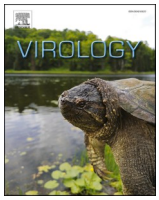
Virology 593 (2024) 110011

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/virology



Meta-transcriptomic identification of groundnut RNA viruses in western Kenya and the novel detection of groundnut as a host for Cauliflower mosaic virus

Dennis Obonyo [a,b], George Ouma [c], Rachel Ikawa [b], Damaris A. Odeny [d,*]

[a] Department of Biotechnology, University of Eldoret, Kenya, P.O Box 1125-30100, Eldoret, Kenya
[b] Centre for Biotechnology and Bioinformatics, University of Nairobi, P.O Box 30197-00100, Nairobi, Kenya
[c] Institute for Climate Change and Adaptation, University of Nairobi, P.O Box 30197-00100, Nairobi, Kenya
[d] International Crops Research Institute for the Semi-Arid Tropics, Eastern and Southern Africa, P.O Box 39063-00623, Nairobi, Kenya

ARTICLE INFO

Handling Editor: Dr. Jasmine Tomar

Keywords:
RNA viruses
Arachis hypogaea
GRD
CaMV
CPPV2
Meta-transcriptomics
GRAV
GRV
PeMOV

ABSTRACT

Background: Groundnut (*Arachis hypogaea* L.) is the 13th most important global crop grown throughout the tropical and subtropical regions of the world. One of the major constraints to groundnut production is viruses, which are also the most economically important and most abundant pathogens among cultivated legumes. Only a few studies have reported the characterization of RNA viruses in cultivated groundnuts in western Kenya, most of which deployed classical methods of detecting known viruses.

Methods: We sampled twenty-one symptomatic and three asymptomatic groundnut leaf samples from farmers' fields in western Kenya. Total RNA was extracted from the samples followed by First-strand cDNA synthesis and sequencing on the Illumina HiSeq 2500 platform. After removing host and rRNA sequences, high-quality viral RNA sequences were *de novo* assembled and viral genomes annotated using the publicly available NCBI virus database. Multiple sequence alignment and phylogenetic analysis were done using MEGA X.

Results: Bioinformatics analyses using as low as ~3.5 million reads yielded complete and partial genomes for Cauliflower mosaic virus (CaMV), Cowpea polerovirus 2 (CPPV2), Groundnut rosette assistor virus (GRAV), Groundnut rosette virus (GRV), Groundnut rosette virus satellite RNA (satRNA) and Peanut mottle virus (PeMoV) falling within the species demarcation criteria. This is the first report of CaMV and the second report of CPPV2 on groundnut hosts in the world. Confirmation of the detected viruses was further verified through phylogenetic analyses alongside reported publicly available highly similar viruses. PeMoV was the only seed-borne virus reported.

Conclusion: Our findings demonstrate the power of Next Generation Sequencing in the discovery and identification of novel viruses in groundnuts. The detection of the new viruses indicates the complexity of virus diseases in groundnuts and would require more focus in future studies to establish the effect of the viruses as sole or mixed infections on the crop. The detection of PeMoV with potential origin from Malawi indicates the importance of seed certification and cross-boundary seed health testing.

1. Introduction

Groundnut (or peanut) (*Arachis hypogaea* L.) is a self-pollinated (Lim and Gumpil, 1984) allotetraploid (AABB; 2n = 4x = 40) grain legume with a genome size of 2.7 Gb (Moretzsohn et al., 2013). It is the 6th most valuable vegetable oilseed crop and 13th most important crop globally, grown throughout the tropics and subtropical regions in more than 100 countries across six continents lying between latitudes 40°N and 40°S (Abate et al., 2012; Naidu et al., 1999; Okello et al., 2010). The estimated area under groundnut production worldwide is 22.6 million hectares resulting in 36.4 MM t of kernel production. The global average yield is 1600 kg ha^{-1}, which is higher than 1000 kg ha^{-1} average yield in sub-Saharan Africa (Abate et al., 2012). Groundnut production in Kenya is concentrated along the coastal and western parts of the country, where it is both a principal source of protein and a major source of cash income to smallholder growers, who are preponderantly women (Naidu

* Corresponding author.
E-mail address: Damaris.Odeny@icrisat.org (D.A. Odeny).

https://doi.org/10.1016/j.virol.2024.110011
Received 16 October 2023; Received in revised form 22 January 2024; Accepted 29 January 2024
Available online 6 February 2024

et al., 1999; Mukoye et al., 2015; Masira, 2017). The estimated area under groundnut cultivation in Kenya is 11,627 ha with a total production of 27,751 t (FAOSTAT, 2018).

Groundnut yield in Kenya has been on the decline with farmers obtaining even less than 50% of the potential output (Kipkoech et al., 2007; Mukoye et al., 2015). This drop in yield is attributed to both abiotic and biotic factors such as fungi, bacteria, nematodes, and viruses (Prasad et al., 2009). Within the group of viruses, riboviruses (RNA viruses) are significant due to their extremely adaptive nature to diverse environments resulting from their high mutation rates and formation of quasi-species (Roossinck, 2012; Watkiss, 2009). More than 30 viruses found in 14 genera have been reported to infect groundnuts naturally worldwide (Sreenivasulu et al., 2008). About 25 of these viruses are of economic importance, nineteen of which were first reported in groundnut, while the remaining were isolated from other alternative hosts (Sreenivasulu et al., 2008). A total of 11 groundnut viruses have been reported across Africa with only four of those reported in Kenya (Sastry et al., 2019). Groundnut Rosette Disease (GRD), which is a complex synergistic interaction of three viruses/agents (Deom et al., 2000), is endemic to sub-Saharan Africa and is the most devastating viral disease of groundnut in Africa. The three agents include Groundnut rosette assistor luteovirus (GRAV); Groundnut rosette umbravirus (GRV) and satellite RNA (satRNA) (Deom et al., 2000; Naidu et al., 1999).

The transmission of GRD is through an aphid, Aphis craccivora Koch, a polyphagous vector, which also feeds on alternative hosts (Mangeni et al., 2020; Singh and Singh, 2017), usually other cultivated legumes. Recent studies in Kenya reported the transfer of cowpea (Vilna unguiculata L. Walp.) viruses into groundnut (Mukoye et al., 2015; Orakha et al., 2019), possibly through the same vector or mechanical transfer. Mixed farming and/or intercropping, which is common among smallholder legume farmers in Africa (Desmae and Sones, 2017), enhances further the transfer of inoculum across different hosts. Without robust virus detection methods, such cross-host viral transfer events go unreported making disease management extremely cumbersome.

The groundnut viruses detection methods reported in Kenya so far have involved techniques such as serological assays, virus-sensitive indicator bioassays, and nucleic acid detection assays (i.e. nucleic acid hybridization and polymerase chain reaction) (Mukoye et al., 2015; Naidu et al., 1999; Sreenivasulu et al., 1991, 2008). Though these techniques are excellent at checking for known viruses, they have the restrained capability for characterizing unknown and novel viruses. Such assays are based on known sequences or reference information (CEFA, 2011), and therefore, cannot detect several viruses simultaneously. More importantly, they do not reveal the presence of mixed infections or co-infecting viruses in the same plant (Syller, 2012). There is an urgent need to use advanced methods such as *meta*-transcriptomics (Shi et al., 2018) to establish the true picture of the viral diversity in groundnut production areas in Kenya. Groundnut has been reported as one of the legumes with the highest number of viruses detected using different methods (Sastry et al., 2019).

Meta-transcriptomics allows for the rapid and relatively inexpensive assembly of viruses within a host sample (Shi et al., 2018) and can therefore greatly expand our knowledge on possible RNA virus diversity (Roossinck et al., 2015) in groundnut. *Meta*-transcriptomics provides an avenue for unearthing the different categories of RNA viruses of economic importance or otherwise (Gutiérrez Sánchez et al., 2016; Kreuze et al., 2009; Roossinck et al., 2015; Webster et al., 2015; Wu et al., 2015). The *Meta*-transcriptomics method has demonstrated reliability, whereas the gold standard method of RT-PCR assays alone could not (Marais et al., 2014), or could only be poorly detected (Amayo et al., 2012).

This study used *meta*-transcriptomics to investigate viruses in groundnut leaf samples obtained directly from farmers' fields in two Sub-counties in western Kenya. We hypothesized that there were abundant groundnut RNA viruses in these sub-counties owing to the common phenomenon of emergence and re-emergence of plant viruses. We collected both symptomatic and asymptomatic groundnut leaf samples from 24 sites and used a *meta*-transcriptomics approach to identify the RNA viruses.

2. Materials and methods

2.1. Plant material

Sampling was done in Gem, Siaya County, and Matayos, Busia County (Fig. 1) in western Kenya, in 2017. The two counties lie between latitudes 0° 28′N to 0° 42′ N and longitudes 33° 58′E to 34° 42′E with an average altitude of 1270 m (Climate Data - https://en.climate-data.org/location/11165-8/). This region has a mixture of rhodic-ferralsol and haplic-lixisol soil types (Omuto, 2013). It also experiences a bimodal rainfall pattern between January and December with the highest average of 1775 mm and a temperature range of 21–25 ^{O}C (Masira, 2017). Prior approval was obtained from all farmers whose fields were sampled. Leaf samples were collected from 24 farms, 12 each in Gem and Matayos Sub-counties. Each farm was surveyed using a W-pattern (Domola et al., 2008) and at least four leaf samples from different parts of the field were obtained and subsequently bulked as one sample. The sample collection was indiscriminate of any particular varieties of groundnut but based solely on planted varieties at the time. The leaf samples were collected from both asymptomatic and symptomatic plants. The samples were stored in RNA*later*® reagent (Sigma-Aldrich, Saint Louis, MO 63103 USA) and transported to the laboratory for processing. All 24-leaf tissues were used for total RNA extraction.

2.2. Total RNA isolation, purification, and quantification

Total RNA was extracted from all the 24 groundnut leaf samples that had been kept in RNA*later*® reagent using Direct-Zol MiniPrep kit (Zymo Research, Irvine, CA, USA) with TRIzol reagent (Thermo Fisher Scientific Inc., Carlsbad, CA 92008 USA) according to the manufacturer's instructions. Depletion of rRNA was done by adding 1 μl of thermostable RNase H (Thermo Fisher Scientific Inc., Carlsbad, CA 92008 USA) to each reaction tube followed by 20 min incubation at 37 °C to enrich for viral RNA. RNA quality was evaluated using agarose gel electrophoresis and the concentration was measured with Qubit® RNA BR Assay Kit (Life Technologies, USA). First-strand cDNA was synthesized using RevertAid First-strand cDNA Synthesis Kit (Thermo Fisher Scientific Inc., USA). Briefly, 100–500 ng of total RNA was mixed with 2–4 μl of random hexamer primer (Thermo Fisher Scientific Inc., Carlsbad, CA USA) in either 20 l or 40 l volumes. The reverse transcription was conducted at 25 °C for 5 min, followed by 45 °C for 60 min. The reaction was terminated by heating at 70 °C for 15 min. The First-strand cDNA samples were sent to Xcelris Labs Limited (Gujarat, India) for a paired-end Illumina TruSeq stranded RNA library preparation and sequencing. Concentrations of each of the samples and the barcodes used for sequencing are provided in Supplementary Table S1. Sequencing was done on Illumina HiSeq 2500 (Illumina, platform) using a 2 × 150 bp configuration.

2.3. Processing of the transcript reads

Raw Illumina reads were trimmed of sequencing adapters and low-quality bases using Trimmomatic v0.36 (Bolger et al., 2014). The quality of the trimmed reads was visualized using FASTQC v.0.11.5 (https://www.bioinformatics.bbsrc.ac.uk/projects/fastqc). The trimmed reads were first mapped to the groundnut reference genome, *Arachis hypogaea* cv. Tifrunner that had been retrieved from the PeanutBase database (https://v1.legumefederation.org/data/v2/Arachis/hypogaea/genomes/Tifrunner.gnm2.J5K5/) using Bowtie2 (Langmead and Salzberg, 2013) to remove groundnut host sequences. Ribosomal RNAs in the reads were removed by SortMeRNA software (Kopylova

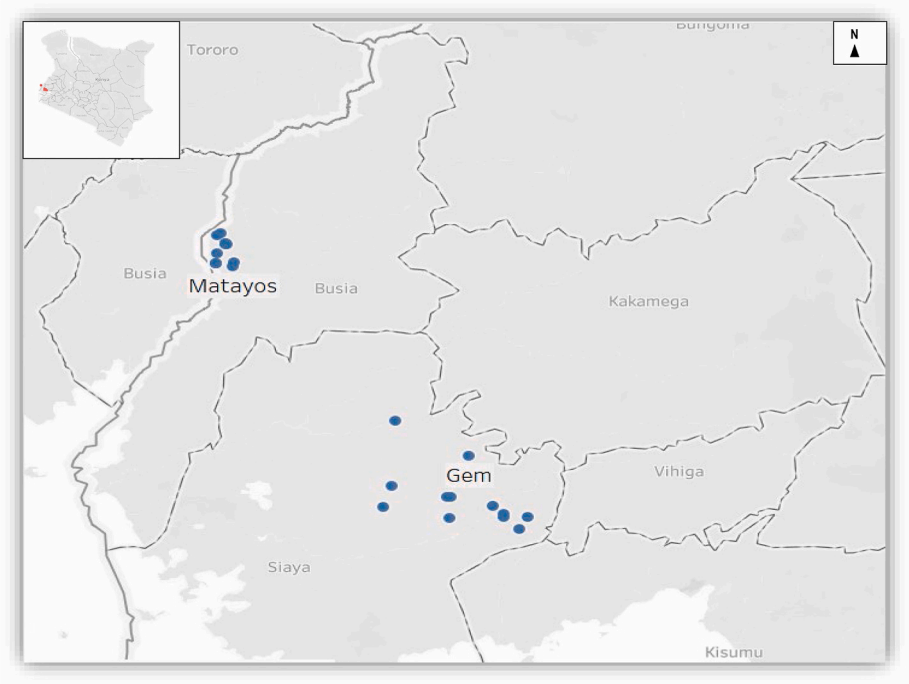


Fig. 1. Map of Kenya showing the sampling sites in Matayos sub-County (Busia) and Gem sub-County (Siaya). Source: Open Street Map drawn using Tableau Software.

et al., 2012). To further remove any additional bacterial contamination, the ribosomal RNA discriminated reads were retrieved and mapped to the NCBI RefSeq-based bacterial genomes (https://www.ncbi.nlm.nih.gov/genome/microbes/) before proceeding with *de novo*-assembly of virus reads using metaSPAdes (Nurk et al., 2017) and megaHIT (Li et al., 2016) assemblers. The Genome Detective (Vilsker et al., 2019), a pipeline that employs a combined assembly, was also used to do the assembly to compare the sensitivity of the assemblers involved.

2.4. Reads error correction and meta-transcriptomic assembly

The host-free and bacterial contamination-free reads obtained were error-corrected using the inbuilt module within the metaSPAdes assembler (Nurk et al., 2017). The reads were then assembled into viral contigs using the metaSPAdes (Nurk et al., 2017) and the megaHIT assembler (Li et al., 2016). The fill module within the MindTheGap pipeline (Guyomar et al., 2019) was used for finishing possible gaps between adjoining contigs. To confirm the integrity of the viral contigs obtained, the clean reads were first mapped to an adopted exhaustive local viral database (https://www.ncbi.nlm.nih.gov/genome/viruses/) and then the mapped reads were used to repeat the *de novo* assembly. The quality of assembly was assessed and visualized using the MetaQUAST (Mikheenko et al., 2016), which compares the statistics of the assemblies' software outputs and provides a basis for deciding on the best assembly outcome based on the parameters used.

2.5. Virus discovery, genome annotation and viral prevalence

The obtained putative viral contigs were used to determine sequence identity (%), coding sequences (Cds), Open reading frames (ORFs) and conserved protein domains. For sequence identity determination, the *de novo* assembled contigs were screened against the NCBI virus database (https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/) with an e-value cutoff of $\leq 1 \times 10^{-3}$ using BlastN. A BlastX (Altschul et al., 1990) was also performed against NCBI using the individual viral contigs to obtain the Cds (https://blast.ncbi.nlm.nih.gov/Blast.cgi?). The search for ORFs was done using the ORF finder (https://www.ncbi.nlm.nih.gov/orffinder/).

Finally, the assembled virus contigs were screened against the Conserved Domain Database (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi), with an e-value cutoff of $\leq 1 \times 10^{-3}$ to identify viral gene segments. The virus associations among the identified viruses were determined based on the proportion of samples with at least a viral contig detection with a high degree of similarity to that of the identified virus as per the demarcation criterion spelt out on ICTV protocol (https://www.microbiologyresearch.org/content/ictv-virus-taxonomy-profiles).

2.6. Multiple sequence alignment and phylogenetic analyses

The longest non-redundant *de novo* assembled viral contigs were used as queries to retrieve existing similar viruses from the NCBI virus database. The top 10–50 hits were retrieved from each group based on the extent of diversity and regional representation for use in further phylogenetic analysis. Multiple sequence alignment was done using the MUSCLE module within MEGA X with default settings (Kumar et al., 2018; Stecher et al., 2020). Maximum likelihood phylogenetic tree diagrams were generated for each of the six viral groups identified using the JTT matrix-based model (Jones et al., 1992) implemented in MEGA version X with 1000 bootstrap replicates as a test for the support of branches.

3. Results

3.1. Symptoms observed in the sampling fields

Of the twenty-four samples collected, twenty were from symptomatic groundnut plants and four were from asymptomatic plants (Table 1). The symptoms observed were typical virus symptoms including leaf chlorosis (LC), mottling (MO), necrotic spots (NR), puckering (P), shriveling (SH), leaf distortion (LD) and stunting (ST) (Table 1; Fig. 2). All the symptomatic groundnut plants, except one (M7 that showed chlorosis), displayed more than one symptom on the same plant (Table 1).

Table 1
A summary of symptoms observed from groundnut plants that were samples.

Sample ID	Symptoms	Sample ID	Symptoms
1. M1	LC, NR	13. G1	LC, NR, SH
2. M2	LC, MO	14. G7	LD, ST
3. M3	Asymptomatic	15. G8	PU, MO
4. M4	LC, MO	16. G12	LC, MO
5. M5	MO, ST	17. G13	MO, SH, LC
6. M7	LC	18. G23	MO, LC
7. M8	MO, SH, LC	19. G24	MO, LC, NR
8. M9	Asymptomatic	20. G25	NR, MO, LC
9. M10	NR, MO	21. G27	LC, NR
10. M11	NR, MO	22. G28	Asymptomatic
11. M15	LC, NR, MO	23. G30	LC, NR
12. M16	NR, MO, LC	24. G33	Asymptomatic

LC: Leaf chlorosis; MO: Mottling; NR: Necrotic spots; P: Puckering; SH: Shriveling; LD: Leaf distortion; ST: Stunted Plant.

3.2. RNA-sequencing and data output

A total of 295, 904, 168 raw reads were generated from 24 groundnut leaf samples with a minimum of 3,591,790 reads for sample G7 and a maximum of 29, 074, 936 for sample G13 (Table 2). After trimming and removal of host sequences and microbial contamination, about 30% (89, 639, 878) of the original raw reads were retained for further analysis (Table 2).

3.3. Viral genome assemblies and annotation

Eighty contigs were obtained from the combined viral genome assemblies of lengths ranging from 101 bp to 5671 bp (Table 3; Supplementary Table S2). The contigs retrieved covered varying proportions of the virus ranging from 8 to 100% viral genomes (Supplementary Table S2). Forty-one contigs were putatively annotated as known groundnut viruses (Table 3) with BLAST similarity ranges of 88–98%. The groundnut viruses and a satellite RNA identified were GRAV, GRV, and satRNA – the three viruses causing groundnut rosette disease and peanut mottle virus (PeMoV) (Table 3). An additional 39 contigs were positively annotated as Cauliflower mosaic virus (CaMV) and cowpea polerovirus 2 (CPPV2) with a BLAST similarity range of 92–99% (Table 3). In general, CPPV2 was the most predominant virus in the groundnut samples analyzed as it was detected in eleven out of the 16 positive samples. The majority of contigs assembled in the study represented partial genomes of the six detected viruses. We recovered one complete contig sequence for the satRNA virus genome (892 nt). Of the six identified viruses, four were detected across the two Sub-counties, while PeMoV was confirmed in Matayos, Busia County and CaMV was established only in Gem, Siaya County. PeMoV was detected in just one sample, which was from the Matayos sub-county in Busia County. All three samples from which CaMV was detected were from Gem, Sub-county in Siaya County (Table 3). Five symptomatic samples (G7, M1, M2, M7, M15) and three asymptomatic samples (G28, G33, and M9) had no detectable virus contig(s), while sample M3 was asymptomatic but tested positive for GRV contig (Table 1; Table 3).

3.4. Virus associations

Virus associations in the study were single, co-infections and multiple infections: four instances of single infections, five cases of co-infections and seven of multiple infections in the Gem sub-county and Matayos sub-county (Fig. 3). The single infections were for CPPV2, GRV and satRNA (Fig. 3). CPPV2 virus infection was further observed in eleven samples either in co-infection or in multiple infections, turning out to be the most prevalent virus detected in the study (Fig. 3). PeMoV virus was detected once as a co-infection of CPPV2 virus, in Matayos, Busia County (Fig. 3). All the CaMV virus cases detected in the study were from Gem, Siaya County and were always part of a multiple infection of at least one of the Groundnut rosette complex viruses, and CPPV2 virus (Fig. 3). There was a single case of GRAV-satRNA present

Table 2
Summary of reads generated and processed.

Sample ID	Raw reads	Trimmed reads	Host subtracted reads	Clean error-corrected reads
M1	8,709,082	7,659,238	3,104,972	3,037,524
M2	14,025,988	12,842,576	1,171,228	1,138,390
M3	10,901,674	9,719,754	542,266	513,654
M4	15,145,008	13,401,288	763,570	740,814
M5	9,195,594	8,336,288	1,167,580	1,133,384
M7	10,195,430	9,232,160	1,529,049	997,034
M8	16,674,828	15,051,646	2,185,994	2,152,998
M9	8,522,016	7,693,678	1,085,332	1,056,926
M10	5,707,668	4,637,402	371,410	363,120
M11	20,123,234	17,483,048	5,485,462	5,441,030
M15	13,838,288	11,956,266	2,416,092	2,353,552
M16	17,291,326	15,318,282	3,269,702	3,221,302
G1	5,585,458	4,830,708	1,438,472	1,391,908
G7	3,591,790	3,161,988	1,207,396	1,167,246
G8	8,284,306	7,272,324	828,172	768,500
G12	4,617,446	3,990,550	551,296	527,764
G13	29,074,936	25,127,114	4,423,918	4,328,766
G23	6,369,114	5,323,634	1,578,920	1,515,976
G24	18,399,678	15,424,836	14,812,090	14,473,454
G25	20,262,388	16,849,324	15,735,684	15,569,976
G27	11,028,814	9,635,140	3,048,854	2,970,226
G28	14,714,052	12,760,382	11,967,884	11,671,106
G30	12,393,254	10,608,252	6,285,670	6,132,304
G33	11,252,796	9,505,080	7,125,230	6,972,924
Total	**295,904,168**	**257,820,898**	**92,096,243**	**89,639,878**

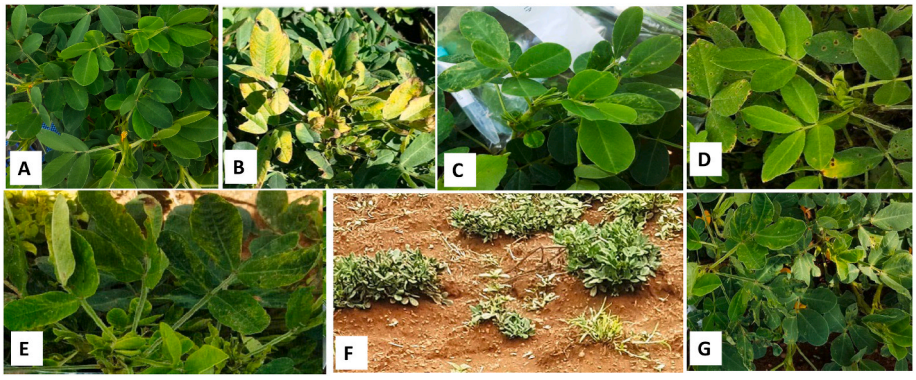


Fig. 2. Symptoms observed from the **groundnut** plants sampled. A. Asymptomatic sample. B. Yellowing of leaves. C. Mottling of the leaves. D. Necrotic spots are visible on the leaves. E. Leaf curling and mottling. F. Green rosette with stunting. G. A mixture of leaf curling, mottling and leaf distortion symptoms.

Table 3
Contig details for the viruses plus a satellite RNA identified on groundnut plants in the study.

Virus annotation	Known hosts	# Contigs	Length Range	Similarity %	E-value	[a]Groundnut Sample IDs
GRAV	Groundnut	9	101–1037	98	2.6e-61	M8, M10, M11, G8, G13, G24, G25, G27, G30
GRV	Groundnut	18	196–3964	88	2.5e-60	M3, M8, G1, G12, G13, G24, G25, G30
SatRNA	Groundnut	11	205–892	90	1.1e-45	M4, M8, M11, G1, G12, G13, G24, G25, G30
PeMoV	Groundnut	3	213–379	96	1e-68	M5
CaMV	*Brassicaceae*	4	214–501	99	3e-92	G8, G13, G24
CPPV2	Cowpea	35	168–5671	92	4.0e-79	M5, M8, M10, M11, M16, G8, G13, G23, G24, G25, G27

[a] All samples coded with M are from Matayos, Busia County while those coded with G are from Gem, Siaya County.

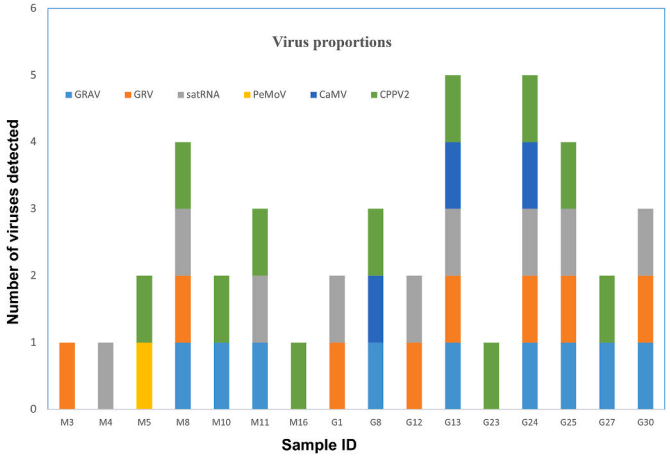


Fig. 3. A graphical representation of the virus proportions detected in each sample.

together, two instances of GRV-satRNA and five instances of association of GRAV-GRV-satRNA. In seven samples across the two sub-counties, at least two of the viruses causing groundnut rosette disease were confirmed. Overall, GRAV (in 9 samples out of 24), satRNA (9 samples out of 24) and GRV (in 8 out of 24 samples) in the study (Fig. 3).

3.5. *Phylogenetic analyses of the detected viruses and a satellite RNA*

Near complete and partial genomes of CaMV, CPPV2, GRAV, satRNA and PeMoV obtained from the combined assembly were used for phylogenetic analysis of each virus strain identified. Distinct phylogenetic trees assigned to each virus category were shown in each case.

3.5.1. *Groundnut rosette assistor virus*

Clustering the 11 distinct GRAV coat protein contigs identified in the current study with 29 other publicly available isolates revealed a monophyletic grouping containing two sub-clusters, mainly corresponding to the regions where the groundnut hosts were cultivated (Fig. 4). The West African sub-cluster was predominated by Ghanaian and Nigerian isolates, although 2 each of Kenyan and Malawian isolates from previous studies also clustered with the West African isolates. The largest cluster comprised mainly of Kenyan isolates with just one exception from Malawi (Fig. 4). Nine of the 11 contigs identified in the current study clustered together with the Kenyan isolates. One isolate from the Matayos sub-County (M8) was an outgroup (Fig. 4). The phylogenetic tree further indicated that GRAV strains in West and East Africa are closely related or are of a similar strain (Fig. 4).

3.5.2. *Groundnut Rosette Virus*

Phylogeny based on the overlapping ORF3/ORF4 - GRV hypothetical protein sequence revealed four major clusters representing Kenyan, Malawian, Nigerian and Ghanaian isolates (Fig. 5). All the isolates from the current study grouped with Kenyan isolates except one, which fell

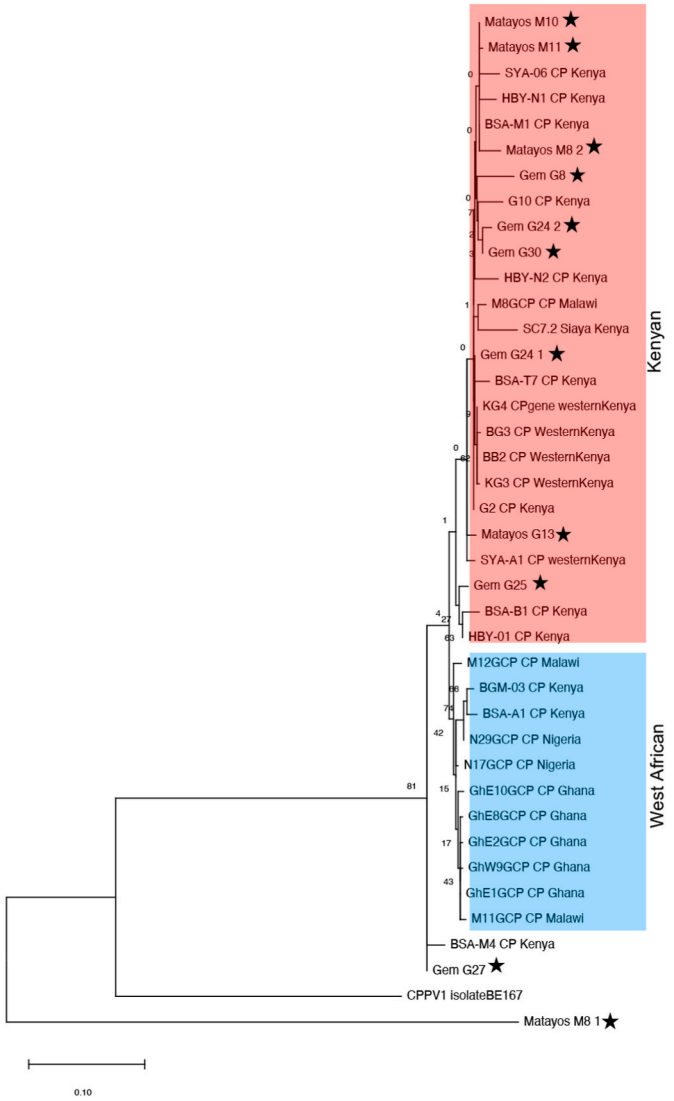


Fig. 4. Clustering of the coat protein of GRAV isolates identified in the current study with those from publicly available databases. Two major clusters were observed within a monophyletic grouping: West African (in blue) and Kenyan (in Red). All isolates from the current study are highlighted with a black star.

into the Malawian group. Malawian and Kenyan isolates are monophyletic groups, similar to the Ghanaian and Nigerian groups. The eastern African (Kenyan and Malawian) isolates appeared more closely related in comparison to the western African (Ghanaian and Nigerian) (Fig. 5).

3.5.3. *Groundnut Rosette Virus-satellite RNA*

Phylogenetic analysis of Groundnut Rosette Virus-satellite RNA was

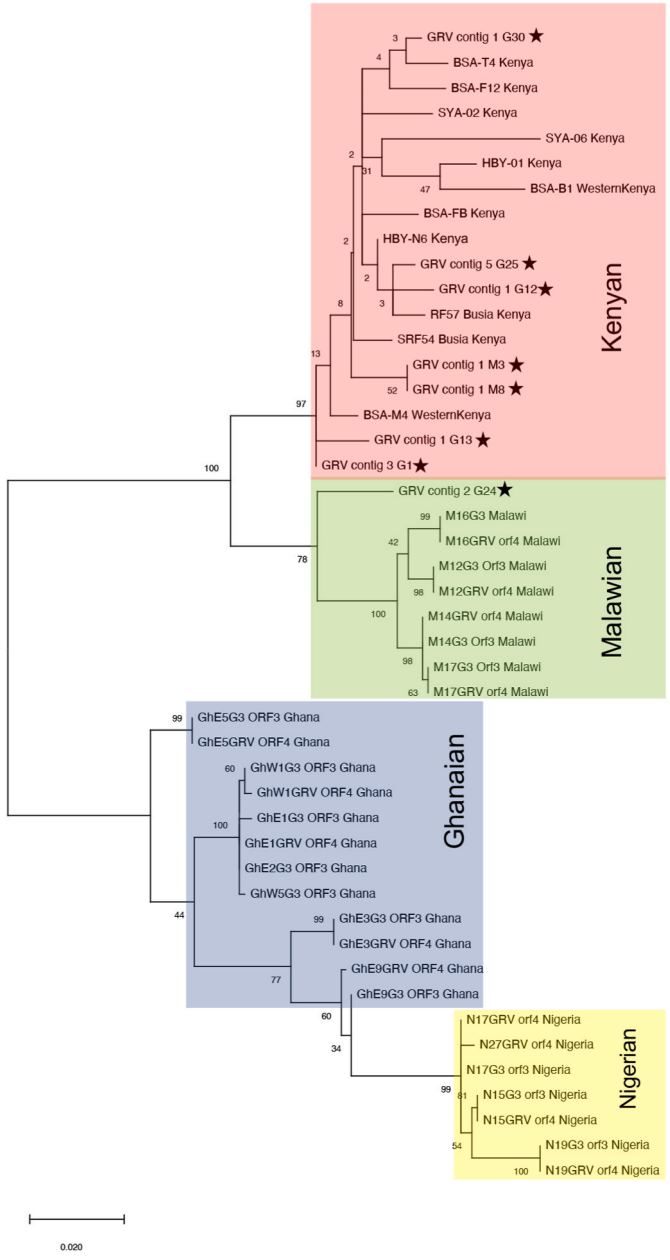


Fig. 5. Phylogenetic tree denoting relationships among GRV isolates constructed based on GRV overlapping ORF3/ORF4 hypothetical protein. All the isolates from the current study are highlighted using black stars. The four major groupings are further distinguished using four different colour backgrounds.

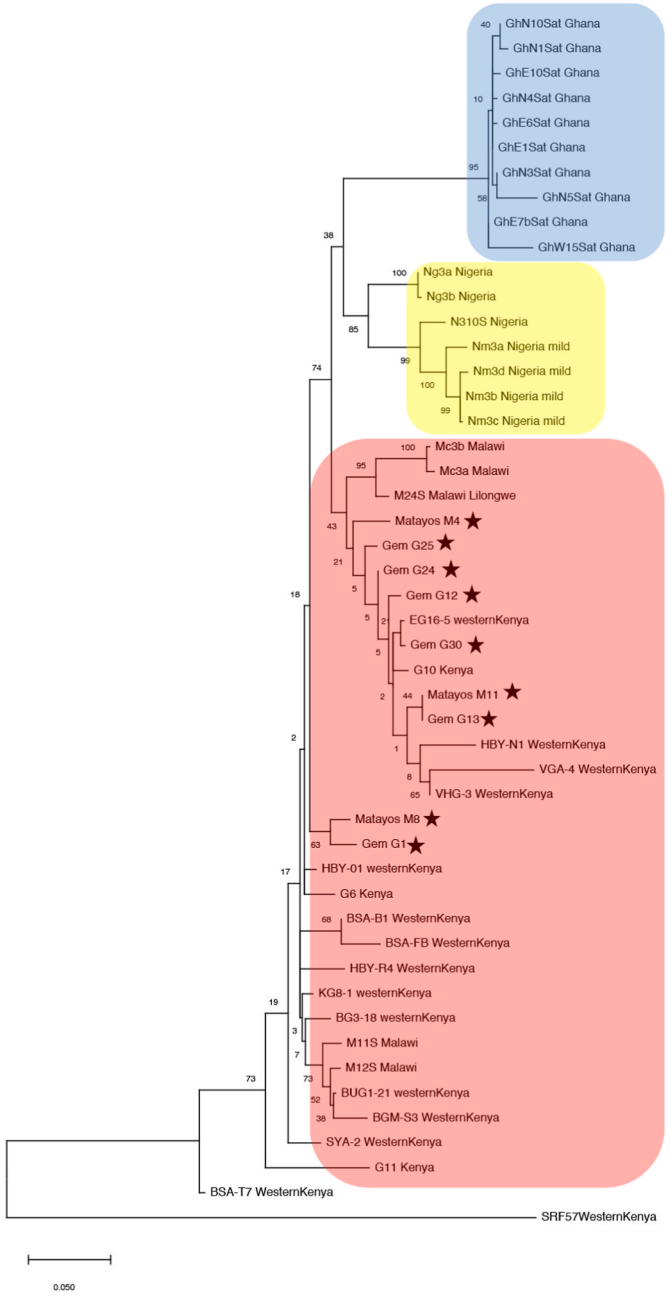


Fig. 6. Maximum-likelihood tree for Groundnut Rosette Virus – satellite RNA isolates based on nucleotide sequences. The clusters highlighted in blue, yellow and red comprise isolates from Ghana, Nigeria, Kenya **and** Malawi respectively. Black stars highlight isolates from the current study.

based on the virus nucleotide sequences generated from our study alongside 39 other publicly available homologous sequences. We established three distinct polyphyletic groups representing isolates from Kenya, Malawi, Nigeria and Ghana (Fig. 6). All the isolates from the present study were grouped with the Kenyan-Malawian group confirming their closer relationship with those from Malawi and Kenya than those from West Africa (Fig. 6).

3.5.4. *Peanut mottle virus*

Phylogenetic analysis of PeMoV was undertaken using polyprotein sequences. The sole PeMoV isolate identified in the current study was compared with 21 other isolates selected from the public databases. The phylogenetic tree portrayed two major clusters, one cluster of mottle viruses isolated from soybean {*Glycine* max (L.) Merr.}, and the second of mottle viruses isolated from *P. vulgaris* or peanut (Fig. 7). Although a complete genome of the Kenyan PeMoV was included in the phylogenetic analysis, our isolate PeMoV_M5 grouped with four other isolates from diverse hosts, from Mexico (SJ 8

5), USA (*Nicotiana clevelandii*), Tanzania (*Phaseolus vulgaris*) and Zambia (*Phaseolus vulgaris*) (Fig. 7). The existing Kenyan PeMoV complete genome clustered with the soybean isolates from South Korea (Fig. 7). The phylogenetic tree further showed a wide geographical distribution of the PeMoV with a diverse number of hosts (Fig. 7).

3.5.5. *Cauliflower mosaic virus*

The phylogeny of CaMV was based on ORF2 sequences. The three isolates identified in the current study clustered together with the publicly available CaMV complete genomes with 100% nodal support (Fig. 8). We obtained two major clusters, one cluster predominantly of

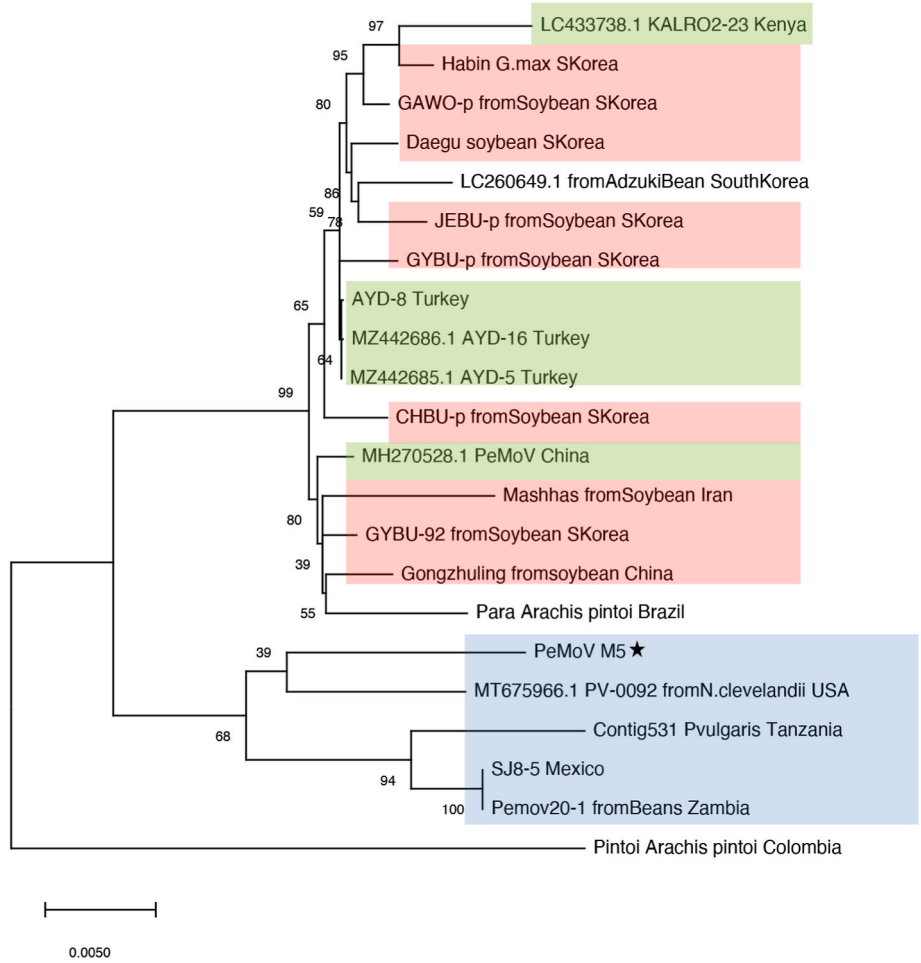


Fig. 7. Maximum-likelihood tree for Peanut Mottle virus isolates based on the polyprotein sequence. Distinct coloured backgrounds denote the major grouping in the phylogenetic diagram, while the black star highlights the study isolate.

mosaic viruses (6 out of 10), and the second cluster-comprised viruses from Dahlia, *Lactuca sativa* and *Tanacetum cinerariifolium* (Fig. 8).

3.5.6. *Cowpea polerovirus 2*

Phylogenetic analysis for the 11 CPPV2 isolates identified in the current study was based on polerovirus, CPPV2 CDC, read-through protein sequences and was done alongside 21 additional isolates from public databases. Three major clusters were identified, the first cluster predominantly of Cucurbit aphid-borne yellows virus (CABYV), the second of Phasey bean mild yellows virus (PBMYV) and the third of Cowpea Polerovirus 2 (CPPV2). Two other minor groupings were obtained for members of the family *Luteoviridae*: GRAV isolates (GRAV SC7.1, GRAV SC7.2) and Chickpea chlorotic stunt virus (CpCSV) (Fig. 9). All isolates from the current study clustered with the publicly available Cowpea polerovirus, except Gem G23, which clustered with the PBMYV isolates. We also noted that our study isolates did not cluster based on the area of collection (Fig. 9).

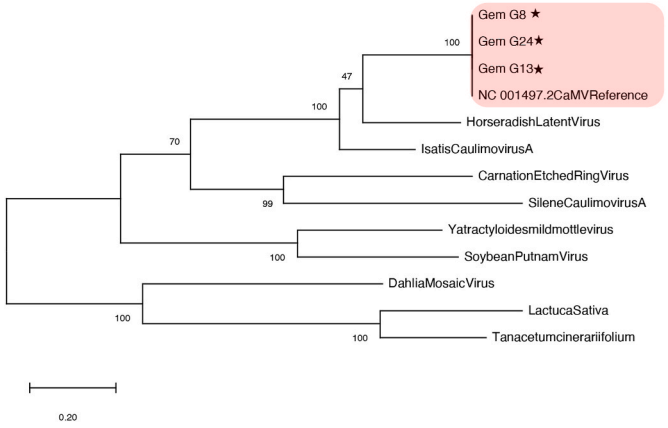


Fig. 8. Phylogenetic tree denoting relationships among CaMV isolates constructed using the CaMV ORF2 sequences. All the isolates from the current study are highlighted using black stars. The cluster highlighted in red comprises isolates from the current study.

4. Discussion

We explored groundnut leaf virome through a *meta*-transcriptomics approach and identified near complete and partial genomes of five viruses and a satellite RNA from symptomatic plant samples and an asymptomatic samples collected directly from groundnut fields in Western Kenya; Four of the viruses and a satellite RNA have RNA genomes (CPPV2, GRAV, GRV, satRNA, PeMoV) while one has DNA genome (CaMV). To the best of our knowledge, this is the first report of CaMV on groundnut samples in the world; and the second report of CPPV2 on groundnut fields in Western Kenya (Were et al., 2019).

The total number of paired-end sequence reads generated in the study ranged from ~3.5 m to 29 m and averaged 12 m reads per sample. This data output is comparable to that generated in other studies involving groundnut virus surveys (Li et al., 2022) and soybean (Elmore

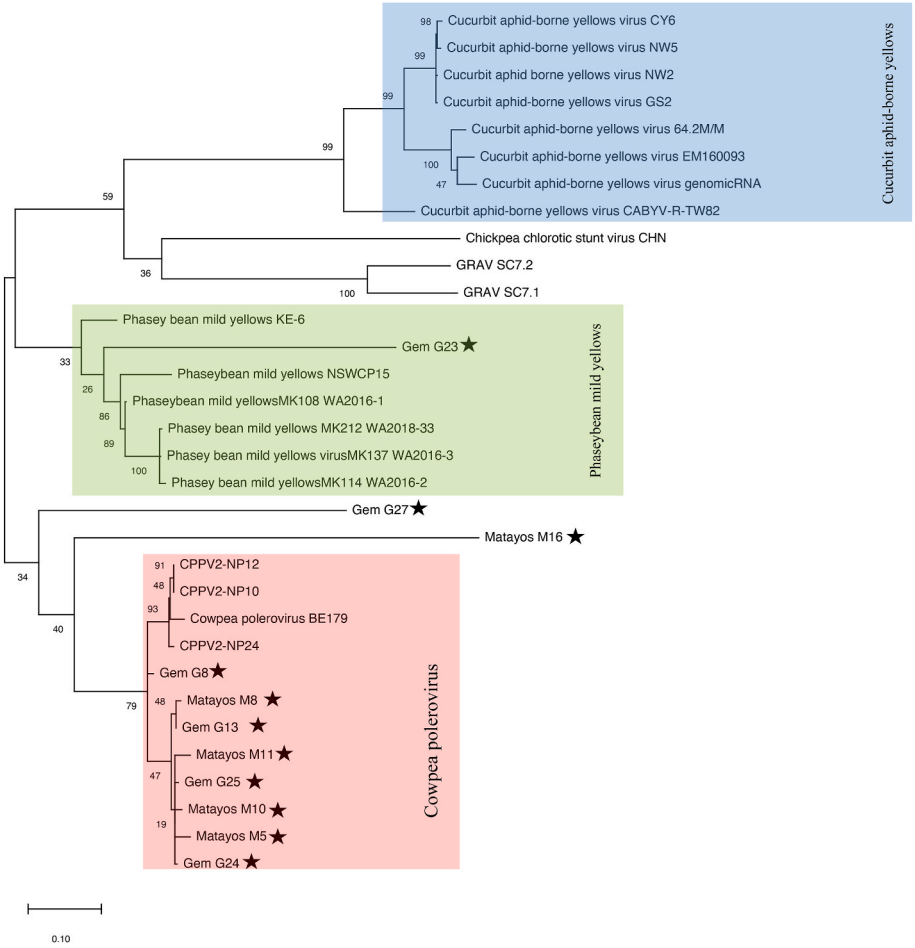


Fig. 9. Maximum-likelihood tree for CPPV2 virus isolates based on the CPPV2 read-through protein sequence. The distinct coloured backgrounds denote the major grouping in the phylogenetic diagram, while the black stars highlight the study isolates.

et al., 2022) but lower than reported in alfalfa (*Medicago sativa* L.) (Nemchinov et al., 2022). While conducting a more extensive groundnut viruses' survey in the same region, Were et al. (2019) identified up to 20 viruses using both ELISA and NGS approaches. Therefore, we cannot rule out the likelihood of additional or new viruses that may have gone undetected by our study because of lower sequence reads. However, our ability to detect the common and novel groundnut viruses from as low as 3.5 million raw reads per sample suggests that the optimum number of sequence reads to recommend for any given study would depend upon the objective of the study, the host and the virus species (Mollov and Malapi-Wight, 2016).

The first detection of CaMV and only the second detection of CPPV2, is significant and has several implications for future management of viral diseases of groundnut, ranging from the likely emergence of virus strains, the unreliability of existing virus detection methods and the potential widening of geographical distribution and increased host ranges for the viruses (Burrell et al., 2017; Dennehy, 2017). Though CaMV was not previously reported in groundnut, aphids transmit the virus, the same vector that transmits most groundnut viruses (NG and Perry, 2004). Furthermore, the most common vegetable in subsistence farms in Kenya is "Sukuma wiki", Kale (*Brassica oleraceae* var. Acephala), which is a host of CaMV (Spence et al., 2007). The groundnut samples used in the current study had been grown alongside other crops in an intercropping system having "Sukuma wiki" and other weedy plants, making cross-host infection by the same vector possible. The same case applies to CPPV2, which is also transmitted by aphids, and the main host, cowpea, is often grown alongside groundnut crops on farmers' fields in western Kenya.

Certain CaMV strains have been reported to infect members of the *Solanaceae* species in addition to the known Brassicaceae family (Haas et al., 2002). This is also not the first member of the *Caulimovirus* genus to infect groundnuts. The Peanut chlorotic streak virus (PCISV) was the first *Caulimovirus* isolated in a groundnut host (Reddy et al., 1993). Other than, in Kenya where CPPV2 has been reported twice, first by Were et al. (2019), elsewhere, CPPV2 has been detected only in cowpea plants (Palanga et al., 2017; Kwak et al., 2022). Given that groundnut is a host to several virus groups, we expect that more extensive viral metagenomics studies in the future will continue to detect additional novel viruses in groundnut and other hosts as well.

We chose to collect samples from both symptomatic and asymptomatic samples. Given the affliction of groundnut plants by several virus groups, it would be difficult to rely on a specific symptom for the detection of a virus infection. Coupled by the mixed infections from other not-so-obvious viruses such as CaMV and CPPV2, which are novel and likely to present different symptoms in the new hosts (Palanga et al., 2017; Orakha et al., 2019; Bak and Emerson, 2020), our study confirms the ability of Next Generation Sequencing (NGS) for a more reliable diagnosis of virus infections in most crops. We also detected a viral contig (GRV) in one of the asymptomatic samples further supporting Okello et al. (2014) on the unreliability of using solely symptoms to identify virus infections in groundnuts. The lack of symptoms in a GRV-containing sample is not surprising, as we know that the presence of GRAV or GRV solely could result in mild to no symptoms (Waliyar et al., 2007).

Virus diseases tend to be complex owing to different subtle interactions that occur between/or among the concerned viruses, and the

host, in addition to other factors such as adaptation to the host, within-host interactions and the interaction between or among the viruses (Moreno and López-Moya, 2020). Our study reported different virus associations ranging from single to as many as five viruses in one sample. Multiple/mixed virus infections are remarkably common (Anitha et al., 2014; Ferriol et al., 2020; Moreno and López-Moya, 2020; Singhal et al., 2021; Were et al., 2019; Xu et al., 2022). GRAV, GRV and satRNA are known to exhibit synergistic interaction among themselves with GRAV contributing to the encapsidation of both GRV and satRNA, while satRNA is responsible for variance in symptomatology (Deom et al., 2000). It is well known that satRNA is packaged with GRV and depends on GRV for its replication, (Taliansky et al., 2000). The sole occurrence of satRNA in one of our symptomatic samples is highly likely to be a case of missed detection of GRV, or GRV and GRAV.

The occurrence of CaMV, CPPV2, GRAV, GRV and satRNA in mixed infections is not unique to our study and can be attributed to the fact that these viruses and satRNA are all transmitted by polyphagous aphid species (Bak and Emerson, 2020; Naidu et al., 1998, 1999; Okello et al., 2014; Palanga et al., 2017). Cowpea polerovirus 2 belongs to the *Polerovirus* genus, which has a wide host range, from cucurbits to cereals, and is transmitted by a wide range of aphid species found in four genera (Latourrette et al., 2021). Poleroviruses have often been found in mixed infections with other viruses (Moreno and López-Moya, 2020). Notably, CaMV, a *Caulimovirus*, has been reported in mixed infections with Turnip mosaic virus in other Brassicaceae hosts (Twardowicz-Jakusz et al., 1999). The interactions in co-infection/mixed infection can either be synergistic antagonistic or neutral (Moreno and López-Moya, 2020; Syller, 2012). The specific interactions among the mixed infections remain to be determined and should be a priority in future studies.

Peanut mottle virus was the only seed-borne virus (Demski, 1975) identified in our study, which further confirms the findings of a previous study (Were et al., 2019) in the same region. Groundnut farmers in western Kenya do not typically use certified groundnut seed, so the result is not surprising. Future studies will need to quantify and establish the actual cost of investment in certified seed by taking into consideration the prevalence and extent of yield loss attributed to this seed-borne virus.

The detection of five viruses in groundnut leaf samples collected from western Kenya further supports the suggestion that groundnut is one of the legumes with the highest number of infecting viruses (Sastry et al., 2019). Although the actual economic importance remains to be determined, the discovery of CaMV and CPPV2 in groundnut will likely add to the pressure on groundnut production. We suggest an extensive study in the same region to establish the effect of the viruses on the crop, either as sole or mixed infections. The nature of the various virus associations detected in our study could reveal a greater intricate interaction and possibly reveal more viruses than reported. A recent review by Trebicki (2020) predicted increases in the severity of viruses through changes in hosts and vectors.

The GRAV isolates in the study exhibited the closest identities with the majority of previously identified Kenyan isolates and were distinct from West African (Nigeria and Ghana) isolates. Our findings support those of Mabele et al. (2021) and Appiah et al. (2017) who reported similarities within specific geographical regions but could also be indicative of germplasm and seed sharing across specific regions. GRV isolates, on the other hand, formed distinct groups representing countries of origin and suggesting a different and potentially faster evolution pattern in comparison to GRAV. Earlier studies reported region-specific clustering of GRV and sat-RNA isolates in comparison to GRAV (Deom et al., 2000). Indeed, the satRNA clustering pattern was also country-specific in our study and mirrored that of GRV more than GRAV. These results indicate that a thorough characterization of the GRD agents will be necessary for a better understanding of their interactions towards disease symptoms across different environments.

The clustering of different PeMoV isolates by host and by region further confirmed earlier results (Beikzadeh et al., 2015) and reflected the seed-borne transmission nature of the virus. The CaMV isolates from our study were highly similar (99%) in the nucleotide sequence and 100% nodal support clustering with the reference CaMV (Franck et al., 1980). The absence of major genetic variations among the CaMV isolates is evidence of CaMV genomic stability despite the change of host plants. However, a more exhaustive study would need to be done with significantly more isolates from Kenya and other countries as well as other hosts to confirm the stability of their genome. The clustering of CPPV2 virus contigs suggested that one of the viruses detected was potentially a PBMYV but will need to be further validated with complete genome sequences and more samples. According to Palanga et al. (2017), both PBMYV and CPPV2 have high genetic similarity and can form clusters with 100% bootstrap support in a phylogenetic tree drawn using the P1–P2 protein sequence.

The *meta*-transcriptomics approach demonstrated in this study offers not only an easy but also a rapid way for the initial diagnosis of these groundnut viruses, shortening their diagnosis by pointing to the possible viral pathogen, which can be validated by other virus diagnostic methods. These findings offer a basis for the universal diagnosis of the six viruses together by providing an opportunity for designing primers for the development of suitable diagnostic assays for the simultaneous identification of the viruses.

5. Conclusion

In this study, we have demonstrated the use of the *meta*-transcriptomics approach in the discovery and identification of six viruses on cultivated groundnut fields starting with as low as ~3.5 million raw paired-end sequence reads. The study further asserts that various virus associations are commonplace in groundnut plants grown in Western Kenya implying more intricate virus interactions not reported before. The majority of isolates for various viruses identified in the study have shown clustering reflecting geographical correlation and having close genetic identities. We proposed the use of the *meta*-transcriptomics procedures as an initial diagnostic procedure for the detection and identification of ground viruses. This approach is suitable for groundnut virus disease monitoring and surveillance in seed system development, exchange and breeding works given its capability for generic use.

Funding

This work was supported by NUFFIC (The Netherlands Organization for International Co-operation in Higher Education through The Netherlands Initiative for Capacity Building in Higher Education (NICHE)), Grant No. CF9383 awarded to the University of Eldoret, Kenya; and partly by the grant from the USAID , FEED THE FUTURE - KENYA, awarded to the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Nairobi, Kenya, Grant No. 081.

Data availability

Raw data are available at NCBI, Bioproject ID: PRJNA788763. The sequences for the six viral isolates in the study were deposited under the following GenBank Accession Numbers: OL999574-OL999577, OL999579-OL999580, OL999582, OL999584-OL999597, OL999601-OL999602, OL999604-OL999606, OL999609-OL999610, and OL999612- OL999615.

CRediT authorship contribution statement

Dennis Obonyo: Conceptualization, Data curation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **George Ouma:** Supervision. **Rachel Ikawa:** Conceptualization, Investigation, Methodology, Supervision. **Damaris A. Odeny:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing

– review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors are greatly indebted to the ICRISAT field officers in Gem, Siaya County and Matayos, Busia County, Kenya, and all the farmers in the two sub-counties for their warm welcome and permission to collect groundnut leaf samples from their fields. We are sincerely thankful to Accadius Lunayo for his useful contributions to the discussions of this manuscript and Samuel Manthi for technical assistance with the laboratory work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.virol.2024.110011.

References

Abate, T., Alene, A.D., Bergvinson, D., Shiferaw, B., Silim, S., Orr, A., Asfaw, S., 2012. Tropical Grain Legumes in Africa and South Asia: Knowledge and Opportunities; *International Crops Research Institute for the Semi-arid Tropics (ICRISAT), Order Code: BOE 056*.

Altschul, S.F., Gish, W., Pennsylvania, T., Park, U., 1990. Basic local alignment search tool. Mol. Biol. 215, 403–410.

Amayo, R., Arinaitwe, A., Mukasa, S., Tusime, G., Kyamanywa, S., Rubaihayo, P., Edema, R., 2012. Prevalence of viruses infecting cowpea in Uganda and their molecular detection. Afr. J. Biotechnol. 11 (77), 14132–14139. https://doi.org/10.5897/ajb11.398.

Anitha, S., Monyo, E.S., Okori, P., 2014. Simultaneous detection of groundnut rosette assistor virus (GRAV), groundnut rosette virus (GRV) and satellite RNA (satRNA) in groundnuts using multiplex RT-PCR. Arch. Virol. 159 (11), 3059–3062. https://doi.org/10.1007/s00705-014-2139-7.

Appiah, A.S., Sossah, F.L., Tegg, R.S., Offei, S.K., Wilson, C.R., 2017. Assessing sequence diversity of Groundnut rosette disease agents and the distribution of Groundnut rosette assistor virus in major groundnut-producing regions of Ghana. Tropical Plant Pathology 42 (2), 109–120. https://doi.org/10.1007/s40858-017-0140-x.

Bak, A., Emerson, J.B., 2020. Cauliflower mosaic virus (CaMV) biology, management, and Relevance to GM plant detection for Sustainable Organic agriculture. Front. Sustain. Food Syst. 4 (March), 1–8. https://doi.org/10.3389/fsufs.2020.00021.

Beikzadeh, N., Hassani-Mehraban, A., Peters, D., 2015. Molecular identification of an isolate of peanut mottle virus (Pemov) in Iran. J. Agric. Sci. Technol. 17 (3), 765–776.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30 (15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Burrell, C.J., Howard, C.R., Murphy, F.A., 2017. Laboratory diagnosis of virus diseases. In: Fenner and White's Medical Virology. Elsevier, pp. 135–154. https://doi.org/10.1016/B978-0-12-375156-0.00010-2.

CEFA, 2011. Good Agronomic Practices for Groundnut in Western Kenya Good Agronomic Practices for Groundnut in Western Kenya Training Manual for Trainers. Au.

Climate Data. https://en.climate-data.org/location/11165-8/[Accessed 18 July 2021].

Demski, J.W., 1975. Source and spread of peanut mottle virus in soybean and peanut. Phytopathology 65 (8), 917. https://doi.org/10.1094/phyto-65-917.

Dennehy, J.J., 2017. Evolutionary ecology of virus emergence. Ann. N. Y. Acad. Sci. 1389 (1), 124–146. https://doi.org/10.1111/nyas.13304.

Deom, C.M., Naidu, R.A., Chiyembekeza, A.J., Ntare, B.R., Subrahmanyam, P., 2000. Sequence diversity within the three agents of Groundnut rosette disease. Phytopathology 90 (3), 214–219. https://doi.org/10.1094/PHYTO.2000.90.3.214.

Desmae, H., Sones, K., 2017. Groundnut Cropping Guide. Africa Soil Health Consortium (ASHC). http://africasoilhealth.cabi.org/wpcms/wp-content/uploads/2017/05/ASHC-English-Groundnut-A5-colour-lowres.pdf.

Domola, M.J., Thompson, G.J., Aveling, T.A.S., Laurie, S.M., Strydom, H., Berg van den, A.A., 2008. Sweet potato viruses in South Africa and the effect of viral infection on storage root yield. Afr. Plant Protect. 14, 15–23.

Elmore, M.G., Groves, C.L., Hajimorad, M.R., Stewart, T.P., Gaskill, M.A., Wise, K.A., Sikora, E., Kleczewski, N.M., Smith, D.L., Mueller, D.S., Whitham, S.A., 2022. Detection and discovery of plant viruses in soybean by metagenomic sequencing. Virol. J. 19 (1), 1–24. https://doi.org/10.1186/s12985-022-01872-5.

FAOSTAT, 2018. Food and Agriculture Organization of the United Nations.

Ferriol, I., Chase, O., Domingo-Calap, M.L., López-Moya, J.J., 2020. Mixed infections of plant viruses in crops: solo vs. Group game. Proceedings 50 (94). https://doi.org/10.3390/proceedings2020050094.

Franck, A., Guilley, H., Jonard, G., Richards, K., Hirth, L., 1980. Nucleotide sequence virus DNA of cauliflower mosaic. Cell 21 (August).

Gutiérrez Sánchez, P.A., Jaramillo Mesa, H., Marin Montoya, M., 2016. Next-generation sequence analysis of the forage peanut (Arachis pintoi) virome. Revista Facultad Nacional de Agronomía 69 (2), 7881–7891. https://doi.org/10.15446/rfna.v69n2.59133.

Guyomar, C., Delage, W., Legeai, F., Mougel, C., Simon, J., Lemaitre, C., Guyomar, C., Delage, W., Legeai, F., Mougel, C., Simon, J., 2019. Reference-guided genome assembly in metagenomic samples. JOBIM 2019 - Journées Ouvertes Biologie, Informatique et Mathématiques 1–8.

Haas, M., Bureau, M., Geldreich, A., Yot, P., Keller, M., 2002. Cauliflower mosaic virus: still in the news. Mol. Plant Pathol. 3 (6), 419–429. https://doi.org/10.1046/j.1364-3703.2002.00136.x.

ICTV Virus Taxonomy Profiles. https://www.microbiologyresearch.org/content/ictv-virus-taxonomy-profiles [Accessed 17 July 2023].

Jones, D.T., Taylor, W.R., Thornton, J.M., 1992. The rapid generation of mutation data matrices. Bioinformatics 8 (3), 275–282.

Kipkoech, A., Okiror, M., Okalebo, J., Maritim, H.K., 2007. Production efficiency and economic potential of different soil fertility management strategies among groundnut farmers of Kenya. World 2 (1), 15–21.

Kopylova, E., Noé, L., Touzet, H., 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28 (24), 3211–3217. https://doi.org/10.1093/bioinformatics/bts611.

Kreuze, J.F., Perez, A., Untiveros, M., Quispe, D., Fuentes, S., Barker, I., Simon, R., 2009. Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: a generic method for diagnosis, discovery and sequencing of viruses. Virology 388 (1), 1–7. https://doi.org/10.1016/j.virol.2009.03.024.

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. Mega X: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35 (6), 1547–1549. https://doi.org/10.1093/molbev/msy096.

Kwak, H.K., Byun, H.S., Hong, S.B., Choi, H.S., Manandhar, H.K., Timila, R.D., Joshi, S., Baidya, S., 2022. First report of cowpea polerovirus 2 and southern cowpea mosaic virus infecting cowpea in Nepal. Plant Dis. 106 (8), 1–2.

Langmead, Salzberg, S.L., 2013. Bowtie2. Nat. Methods 9 (4), 357–359. https://doi.org/10.1038/nmeth.1923.Fast.

Latourrette, K., Holste, N.M., Garcia-ruiz, H., 2021. Polerovirus genomic variation. Virus Evolution 7 (December), 1–18.

Li, D., Liu, C.-M., Luo, R., Sadakane, K., Lam, T.-W., 2016. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Application Note 1–2. https://doi.org/10.1093/bioinformatics/btw419.

Li, M., Lin, Q., Chen, Y., Xu, F., Peng, J., Zheng, H., Wu, G., Rao, S., Chen, J., Lu, Y., Yan, F., 2022. First report of Turnip mosaic virus in peanut (Arachis hypogaea) in China. Plant Dis. 2–4.

Lim, E.S., Gumpil, J.S., 1984. The flowering, pollination and hybridization of groundnuts (*Arachis hypogaea* L.). Pertanika 7 (2), 61–66.

Mabele, Anthony, Simiyu, Were, H.K., Owuor Ndong'a, M.F., Mukoye, B., 2021. Occurrence and genetic diversity of groundnut rosette assistor virus in western Kenya. Crop Protect. 139 (2021), 1–7. https://doi.org/10.1016/j.cropro.2020.105381.

Mangeni, B.C., Were, H.K., Ndong'a, M., Mukoye, B., 2020. Incidence and severity of bean common mosaic disease and resistance of popular bean cultivars to the disease in western Kenya. J. Phytopathol. 1–15. https://doi.org/10.1111/jph.12928. June 2019.

Marais, A., Faure, C., Couture, C., Bergey, B., Gentit, P., Candresse, T., 2014. Characterization by deep sequencing of divergent plum bark necrosis stem pitting-associated virus (PBNSPaV) isolates and development of a broad-spectrum PBNSPaV detection assay. Phytopathology 104 (6), 660–666. https://doi.org/10.1094/PHYTO-08-13-0229-R.

Masira, I., 2017. AGRICULTURE and FOOD AUTHORITY: Nuts and Oil Crops Directorate- *Busia County Production And Market Research* (Issue February).

Mikheenko, A., Saveliev, V., Gurevich, A., 2016. MetaQUAST: evaluation of metagenome assemblies. Bioinformatics 32 (7), 1088–1090. https://doi.org/10.1093/bioinformatics/btv697.

Mollov, D., Malapi-Wight, M., 2016. Next Generation Sequencing: a useful tool for detection of sugarcane viruses in quarantine programs. Proc. Int. Soc. Sugar Cane Technol. 29 (29), 1631–1635. January 2016.

Moreno, A.B., López-Moya, J.J., 2020. When viruses play team sports: mixed infections in plants. Phytopathology 110 (1), 29–48. https://doi.org/10.1094/PHYTO-07-19-0250-FI.

Moretzsohn, M.C., Gouvea, E.G., Inglis, P.W., Leal-Bertioli, S.C.M., Valls, J.F.M., Bertioli, D.J., 2013. A study of the relationships of cultivated peanut (Arachis hypogaea) and its most closely related wild species using intron sequences and microsatellite markers. Ann. Bot. 111 (1), 113–126. https://doi.org/10.1093/aob/mcs237.

Mukoye, B., Mangeni, B.C., Leitich, R.K., Wosula, D.W., Omayio, O.D., Nyamwamu, P.A., Arinaitwe, W., Winter, S., A, M., W, K., 2015. First report and biological characterization of cowpea mild mottle virus (CPMMV) infecting groundnuts in western Kenya. Journal of Agri-Food and Applied Sciences 3 (1), 1–5.

Naidu, R.A., Bottenberg, H., Subrahmanyam, P., Kimmins, F.M., Robinson, D.J., Thresh, J.M., 1998. Epidemiology of groundnut rosette virus disease: current status and future research needs. Ann. Appl. Biol. 132 (3), 525–548. https://doi.org/10.1111/j.1744-7348.1998.tb05227.x.

Naidu, R.A., Kimmins, F.M., Deom, C.M., Subrahmanyam, P., Chiyembekeza, A.J., van der Merwe, P.J., 1999. Groundnut rosette: a virus disease affecting groundnut production in sub-saharan Africa. Plant Dis. 83 (8), 700–709.

Nemchinov, L.G., Irish, B.M., Grinstead, S., Shao, J., Vieira, P., 2022. Diversity of the virome associated with alfalfa (Medicago sativa L.) in the U.S. Pacific Northwest. Sci. Rep. 12 (1), 1–14. https://doi.org/10.1038/s41598-022-12802-4.

Ng, J.C.K., Perry, K.L., 2004. Transmission of plant viruses by aphid vectors. Mol. Plant Pathol. 5, 505–511. https://doi.org/10.1111/J.1364-3703.2004.00240.X.

Nurk, S., Meleshko, D., Korobeynikov, A., Pevzner, P.A., 2017. metaSPAdes: a new versatile metagenomic assembler. Genome Res. 1 (27), 30–47. https://doi.org/10.1101/gr.213959.116.4.

Okello, D.K., Biruma, M., Deom, C., 2010. Overview of groundnut research in Uganda: past, present and future. Afr. J. Biotechnol. 9 (20), 2843–2850. https://doi.org/10.5897/AJB09.013.

Okello, D., Akello, L.B., Tukamuhabwa, P., Odong, T.L., Ochwo-Ssemakula, M., Adriko, J., Deom, C.M., 2014. Groundnut rosette disease symptoms types distribution and management of the disease in Uganda. Afr. J. Plant Sci. 8 (3), 153–163. https://doi.org/10.5897/AJPS2014.1164.

Omuto, C.T., 2013. Major soil and data types in Kenya. In: Developments in Earth Surface Processes, first ed., Vol. 16. Elsevier B.V. https://doi.org/10.1016/B978-0-444-59559-1.00011-6.

Orakha, P.O., Were, H., Ndonga, M., Benard, M., 2019. First report of cowpea polero virus 1 (CPPV1) infecting cowpea in Kenya. Int. J. Genet. Genom. 7 (4), 123. https://doi.org/10.11648/J.IJGG.20190704.16.

Palanga, E., Martin, D.P., Galzi, S., Zabré, J., Bouda, Z., Neya, J.B., Sawadogo, M., Traore, O., Peterschmitt, M., Roumagnac, P., Filloux, D., 2017. Complete genome sequences of cowpea polerovirus 1 and cowpea polerovirus 2 infecting cowpea plants in Burkina Faso. Arch. Virol. 162 (7), 2149–2152. https://doi.org/10.1007/s00705-017-3327-z.

PeanutBase database. https://v1.legumefederation.org/data/v2/Arachis/hypogaea/genomes/Tifrunner.gnm2.J5K5/[Accessed 27 August 2022].

Prasad, P.V.V., Kakani, Vijaya Gopal, Upadhyaya, H.D., 2009. Growth and production of groundnut. In: Verheye, Willy H. (Ed.), Soils, Plant Growth and Crop Production, Encyclopedia of Life Support Systems (EOLSS), Developed under the Auspices of the UNESCO, vol. 289. Eolss Publishers, Oxford, UK, pp. 1–26. http://www.eolss.net.

Reddy, D.V., Richins, R.D., Rajeshwari, R., Iizuka, N., Manohar, S., Shepherd, R.J., 1993. Peanut chlorotic streak virus, a new Caulimovirus infecting peanuts (*Arachis hypogaea*) in India. Phytopathology 83 (2), 129. https://doi.org/10.1094/phyto-83-129.

Roossinck, M.J., Martin, D.P., Roumagnac, P., 2015. Plant virus metagenomics: advances in virus discovery. Phytopathology, Handelsman. https://doi.org/10.1094/PHYTO-12-14-0356-RVW, 2004, PHYTO12140356RVW.

Roossinck, Marilyn J., 2015. Plants, viruses and the environment: ecology and mutualism. Virology 479–480, 271–277. https://doi.org/10.1016/j.virol.2015.03.041.

Roossinck, Marilyn J., 2012. Plant virus metagenomics: biodiversity and ecology. Annu. Rev. Genet. 46, 359–369. https://doi.org/10.1146/annurev-genet-110711-155600.

Sastry, K.S., Mandal, B., Hammond, J., Scott, S.W., Briddon, R.W., 2019. *Arachis hypogaea* (Peanut/groundnut). In: Encyclopedia of Plant Viruses and Viroids. Springer, New Delhi. https://doi.org/10.1007/978-81-322-3912-3. Encyclopedia of Plant Viruses and Viroids.

Shi, M., Zhang, Y., Holmes, E.C., 2018. Meta-transcriptomics and the evolutionary biology of RNA viruses. Elsevier Virus Res. 83–90. January.

Singh, G., Singh, R., 2017. Distribution and economic importance of Aphis (Aphis) craccivora Koch, 1854 (Aphidini: aphidinae: Aphididae: Hemiptera) and its food plants in India. Int. J. Recent Adv. Multidiscip. 4 (2), 2274–2286.

Singhal, P., Nabi, S.U., Yadav, M.K., Dubey, A., 2021. Mixed infection of plant viruses: diagnostics, interactions and impact on the host. J. Plant Dis. Prot. 128 (2), 353–368. https://doi.org/10.1007/s41348-020-00384-0.

Spence, N.J., Phiri, N.A., Hughes, S.L., Mwaniki, A., Simons, S., Oduor, G., Chacha, D., Kuria, A., Ndirangu, S., Kibata, G.N., Marris, G.C., 2007. Economic impact of Turnip mosaic virus, Cauliflower mosaic virus and Beet mosaic virus in three Kenyan vegetables. Plant Pathol. 56 (2), 317–323. https://doi.org/10.1111/j.1365-3059.2006.01498.x.

Sreenivasulu, P., Subba Reddy, C.H.V., Ramesh, B., Lava Kumar, P., 2008. Virus Diseases of Groundnut.

Sreenivasulu, P., Kuhn, C.W., Naidu, H.A., Demski, J.W., Reddy, V.R., Nayudu, M.V., 1991. Viruses Infecting Peanuts (Arachis hypogaea). Taxonomy, Identification, and Disease Management. Area, December.

Syller, J., 2012. Facilitative and antagonistic interactions between plant viruses in mixed infections. Mol. Plant Pathol. 13 (2), 204–216. https://doi.org/10.1111/j.1364-3703.2011.00734.x.

Taliansky, M.E., Robinson, D.J., Murant, A.F., 2000. Groundnut rosette disease virus complex: biology and molecular biology. Adv. Virus Res. 55, 357–400. https://doi.org/10.1016/s0065-3527(00)55008-8.

Trebicki, P., 2020. Climate change and plant virus epidemiology. Virus Res. 286, 198059 https://doi.org/10.1016/j.virusres.2020.198059.

Twardowicz-Jakusz, A., Zielinska, L., Kaniewski, W., 1999. Occurrence and identification of viruses affecting Brassica vegetable crops in Poland. III. Virus identification and characterization. J. Plant Protect. Res. 39 (1).

Vilsker, M., Moosa, Y., Nooij, S., Fonseca, V., Ghysens, Y., Dumon, K., Pauwels, R., Alcantara, L.C., Vanden Eynden, E., Vandamme, A.M., Deforche, K., De Oliveira, T., 2019. Genome Detective: an automated system for virus identification from high-throughput sequencing data. Bioinformatics 35 (5), 871–873. https://doi.org/10.1093/bioinformatics/bty695.

Waliyar, F., Kumar, P., N, B.R., M, E., Nigam, S., R, A.S., Osiru, M., Diallo, A.T., 2007. A century of research on groundnut rosette disease and its management. Inf. Bull. 75.

Watkiss, E., 2009. RNA viruses: strategies to maximize coding potential. Advanced Virology 1–22. https://www.klinikum.uni-heidelberg.de/fileadmin/inst_hygiene/virologie/vorlesungen/bio2007/viral_strategies08_hgk.pdf.

Webster, C.L., Waldron, F.M., Robertson, S., Crowson, D., Ferrari, G., Quintana, J.F., Brouqui, J.M., Bayne, E.H., Longdon, B., Buck, A.H., Lazzaro, B.P., Akorli, J., Haddrill, P.R., Obbard, D.J., 2015. The discovery, distribution, and evolution of viruses associated with drosophila melanogaster. PLoS Biol. 13 (7), 1–33. https://doi.org/10.1371/journal.pbio.1002210.

Were, H.K., Mukoye, B., Stomeo, F., Were, M.N., Joyce, N., Torrance, R., 2019. Occurrence and distribution of viruses infecting groundnuts (Arachis hypogaea L) in Kenya. RUFORUM Working Document Series 18 (18), 232–238.

Wu, Q., Ding, S.-W., Zhang, Y., Zhu, S., 2015. Identification of viruses and viroids by next-generation sequencing and homology-dependent and homology-independent algorithms. Annu. Rev. Phytopathol. 53 (1), 150605182533006 https://doi.org/10.1146/annurev-phyto-080614-120030.

Xu, Y., Ghanim, M., Liu, Y., 2022. Editorial: mixed infections of plant viruses in nature and the impact on agriculture. Viruses 8 (11), 1–2. https://doi.org/10.3390/v8110303.