is not a CMD2 cultivar, does not have a mutation at either position.

SUPPLEMENTARY TABLE 1

Primers for viral qPCR, SEGS-1 episomes detection, and site-directed mutagenesis.

SUPPLEMENTARY TABLE 2

ANOVA of the effects of SEGS-1 and CMB treatment on log_{10} viral DNA copy number in CMD2 and CMD1 cassava cultivars.

SUPPLEMENTARY TABLE 3

Primers for cloning and sequencing cassava genomic DNA.

References

Aimone, C., De León, L., Dallas, M., Ndunguru, J., Ascencio-Ibáñez, J., and Hanley-Bowdoin, L. (2021a). A new type of satellite associated with cassava mosaic begomoviruses. *J. Virol.* 95 (21), e00432-21. doi: 10.1128/jvi.00432-00421

Aimone, C., Hoyer, J. S., Dye, A. E., Deppong, D. O., Duffy, S., Carbone, I., et al. (2022). An experimental strategy for preparing circular ssDNA virus genomes for next-generation sequencing. *J. Virological Methods* 300, 114405. doi: 10.1016/j.jviromet.2021.114405

Aimone, C., Lavington, E., Hoyer, J. S., Deppong, D. O., Mickelson-Young, L., Jacobson, A., et al. (2021b). Population diversity of cassava mosaic begomoviruses increases over the course of serial vegetative propagation. *J. Gen. Virol.* 102, 001622. doi: 10.1099/jgv.0.001622

Akano, A., Dixon, A., Mba, C., Barrera, E., and Fregene, M. (2002). Genetic mapping of a dominant gene conferring resistance to cassava mosaic disease. *Theor. Appl. Genet.* 105, 521–525. doi: 10.1007/s00122-002-0891-7

Ascencio-Ibáñez, J., Dallas, M., and Hanley-Bowdoin, L. (2023). "Begomovirus inoculation in arabidopsis and cassava," in *Plant-Virus Interactions* (New York: Springer), 71–79.

Ascencio-Ibáñez, J. T., Sozzani, R., Lee, T.-J., Chu, T.-M., Wolfinger, R. D., Cella, R., et al. (2008). Global analysis of Arabidopsis gene expression uncovers a complex array of changes impacting pathogen response and cell cycle during geminivirus infection. *Plant Physiol.* 148, 436–454. doi: 10.1104/pp.108.121038

Bayata, A. (2019). Review on nutritional value of cassava for use as a staple food. *Sci. J. Anal. Chem.* 7, 83–91. doi: 10.11648/j.sjac.20190704.12

Bredeson, J. V., Lyons, J. B., Prochnik, S. E., Wu, G. A., Ha, C. M., Edsinger-Gonzales, E., et al. (2016). Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. *Nat. Biotechnol.* 34, 562–570. doi: 10.1038/nbt.3535

Crespo-Bellido, A., Hoyer, J., Dubey, D., Jeannot, R., and Duffy, S. (2021). Interspecies recombination has driven the macroevolution of cassava mosaic begomoviruses. *J. Virol.* 95 (17), e00541-21. doi: 10.1128/jvi.00541-00521

Deng, D., Otim-Nape, W., Sangare, A., Ogwal, S., Beachy, R., and Fauquet, C. (1997). Presence of a new virus closely related to East African cassava mosaic geminivirus, associated with cassava mosaic outbreak in Uganda. *Afr JRoot Tuber Crops.* 2, 23–28.

Duffy, S., and Holmes, E. (2009). Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. *J. Gen. Virol.* 90, 1539–1547. doi: 10.1099/vir.0.009266-0

FAOSTAT. (2021). *Food and Agriculture Organization of the United Nations statistical database.* Available online at: https://www.fao.org/faostat/en/data/QCL (Accessed March, 2024).

Fiallo-Olivé, E., and Navas-Castillo, J. (2020). Molecular and biological characterization of a New World mono-/bipartite begomovirus/deltasatellite complex infecting Corchorus siliquosus. *Front. Microbiol.* 11, 1755. doi: 10.3389/fmicb.2020.01755

Fondong, V., Pita, J., Rey, M., De Kochko, A., Beachy, R., and Fauquet, C. (2000). Evidence of synergism between African cassava mosaic virus and a new double-recombinant geminivirus infecting cassava in Cameroon. *J. Gen. Virol.* 81, 287–297. doi: 10.1099/0022-1317-81-1-287

Gong, P., Tan, H., Zhao, S., Li, H., Liu, H., Ma, Y., et al. (2021). Geminiviruses encode additional small proteins with specific subcellular localizations and virulence function. *Nat. Commun.* 12, 4278. doi: 10.1038/s41467-021-24617-4

Gui, X., Liu, C., Qi, Y., and Zhou, X. (2022). Geminiviruses employ host DNA glycosylases to subvert DNA methylation-mediated defense. *Nat. Commun.* 13, 575. doi: 10.1038/s41467-022-28262-3

Hanley-Bowdoin, L., Bejarano, E., Robertson, D., and Mansoor, S. (2013). Geminiviruses: masters at redirecting and reprogramming plant processes. *Nat. Rev. Microbiol.* 11, 777–788. doi: 10.1038/nrmicro3117

Hoyer, J. S., Fondong, V. N., Dallas, M. M., Aimone, C. D., Deppong, D. O., Duffy, S., et al. (2020). Deeply sequenced infectious clones of key cassava begomovirus isolates from Cameroon. *Microbiol. Resource Announcements* 9 (46), e00802-20. doi: 10.1128/mra.00802-00820

Legg, J., Jeremiah, S., Obiero, H., Maruthi, M., Ndyetabula, I., Okao-Okuja, G., et al. (2011). Comparing the regional epidemiology of the cassava mosaic and cassava brown streak virus pandemics in Africa. *Virus Res.* 159, 161–170. doi: 10.1016/j.virusres.2011.04.018

Legg, J., Kumar, P., Makeshkumar, T., Tripathi, L., Ferguson, M., Kanju, E., et al. (2015). "Cassava virus diseases: biology, epidemiology, and management," in *Advances in virus research* (San Diego/London: Elsevier), 85–142.

Lim, Y.-W., Mansfeld, B. N., Schläpfer, P., Gilbert, K. B., Narayanan, N. N., Qi, W., et al. (2022). Mutations in DNA polymerase δ subunit 1 co-segregate with CMD2-type resistance to Cassava Mosaic Geminiviruses. *Nat. Commun.* 13, 3933. doi: 10.1038/s41467-022-31414-0

Nagar, S., Hanley-Bowdoin, L., and Robertson, D. (2002). Host DNA replication is induced by geminivirus infection of differentiated plant cells. *Plant Cell* 14, 2995–3007. doi: 10.1105/tpc.005777

Ndunguru, J., De León, L., Doyle, C., Sseruwagi, P., Plata, G., Legg, J., et al. (2016). Two novel DNAs that enhance symptoms and overcome CMD2 resistance to cassava mosaic disease. *J. Virol.* 90, 4160–4173. doi: 10.1128/JVI.02834-15

Okogbenin, E., Setter, T., Ferguson, M., Mutegi, R., Ceballos, H., Olasanmi, B., et al. (2013). Phenotypic approaches to drought in cassava. *Front. Physiol.* 4, 93. doi: 10.3389/fphys.2013.00093

Peng, H., Mirouze, M., and Bucher, E. (2022). Extrachromosomal circular DNA: a neglected nucleic acid molecule in plants. *Curr. Opin. Plant Biol.* 69, 102263. doi: 10.1016/j.pbi.2022.102263

Pita, J., Fondong, V., Sangare, A., Otim-Nape, G., Ogwal, S., and Fauquet, C. (2001). Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *J. Gen. Virol.* 82, 655–665. doi: 10.1099/0022-1317-82-3-655

Pulido, M., and Casacuberta, J. M. (2023). Transposable element evolution in plant genome ecosystems. *Curr. Opin. Plant Biol.* 75, 102418. doi: 10.1016/j.pbi.2023.102418

Rabbi, I., Hamblin, M., Kumar, P., Gedil, M., Ikpan, A., Jannink, J., et al. (2014). High-resolution mapping of resistance to cassava mosaic geminiviruses in cassava using genotyping-by-sequencing and its implications for breeding. *Virus Res.* 186, 87–96. doi: 10.1016/j.virusres.2013.12.028

Rajabu, C. A., Dallas, M. M., Chiunga, E., De León, L., Ateka, E. M., Tairo, F., et al. (2023). SEGS-1 a cassava genomic sequence increases the severity of African cassava mosaic virus infection in Arabidopsis thaliana. *Front. Plant Sci.* 14, 1250105. doi: 10.3389/fpls.2023.1250105

Stanley, J., and Gay, M. R. (1983). Nucleotide sequence of cassava latent virus DNA. *Nature* 301, 260–262. doi: 10.1038/301260a0

Vanitharani, R., Chellappan, P., Pita, J. S., and Fauquet, C. M. (2004). Differential roles of AC2 and AC4 of cassava geminiviruses in mediating synergism and suppression of posttranscriptional gene silencing. *J. Virol.* 78, 9487–9498. doi: 10.1128/JVI.78.17.9487-9498.2004

Yang, X., Wang, B., Li, F., Yang, Q., and Zhou, X. (2016). Research advances in geminiviruses. *Curr. Res. Topics Plant Virol.*, 251–269. doi: 10.1007/978-3-319-32919-2_11

Zhou, X., Liu, Y., Calvert, L., Munoz, C., Otim-Nape, G., Robinson, D., et al. (1997). Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. *J. Gen. Virol.* 78, 2101–2111. doi: 10.1099/0022-1317-78-8-2101

Zhuang, J., Zhang, Y., Zhou, C., Fan, D., Huang, T., Feng, Q., et al. (2024). Dynamics of extrachromosomal circular DNA in rice. *Nat. Commun.* 15, 2413. doi: 10.1038/s41467-024-46691-0