DISTRIBUTION AND CHARACTERIZATION OF GROUNDNUT ROSETTE ASSOCIATED VIRUSES IN WESTERN KENYA

Benard Mukoye

A thesis submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy in Crop Protection of Masinde Muliro University of Science and Technology

May 2020

DECLARATION

This thesis is my original work prepared with no other than the indicated sources and support and has not been presented elsewhere for a degree or any other award.

Signature…….. Date……

Benard Mukoye

SCP/H/02/2015

CERTIFICATION

The undersigned certify that they have read and hereby recommend for acceptance of Masinde Muliro University of Science and Technology a thesis entitled *"Distribution and characterization of groundnut rosette associated viruses in western Kenya".*

Signature…….. Date……

Prof Hassan K. Were

Department of Agriculture and Land Use Management

Masinde Muliro University of Science and Technology

Signature…….. Date……

Dr. Millicent F.O. Ndonga

Department of Biological Sciences

Masinde Muliro University of Science and Technology

COPYRIGHT

This thesis is copyright materials protected under the Berne Convection, the copyright Act 1999 and other international and national enactments in that behalf, on intellectual property. It may not be reproduced by any means in full or in part except for short extracts in fair dealing so for research or private study, critical scholarly review or discourse with acknowledgment, with written permission of the Director School of Graduate Studies on behalf of both the author and Masinde Muliro University of Science and Technology

DEDICATION

This thesis is dedicated to my wife Joslyne Jepkemboi and my daughter Lynn Mwavishi for their patience and support during my research.

ACKNOWLEDGEMENT

I greatly thank The Almighty God who gave me good health and strength to accomplish this work.

I sincerely thank my supervisors, Prof. Hassan K. Were and Dr. Millicent F.O. Ndong'a for their guidance, mentorship, support and encouragement that led to the success of this work. Extended appreciation goes to Prof Lesley Torrance (The James Hutton Institute – JHI, UK) for facilitating training and analysis in bioinformatics.

I sincerely acknowledge the financial support towards this study by The National Research Fund (NRF-Kenya), The Royal Society of The United Kingdom, International Foundation for Science and the ILRI BecA hub. I am grateful to Dr. Wellington Ekaya and the entire BecA capacity building team for allowing part of this work to be done at the BecA labs.

Appreciation to my colleagues Bonphace Collins Mangeni, Anthony Simiyu Mabele and Patrick Orakha Odhiambo for giving me extensive support to accomplish data collection. My sincere gratitude goes to the entire Biological Sciences Department staff of Masinde Muliro University of Science and Technology (MMUST) for their comments, support, encouragement and advice that contributed to the success of this work and by extension the MMUST administration.

ABSTRACT

Groundnut (*Arachis hypogaea* L.) is an economically important edible oilseed legume in Sub-Saharan Africa (SSA). Smallholder farmers, who account for 75% of producers, depend on it for food and income. However the yields are far below the world averages. Groundnut rosette disease (GRD) is a major constraint of groundnuts in Sub-Saharan Africa (SSA) causing up to 100% yield losses. The disease is caused by two synergistic viruses; groundnut rosette assistor virus (GRAV, genus *Luteovirus*) and groundnut rosette virus (GRV, genus *Umbravirus*) associated with a satellite-ribonucleic acid (Sat-RNA). Some of the setbacks in the epidemiological studies of GRD associated viruses include the complex etiology of the disease and lack of specific diagnostic tools. Simultaneous detection of the causal agents is possible by multiplex RT-PCR but this depends on the availability of specific primers to known agents that occur in a specific area. Information on occurrence and distribution of GRD in western Kenya was not documented and little was known about the characteristics of associated viruses. This study determined the distribution and characterized GRD associated viruses in western Kenya. Two surveys were conducted (2016/2017) in six counties; Bungoma, Busia, Homabay, Kakamega, Siaya and Vihiga. Symptomatic and asymptomatic groundnut and some bean leafy samples were collected for laboratory analysis. Total RNA was extracted from the leaf samples using RNeasy Mini Kit **(**Qiagen**)** according to the manufacturers' instructions and used for double stranded cDNA synthesis using the SuperScript II kit. The cDNA was column-purified with the DNA Clean & ConcentratorTM-5 – DNA kit. The samples were then processed with the transposon-based chemistry library preparation kit (Nextera XT, Illumina) following manufacturer's instructions. The fragment sizes structure of the DNA libraries was assessed using the Agilent 2100 Bioanalyzer. The indexed denatured DNA libraries were sequenced (200-bp paired-end sequencing) on the Illumina MiSeq platform (Illumina). Reads quality check was done using FastQC. Trimmed reads were used for denovo assembly and contigs aligned to the viral genomes database using CLC Genomics Workbench 10.1.2. The assembled contigs were subjected to a BLASTn search against the GenBank database. Phylogenetic analyses and comparisons were performed using MEGA X. Primers were designed using Primer3Plus from consensus sequences. Biological characterization of GRD was done through sap inoculation on leguminous hosts. Average incidence was 53% and 41% in the short and long rain seasons, respectively. Chlorotic rosette was the dominant symptom followed by Green rosette and Mosaic. Most farmers (65%) sourced groundnut seeds from open air market. Complete nucleotide sequences of Sat-RNA revealed identities of 88-100% with those from Malawi, Nigeria and Ghana. Isolate EG16-5 clustered together with chlorotic M24S, all chlorotic isolates and yellow blotch. The GRV isolates shared 84-98% sequence identity with those available GeneBank. The GRAV coat protein (GRAV-CP) gene sequences revealed 97-100% identity with GeneBank isolates. Complete GRAV sequences clustered closest with *Luteoviruses* in phylogenetic analysis. Leguminous hosts showed varied symptoms and tested positive for Sat-RNA and GRAV using the designed primers. The variations of GRD symptoms observed on groundnuts were due to the existence of different variants of Sat-RNA. Sat-RNA and GRV are more diverse than GRAV. The GRD viruses have hosts among the commonly grown legumes and this enhance the perpetuation of the disease. The study recommends an urgent need to curb GRD, possibly through the exploitation of pathogen derived resistance (PDR).

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

AEZs - Agro-Ecological Zones.

ANOVA - Analysis of Variance.

bp - Base Pair.

dsRNA - Double Stranded Ribonucleic Acid.

FAO - Food and Agricultural Organization.

FAOSTAT - Food and Agricultural Organization Statistics.

GRAV - Groundnut Rosette Assistor Virus.

GRD - Groundnut Rosette Disease.

GRV - Groundnut Rosette Virus.

ICRISAT - International Crops Research Institute for the Semi-Arid Tropics.

ICTV - International Committee on Taxonomy of Viruses.

LM - Lower Midland.

MMLV RT - Moloney murine leukemia virus reverse transcriptase enzyme.

NGS - Next Generation Sequencing.

nt - Nucleotide

ORF - Open Reading Frame.

PDR - Pathogen-derived resistance.

RT-PCR - Reverse Transcription Polymerase Chain Reaction.

SADC - Southern African Development Community.

Sat-RNA - Satellite Ribonucleic Acid.

SSA - Sub-Saharan Africa.

ssRNA - Single Stranded Ribonucleic Acid

CHAPTER ONE

INTRODUCTION

1.1 Groundnut production and importance

Groundnut (*Arachis hypogaea L.)*, which is native to southern America, belongs to the family *Fabaceae* (Usman, 2013). It is the fifth most important annual oilseed and food legume crop. It is grown in diverse environments throughout the semi-arid and sub-tropical regions, in nearly 100 countries in the six continents of the world (Kumar *et al.*, 2007). The most important groundnut producing countries are Argentina, Chad, China, India, Indonesia, Myanmar, Nigeria, South Africa, Senegal, Sudan, USA, and Vietnam (Kumar *et al*., 2007). In Africa, the area planted to groundnuts represents 40% globally. It is only South Africa that recorded 26% of the highest averages while East Africa was among the lowest (ICRISAT, 2012; World Bank, 2015; FAOSTAT, 2016). In Kenya, the crop is mainly grown in western Kenya by smallholder farmers for food and sale. The two main groundnut types in Kenya are the bunch type (Red Valencia) maturing within 90-100 days, and the runner type (Homabay), maturing in 120-150 days. Other varieties grown include: Manipita, Makulu Red, Bukene, Asyria Mwitunde, Texas Peanut, Serere 116 (white) and Alika. The current growers yield in Kenya is 450-700kg/ha (Kayondo *et al.*, 2014).

Groundnut production is of great value in terms of income and nutrition for smallholder farmers in East Africa (Kidula *et al*., 2010; Okello *et al*., 2010). Groundnut seeds (raw, sun dried and roasted) contain moisture content of 7.4%, 3.4%, 1.1% ; crude protein of 24.7%, 21.8%, 18.4%; ash content of 1.5%, 1.4%, 1.4%; crude fat of 46.1%, 43.8%, 40.6%; crude fiber of 2.8%, 2.4%, 2.4% and carbohydrate of 17.4%, 27.2%, 36.1% respectively. Groundnut mineral ions include; Sodium (0.71%, 0.69%, 0.57%), Phosphorus (0.68%, 0.65%, 0.69%), Potassium

(0.47%, 0.51%, 0.55%), Zinc (0.44%s, 0.42%, 0.50%), and Iron (0.40%, 0.47%, 0.43%), respectively (Ayoola *et al.,* 2012). Groundnut seed can be used as poultry feed, complete diet for elderly people who need much protein but less carbohydrates and as an antidote in cases of malnutrition in children (Ayoola *et al.,* 2012). The haulms and groundnut cakes are fed to livestock as hay, while the groundnut seed can be processed as snacks or consumed as a whole seed. It is also a source of vitamins like niacin, falacin, riboflavin, and thiamine. Groundnuts as a legume helps fix nitrogen in soil which enhance productivity in the cereal cropping systems (Smartt. 1994).

1.2 Constraints to groundnut production

Resource poor smallholder farmers grow nearly 75 - 80% of the world's groundnuts in developing countries obtaining yields of 500-800kg/ha, as opposed to the potential yield of >2.5t/ha (Kayondo *et al.,* 2014). In western Kenya, an average of 600 – 700 kg/ha is achieved which is less than 30-50% of the potential yield (Kidula *et al*., 2010). Low yields are mainly attributed to poor quality seeds, drought, poor agronomic practices, numerous pests and diseases caused by numerous pathogenic viruses, fungi, bacteria and nematodes (Mutegi, 2010; Okello *et al.,* 2010). Worldwide, nearly 31 viruses infect groundnuts in nature (Kumar *et al.,* 2007). These viruses belong to various genera including *Potyvirus, Tospovirus, Cucumovirus, Pecluvirus, Soymovirus Umbravirus, Begomovirus, Bromovirus, Carlavirus, Ilarvirus, Luteovirus, Potexvirus, Rhabdovirus* and *Tymovirus*. Nineteen of these viruses were first isolated from groundnuts, while the rest from other hosts, but they commonly occur on groundnuts (Salem *et al*., 2010). The most economically important viruses of groundnuts are *Groundnut rosette virus* (GRV), *Cucumber Mosaic Virus* (CMV), *Peanut mottle virus* (PeMoV), *Groundnut bud necrosis virus*

(GBNV), *Indian peanut clump virus* (IPCV), *Groundnut rosette assistor virus* (GRAV), *Peanut stripe virus* (PStV), *Peanut clump virus* (PCV), *Tomato spotted wilt virus* (TSWV), *Tobacco streak virus* (TSV) (Okello *et al.,* 2014) and *Cowpea mild mottle virus* (CPMMV) (Mukoye *et al*., 2015). Among the viral diseases, Groundnut rosette disease (GRD) is the most devastating in Sub-Saharan Africa that causes an estimated loss of 156 million USD every year (Waliyar *et al*., 2007).

1.3 Statement of the problem

Despite the importance of groundnuts in terms of income, food and nutritional security, in western Kenya, farmers continue to experience very low yields. This is mainly due to Groundnut rosette disease (GRD) which is endemic and most destructive viral disease in SSA (Wangai *et al.,* 2001; Waliyar *et al*., 2007; Okello *et al*., 2014). Some of the setbacks in the epidemiological studies of GRD associated viruses include the complex etiology of the disease and lack of specific diagnostic tools. This affect the development of appropriate management strategies for the disease. Limited documented information on the distribution and diversity of GRD associated viruses (Wangai *et al*., 2001; Kidula *et al*., 2010; Thuo *et al*., 2014), has led to continued increase in yield losses (over 50%) amongst groundnut farmers. In western Kenya, very severe and highly variable GRD symptoms were observed in groundnut farms (Mukoye *et al*., 2018). The underlying cause possibly lies in the genetic variability in one or all of the GRD associated agents, mainly the Sat-RNA of GRV (Murant and Kumar, 1990). Since 1998, when the last survey on GRD was done (Wangai *et al*., 2001), there was no new information about the disease and its associated viruses. This hinders accurate and robust diagnosis of GRD and development of management strategies. Simultaneous detection of the GRD causal agents is possible by multiplex PCR (Anitha *et al*., 2014) but this depends on the

availability of specific primers to known agents that occur in a specific area. This information was limited for GRD causal agents in western Kenya and therefore, a robust detection method which could single out all the GRD agents and their variants was necessary. The variants of the three GRD agents have potential permutations and therefore able to form viable alternatives that can adapt to diverse and changing econiches. Over time and under high selection pressure, such "evolution" in the associated viruses can easily result into new disease patterns (Okello *et al.,* 2014).

1.4 Justification of the study

Observations made by Wangai *et al*., (2001), showed that GRD incidence ranged between 24 - 40% in areas of western Kenya surveyed in the groundnut growing seasons of 1997-1998. This is a long time ago and the dynamics of the disease might have changed and therefore the need for current study.

In Kenya, the diversity of GRV has been done only basing on the sequences of ORF3 and 4 and that of GRAV only by the coat protein sequences obtained by PCR using primers of already characterized viruses (Wangai *et al*., 2001). Next Generation Sequencing (NGS) can detect all the GRD causal agents including their variants in a single run. This will unveil the GRD causal agents available in western Kenya for molecular characterization and diversity studies. Additionally, no genomic sequences (partial or complete) of GRD associated viruses from western Kenya existed in the GeneBank.

Taliansky *et al.,* (2000) reported that single infections of GRAV or GRV in groundnuts, have insignificant impact on plant growth and yield expressing only transient mottle symptoms. These results have however, been contradicted by Naidu and Kimmins (2007) who reported that, in susceptible groundnuts cultivars, infection

by GRAV alone affects plant growth and contributes to significant yield losses. The host range of Kenyan GRD virus isolates had not been determined especially on common cultivated legumes. Thus, the need for biological characterization of GRD causal virus isolates from western Kenya to establish host range and symptomology on common legumes in the region.

Several methods have been used to manage GRD viruses. They include application of pesticides to reduce vector populations, various cropping patterns to delay onset and spread of both vector and disease, and cultural practices. However, very little success has been achieved with each of these approaches (Naidu *et al.*, 1999a). The limitation in the documentation of the impact of GRD, in Kenya, could be due to misdiagnosis, as a result of a lack of in depth knowledge of the GRD causal agents. There was, therefore, need to document the occurrence and molecular characteristics of GRD agents in western Kenya to facilitate the development of accurate rosette disease identification, monitoring and recommend appropriate management and control methods.

1.5 General objective

The general objective of this study was to determine the distribution of groundnut rosette disease (GRD and diversity of associated viruses in western Kenya.

1.5.1 Specific objectives

i. To determine the distribution of GRD in western Kenya.

ii. To determine the genetic diversity of GRD associated viruses in western Kenya.

iii. To determine the biological characteristics of GRD in western Kenya.

iv. To develop molecular diagnostic tools for GRD associated viruses in western Kenya.

1.6 Hypothesis

i. Groundnut rosette disease do not occur in all main groundnuts growing regions of western Kenya.

ii. The GRD associated viruses in western Kenya are not diverse genetically.

iii. The GRD in western Kenya has no similar biological characteristics with those from other regions of SSA.

iv. The available GRD diagnostic primers cannot detect GRD associated viruses from western Kenya.

CHAPTER TWO

LITERATURE REVIEW

2.1 Occurrence and distribution of groundnut rosette disease

The initial report of Groundnut rosette disease (GRD) was in Tanzania (formerly Tanganyika) in 1907 (Waliyar *et al.*, 2007). Since then, reports of the disease have been documented in many other countries within the SSA. These include Kenya, Uganda, Malawi, Angola, Madagascar, Swaziland, South Africa, Ivory Coast, Burkina Faso, Ghana, Nigeria, Gambia, Niger, Senegal and the Democratic Republic of Congo (DRC) (Wangai *et al.,* 2001; Kidula *et al*., 2010; Thuo *et al*., 2014). Even though groundnuts are grown in many other countries outside Africa, the GRD associated viruses have only been detected in SSA. Additionally, the *Aphis craccivora,* the vector of the GRD viruses, is found in many of the groundnuts growing regions.

Symptoms similar to those induced by GRD were reported in some Asian and South American countries, although, diagnostic tests were not conducted to confirm the presence of the disease (Reddy, 1991). The disease is therefore endemic to groundnut producing countries in SSA. The reason for this is speculated that the Portuguese, in the 16th Century, brought in groundnuts from South America that was infested by a pathogen that is endemic to SSA. The pathogen then established and spread in the region. Such a phenomenon is referred to as a new encounter phenomenon (Olorunju *et al.,* 2001). The phenomenon occurs when a pest which evolved with other host species in a certain geographical region, infect a newly introduced crop (Deom *et al*., 2000).

The rural economy in many groundnut producing countries in SSA is usually crippled in the event of GRD epidemics. Despite the fact that GRD epidemics do not

occur every year, devastating losses are experienced in the event of an epidemic. Eastern Zambia (1995-1996) lost an estimated US$ 5 million when 43,000 ha of groundnut was infested with GRD. In Central Malawi (1994-1995), farmers abandoned groundnut farms by 23%, following an unpredictable GRD epidemic resulting to an estimated loss of US$ 155 million (SADC/ICRISAT., 1996; Taliansky *et al.,* 2000). In 1975, Northern Nigeria lost 0.5 million tonnes of groundnut estimated at US$ 5 million as a result of GRD affecting 0.7 million ha of groundnut farms (Olorunju *et al.,* 2001).

The yield losses in groundnuts due to GRD viruses, depend on the stage of growth of the plant when infection occurs. Infections that occur before flowering, cause over 90% loss in pod yield. Variable yield losses occur when infection occurs between flowering and pod maturing stage, whereas negligible effects are caused in subsequent infections (Kumar *et al.,* 2007). While the devastating impact of GRD epidemics, was documented in a few instances (Herselman *et al*., 2004), ICRISAT estimates that greater yield losses in groundnuts in the semi-arid tropics of the world are mainly caused by GRD than any other groundnut virus disease (Subrahmanyan *et al.,* 1998).

2.2 Etiology of groundnut rosette disease

The GRD is caused by three agents; Groundnut rosette assistor virus (GRAV), Groundnut rosette *umbravirus* (GRV) and GRV associated Satellite-RNA (Sat-RNA) (Taliasky *et al.,* 2003). These three agents depend on each other intricately, and they all have an important role in the spread and biology of GRD. For vector transmission of GRV by *Aphis craccivora*, GRAV is needed (Naidu *et al*., 1998a).

Groundnut rosette assistor virus (GRAV) is unassigned virus in the family *Luteoviridae* (Deom *et al.* 2000). The GRAV virion are isometric shaped with 28nm

diameter non-enveloped particles of polyhedral symmetry. It is a single stranded positive sense RNA non-segmented genome of 6900 nt that encodes both structural and non-structural proteins (Murant *et al.*, 1990). It is suggested that GRAV encodes six open reading frames (ORFs) just like other *luteoviruses* (Fig. 1). The GRAV virions are composed of 24.5kDa single coat protein (CP) subunits. This virus is anti-genetically related to *Potato leaf roll virus*, *Beet western yellow virus* and *Bean/pea leaf roll virus* (Scott *et al.*, 1996). Replication of GRAV occurs autonomously in the cytoplasm of the phloem tissue. Vector transmission of GRAV is by *Aphis craccivora* in a persistent circulative manner. Mechanical transmission by sap inoculation, pollen, seed or by contact between the plant is not possible. Experimentally, transmission is only possible by grafting (Naidu *et al*., 1998a). The virus occurs wherever GRD has been reported and groundnuts is the only known natural host. Infections by GRAV alone in groundnuts results to symptomless or transient mottling, and can cause substantial yield loss in susceptible cultivars (Waliyar *et al.*, 2007).

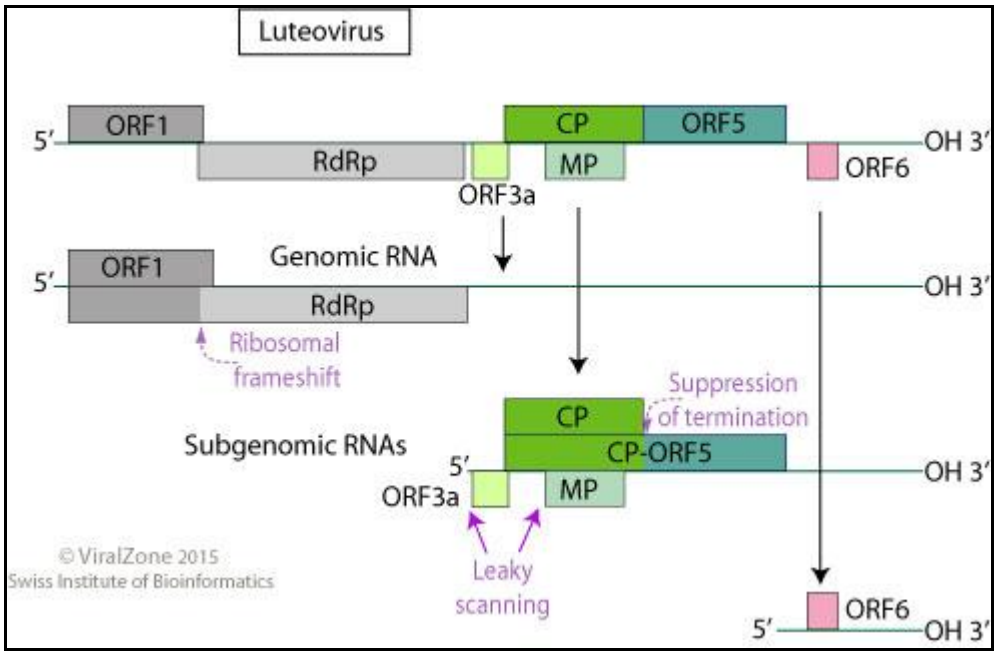


Figure 1: Genomic structure of Luteovirus.

Source: Swiss Institute of Bioinformatics.

Groundnut rosette virus (GRV) belongs to the genus *Umbravirus*. On isolation and characterization, Taliansky *et al.*, (2003), found that the virus has no structural (coat) protein and thus forms no conventional virus particles. The GRV genome is non-segmented single-stranded linear molecule, positive sense RNA of 4019 nt that encodes four ORFs (Taliansky *et al.*, 2003) (Fig. 2). Replication of GRV occurs in the cytoplasm of infected tissue autonomously (Taliansky *et al.*, 2003). The virus alone, causes transient symptoms, but in association with a Sat-RNA, clear rosette symptoms occur (Waliyar *et al.*, 2007). Encapsidation and vector transmission of GRV by *A. cracivora* (in a persistent mode) is dependent on GRAV (Robinson *et al.*, 1999). Transmission of GRV is not possible through pollen, seed or contact between plants, however it is possible by mechanical sap inoculation and grafting (Waliyar *et al.*, 2007). The only natural host of GRV is groundnuts, but some experimental hosts in the *Chenopodiaceae* and *Solanaceae* families have been reported (Waliyar *et al.*, 2007). Only one strain (MC1) of GRV has been reported (Taliansky *et al.*, 1996), and the virus is restricted to SSA and Madagascar (Okello *et al.*, 2013).

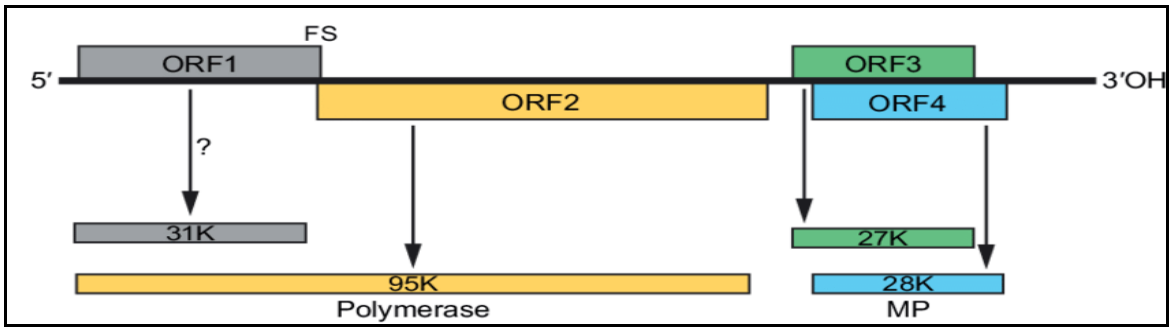


Figure 2: The genomic organization of groundnut rosette virus (GRV).

The genome RNA is represented by the continuous horizontal line while the ORFs are represented by the numbered blocks. The predicted translation products with their sizes are shown on the lower part of the diagram. (Source: International Committee on Taxonomy of Viruses, ICTV, 2012).

The Sat-RNA is sub-viral RNAs of GRV and belongs to the sub-group-2 (small linear) satellite-RNAs. It is of size 895 – 903 nt, single-stranded, linear non-segmented RNA (Blok *et al.*, 1994). Its replication, encapsidation and movement within and between plants is entirely dependent on GRV. Sat-RNA plays a critical role as a helper virus in the transmission of GRV (Taliansky *et al.,* 1997) and GRD symptom expression (Murant and Kumar, 1990; Taliansky *et al.,* 2000). Ten variants of Sat-RNA associated with GRV have been determined (Blok *et al.*, 1994). The different rosette symptoms (chlorotic [yellowing], green and mosaic rosette) are caused by different variants of Sat-RNA (Murant and Kumar., 1990; Olorunju *et al.,* 2001; Kayondo *et al*., 2014). The GRV Sat-RNAs that cause chlorotic and green rosette symptoms in SSA are at least 87% identical. The Sat-RNA contains up to five ORFs in either positive or negative sense but the role of any translational products from these ORFs is unknown (Blok *et al.,* 1994; Taliasky *et al.,* 2003). Vector transmission of Sat-RNA by aphids occurs in presence of GRAV and GRV. Mechanical transmission occurs alongside GRV (Waliyar *et al.,* 2007).

2.3 Symptoms of groundnut rosette disease

Both chlorotic and green rosette symptoms occur throughout the groundnuts growing areas of SSA and sometimes the two can occur in the same field (Mugisa *et al*., 2016). Mosaic rosette, reported in East Africa, is a less common symptom that results from double infection of the groundnut with the chlorotic and green rosette variants of Sat-RNA (Scott *et al.,* 1996; Waliyar *et al.,* 2007). Other symptoms apart from the green and chlorotic rosette may be expressed in infected groundnuts. This suggests wider variability of the visible symptoms of the diseased plants with reduced twisted leaf size resulting in bushy appearance, severe stunting with shortened internodes (Naidu *et al.,* 1998b). Leaves with chlorotic rosette usually show bright yellow with a few

green islands and curled lamina while leaves appear dark green with light to dark green mosaic in case of green rosette (Naidu *et al.*, 1999a). Variation in the GRD symptoms is mainly due to the Sat-RNA variants (Taliansky and Robinson, 1997). Symptom variability under field conditions can be influenced by climatic conditions, genotypic differences of the groundnut cultivars, stage of plant at infection as well as mixed infection with other viruses/agents (Naidu *et al.,* 2007). Green rosette symptoms predominate in eastern Uganda (Okello *et al.*, 2014). This is centrally to the findings by Wangai *et al.* (2001) that chlorotic rosette predominate throughout SSA. This finding is of great importance because eastern Uganda and western Kenya grows more groundnuts in East Africa. Therefore further research to understand the dynamics in GRD symptomology and vector behavior of associated viruses, is needed (Okello *et al.*, 2014).

In Nigeria, over a period of 20 years, a shift from green to chlorotic rosette occurred. The cause of this shift could be as a result of genome changes in the GRD associated viruses, different vector biotypes or cropping patterns (Okello *et al.,* 2014). Such changes therefore suggests the need for routine monitoring and documentation of GRD to support research efforts for its effective control and management.

2.4 Epidemiology and host range of groundnut rosette disease

The hosts of GRD associated viruses are only groundnut and some of its wild relatives (Waliyar *et al*., 2007). GRD epidemiology intricately involves synergistic interaction between and among GRAV, GRV and a Sat-RNA, the aphid vector, the host plant and environment (Naidu *et al*., 1998a). Experimentally, GRAV has been vector transmitted by viruliferous *Aphis craccivora* to *Gomphrena globosa* L., *Montia Perfoliata* L., *Stylosanthes gracilis* Taub, *Pisum sativum* L., S. *mucronata* Wild, *S. hamata* (L.) Taub, S. *sundaica* Taub, *Trifolium incarnatum* L., T. *Pratense* L.,

Spinacia Oleracea L. and *Caspella bursa-Pastoris* (L.) Medicus (Ayoola *et al*., 2012). Symptomless infections were observed in all these plants except *C. bursa-pastoris,* which showed chlorotic symptoms. However, diagnostic assays confirmed the replication of the viruses in samples from these plants (Waliyar *et al*., 2007). Experimental hosts of GRV and Sat-RNA were characterized in various species in *leguminosae, chenopodiaceae* and *solanaceae* through artificial mechanical sap inoculation. Local lesion hosts of GRV include *Chenopodium murale* and *C. amaranticolor*; while systemic hosts include *C. amaranticolor*, *Phaseolus vulgaris, Glycene max, Nicotiana Clevelandii* and *N. benthamiana* (Waliyar *et al.,* 2007). *Gomphrena globosa, Spinacia oleracea, Stylosanthes gracilis, S. Sundaica, S. mucronata, Trifolium repens* and *T. incarnatum* are all experimental hosts of all the three GRD agents (Murant *et al*., 1990).

The cowpea aphid, also known as groundnut aphid (*Aphis craccivora*), transmits all the GRD associated viruses in a persistent circulative manner. To be transmitted by aphids, GRV and Sat-RNA are packaged within the GRAV coat protein. Studies have shown that all the GRAV particles, whether they contain GRAV-RNA or GRV-RNA and Sat-RNA, are acquired by the aphid vector, from phloem sap in 4h and 8h acquisition access feeding, for chlorotic and green rosette, respectively. All the three GRD agents are not always transmitted together by *A. craccivora* (Naidu *et al.,* 1998a). This depends on the feeding patterns of the aphid. In short inoculation feeding (stylet pathway or test probe phase), *A craccivora* probe groundnut leaves, without reaching the phloem thus picks only GRV and Sat-RNA, which multiply in the epidermal and mesophyll cells. The GRAV particles only replicates in the phloem cells and therefore cannot replicate within the mesophyll cells even when deposited there (Naidu *et al.,* 1999b). It is only when the stylets of *Aphis craccivora* penetrate

the sieve elements (salivation phase) of the phloem cells, that all GRD agents are transmitted. This suggests that the chances of transmitting all the three agents together is high when there is long inoculation feeding period, or the number of aphids per plant is increased. Diseased plants infected only with GRV and Sat-RNA are termed as dead-end sources of inoculum as the aphid fails to acquire and transmit these two agents. However, such plants become sources of GRD inoculum upon infection with GRAV due to A. *craccivora* feeding (Deom *et al*, 2000). Damage caused by GRD on groundnut underscores the need for further studies on the epidemiology of the disease and development of appropriate control and management strategies that shrinks the virus inoculum. This will help limit the already resistant/tolerant varieties from succumbing to GRD at high inoculum pressure (Appiah *et al*., 2016).

Aphis craccivora exists in various biotypes that differ in specificity of host plants and transmission efficiency of GRD associated viruses and this has significant implications in the epidemiology of GRD complex (Waliyar *et al.*, 2007). All the three GRD agents are not seed-borne and therefore initial infection of crops is dependent on existence of infected plants, which act as virus sources and the aphids (vectors) (Naidu *et al.*, 1998b). Therefore, between cropping seasons, any surviving infected groundnuts serve as potential virus source for spread of GRD (Waliyar *et al.*, 2007). Influx of viruliferous aphids following prevailing wind currents from regions with infections, can serve as sources of initial infection in regions without GRD (Olorunju *et al*., 2001). The vector *Aphis craccivora* is polyphagous survives on over 142 species of plants and therefore, any of these hosts could be a source of the GRD virus complex (Naidu *et al.*, 1998b). Research efforts to find any alternative natural hosts of the GRD viruses have not yet succeeded (Waliyar *et al.*, 2007). However,

Okello *et al*., (2017), detected GRD causal agents in *Cassia obtusifolia* and recommended transmission studies to validate this finding.

Groundnut rosette disease (GRD) is a polycyclic disease, since surviving diseased plants from previous cropping season, become source of virus complex inoculum for another disease cycle in the field. Primary spread of the disease is facilitated by the winged aphid vectors. Migration of apterate aphids and nymphs are largely responsible for secondary spread of the disease within the field (Naidu *et al.,* 1998b). Initial infections that occur at early stages of plant growth enhance repeated cycles of infections thus increasing the severity of the disease in the groundnut fields. The stage of plant growth, infection time, type of groundnut cultivars, vector (aphid) transmission efficiency, crop density, proximity to inoculum sources and climatic conditions determine the nature and pattern of GRD spread (Waliyar *et al.,* 2007).

2.5 Diversity of GRD causal agents

The Coat Protein (CP) gene is the only region that has been utilized to determine the diversity GRAV of the Kenyan isolates and SSA at large (Wangai *et al*., 2001; Anitha *et al*., 2014). Different isolates from different regions in Kenya showed 97-100% nucleotide identity (Wangai *et al*., 2001). These isolates displayed 96-98% sequence similarity with those of Malawi and Nigeria (Wangai *et al*., 2001). In the same study, Wangai *et al*., (2001), the GRV diversity was determined using the nucleotide sequences of GRV ORF3 and 4. The two GRV ORFs displayed 99% sequence identity among the Kenyan isolates and showed sequence homology of 95-96% with Malawian isolates and 87-88% with the Nigerian isolates. The GRV associated Sat-RNA sequences of Kenyan isolates shared sequence identity of 95% with Malawian isolate (M24S) and 89% with Nigerian isolate (NG3a). However, none of the GRD

associated viruses sequences from Wangai *et al*., (2001) are not available in the GeneBank.

Deom *et al*., (1999), observed that the GRAV CP gene was highly conserved (97-99%) within isolates from the same geographical area (Malawi) but less conserved (88-89%) among isolates from two distant geographical locations (Malawi and Nigeria). Similar observations were reported for the sequences of GRV ORF3 and 4 as well as the Sat-RNA from Malawi and Nigeria (Deom *et al*., 1999).

2.6 Detection of GRD causal agents

Simultaneous detection of the GRD causal agents is possible by multiplex PCR (Anitha *et al*., 2014) and also by single run PCR (Naidu *et al*, 1998).The use of such molecular techniques has increased the speed and accuracy of viral disease diagnosis in crops, however these techniques only allow the detection of known viruses, i.e., viruses that have been characterized (Mumford *et al*., 2006). When such techniques are unavailable, or the viruses are unknown or poorly characterized, then disease diagnosis needs tests done using indicator plants in expensive glasshouses or the use of field indexing, both of which are lengthy and labor intensive. Methods for simultaneous detection of multiple viruses become very useful in such scenarios. Next generation sequencing (NGS) is now one of the principal methods in detection of multiple viruses (both known and unknown) (Boonham *et al*., 2014). Detection of viral RNA and DNA genomes in infected plant material by next generation sequencing (NGS) (Kreuze *et al*., 2009), is possible through the extraction and sequencing of total RNA and DNA (Eichmeier *et al*., 2016). NGS has the ability to sequence whole genomes of known and unknown viruses and the ability to detect multiple viruses from a mixed infection, thus providing a very sensitive diagnostic method for the rapid and routine detection of viruses. NGS being non-specific, can be

used to detect all known and unknown viruses present in a host irrespective of their pathogenicity.

Next generation sequencing (NGS) technologies are currently becoming popular methods to obtain whole plant virus genomes in a relatively short period of time (Boonham *et al*., 2014). Because of the ability to use total RNA extractions, NGS is useful and common in obtaining complete genomes of plant viruses (Adams *et al*., 2009). The challenge faced lies not only in accessing and using NGS technology, but also in analysing and interpreting the very large datasets generated (Boonham *et al*., 2014). Therefore, virus complexes that cause diseases in combination, such as GRD, can quickly be characterized using NGS technologies as all the causal viruses can be sequenced simultaneously. This enhance adequate characterization and further development of diagnostic tools including primers that can detect an entire range of variants/strains of the target viruses.

2.7 Management of GRD

Many efforts for management of GRD have focused on refining cropping practices to delay the onset and spread of both the vector and the disease, and on breeding for host-plant resistance (Olorunju *et al*., 2001). Chemical control of aphids and rogueing of volunteer plants from previous cropping season are likely to provide effective management of the disease (Naidu *et al*., 1998b). However, such practices are rarely feasible for the subsistence farming systems of SSA (Appiah *et al*., 2017).

A single recessive gene has been identified that confer resistance to the aphid vector (van der Merwe and Subrahmanyan, 1997). This gene is mapped on linkage Group-1, at a distance of 3.9 nm from a marker, originating from a susceptible parent (ICGV-SM 93541) (Herselman *et al.,* 2004). This DNA marker if identified, can hasten aphid

resistance screening and breeding process through development and use of marker assisted selection.

Efforts have been made to exploit pathogen-derived resistance (GRAV replicase and CP genes, movement protein genes and Sat-RNA derived sequences) to GRD, in developing broad based agronomically superior, groundnut cultivars (Taliansky *et al.*, 1996). However, this has not been effectively exploited. Only limited field resistance is available for either virus, in popular groundnut cultivars and landraces, which have less than superior agronomic traits. This phenomenon needs further evaluation of the germplasm in popular groundnut genotypes. The lack of adequate information on the occurrence of the GRD and variability in the GRD associated viruses hinders the development of management strategies aimed at reducing the huge losses caused by this viral disease (Appiah *et al*., 2017). Therefore, this study intended to contribute to the efforts in management of GRD by documenting the current status of the disease and the genetic diversity of the associated viruses. These information is useful in breeding for resistance that is appropriate for Kenya and entire SSA.

CHAPTER THREE

MATERIALS AND METHODS

3.1 Rosette disease diagnostic survey

A survey to determine GRD occurrence and distribution was conducted in all major groundnut growing areas of western Kenya. Groundnut fields were sampled during the short rains (October to December 2016) and long rains season (May to July 2017). The following Counties were covered: Bungoma, Busia, Homabay, Kakamega, Siaya and Vihiga. Sampling of groundnut farms was done by stopping at regular predetermined intervals, of 3-8 km along motorable roads that traverses each sampling area. The survey were conducted, by walking through groundnut fields, and visually inspecting groundnut crops for symptomatic leaves. Depending on the farm size, quadrats of $10m^2$ were estimated, disease incidence and severity was scored for each quadrat through random sampling. A questionnaire and disease diagnostic score sheet, was used to record GRD virus incidence and severity in each farm (appendix I). Disease incidence was calculated according to Reddy, (1991), as the percentage of plants showing GRD virus symptoms, to the total number of plants observed in the field as shown in the following equation:

$$\text{Disease incidence} = \frac{\text{Number of GRD virus symptomatic Plants}}{\text{Total number of groundnut plants sampled}} \times 100\%$$

Groundnut rosette disease incidence was scored using a rating scale according to Reddy, (1991) where: low incidence = 1-20%; moderate incidence = 21-49% and high incidence = 50-100%. The GRD severity was scored using a severity scale of 0 – 3, where: 0 = No disease, 1 = Mild, 2 = Moderate and 3 = Severe. The types of GRD symptoms observed were recorded.

Leaf samples showing virus-like symptoms of green mosaic, leaf distortion, downward curling, mottling, chlorotic areas, necrotic spots, local lesions, stunting or a combination of these were collected in RNA*later*® RNA Stabilization Solution and kept at 4^{o}C until further analysis. Some leaf samples of common bean were also collected in situations where they were found intercropped with groundnuts. Geographical Positioning System (GPS) (entrex venture HC GARMIN^{TM}), was used to record the latitude, longitude and altitude of the sampled regions.

3.2 Determination of genome sequence of the GRD agents

The simultaneous detection of GRD agents was done by next generation sequencing (NGS).

3.2.1 Total RNA extraction

Total RNA was extracted from the leaf samples using Qiagen RNeasy Plant Mini Kit (Qiagen, Germany) according to the manufacturers' protocol. The RNA was quantified using Nano-drop and 1.5% agarose gel electrophoresis.

3.2.2 Sequencing

The extracted total RNA was used for double stranded cDNA synthesis using the SuperScript II (Thermo Fisher Scientific, Waltham, USA) kit. The cDNA was column-purified with the DNA Clean & ConcentratorTM-5 – DNA kit (Zymo Research, Irvine, USA) and quantified with the Qubit 2.0 Fluorometer (Thermo Fisher Scientific). The samples were then processed with the transposon-based chemistry library preparation kit (Nextera XT, Illumina) following manufacturer's instructions. The fragment sizes structure of the DNA libraries was assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, USA). The indexed denatured DNA libraries were sequenced (200-bp paired-end sequencing) on the Illumina MiSeq platform (Illumina).

Reads quality check was done using FastQC (version 0.11.5). Reads were then trimmed to remove poor quality sequences with parameters: LEADING:20 TRAILING:20 SLIDINGWINDOW:4:20 and a minimum read length of 20. Trimmed reads (Haas *et al*., 2013) were used for de novo assembly and contigs aligned to the viral genomes database (ftp://ftp. ncbi.nih.gov/genomes/Viruses/all.fna.tar.gz/, downloaded on October 2017) using CLC Genomics Workbench 10.1.2. The assembled contigs were subjected to a BLASTn search against the GenBank database (Altschul *et al*., 1990). Complete and partial GRV, GRAV and Sat-RNA sequences used for comparison and phylogenetic analyses were retrieved from GenBank (http://www.ncbi.nlm.nih.gov/). Phylogenetic analyses and comparisons were performed using the MEGA X (Kumar *et al*., 2018).

3.3 Host range studies

Host range experiments were carried out in a greenhouse, through mechanical inoculation of the major legumes with GRD viruses. Low concentrations of the rosette disease agents in host plants, makes it essential to develop a reliable and sensitive method for their detection (Usman, 2013; Salem *et al*., 2010).

Groundnuts (*Arachis hypogaea),* common beans (*Phaseolus vulgaris),* cowpea (*Vigna unquiculata),* soyabean (*Glycine max),* and green grams (*Vigna radiata),* which are leguminous indicator plants (Mugisa *et al*., 2016), were planted in plastic pots. Three seeds per legume species/variety were planted in each pot. After germination, the seedlings were thinned to remain with two plants per pot.

Some of symptomatic leaf samples from the survey, were ground using a sterilized pestle and mortar, with the aid of dust powdered Carborundum 320 grit. Freshly prepared ice-cold 0.01M Potassium Phosphate buffer (K_2HP0_4 + KH_2P0_4), pH 7.0,

containing 0.2% Sodium Sulfite and 0.01M Mercaptoethanol (1: 6 [w/v] tissue: buffer), was added to the ground tissue, mixed and transferred to a falcon tube, and allowed to stand for 5 min on ice, for debris to settle at the bottom of the tube. The sap was kept on ice, until inoculation was completed. The test plants were dusted with Carborundum abrasive. The inoculum was applied gently on the leaf surfaces, using saturated cotton wool swab. After inoculation, the excess inocula on the leaves were gently washed with sterilized distilled water. The plants were observed on weekly basis for any viral symptoms development for 5 weeks.

Leafy samples were then collected and tested for GRD causal agents by RT-PCR. Groundnuts field samples with chlorotic, mosaic and green rosette were included in the analysis. Total RNA was extracted as described in section 3.2.1. The primers used were designed using Primer3Plus (-http://primer3plus.com/cgi-bin/dev/primer3plus.cgi) using consensus sequences from this study and those from the GeneBank (Table 1). The RT-PCR was done essentially as described by Naidu *et al*., (1998) with some modifications. Two step RT-PCR was done using One Taqman master mix. Two µl of RNA was initially used in cDNA synthesis which was run at 42°C for 1 h followed by denaturation step of 5 min at 80°C. The cDNA synthesis reaction was composed of target virus reverse primer (200 ng), MMLV RT, MMLV buffer, dNTPS, DTTS, RNA (2µl) and water. Five µl of cDNA was then used in the amplification step. The amplification mixture was composed of One Taqman master mix, forward and reverse primers, cDNA and water. Amplifications were carried out in a Eppendorf Cycler using the following temperature regime: a denaturation phase at 94°C for 2 min followed by 35 cycles of amplification (94°C for 1 min, 55°C for 1 min, and 2 min at 72°C) and a final extension at 72°C for 10 min. Ten µl of PCR

products were analyzed by 1.2% agarose gel electrophoresis in TBE buffer, stained with ethidium bromide and finally visualized under UV light.

Table 1: Primers designed and used in detection of causal agents of GRD.

Primers	Sequence (5' > 3')	Specific to
GRVSATF	ATGCAGATTGGTAGCCTTGG	Sat-RNA
GRVSATR	CTGTGTATGCGCCCATTAAG	Sat-RNA
GRAVF	GCAATGGACGAGCTAACAGG	GRAV-CP
GRAVR	ACTTGATGGTGAACCGGAAG	GRAV-CP
GRVKenF	GCAAAATTTTTAGTCGGGGAAG	GRV ORF3 and ORF4
GRVKenR	GGTCTTATGTTCAGGCTGTCAA	GRV ORF3 and ORF4

3.4 Survey data analysis

The collected data on GRD virus incidence and severity, was subjected to analysis of variance (ANOVA), using Statistical Analysis System (SAS) program version 9.3.1 software (SAS Institute, 2013). Pairwise comparisons of means was done using Least Significance Differences (LSD) for multiple-means comparison method at $P \leq 0.05$ confidence level.

CHAPTER FOUR

RESULTS

4.1 Distribution of GRD in western Kenya

A total of 526 farms were surveyed in 6 counties (253 in long rain and 273 in short rain). Groundnut rosette disease was observed in all the 6 Counties surveyed. The disease expressed varied symptoms across the Counties. The incidence and severity of GRD varied across the surveyed Counties.

4.1.1 Major GRD symptoms and their distribution in western Kenya

Rosette infected plants were dwarf with increased tillering though some were tall but expressed other major symptoms associated with GRD. The main symptoms observed across Counties in order of abundance, starting from the most prevalent, were chlorotic rosette, green rosette and severe mosaic. Other symptoms observed include leaf rolling, upward leaf curling and severe leaf bunching (Fig. 3). The distribution of the major GRD symptoms is shown in Fig. 4.

Figure 3: Some of the virus-like symptoms observed in the surveyed fields

A: dwarfed plant with green rosette; **B**: severe chlorosis (yellow) on young leaves and dwarfing; **C**: severe young leaf rolling, chlorosis and bunching on a dwarfed plant; **D**: Mosaic mostly on young leaves.

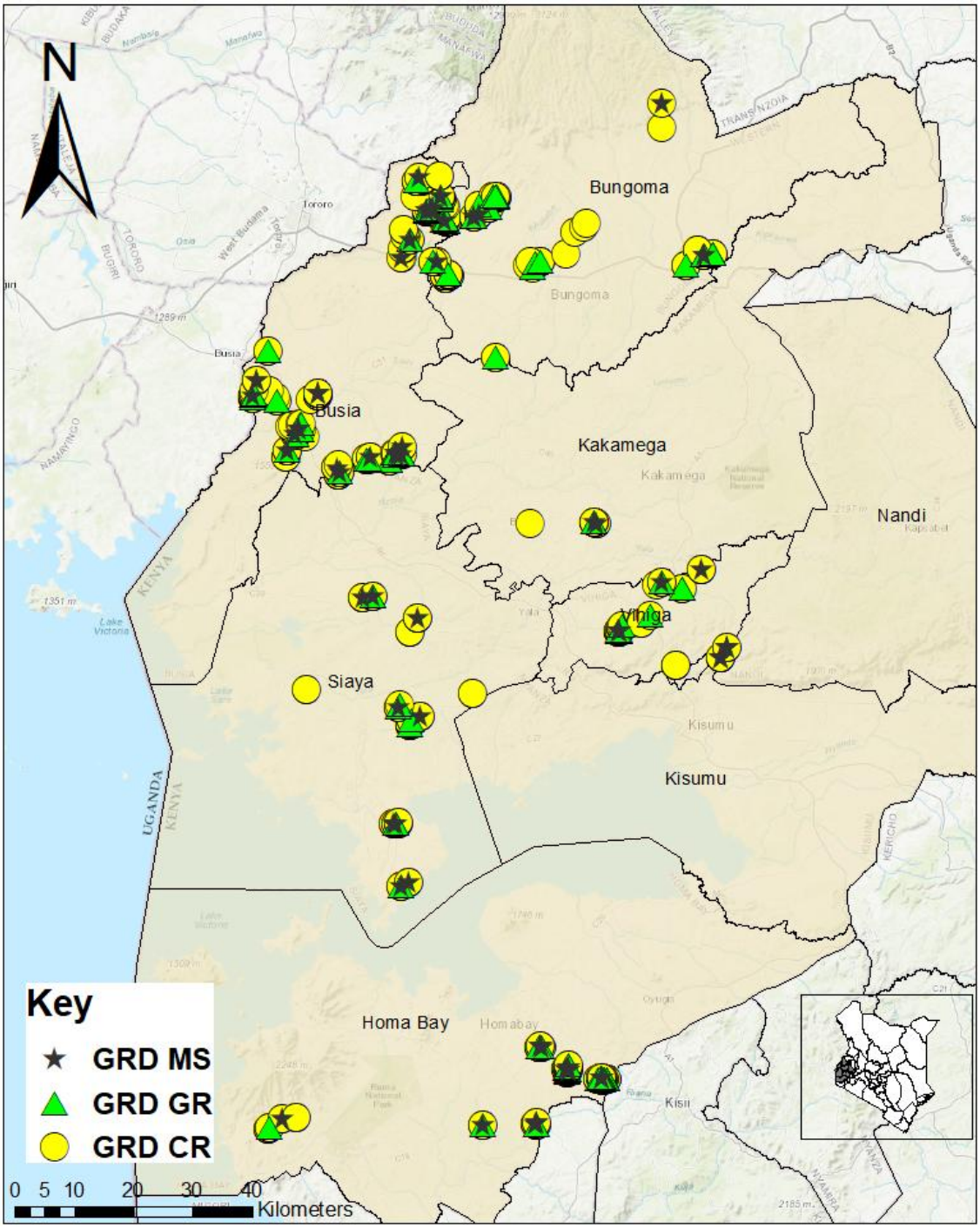


Figure 4: A map of western Kenya showing the distribution of major GRD symptoms in the surveyed Counties.

(MS-**Mosaic,** GR-**Green Rosette,** CR-**Chlorotic Rosette).**

4.1.2 The GRD incidence and severity

Generally, GRD incidence was high during the short rain season than the long rain season in all Counties. Highest mean GRD incidence was recorded in Kakamega in

the short rain season (94.12%) while the lowest was in Bungoma (30.89%) during the long rain season (Table 2). There was a significant difference in GRD incidence among the counties (p=0.011). Overall, Siaya had the lowest incidence which was significantly different from that of Kakamega but did not vary significantly from that of Bungoma, Busia, Homabay and Vihiga.

Table 2: Mean GRD incidence (%) per County.

County	Season	N	Mean (%)	Std. Error of Mean (+/-)
Bungoma	Long rain	45	30.89	4.534
	Short rain	47	66.51	4.295
Busia	Long rain	74	43.36	3.526
	Short rain	108	46.56	2.728
Homabay	Long rain	73	48.60	3.919
	Short rain	55	48.22	4.025
Kakamega	Long rain	30	43.47	5.283
	Short rain	17	94.12	4.779
Siaya	Long rain	31	33.94	4.820
	Short rain	26	43.23	6.645
Vihiga	Short rain	20	47.50	6.412
	Long rain	253	41.51	1.962
	Short rain	273	53.04	1.909

Mean GRD severity ranged from mild (1) to moderate (2) across all the counties and seasons. Short rains season recorded high severity compared to long rains season in all Counties surveyed. Highest mean severity was recorded in Kakamega in the short rains season (2.46) while the lowest was in Siaya during the long rains (1.46) (Table 3). However, the severity was not significantly different between the Counties.

Table 3: Mean GRD severity per County.

County	Season	N	Mean	Std. Error of Mean (+/-)
Bungoma	Long rain	45	1.49	0.10
	Short rain	47	2.21	0.12
Busia	Long rain	74	1.71	0.09
	Short rain	108	2.11	0.08
Homabay	Long rain	73	1.95	0.09
	Short rain	55	2.04	0.10
Kakamega	Long rain	30	1.53	0.12
	Short rain	17	2.46	0.13
Siaya	Long rain	31	1.46	0.14
	Short rain	26	1.96	0.17
Vihiga	Short rain	20	1.98	0.14
Total	Long rain	253	1.69	0.05
	Short rain	273	2.15	0.05

The incidence of GRD seemed to increase with increase in severity. Where severity was high, incidence was high and predicted to increase significantly (Fig. 5).

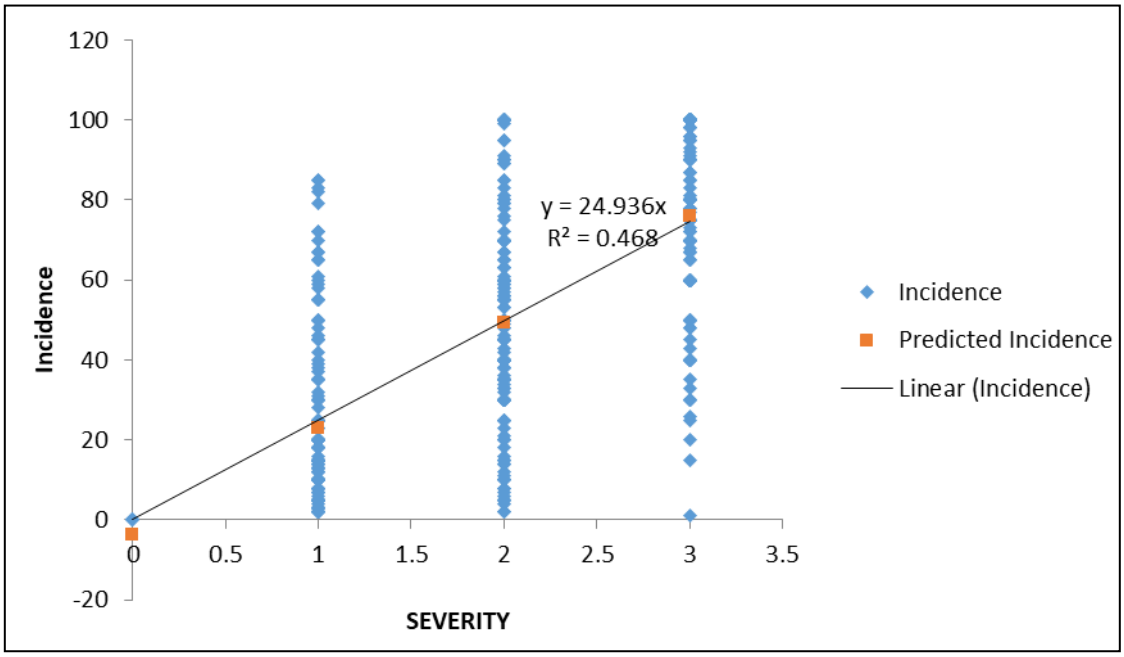


Figure 5: A line graph showing relationship between GRD severity and incidence

4.2 Socio-economic information of the groundnut farmers in western Kenya

Groundnuts were grown in two cropping patterns. These are stand-alone (no intercrop), which was the most common pattern (63%) and intercropping with other legumes (37%), such as cowpeas, soybeans and beans. Most farmers (65%) sourced groundnuts seeds from open air markets, some used own saved seed (28%) while others borrowed from neighbors (7%) (Table 4).

Table 4: Socio-economic data of groundnut farmers in western Kenya

County	**Cropping patterns (%)**		**Seed source (%)**		
	Intercrop with other legumes	No intercrop	Open market	Own saved	Neighbors
Bungoma	18	82	79	16	5
Busia	35	65	73	20	7
Kakamega	48	52	59	37	4
Vihiga	83	17	59	29	12
Homabay	24	76	60	35	5
Siaya	22	88	62	32	6
Overall	**37**	**63**	**65**	**28**	**7**

4.3 Groundnut varieties grown

Various groundnut varieties were planted by farmers across the surveyed areas. The main varieties grown were Red Valencia (48.7%), followed by Uganda red (21.1%), Homabay (12.0%) and CG7 (7.0%). The others, mostly local, were grown by less than 5% of the farmers (Table 5).

Table 5: Rank of groundnut varieties grown in western Kenya.

Variety	Frequency	%
Red Valencia	256	48.7
Uganda red	111	21.1
Homabay	63	12.0
CG7	37	7.0
Local	26	4.9
Loteseto	7	1.3
Local (Purple)	6	1.1
Local white	5	1.0
SM	5	1.0
Local Red	4	0.8
Madiaba	3	0.6
GL2	1	0.2
Local (Teso)	1	0.2
SB3	1	0.2

4.4 The diversity of GRD associated viruses

The genetic diversity of the GRD associated viruses was determined by analysis of the sequence reads obtained by high throughput sequencing (NGS).

4.4.1 Quantification of RNA

The extracted RNA was quantified using Nano-drop and gel electrophoresis (Appendix VI). The RNA quantities ranged between 200-15000 ng/µl.

4.4.2 Description of the sequence raw reads

The number and characteristics of raw sequence reads (FASTQ) obtained are summarized in Table 6. The details of the reads before and after trimming are shown. The raw sequence reads ranged between 700,000 – 7,300,000 bases in both forward (R1) and reverse (R2) directions.

Table 6: Details of the raw sequence reads

Sample ID*	Before Trim			After Trim		
	Total reads	Reads length	%GC	Total reads	Reads length	%GC
KG8-R1	1661211	35-151	48	1646161	20-136	48
KG8-R2	1661211	35-151	48	1646161	20-136	48
EG16-R1	735911	35-151	52	728706	20-136	52
EG16-R2	735911	35-151	52	728706	20-136	52
BG3-R1	755228	35-151	50	747786	20-136	49
BG3-R2	755228	35-151	50	747786	20-136	49
BUG1-R1	849154	35-151	49	839451	20-136	49
BUG1-R2	849154	35-151	49	839451	20-136	49
E3-R1	5799379	35-151	50	4503868	20-136	50
E3-R2	5799379	35-151	50	4503868	20-136	50
E5-R1	3329984	35-151	51	3144874	20-136	50
E5-R2	3329984	35-151	51	3144874	20-136	50
E7-R1	3238295	35-151	51	2993742	20-136	50
E7-R2	3238295	35-151	51	2993742	20-136	50
E8-R1	7263305	35-151	50	6707699	20-136	50
E8-R2	7263305	35-151	50	6707699	20-136	50

*KG8 and EG16 = Samples from Kakamega, BG3 = Sample from Bungoma, BUG1, E3, E5 and E8= Samples from Busia, E7 = Sample from Siaya.

4.4.3 Diversity of GRV-Sat-RNA

Six complete genomes of the Sat-RNA were assembled. The assembled Sat-RNA sequences were from different areas of the surveyed Counties. The sequences varied slightly in number of nucleotides (nt) ranging between 896 – 901 nt (Table 7).

Table 7: Description of the Sat-RNA sequences assembled

Sample ID	Sat-RNA ID	Sequence length (nt)	County of origin
EG16	EG16-5	901	Kakamega
E7	E7	896	Siaya
E8	E8	897	Busia
BUG1	BUG1-21	901	Busia
KG8	KG8-1	898	Kakamega
BG3	BG3-18	901	Bungoma

The six Sat-RNAs from Kenya were then compared with those from the GeneBank. In the phylogenetic tree all Kenyan isolates formed two distinct clusters together with

Malawian isolates. Isolates E7 and E8 grouped with M11S, isolates BUG1-21, BG3-18 and KG8-1 clustered together with M16S while isolate EG16-5 grouped with M24S. All Nigerian isolates grouped together similar to Ghanaian isolates. Sequence identities of between 88-100% of the Kenyan isolates and those from Malawi, Nigeria and Ghana were revealed. Very close identities of between 92-100% were observed between the Kenyan isolates and those from Malawi, followed by Nigerian isolates (90-93%) and least with Ghanaian isolates (86-89%). Isolate BUG1-21 had 100%, 99% and 98% identities with M16S, M12S, M11S respectively, which are all green rosette variants, and 94% with M24S (chlorotic variant). While the other western Kenya isolates (KG8-1, BUG1-21, BG3-18, E7 and E8) had 92-95% identity with Malawian isolate M24S (chlorotic rosette variant), isolate EG16-5 (Kakamega) showed the closest identity (97%) with this isolate. The same isolate EG16-5 was the only that clustered together with M24S, all chlorotic isolates (Z29702.1, Z29703.1) and yellow blotch (Z29710.1, Z29711.1). Isolates E7 and E8 were closest to Malawian isolate M11S with 97% and 99% identity respectively. Isolates BG3-18 and KG8-1 were closest to Malawian isolates M16S displaying 97% identity (Fig. 6).

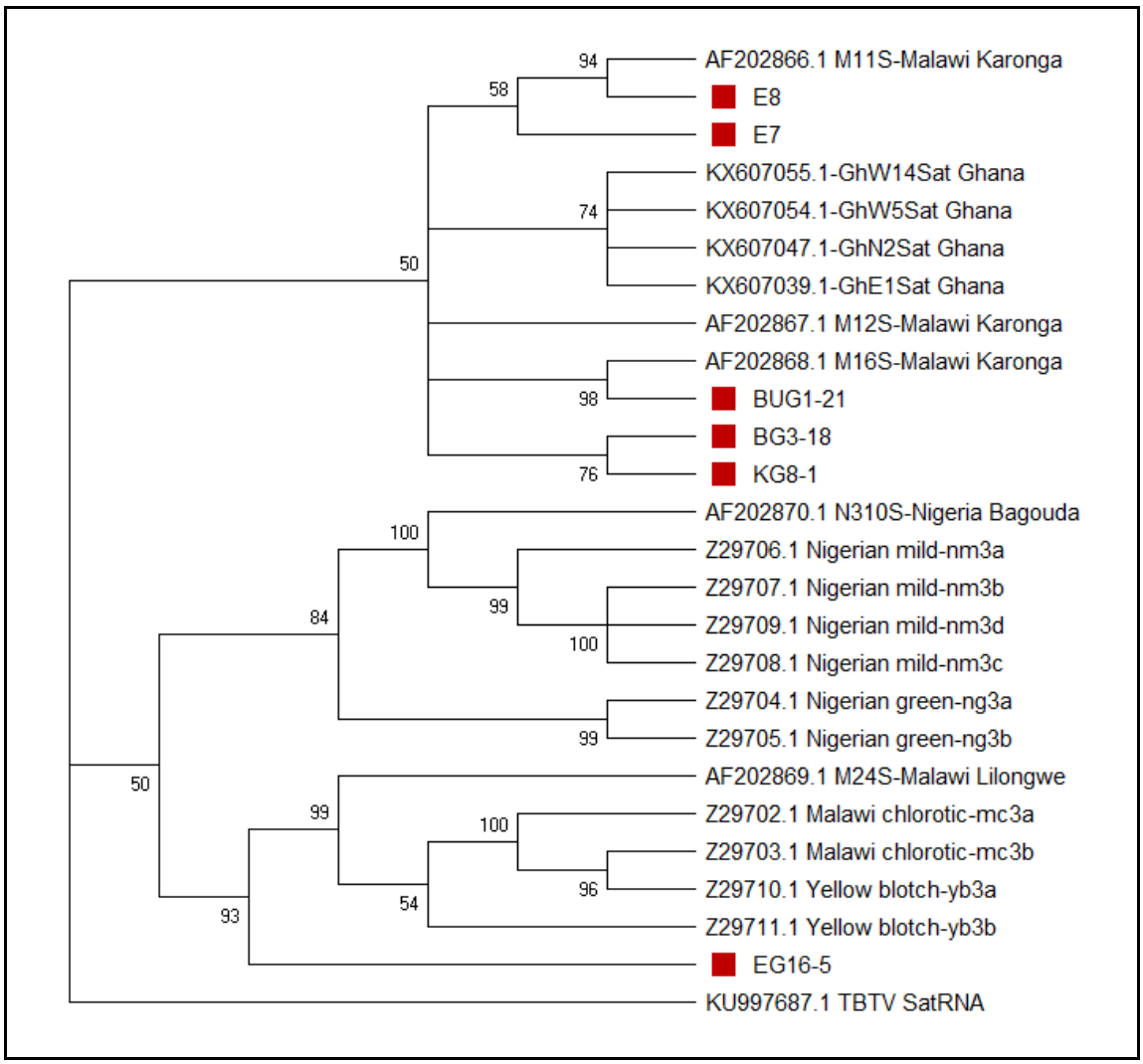


Figure 6: Phylogenetic tree of western Kenya Sat-RNA and GeneBank isolates.

The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model (Tamura and Nei, 1993). The tree is rooted on Sat-RNA of a distantly related *Umbravirus* (Tobacco bushy top virus - KU997687.1 TBTV). Bootstrap confidence values (500 replications) are shown.

4.4.4 Estimates of Evolutionary Divergence between Sat-RNA Sequences

Among the six western Kenya Sat-RNA sequences, Isolate EG16-5, showed more evolutionary divergence (above 0.06 base substitutions) compared with the other Isolates. The least divergence was observed between Isolates BG3-18 and KG8-1 (0.02 base substitutions). This was similar divergence (above 0.06 base substitutions) between M24S (chlorotic) and all the other Malawian Isolates (green) (Table 8).

Table 8: Estimates of Evolutionary Divergence between the Sat-RNA Sequences from Kenya and those in GeneBank

	Isolates	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1	AF202866.1_M11S		0.10	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.00	0.01	0.03	0.03	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
2	**KU997687.1_TBTV_SatRNA**	0.97		0.12	0.10	0.11	0.11	0.12	0.12	0.12	0.12	0.11	0.11	0.54	0.54	0.54	0.54	0.14	0.14	0.11	0.11	0.14	0.14	0.14	0.14	0.14	0.12
3	E7	0.03	1.04		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.03	0.03	0.03	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
4	E8	0.01	0.96	0.04		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.03	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
5	AF202867.1_M12S	0.02	1.00	0.04	0.03		0.00	0.01	0.01	0.01	0.01	0.00	0.01	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
6	AF202868.1_M16S	0.02	1.00	0.04	0.03	0.01		0.01	0.01	0.01	0.01	0.00	0.01	0.03	0.03	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
7	AF202869.1_M24S	0.07	1.09	0.08	0.08	0.07	0.06		0.01	0.01	0.01	0.01	0.01	0.04	0.04	0.04	0.04	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
8	AF202870.1_N310S	0.10	1.11	0.11	0.11	0.10	0.09	0.10		0.01	0.01	0.01	0.01	0.04	0.04	0.04	0.04	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
9	BG3-18	0.04	1.10	0.06	0.05	0.04	0.03	0.06	0.10		0.01	0.01	0.01	0.03	0.03	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
10	KG8-1	0.03	1.09	0.05	0.04	0.03	0.03	0.06	0.09	0.02		0.01	0.01	0.03	0.03	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
11	BUG1-21	0.02	1.00	0.04	0.03	0.01	0.00	0.06	0.09	0.03	0.03		0.01	0.03	0.03	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
12	EG16-5	0.07	1.03	0.08	0.08	0.07	0.07	0.04	0.10	0.07	0.06	0.07		0.03	0.03	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
13	KX607055.1-GhW14	0.05	0.99	0.05	0.05	0.03	0.05	0.09	0.10	0.06	0.06	0.05	0.06		0.00	0.00	0.00	0.04	0.04	0.04	0.04	0.04	0.05	0.05	0.05	0.04	0.04
14	KX607054.1-GhW5	0.05	0.99	0.05	0.05	0.03	0.05	0.09	0.10	0.06	0.06	0.05	0.06	0.00		0.00	0.00	0.04	0.04	0.04	0.04	0.04	0.05	0.05	0.05	0.04	0.04
15	KX607047.1-GhN2	0.05	0.99	0.05	0.05	0.03	0.05	0.09	0.10	0.06	0.06	0.05	0.06	0.00	0.00		0.00	0.04	0.04	0.04	0.04	0.04	0.05	0.05	0.05	0.04	0.04
16	KX607039.1-GhE1	0.05	0.99	0.05	0.05	0.03	0.05	0.09	0.10	0.06	0.06	0.05	0.06	0.00	0.00	0.00		0.04	0.04	0.04	0.04	0.04	0.05	0.05	0.05	0.04	0.04
17	Z29702.1-chlorotic-mc3a	0.10	1.15	0.12	0.11	0.10	0.10	0.05	0.14	0.09	0.09	0.10	0.07	0.10	0.10	0.10	0.10		0.00	0.01	0.01	0.02	0.02	0.02	0.02	0.00	0.01
18	Z29703.1-chlorotic-mc3b	0.10	1.17	0.12	0.11	0.10	0.10	0.04	0.14	0.09	0.09	0.10	0.07	0.10	0.10	0.10	0.10	0.01		0.01	0.01	0.02	0.02	0.02	0.02	0.00	0.01
19	Z29704.1-green-ng3a	0.08	1.03	0.09	0.09	0.08	0.07	0.08	0.08	0.08	0.07	0.07	0.08	0.10	0.10	0.10	0.10	0.11	0.11		0.00	0.01	0.01	0.01	0.01	0.01	0.01
20	Z29705.1-green-ng3b	0.09	1.02	0.10	0.09	0.08	0.07	0.08	0.08	0.08	0.07	0.07	0.08	0.10	0.10	0.10	0.10	0.11	0.11	0.00		0.01	0.01	0.01	0.01	0.01	0.01
21	Z29706.1-mild-nm3a	0.11	1.21	0.11	0.11	0.10	0.10	0.12	0.04	0.10	0.10	0.10	0.12	0.09	0.09	0.09	0.09	0.15	0.15	0.09	0.10		0.01	0.01	0.01	0.02	0.01
22	Z29707.1-mild-nm3b	0.10	1.17	0.11	0.11	0.10	0.09	0.11	0.04	0.10	0.09	0.09	0.11	0.12	0.12	0.12	0.12	0.15	0.15	0.09	0.09	0.02		0.00	0.00	0.02	0.01
23	Z29708.1-mild-nm3c	0.10	1.17	0.11	0.11	0.10	0.09	0.11	0.04	0.10	0.09	0.09	0.11	0.12	0.12	0.12	0.12	0.15	0.15	0.09	0.09	0.02	0.00		0.00	0.02	0.01
24	Z29709.1-mild-nm3d	0.11	1.18	0.12	0.12	0.10	0.10	0.12	0.04	0.10	0.10	0.10	0.11	0.12	0.12	0.12	0.12	0.16	0.15	0.09	0.09	0.03	0.00	0.01		0.02	0.02
25	Z29710.1_Yellow_blotch-yb3a	0.10	1.19	0.12	0.11	0.10	0.10	0.04	0.14	0.09	0.09	0.10	0.07	0.10	0.10	0.10	0.10	0.01	0.00	0.11	0.11	0.15	0.14	0.15	0.15		0.01
26	Z29711.1_Yellow_blotch-yb3b	0.08	1.10	0.09	0.10	0.08	0.08	0.02	0.12	0.08	0.08	0.08	0.05	0.10	0.10	0.10	0.10	0.05	0.05	0.09	0.10	0.13	0.13	0.13	0.14	0.05	

* The number of base substitutions per site from between sequences are shown. Standard error estimate(s) are shown above the diagonal. Analyses were conducted using the Tamura-Nei model (Tamura and Nei, 1993). The rate variation among sites was modeled with a gamma distribution (shape parameter = 1). The analysis involved 26 nucleotide sequences. KU997687.1_TBTV_SatRNA was used as outgroup.

Further alignment of the Kenyan Sat-RNA Isolates alongside the green rosette (M16S, M11S) and chlorotic rosette isolates (M24S) showed major changes in nucleotide sequences at positions 161-211 and 539-546 which were similar to those in M24S. These positions had very minimal nucleotide changes in the rest of the western Kenyan and Malawian Isolates (Fig. 7; Appendix II)

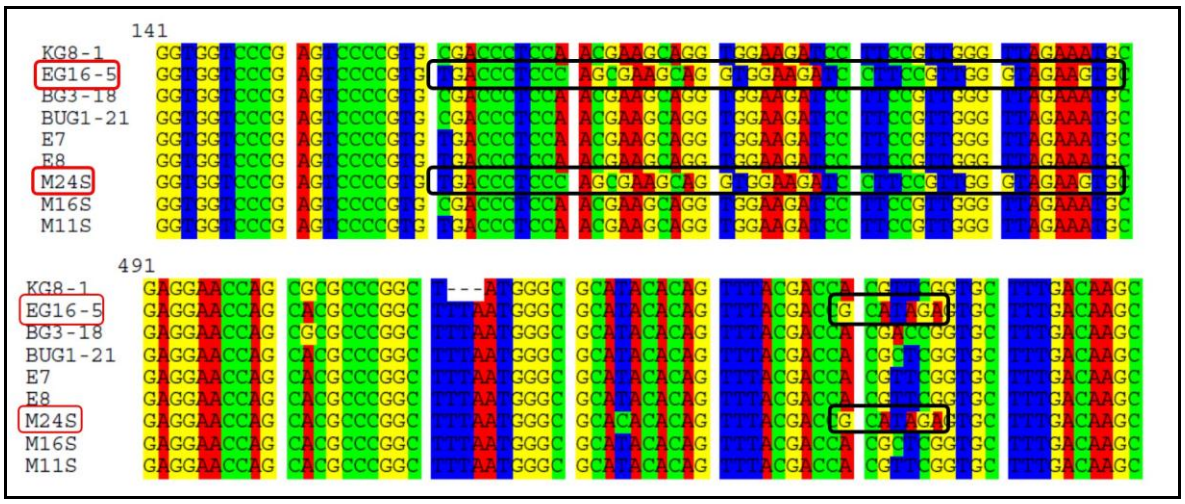


Figure 7: Regions of high divergence between the green rosette and chlorotic rosette isolates

The six Kenyan Sat-RNAs were deposited in the GeneBank with accession numbers LC469779, LC472299, LC472300, LC472301, LC472302 and LC472303.

4.4.5 Diversity of GRV

Four GRV genomes (E3, E5, E7, and E8) were assembled (2401-4171 nt). The assembled genomes were compared with 3 complete genomes available in the GeneBank; MG646923.1 (SRF540), MG646922.1 (SRF57) (from western Kenya) and MC1-Z69910.1 (from Malawi) and GRV ORF3 and 4 complete cds from Malawi and Nigeria. Isolates E3 and E5 clustered closest with SRF54 and SRF57 than E7 and E8. E7 and E8 formed a distinct clade, however they shared at least 84% identity with other GRV genomes. E3 was closest to MC1 with 98% identity followed by E7 (86%) and least with E5 and E8 at 84% each. All isolates shared between 97-98% identity with both SRF54 and SRF57 (Fig. 8).

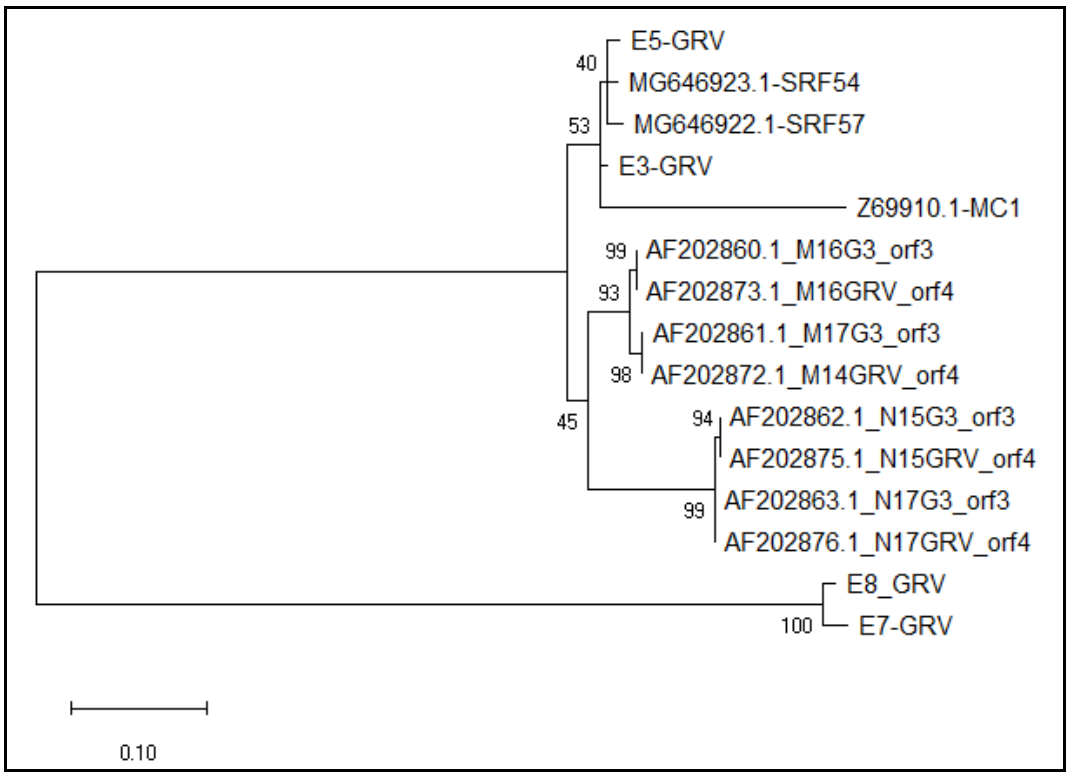


Figure 8: Phylogenetic tree of the Kenyan GRV isolates and those in GeneBank

The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model (Tamura and Nei, 1993). Bootstrap confidence values (500 replications) are shown.

4.4.5 Diversity of GRAV

Four GRAV coat protein (CP) gene sequences were assembled (600 nt). The four were compared with GRAV CP gene sequences from Malawi, Nigeria and Ghana available in the GeneBank. The comparison revealed 97-100% identity with the Kenyan isolates. Isolates GRAV-5 and GRAV-19 each had 100% identity with M16GCP (AF195824.1) and 99% with M8GCP (AF195502.1) then 98% with the other Malawian, Nigerian and Ghanaian Isolates. Isolate GRAV-22 had 99% identity with isolates M16GCP and M8GCP then 98% with the other Malawian, Nigerian and Ghanaian Isolates. Isolate GRAV-12 (isolated from common beans) had 100% identity with M16GCP and 99% with M8GCP from Malawi, then 98% with the rest of Malawian, Ghanaian and Nigerian isolates except N29GCP (AF195828.1) and

N15GCP (AF195825.1) that showed 97% identity. In phylogenetic tree, all Kenyan isolates clustered together with isolate M16GCP. In general all western Kenya isolates exhibited closest identity and grouped together with some Malawian isolates, M16GCP and M8GCP than the rest of Malawian, Nigerian and Ghanaian isolates (Fig. 9).

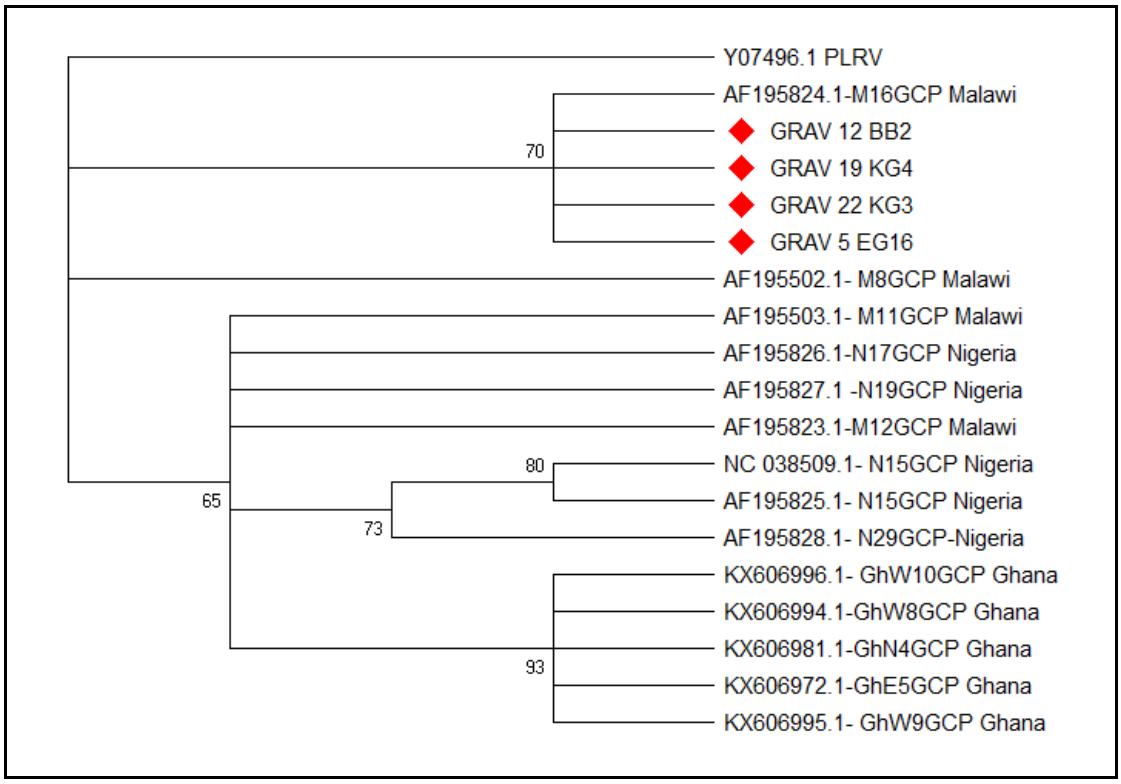


Figure 9: Phylogenetic tree of the 600nt western Kenya GRAV CP and GeneBank isolates.

The evolutionary history was inferred by using the Maximum Likelihood method based on the Kimura 2-parameter model (Kimura, 1980). The tree is rooted on of a distantly related *Luteovirus* (Potato leaf roll virus – Y07496.1 PLRV). Bootstrap confidence values (500 replications) are shown.

The four GRAV sequences were deposited in GeneBank with accession numbers LC480460 (GRAV 12), LC480461 (GRAV 22), LC480462 (GRAV 19) and LC480463 (GRAV 5).

In addition three complete GRAV genomes (E5-GRAV, E7-GRAV and E8-GRAV) were assembled. These three shared 97-98% sequence identity with GRAV complete

and partial sequences from Malawi, Ghana and Nigeria available in the GeneBank. The three were then compared with representatives of other assigned viruses in the family *Luteoviridae*, namely: *Polerovirus*, *Enamovirus* and *Luteovirus*. The Kenyan GRAV isolates clustered closest with the *Luteoviruses* than *Poleroviruses* and *Enamoviruses* in the phylogenetic tree (Fig. 10).

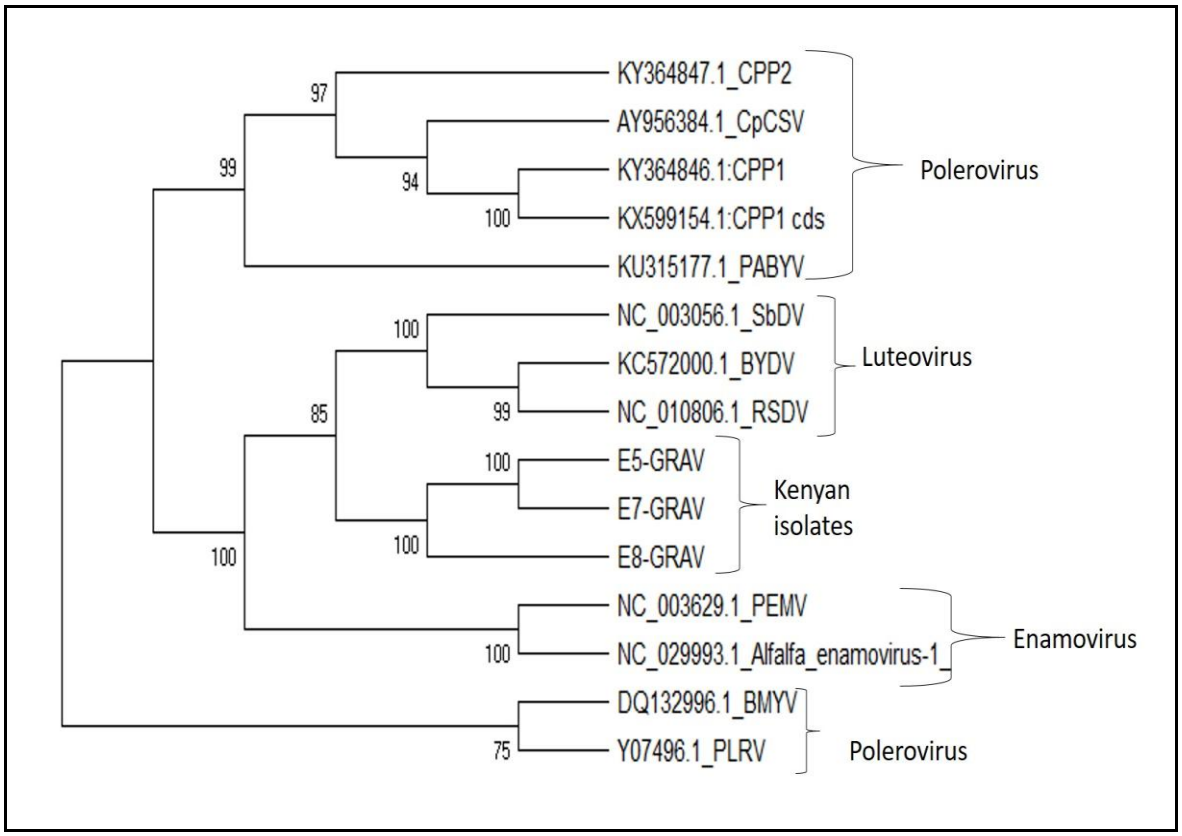


Figure 10: Phylogenetic tree of the three complete genome GRAV Kenyan isolates with other viruses in the family Luteoviridae.

The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model. The percentage of trees in which the associated taxa clustered together is shown next to the branches. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Bootstrap confidence values (500 replications) are shown.

4.5 Host range

The screened plants expressed distinct symptoms of stunted growth, shortened internodes, thickened stems, necrosis, dwarfism with bushy appearance, dark green,

yellowing with chlorosis lesions, mixed mosaic, reduced leaf area with twisted and distorted leaves curling downwards and upwards (Table 9; Fig. 11).

Table 9: Greenhouse test crop symptoms and RT-PCR test results

Test plant	Local Symptoms*	Systemic symptoms*	Sat-RNA	GRAV	GRV
Cowpea	N	SS, CS	+	+	-
Groundnuts	N	SS, CS, VC	+	+	+
Soybean	N	SS, CS, BN	+	+	-
Common beans	N	SS, DC, CS	+	+	-
Green grams	N	SS, D, CS	+	+	+
Physalis peruviana L. (golden berry)	N	DC, CB	+	+	+

*Key: N – necrosis, SS-shiny leaf surface, CS-chlorotic spots, VC-veinal chlorosis, DC-downward leaf curling, CB-chlorotic blotches, BN-Back necrosis, D-Dwarfing.

Figure 11: Symptoms expressed by plants inoculated with GRD associated viruses.

1a: Stunting and chlorosis on groundnut, **1b**: Healthy groundnut; **2a**: Shiny chlorosis in cowpea, **2b**: Healthy cowpea.

4.5.1 RT-PCR for the green house and field samples and validation of GRD diagnostic primers

Ten samples, seven from the green house plants inoculated with GRD viruses and three collected during survey in farms, were tested by RT-PCR to detect GRAV, GRV and Sat-RNA using the designed primers shown in Table 1. Total RNA eluted typically ranged between 30 – 55 ng/µl (Table 10; Fig. 12).

Table 10: Quantities of total RNA eluted for samples used in RT-PCR.

Sample ID	RNA (ng/µl)
2	38.7
1	45.5
3	49.6
4	51.8
5	31.9
6	54.2
7	42.6
8	48.7
9	30.0
10	55.0

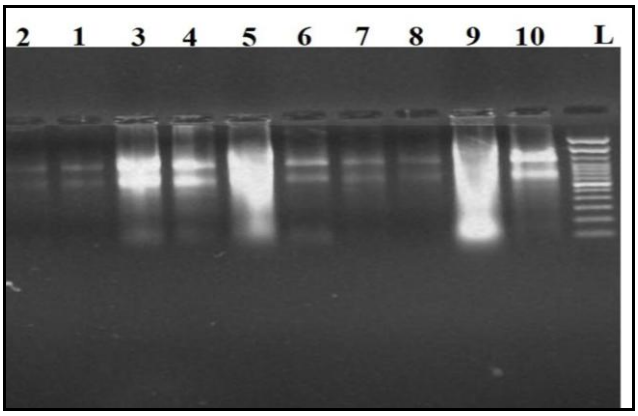


Figure 12: Gel quantification of total RNA eluted.
Lanes 1-10 corresponds to sample codes: 1- inoculated bean, 2- inoculated soybean, 3- green rosette groundnut, 4- inoculated golden berry, 5- chlorotic rosette groundnut, 6- inoculated ground nut, 7- inoculated cowpea, 8- chlorotic rosette groundnut, 9- mosaic rosette groundnut, 10- inoculated green gram, L – 100 bp ladder (Fermentas).

All ten samples tested positive for Sat-RNA and GRAV while three for GRV. Band sizes of approx. 900 bp 567 bp and 860 bp were observed on gel for Sat-RNA, GRAV

and GRV respectively (Fig. 13, 14 and 15). The designed diagnostic primers were able to detect all the three GRD causal agents.

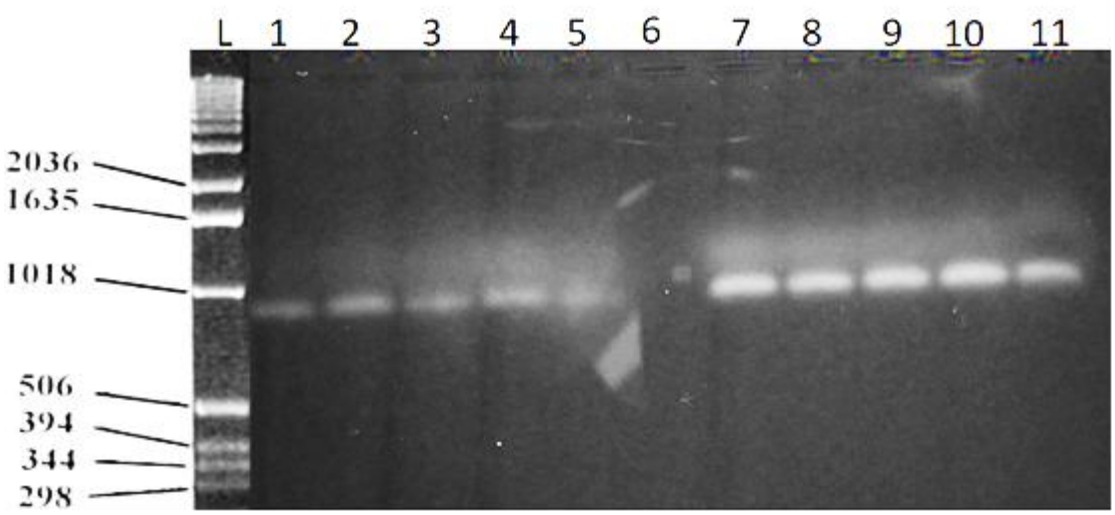


Figure 13: Gel electrophoresis of RT-PCR amplified RNA using primers specific for Sat-RNA for samples inoculated with GRD viruses and GRD symptomatic field samples.

Expected band size was 900 bp. Lane L- 1 kb ladder, 1- inoculated bean, 2- inoculated soybean, 3- green rosette groundnut, 4- inoculated golden berry, 5- chlorotic rosette groundnut, 6- negative control (molecular grade water), 7- inoculated cowpea, 8- chlorotic rosette groundnut, 9- mosaic rosette groundnut, 10- inoculated green gram, 11- inoculated ground nut.

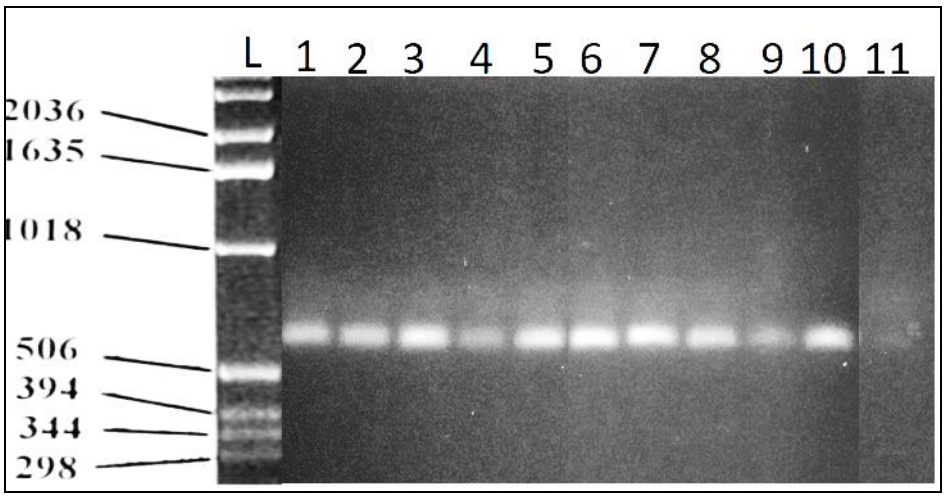


Figure 14: Gel electrophoresis of RT-PCR amplified RNA using primers specific for GRAV-CP gene for samples inoculated with GRD viruses and GRD symptomatic field samples.

Expected band size was 597 bp. Lane L- 1 kb ladder, 1- inoculated bean, 2- inoculated soybean, 3- green rosette groundnut, 4- inoculated golden berry, 5- chlorotic rosette groundnut, 6- inoculated ground nut, 7- inoculated cowpea, 8- chlorotic rosette groundnut, 9- mosaic rosette groundnut, 10- inoculated green gram, 11- negative control (molecular grade water).

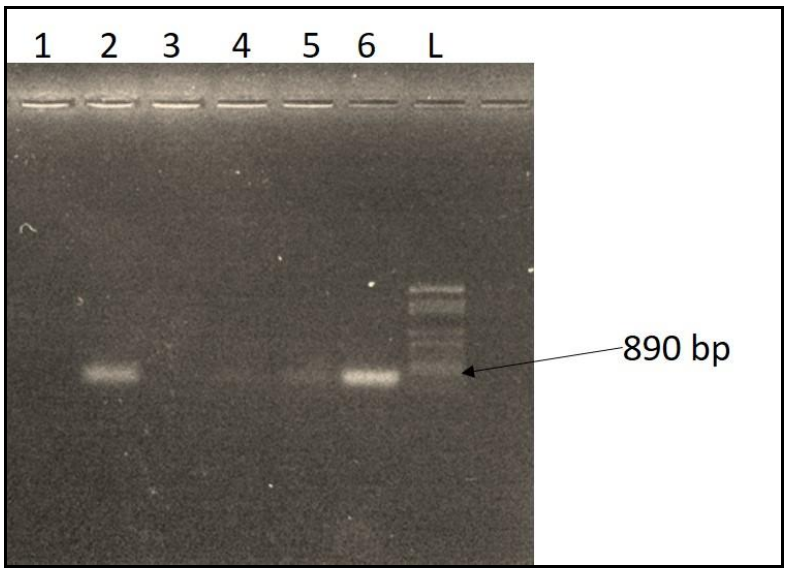


Figure 15: Gel electrophoresis of RT-PCR amplified RNA using primers specific for GRV for samples inoculated with GRD viruses and GRD symptomatic field samples.

Expected band size was 860 bp. Lane 1-inoculated cowpea, 2- green rosette groundnut, 3-inoculated soybean, 4-inoculated beans, 5-inoculated green grams, 6-inoculated golden berry, L-100bp ladder (Fermentas).

CHAPTER FIVE

DISCUSSION

Groundnut rosette is the most prevalent disease of groundnuts in western Kenya. The disease was recorded in every County that was surveyed with incidences of up to 100%. The short rain season recorded higher incidence (53%) than the long rains (41%). This could be attributed to the high vector pressure during the short rains as compared to the long rains season when the aphid pressure is low as a result of heavy rains that wash the insects away. A study by Mugisa *et al*., (2016) found that periods of long rains negatively affected GRD progression as aphid vector pressure was low. Were *et al*., (2013) reported a positive correlation between potato disease incidence and aphid numbers. This further supports the implication that virus disease incidence variations between the seasons contributed to by differences in vector pressure. Incidence increased with increase in severity due to early infection leading to intensification of the viruses as the plant grows and build-up of inoculum for vectors to spread to nearby plants. Groundnut rosette is a polycyclic disease whereby diseased plants from previous cropping season serves as inoculum sources for initiating subsequent disease spread (Naidu *et al*., 1998a). In western Kenya, groundnuts are grown in two cropping seasons (long rains and short rains) and due to limitation in land to practice shift cultivation, the same piece of land is continuously used to grow the same or related host crops in the subsequent cropping season. Therefore, GRD infected groundnuts and possibly hosts of any of the GRD associated viruses remaining from the long rains season serves as immediate sources of the GRD agents beginning the disease cycle at early stages of crop development in the short rains cropping season. Such initial infections that occur at early stages of plant growth enhance repeated cycles of infections thus increasing the severity of the disease in the

groundnut fields (Waliyar *et al*., 2007). Sources of groundnuts seeds were mainly open air market and own saved seed. There was no single record of groundnuts seeds from agro-dealer. This implies that there is lack of a seed system for groundnuts that is reliable in terms of certification to ascertain the quality of the seed. Although GRD associated viruses are not seed-borne, other seed-borne viruses of groundnuts such as CPMMV can easily be spread when farmers use seed from uncertified sources.

All major GRD symptoms were observed in the surveyed region with chlorotic rosette being most prevalent followed by green rosette. This supports the findings of Wangai *et al*., (2001) who reported chlorotic rosette to be the most prevalent GRD symptoms in the region. The high prevalence of the chlorotic rosette could also be attributed to its higher transmission efficiency compared to green rosette. This observation concurs with that of Misari *et al.*, (1988a), who reported minimum acquisition feeding periods of 4 h and 8 h for chlorotic and green rosette respectively and the median latent periods of 26.4 h, 38.4 h respectively, for chlorotic and green rosette. The mosaic symptom has not been previously reported but was distributed in most of the surveyed region. This suggests that there is evolution of new variants of Sat-RNA in western Kenya that might be causing these new symptoms or the mosaic was due to another causal agent. A total of 10 variants of Sat-RNA have been reported to be associated with the various GRD symptoms (Blok *et al*., 1994). A mixture of either variants, especially the chlorotic and green rosette and/or the mild ones, are likely to induce the mosaic symptoms (Naidu *et al*., 1998a). It is therefore possible that some of these variants occur in western Kenya in mixed infections, thus causing the varied symptom observed, especially the mosaic. Apart from the typical rosette symptoms, other symptoms including severe leaf curling and bunching were observed. This suggests that there is wider variability in expression of GRD and could be due to more severe

variants of associated viruses or other agents. It is worth noting that from the Next generation Sequences (NGS) of this study, other than GRD associated viruses, other viruses were detected (data not shown) and could be the reason for some of the new symptoms observed on groundnuts (Mukoye *et al*., 2018).

The Sat-RNAs assembled were all complete genomes as they ranged between 896-901 nucleotides in length. The length of Groundnut rosette virus associated Sat-RNA range between 895-903 nucleotides (Blok *et al*., 1994). The western Kenya Sat-RNAs showed close identity (92-100%) to Malawian isolates than those from Ghana and Nigeria (88-93%). This implies that the genetic diversity of the Sat-RNA become more varied with wide geographical distance. Kenya and Malawi are located in Eastern Africa while Ghana and Nigeria are in West Africa thus a wider geographical separation. This finding concurs with that of Wangai *et al*., (2001) who observed a closer sequence relationship between Kenyan Sat-RNA isolates and those from Malawi. However, this study has reported sequence identity of up to 100% with Malawian isolates as opposed to 95% reported by Wangai *et al*., (2001). This suggests that more variants of Sat-RNA exist in western Kenya that could be contributing to the diverse symptoms expressed by GRD. This study used NGS which has been demonstrated to be more reliable in detection of new or poorly characterized viruses (Rott *et al*., 2017). There were variations among the western Kenya Sat-RNA isolates similar to Malawian isolates where they formed distinct clusters in the phylogenetic tree. Isolate EG16-5 was the most distinct and clustered together with chlorotic and yellow blotch Sat-RNA variants. This suggests that this isolate is associated with the chlorotic rosette symptoms that were most prevalent in the surveyed areas. Further analysis of evolutionary divergence between the Sat-RNA isolates revealed Isolate EG16-5 to be the most diverse among the western Kenya Isolates (above 0.06) while

the rest were less diverse (below 0.02). Moreover, when compared with Malawian isolates, the same isolate (EG16-5) had the least evolutionary divergence with M24S (chlorotic rosette) than the rest of the Malawian isolates (green rosette). In addition, major changes in nucleotide sequences in isolate EG16-5 were same as those observed in Isolate M24S. These observations further supports that Isolate EG16-5 is a chlorotic variant of Sat-RNA.

The GRV isolates from western Kenya shared 97-98% identity with those previously described by Wainaina *et al*., (2018). This implies that there was close identity among GRV genomes from western Kenya. However, isolate E3 was the closest (98%) to MC1 (Malawian isolate) than the rest of the isolates (84-86%) implying E3 had less divergence from Malawian GRV as was observed by Wangai *et al*., (2001) using ORF3 and 4 who reported identity of 95-96%. This study, however, has also found existence of more diverse GRV isolates (84-86%) in comparison to Malawian MC1. Similar observation was reported by Wainaina *et al*., 2018 where sequence similarity of 84% was observed between Kenyan and Malawian GRV isolates. The use of more genome regions of GRV, in addition to ORF3 and 4, therefore gave more genomic characteristics of GRV isolates leading to identification of more diversity between Kenyan and Malawian GRV isolates. All the four GRv isolates in this study shared 84-98% identity with other GRV isolates in the GeneBank confirming that they are not new viruses but GRV (King *et al*., 2012).

The four GRAV CP gene sequences from western Kenya clustered together and had 97 – 100% identity with those from Malawi, Ghana and Nigeria implying that there was no much difference among the western Kenya GRAV CP gene isolates. Kenyan GRAV CP isolates exhibited closest identities with Malawian isolates than Nigerian and Ghanaian isolates. This findings concurs with Wangai *et al*., (2001) and Appiah *et*

al., (2017) who observed closer identity between sequences from the same geographical region as compared to those from separate geographical regions. In the study, Wangai *et al*., (2001) found that Kenyan isolates of GRAV CP gene shared 98% nucleotide identity with Malawian isolates as compared to 96-97% with those from Nigeria. Appiah *et al.,* (2017) observed that Ghanaian GRAV CP gene sequence isolates had 98-99% nucleotide identity as compared to 97-99% with Malawian isolates. Such differences due to geographical distances could be as a result of differences in environmental conditions that bring about variations in evolution of the viruses. All western Kenya GRAV CP isolates were closest to Malawian isolates M16GCP and M8GCP (99-100%) than the other isolates from Malawi, Nigeria and Ghana. A similar observation was noted by Wangai *et al*., (2001) where two of the Kenya isolates in the study (K1 and K2), specifically from western Kenya were closest to M16GCP and M8GCP than with the rest of her isolates from other regions in Kenya. This could imply that the GRAV CP gene from western Kenya have not evolved for at least the last 20 years. However variation could exist in GRAV from other regions in Kenya. It is worth noting that our studies (not published) found GRAV in common beans (*Phaseolus vulgaris*). This is a new finding as only groundnuts were the only known natural hosts of GRAV (Waliyar *et al*., 2007). This suggests that GRAV has other natural hosts other than groundnuts and therefore being an important agent of GRD it can survive in such hosts which can serve as sources of infection when picked by its vector (aphids). In general all GRAV CP gene sequences both in this study and those in GeneBank shared 97-100% nucleotide identity. This implies that GRAV CP gene is highly conserved across the wide geographical region in Sub-Saharan Africa. It can thus be targeted as a suitable candidate for development

of pathogen-derived resistance (PDR) through genetic engineering that can be used across Sub-Saharan Africa (Deom *et al*., 2000; Appiah *et al*., 2017).

The three complete genomes of GRAV clustered closest with the *Luteoviruses* than *Poleroviruses* and *Enamoviruses*. This gives an indication that GRAV could be having genomic characteristics similar to the *Luteoviruses*. This finding contradicts that of Jones *et al*., (2020) who reported two GRAV isolates that grouped together with *Poleroviruses* and suggested that the unassigned GRAV to be assigned to the genus *Polerovirus*. This could be possibly due the fact that the comparison was only made using the protein sequences of the coat protein gene alone. The comparison in this study involved the entire genome of GRAV from a metagenomics study which therefore compared all regions of the entire genome allowing precise grouping (Simmonds & Aiewsakun, 2018). This study therefore suggest that GRAV be classified as a member of the genus *Luteovirus*.

All major legumes screened as hosts of GRD agents developed both local and systemic symptoms and PCR confirmed presence of the GRD causal agents. This is an indication that the major legumes grown in western Kenya can serve as alternative hosts of one or all of the GRD agents. Using mechanical sap inoculations, several species in *leguminosae, chenopodiaceae* and *solanaceae* have been identified as experimental hosts of GRV and Sat-RNA. In the same families *Glycene max* and *Phaseolus vulgaris* are among the systematic hosts of GRV (Waliyar *et al.,* 2007). These can therefore become sources of inoculum when the main natural host is planted adjacent to or intercropped with such infected alternative hosts. In this study, sequences from a field sample revealed the presence of GRAV in common bean (*Phaseolus vulgaris*). This further confirms common beans as an alternative host of GRD causal agent in nature. One of the cropping systems used in western Kenya of

mixing all legumes in the same piece of land could therefore enhance the spread of GRD among the host plants.

All the sets of primers designed in this study were able to detect all the GRD causal viruses. Diagnostic primers need to be able to detect a whole range of virus variants and strains for use in making plant health decisions. With proper characterization of the GRD viruses in Kenya and designing primers from consensus sequences across SSA, these primers can be utilized in routine diagnosis of GRD.

Conclusion

This study concludes that:

- Groundnut Rosette (GRD) is still the major disease of groundnuts and is present whenever groundnuts are grown in western Kenya. Chlorotic rosette is the most prevalent form of symptom on groundnuts in western Kenya. The mosaic rosette is an emerging symptom in groundnuts and could be due to dual infection by Sat-RNA variants or other agents.

- Genetic diversity of the GRV and Sat-RNA become more varied with wide geographical distance. The western Kenya Sat-RNA variants were closely identical to those of Malawi than those from Nigeria and Ghana. New variants of Sat-RNA exists in western Kenya that are contributing to the diverse symptoms expressed by GRD. A new chlorotic rosette variant of Sat-RNA in Kenya was unveiled in this study (EG16-5).

- The GRAV CP gene is less diverse even with wide geographical distance. All the western Kenya isolates showed close identity of 97-100% when compared with those from Malawi, Ghana and Nigeria. Common bean is a new natural host of GRAV in addition to groundnuts.

- GRAV is most likely to be a member of the genus *Luteovirus* in the family *Luteoviridae*.

- All the major legumes grown in western Kenya are susceptible to GRD agents through mechanical inoculation and therefore can serve as alternative hosts of one or all of the GRD agents.

- The designed primers can detect the GRD associated viruses and thus can be utilized in routine diagnosis of GRD.

- The use of NGS is essential in discovery of new plant viruses and characterization of those that are poorly characterized.

Recommendations

This study recommends the following:

- There is need for urgent measures to manage GRD in western Kenya possibly through the exploitation of pathogen-derived resistance (PDR).

- Crop rotation of groundnuts with non-hosts of GRD be adopted as a cultural measure to break the cyclic nature of the disease. It is also important to check the soil to ascertain whether the GRD associated viruses are soil borne.

- Volunteer leguminous crops from previous cropping season be rogued before planting new crop to reduce the chances of acting as immediate initial sources of GRD inoculum.

- There is need for a reliable seed production and certification for groundnuts in western Kenya.

REFERENCES

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

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990). Basic local alignment search tool. *J Mol Biol* 215:403– 410. https://doi.org/10.1016/ S0022-2836(05)80360-2.

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. *Arc.virol*.159:3059-3062,doi: 10.1007/s00705-014-2139-7

Appiah, A. S., Offei, S. K., Tegg, R. S., & Wilson, C. R. (2016). Varietal response to groundnut rosette disease and the first report of *Groundnut ringspot virus* in Ghana. *Plant Dis.* 100(5):946-952. http://dx.doi.org/10.1094/PDIS-07-15-0838-RE

Appiah, A. S., Sossah, L. F., Tegg, S. R., Offei, K. S. & Wilson, R. C. (2017). Assessing sequence diversity of goundnut rosette disease agents and the distribution of *groundnut rosette assistor virus* in major groundnut-producing regions of Ghana. *Trop. Plant Pathol*. Doi: 10.1007/s40858-017-0140-x

Ayoola, P. B., Adeyeye, A. & Onawumi, O. O. (2012). Chemical evaluation of food value of groundnut *(Arachis hypogaea)* seeds. *American journal of food and Nutrition.* Online:ISBN 2157-1317,doi:10.5251/ajfn.2012.2.3.55.57.

Blok,V. C., Ziegler, A., Robinson, D. J. & Murant, A. F. (1994). Sequences of 10 variants of the satellite - like RNA -3 of groundnut rosette virus. *Virology* 202: 25-32.

Boonham, N., Kreuze, J., Winter, S., van der Vlugt R, Bergervoet, J., *et al*. (2014) Methods in virus diagnostics: from ELISA to next generation sequencing. *Virus Research,* 186: 20 - 31.

Bucheyeki, T. L., Shenkalwa, E. M., Mapunda, T. X. & Matata, L. W. (2008). On-farm evaluation of promising groundnut varieties for adaptation and adoption in Tanzania. *African Journal of Agricultural research*, 3:531-600.

Deom, C. M., Naidu, R. A., Chiyembekeza, A. J., Ntare, B. R. & Subrahmanyam, P. (2000). Sequence diversity with the three agents of groundnut rosette disease.*Phytopathol.* 90:214-219.doi:10.1094/PHYTO.2000.90.3.214

Eichmeier, A., Kominkora, M., Kominek, P. & Baranek, M. (2016). Comprehensive virus detection using next generation sequencing in grapevine vascular tissues of plants obtained from the wine region of Bohemia and Moravia (Czech Republic). *PloS ONE* 11(12):e0167966.doi:10.1371/journal.pone0167966

FAOSTAT. (2013). Global groundnut yield. Available at *http://faostat.fao.org/* Groundnut /Arachis/Start.htm (Accessed on 26th July, 2017).

Haas B.J, Papanicolaou A, Yassour M, Grabherr M, Blood P.D, Bowden J, Couger M.B, Eccles D, Li B, Lieber M, MacManes M.D, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey C.N, Henschel R, LeDuc R.D, Friedman N. & Regev A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc 8:1494 – 1512. https://doi.org/10.1038/nprot.2013.084.

Herselman, L., Thwaites, R., Kimmins, F. M., Curtois, B., van der Merwe, P. J. A. & Seal, S. E. (2004). Identification and mapping of AFLP markers linked to peanut (*Arachis hypogaea L.*) resistance to the aphid vectorof groundnut rosette disease. *Theor. Appl. Genet.* 109:1426-1433.

Hong, J., Feng, H., Wang, F., Ranjan, A., Cheng, J., Jiang, J., Ghirlando, R., Xiao, T. S., Wu, C. & Bai, Y. (2014). The catalytic subunit of the SWR1 remodeler is a histone chaperone for the H2A.Z-H2B dimer. *Mol. cell* 53(3):498-505.

ICRISAT. (2012). Groundnut production information. Available at *http://www.icrisat-org/*Icrisat-crops.html (Accessed on 1st Agust, 2017).

Jones, S., Cowan, G., MacFarlane, S., Mukoye, B., Mangeni, B.C., Were, H. & Torrance, L. (2020). RNA sequence analysis of diseased groundnut (*Arachis hypogaea*) reveals the full genome of groundnut rosette assistor virus (GRAV). *Virus Research*; 277 (2020) 197837.

Kayondo, S. I., Rubaihayo, P. R., Ntare, B. R., Gibson, P. I., Edema, R., Ozimati, A. & Okello, D. K. (2014). Genetics of resistance to groundnut rosette virus disease*: African crop science journal,* 22:21-29. ISSN:1021-9730/2014.

Kidula, N., Okoko, N., Bravo-Ureta, B. E., Thuo, M. & Wasilwa, L. (2010). A preliminary analysis of yield differences in groundnuts between research and non-research farmers in Kenya. In paper presented at the 12th KARI biennial scientific conference, 8-12 November 2010, Naiobi Kenya.

King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J. (2012) Virus taxonomy. Ninth report of the international committee on taxonomy of viruses. Elsevier/Academic Press, London

Kreuze, J. F., Perez, A., Untiveros, M., Quispe, D., Fuentes, S. & Barker, I. (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* 388:1-7.doi:10.1016/j.virol.2009.03.024.PMID:19394993

Kumar, P. L. & Waliyar, F., (Ed). (2007). Diagnosis and detection of viruses infecting ICRISAT mandate crops:*Methods Manual*.Patancheru 502 324, Andhra Pradesh, India*;International Crops Research Institute for the Semi-Arid Tropics.*133pp.

Kumar, S., Stecher, G., Li M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* 35:1547-1549.

Misari, S. M., Abraham, J.M., Demski J.W., Ansa, O. A., Kuhn, C.W, Casper, R. & Breyel, E. (1988a). Aphid transmission of the viruses causing chlorotic and green rosette diseases of peanut in Nigeria. *Plant Disease* 72: 250-253.

Mugisa, I.O., Karungi, J., Akello, B., Ochwo-Ssemakula, M.K.N., Biruma,M., Okello,D.K. & Otim,G. (2016). Determinants of groundnut rosette virus disease occurrence in Uganda.*Elseviercropprotectionjournal*.http://dx.doi.org/10.1016/j.cropro.2015.10.019.

Mukoye, B., Mangeni, B.C., Leitich, R.K., Wosula, D. W., Omayio, D.O., Nyamwamu P.A., Arinaitwe, W., Winter, S., Abang, M.M. & Were, H.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: 1-5.

Mukoye, B., Mangeni, B.C., Sue, J., Ndonga, M.F.O., & Were, H.K. (2018). Next Generation Sequencing as a tool in modern pest diagnosis. A case study of groundnuts (*Arachis hypogaea*) as a potential host of new viruses in western Kenya.

Conference proceedings: The 2nd Phytosanitary Conference, 4th – 8th June, 2018 at KEPHIS-Nairobi, Kenya.

Mumford, R., Boonham, N., Tomlinson, J., & Barker, I. (2006). Advances in molecular phytodiagnostics - new solutions for old problems. *European Journal of Plant Pathology*. 116, 1‑19. doi: 10.1007/s10658-006-9037-0

Murant, A. F., & Kumar, I. K. (1990). Different variants of the satellite RNA of groundnuts rosette virus are responsible for the chlorotic and green forms of groundnut rosette disease.*Ann.Appl.Biol.*117:85-92.

Mutegi, C. K. (2010). The extend of aflatoxin and aspergillus section flavi,penicillium spp.. and Rhizopus spp. contamination of peanuts from households in Western Kenya and the causative factors of contamination. *PhD dissertation , University of Kwazulu-Natal,Pietermaritzburg* . South Africa.

Naidu, R. A., Robinson, D. J., & Kimmins, F. M. (1998a). Detection of each of the causal agents of groundnut rosette disease in plants and vector aphids by RT-PCR. *J.Virol.Methods* 76:9-18.

Naidu, R. A, Bottenberg, H., Subrahmanyam, P., Kimminns, F. M., Robinson, D. J. & Thresh, J. M. (1998b). Epidemiology of groundnut rosette virus deasease:*Current status and future research needs.Ann. Appl. Biol.*132:525-548.

Naidu, R. A, Kimmins, F. M., Deom, C. M., Subrahimanyam, P., Chiyeubekeza, A. J. & van der Merwe P. J. A. (1999a). Groundnut rosette: Avirus disease affecting groundnut production in Sub-Saharan Afric*a. Plant Dis*.83:700-709.

Naidu, R. A., Kimmins, F., Holt, J., Robinson, D. J., Deom, C. M. & Subrahmanyam, P. (1999b). *Spatiotemporal separation of groundnut rosette disease agents. Phytopathol.*89:934-941.

Naidu, R. A. & Kimmins, F. M. (2007). The effect of groundnut rosette assistor virus on the agronomic performance of four groundnut (*Arachis hypogaea*) genotypes. *Journal of phytopathology* 155:350-356.doi:10.1111/j.1439-0434.2007.01234.x

Ntare, B. R., Olorunju, P. E. & Hildebrand, G. L. (2002). Progress in breeding early maturing peanut cultivars with resistance to groundnut rosette disease in West Africa.*Peanut Science* 29:17-23.

Okello, D. K.,Birima, M. & Deom, C. M.,. (2010). Overview of groundnut research in Uganda: Post,present and future. *Afr. J. Biotechnol.*9:6448-6459.

Okello, D. K., Monyo, E., Deom, C. M., Ininda, J. & Oloka, H. K. (2013). Groundnuts production guide for Uganda: Recommended practices for farmers. *National Agricultural Research Organization, Entebbe.*ISSN:978-9970-401-06-2

Okello, D. V, Akello, L. B, Tukamuhabwa, P., Odongo, T. L, Ochwo-Ssemakula, M., Adriko, J. & Deom, C. M. (2014). Groundnut rosette disease symptom types, distribution and management of the disease in Uganda. *African journal of plant science.* 8: 153-163.

Okello, D. V., Ugen, M. A., Tukamuhabwa, P., Ochwo-Ssemakula, M., Odong, T. L., Adriko, J., Kiconco, F., Male, A., and Deom, C. M. (2017). Molecular diagnostics of groundnut rosette disease agents in Uganda: Implications on epidemiology and management of groundnut rosette disease. *Journal of plant breeding and crop science*, 9: 63-70.

Olorunju, P. E., Ntare, B. R., Pande, S. & Reddy, S. V. (2001). Additional sources of resistance to groundnut rosette disease in groundnut germplasm and breeding lines.*Ann.Appl.Biol.*159:259-268.

Reddy, D. V. R. (1991). Groundnut viruses and virus diseases;Distribution,identification and control. *Rev.Plant Pathol.*70:665-678.

Robinson, D. J., Ryabov, E. V., Raj, S. K., Roberts, I. M. & Taliansky, M. E. (1999). Satellite RNA is essential for encapsidation of groundnut rosette *umbravirus* RNA by groundnut rosette assistor luteovirus coat protein. *Virol.* 254:104-114.

Roossinck, M. J. (1997). Mechanism of plant virus evolution. *Ann.Rev.Phytopathol.*35:191-209.

Scott, K. P., Farmer, M. J., Robinson, D. J., Torrence, L. & Murant, A. F. (1996). Comparison of the coat protein of groundnut rosette assistor virus with those of other *luteovirus*. *Ann.Appl.Biol.*128:77 - 83.

SADC/ICRISAT Groundnut Project Annual Progress Report for 1996. *Chitedze Research Station,* PPO Box1096, Lilongwe, Malawi.

Saitou N. and Nei M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4:406-425.

Salem, N. M., Ehlers, J. D., Roberts, P. A. & Ng, J. C. K. (2010). Biological and molecular diagnosis of seedborne viruses in cowpea germplasm of geographical diverse sub-Saharan origins. *Plant Pathology* 59:773-784.

SAS Institute. (2013).The SAS System for Windows. *Release* 9.3.1. SAS Inst.Cary,NC.

Simmonds, P. & Aiewsakun, P. (2018). Virus classification – where do you draw the line? *Archives of virology* 163:2037-2046.

Smartt. J. (1994). The groundnut in farming systems and the rural economy:A *global view*.Pages 664-699 in*:The Groundnut Crop: Ascientific basis for improvement.* J.Smartt, ed.Chapman & Hall, London.

Subrahmanyan, P., Hildebrand, G. L., Naidu, R. A., Reddy, L. J. & Singh, A. K. (1998). Sources of resistance to groundnut rosette disease in global groundnut germplasm.*Ann Appl.Biol.*132:473-485.

Tajima F. and Nei M. (1984). Estimation of evolutionary distance between nucleotide sequences. *Molecular Biology and Evolution* 1:269-285.

Taliansky, M. E., Robinson, D. J. & Murant, A. F. (1996). Complete nucleotide sequence and organisation of the RNA genome of groundnut rosette umbravirus. *J.Gen.Virol.*77:2335-2345 .

Taliansky, M. E. & Robinson, D. J. (1997). Trans-acting untranslated elements of groundnut rosette virus satelitte RNA are involved in symptom production. *J.Gen.Virol.*78: 1277-1285.

Taliansky, M. E., Robinson, D. J. & Murant, A. F. (2000). Groundnut rosette disease virus complex: Biology and Molecular Biology. *Advances in virus research,* 55:357-400.

Taliansky, M. E. & Robinson, D. J. (2003). Molecular Biology of umbraviruses*:Phantom warriors. J.Gen.Virol.*84:1951-1960.

Tamura K., Nei M., and Kumar S. (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences* (*USA*) 101:11030-11035.

Thuo, M., Bell, A. A., Bravo-Ureta, B. E., Lachaud, M. A., Okello, D. K., Okoko, N. E., Kidula, N. L., Deom, C. M. and Puppala, N. (2014). Effects of social network factors

on information acquisition and adoption of improved groundnut vaarieties: the case of Uganda and Kenya. *Agric. Hum. Values,* 31: 339-353.

Tillman, B. L. & Stalker, H. T. (2009). Peanut: In Johann Volmann and Istvan Rajcan Handbook of Plant Breeding, *Oil Crops* 4: 287-316.

Usman, A. (2013). Genetic analysis of resistance to rosette disease of groundnut (*Arachis hypogaea* L.). *A thesis submitted to the University of Ghana, Legon,*ISSN:10293978.

van der Merwe, P. J. A. & Subrahmanyan, P. (1997). Screening of rosette resistant short-duration groundnut breeding lines for yield and other characteristics.*Int Arachis.Newsl.*17:23-24.

Varshney, R.K., Kudapa, H., Roorkiwal, M., Thudi, M., Pandey, M. K., Saxena, R. K., *et al.* (2012b). Advances in genomics research and molecular breeding applications in SAT legume crops by using next generation sequencing and high-throughput genotyping technologies. *J Biosci* 37:811-20.

Varshney, R. K., Mohan, M. S., Gaur, M. P., Gangarao, R. P. N. V., Pandey, K. M., Bohra, A., et al. (2013). Achievements and prospecte of genomics-assisted breeding in three legume crops of the semi-arid tropics. *Elsevier Biotechnology advances* 31:1120-1134.

Wainaina, J.M., Harvey, J., Ateka, E., Makori, T., Karanja, D., Kehoe, M.A., Boykin, L.M. (2018). Genomic characterisation and evolutionary relationships of groundnut rosette virus from the western highlands of Kenya. *Tropical Plant Pathology* 43:583–585.

Waliyar, F., Kumar, P. L., Ntare, B. R., Monyo, E., Nigam, S. N., Reddy, A. S., Osiru, M. & Diallo, A. T. (2007). A Century of Research on Groundnut Rosette Disease and its Management.*Information Bulletin* no.75.Patancheru 502 324, Andhra Pradesh,

India.*International Crops Research Institute for the Semi-Arid Tropics*,40 pp.ISBN 978-92-9066-501-4.

Wangai, A. W., Pappu, S. S., Pappu, H. R., Okoko, N., Deom, C. M. & Naidu, R. A. (2001). Distribution and characteristics of groundnut rosette disease in Kenya.*Plant Disease*,85:470-474.

Were, H. K., Kabira, J. N., Kinyua, Z. M., Olubayo, F. M., Karinga, J. K., Aura, J., Lees, A. K., Cowan, G. H. & Torrance, L. (2013). Occurrence and Distribution of Potato Pests and Diseases in Kenya. *Potato Research* 56:325–342.

World Bank. (2015). Statistical information. Kenya and Uganda. Available at *http://data.worldbank.org* (Accessed on 24th June 2017).

APPEDICES

Appendix I: Disease survey score sheet

SURVEY DISEASE SCORE SHEET

CROP…….VARIETY……..…

Farmer's name……County……..…

District……Division……

Location……Sub-Location…….….

Village…….Date……

GPS readings; Altitude (Meters)……Longitude……..

Latitude (North or South)…….AEZ…….……

Groundnut variety grown ……..

Cropping pattern (*Tick the appropriate option*):

- No intercrop (stand-alone) ……
- Intercrop with other legume ……

Seed source (*Tick the appropriate option*):

- Agro-dealer ……..
- Own saved seed …….
- Open air market …….
- Neighbors ……

Disease score sheet

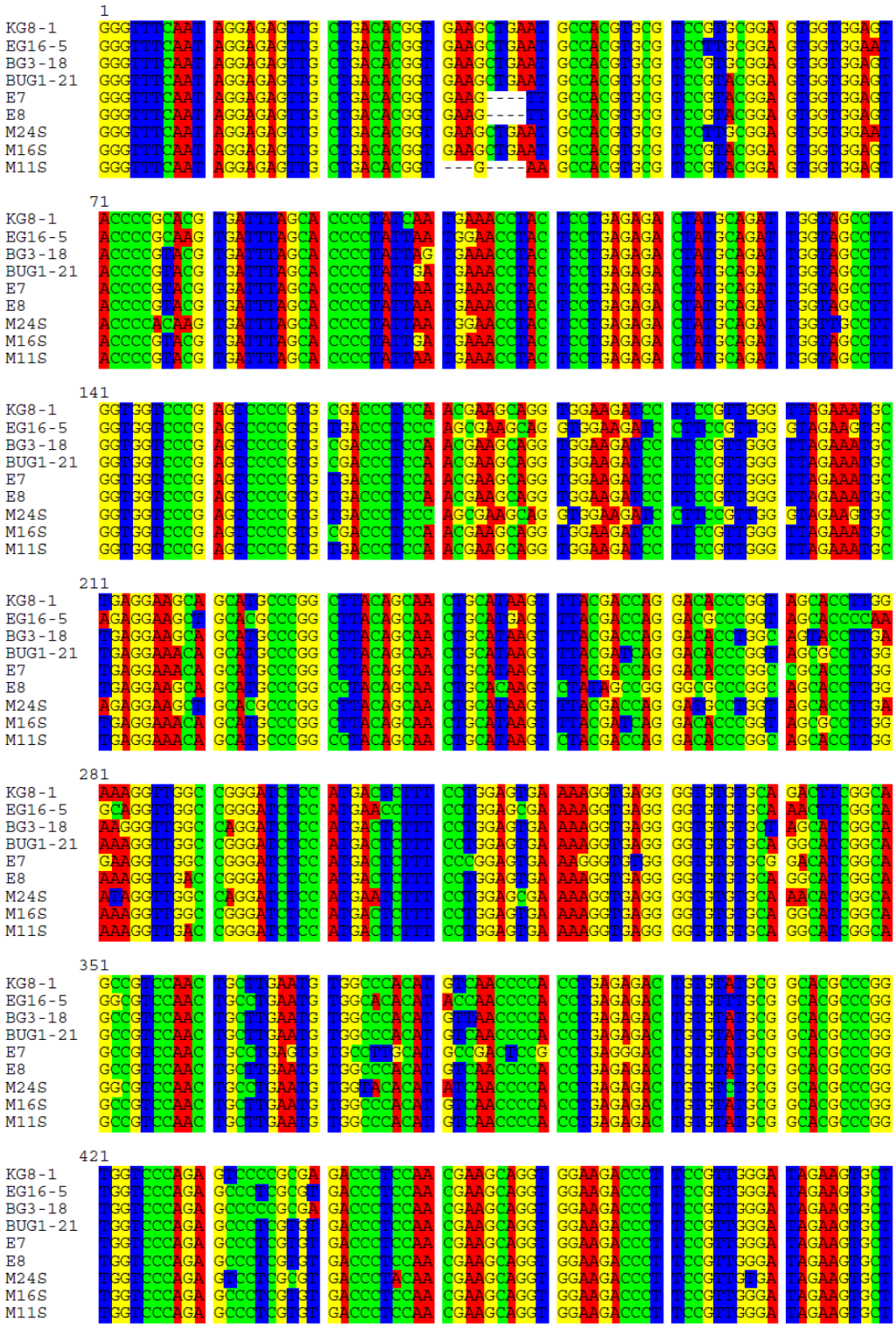
	Disease name…….			
	No. of plants affected per $10m^2$ quadrat	Symptoms	Distribution (whole field, spots)	Severity 0-3

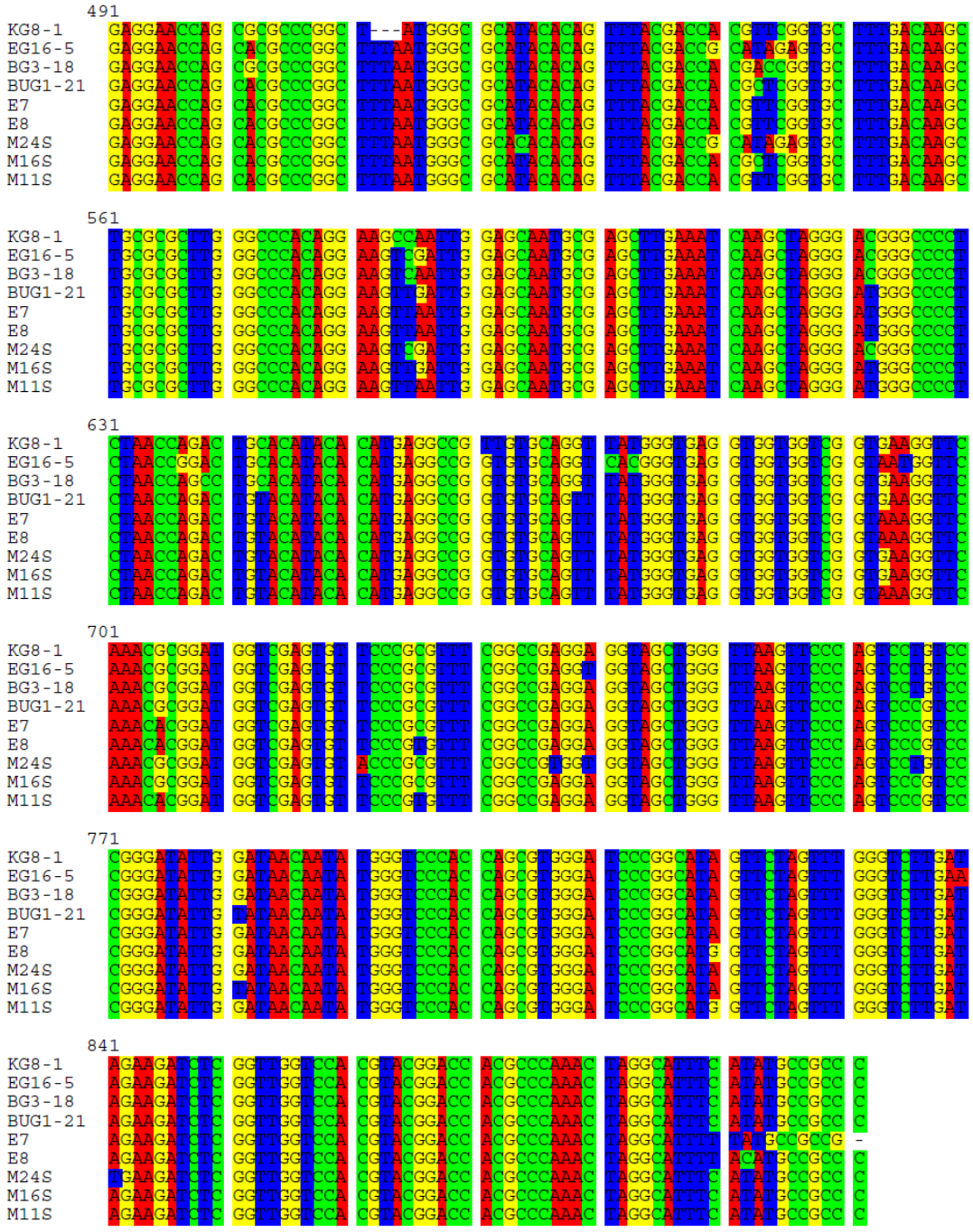
1				
2				
3				
4				

*Severity: 0= No disease; 1=Mild; = Moderate; 3=Severe.

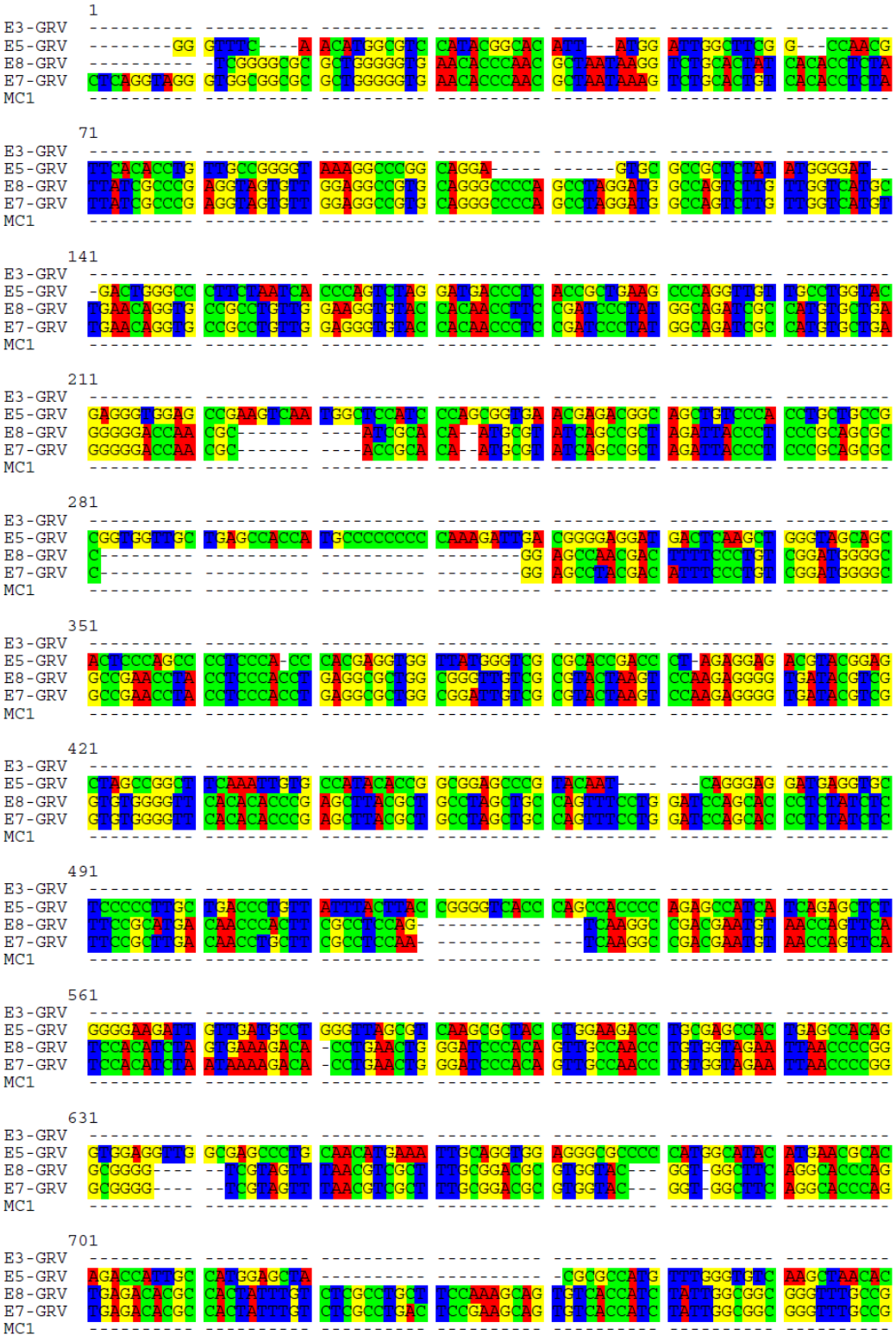
Number of plants affected per $10m^2$: select the area most affected, 10 steps square quadrat, count infected and total plants, (e.g. $^{20}/_{50}$ indicates 20 plants infected out of 50 plants in the 10x10 steps square quadrat).

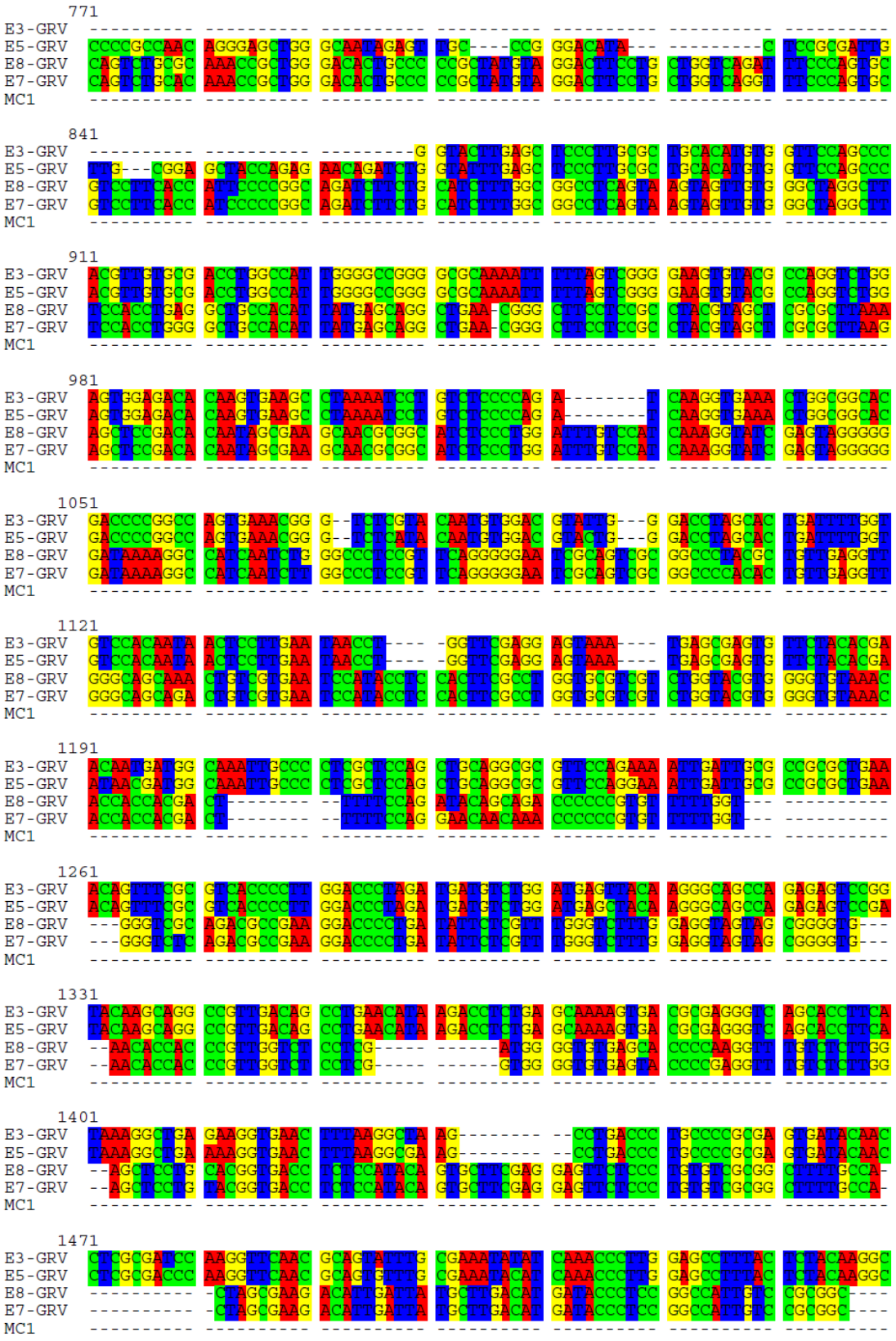
Appendix II: Sat-RNA alignment with green and chlorotic rosette Malawian Isolates

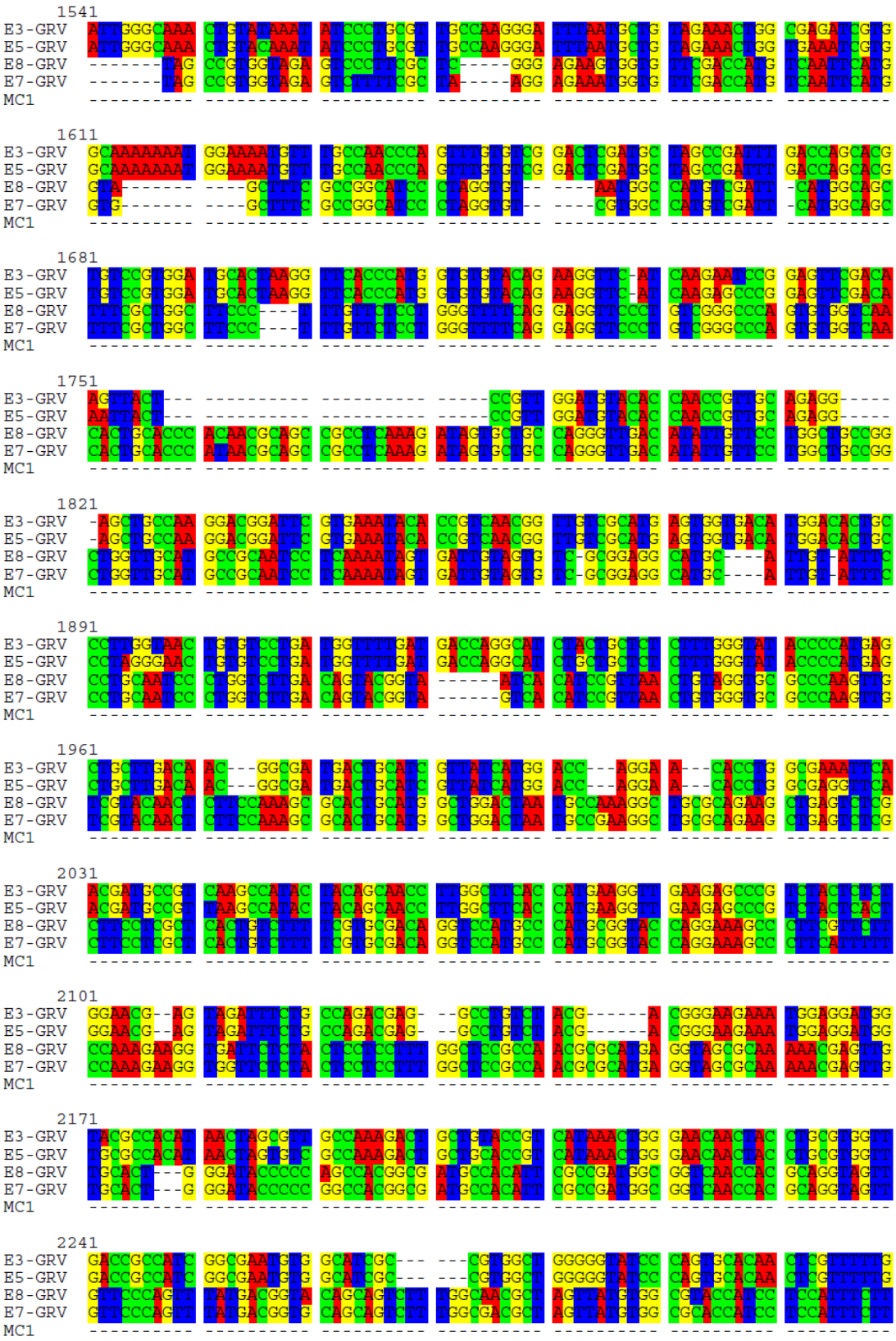


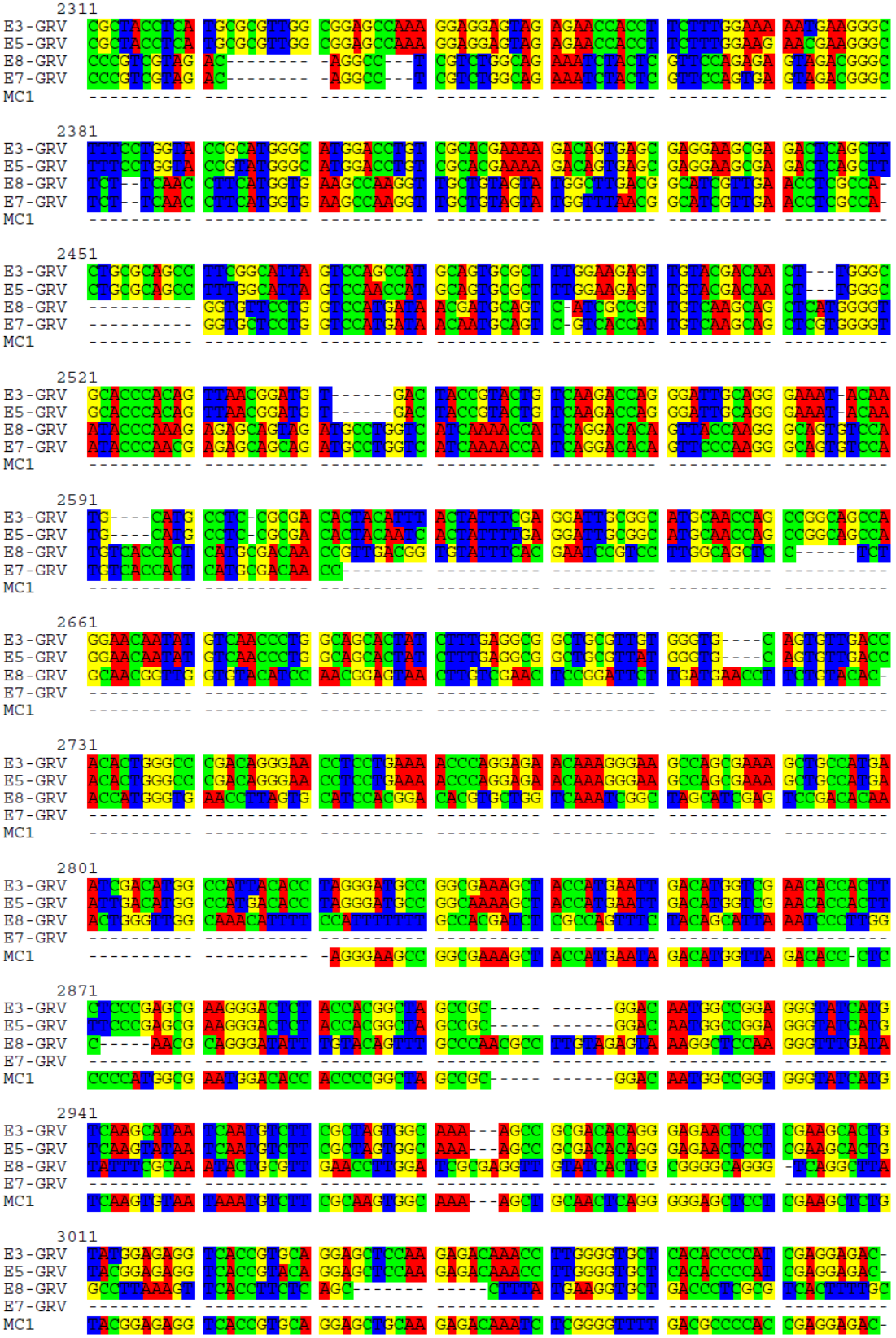

491
KG8-1
EG16-5
BG3-18
BUG1-21
E7
E8
M24S
M16S
M11S
561
631
701
771
841

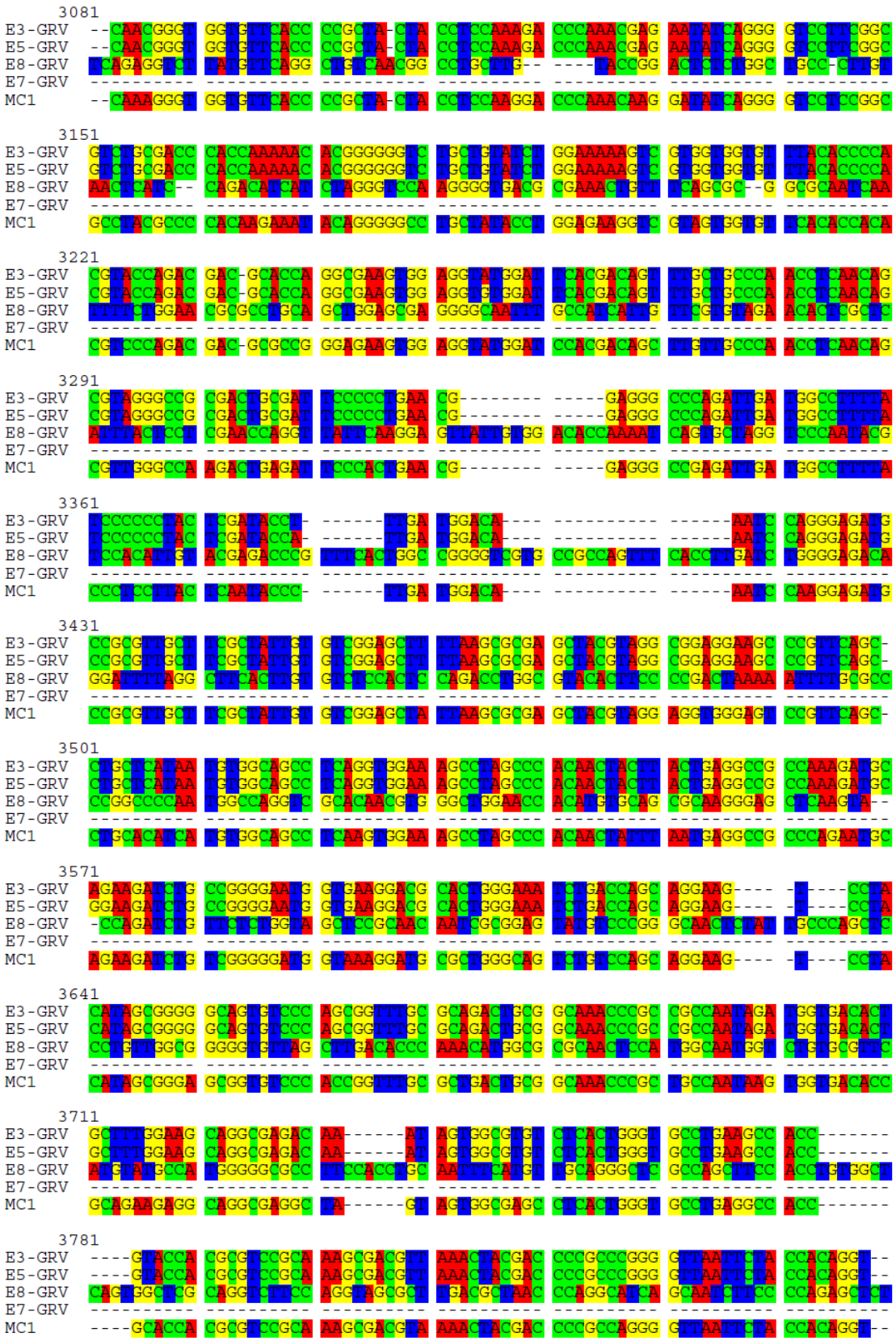
Appendix III: Aligned GRV sequences

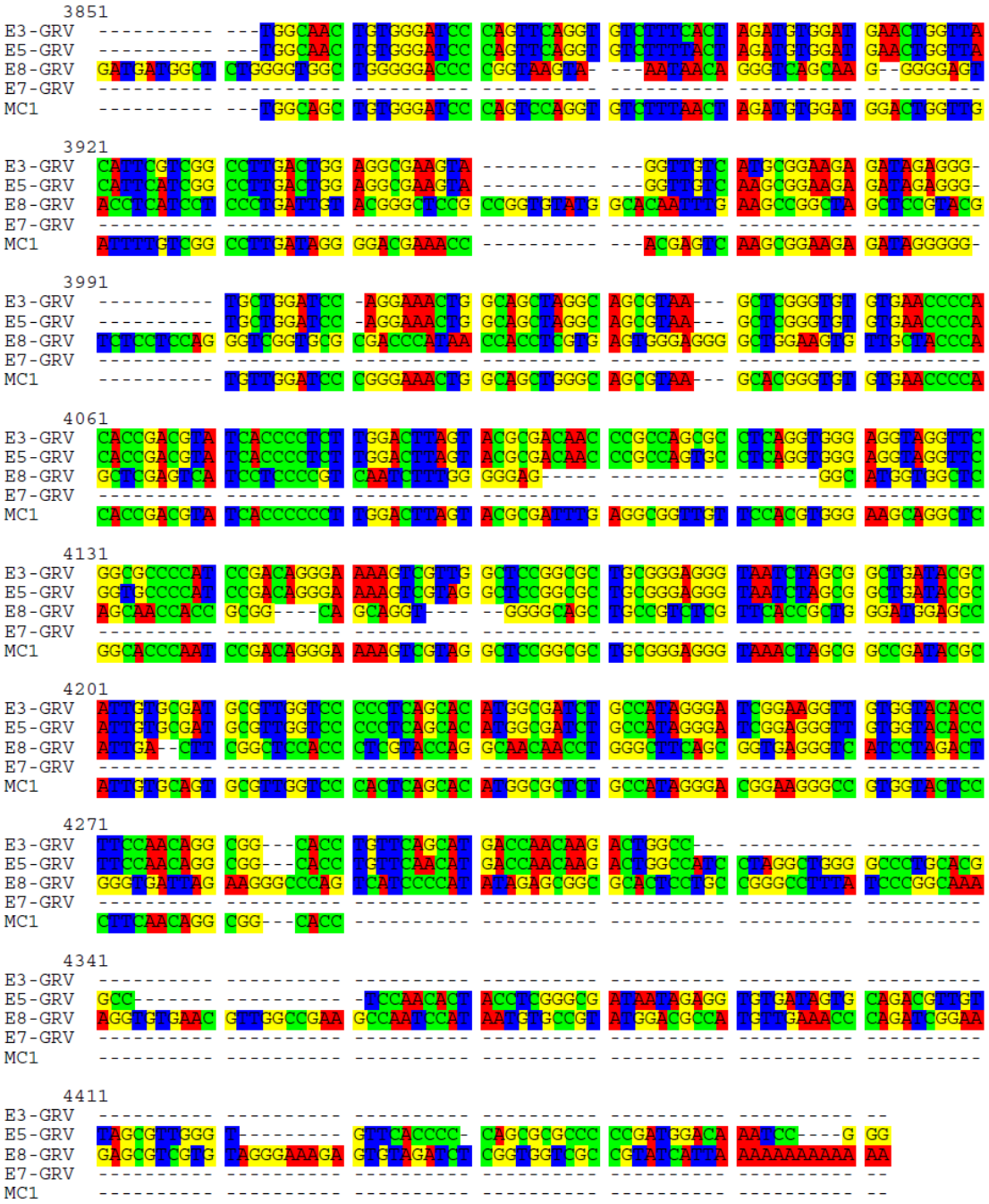



771
841
911
981
1051
1121
1191
1261
1331
1401
1471
E3-GRV
E5-GRV
E8-GRV
E7-GRV
MC1

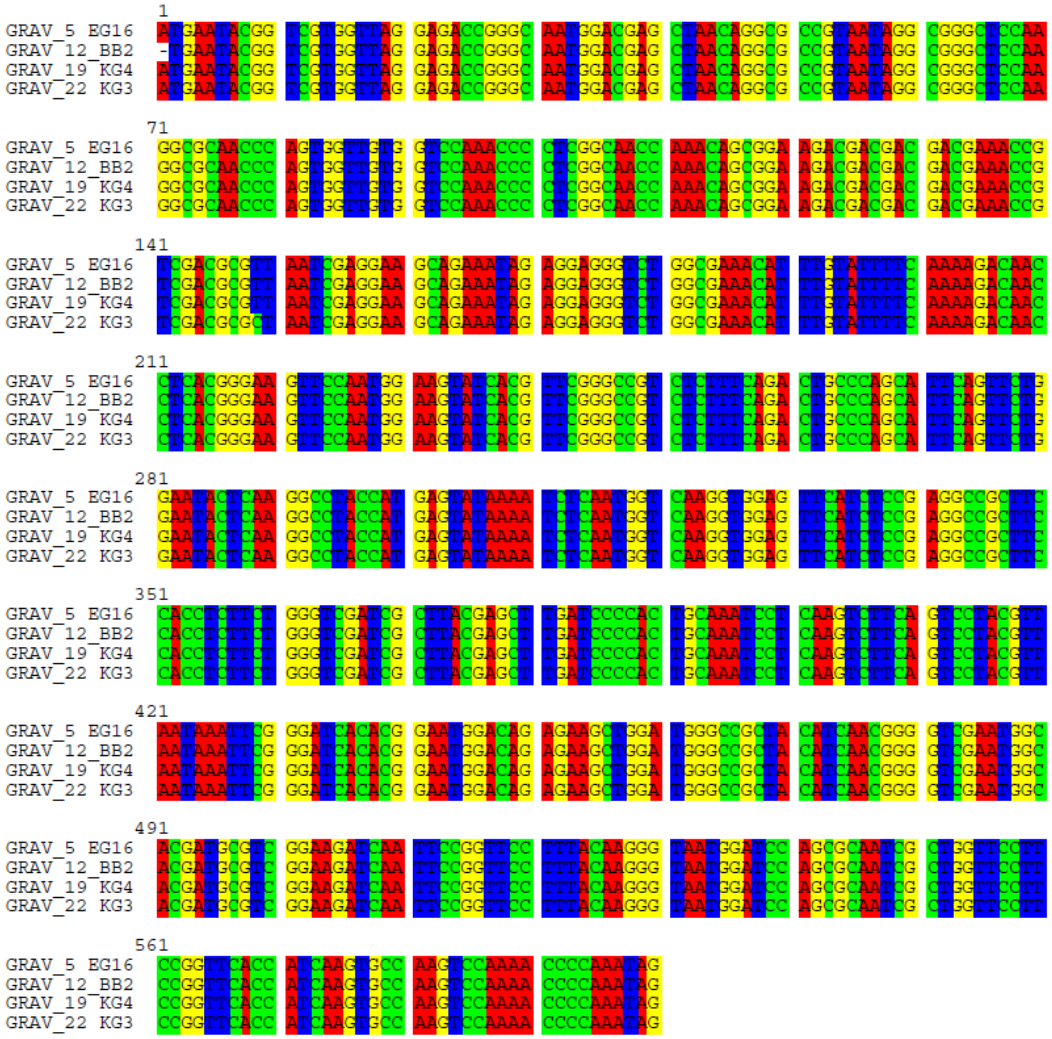

1541
E3-GRV
E5-GRV
E8-GRV
E7-GRV
MC1
1611
1681
1751
1821
1891
1961
2031
2101
2171
2241


2311
E3-GRV
E5-GRV
E8-GRV
E7-GRV
MC1
2381
2451
2521
2591
2661
2731
2801
2871
2941
3011

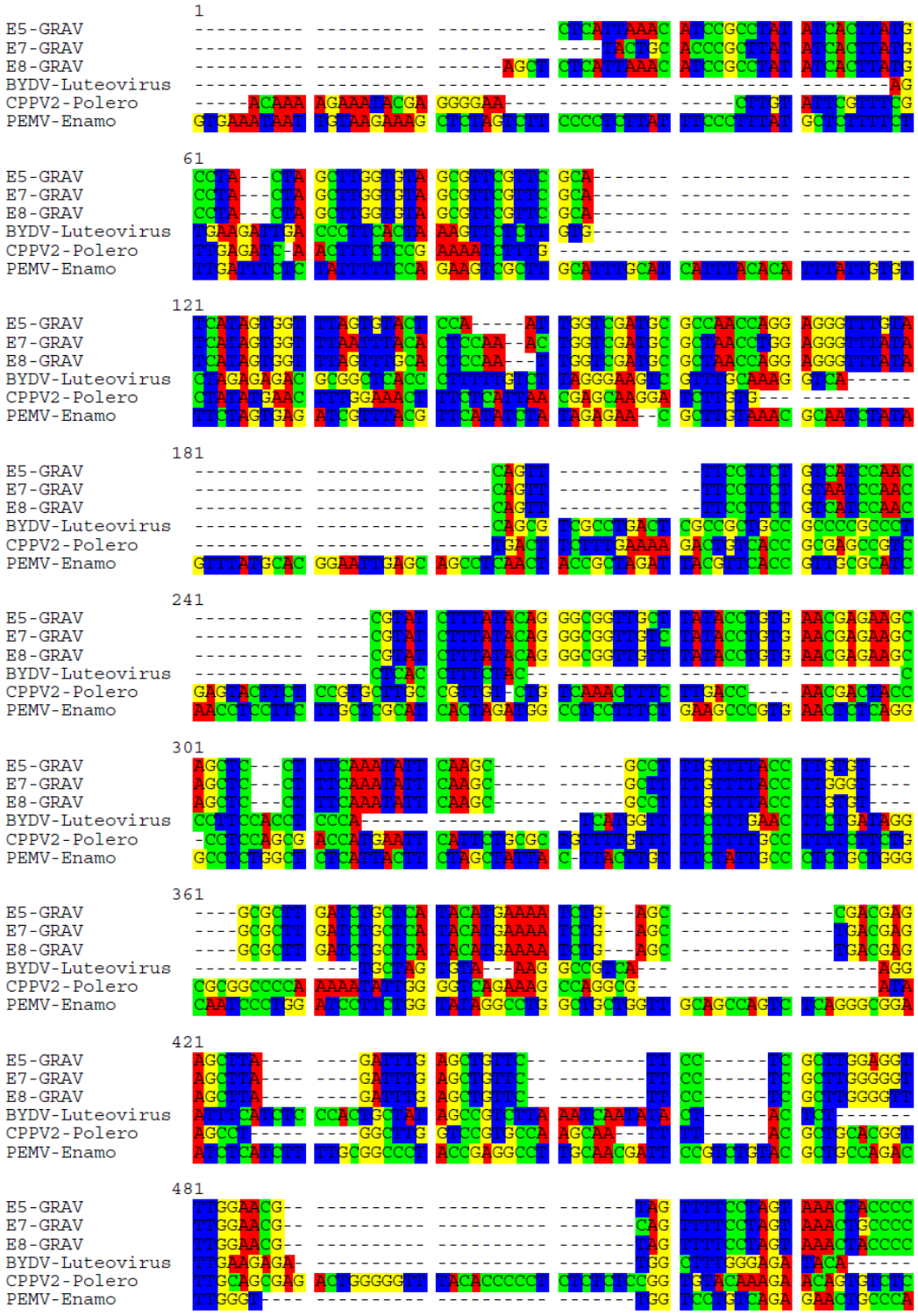

3081
E3-GRV
E5-GRV
E8-GRV
E7-GRV
MC1
3151
3221
3291
3361
3431
3501
3571
3641
3711
3781

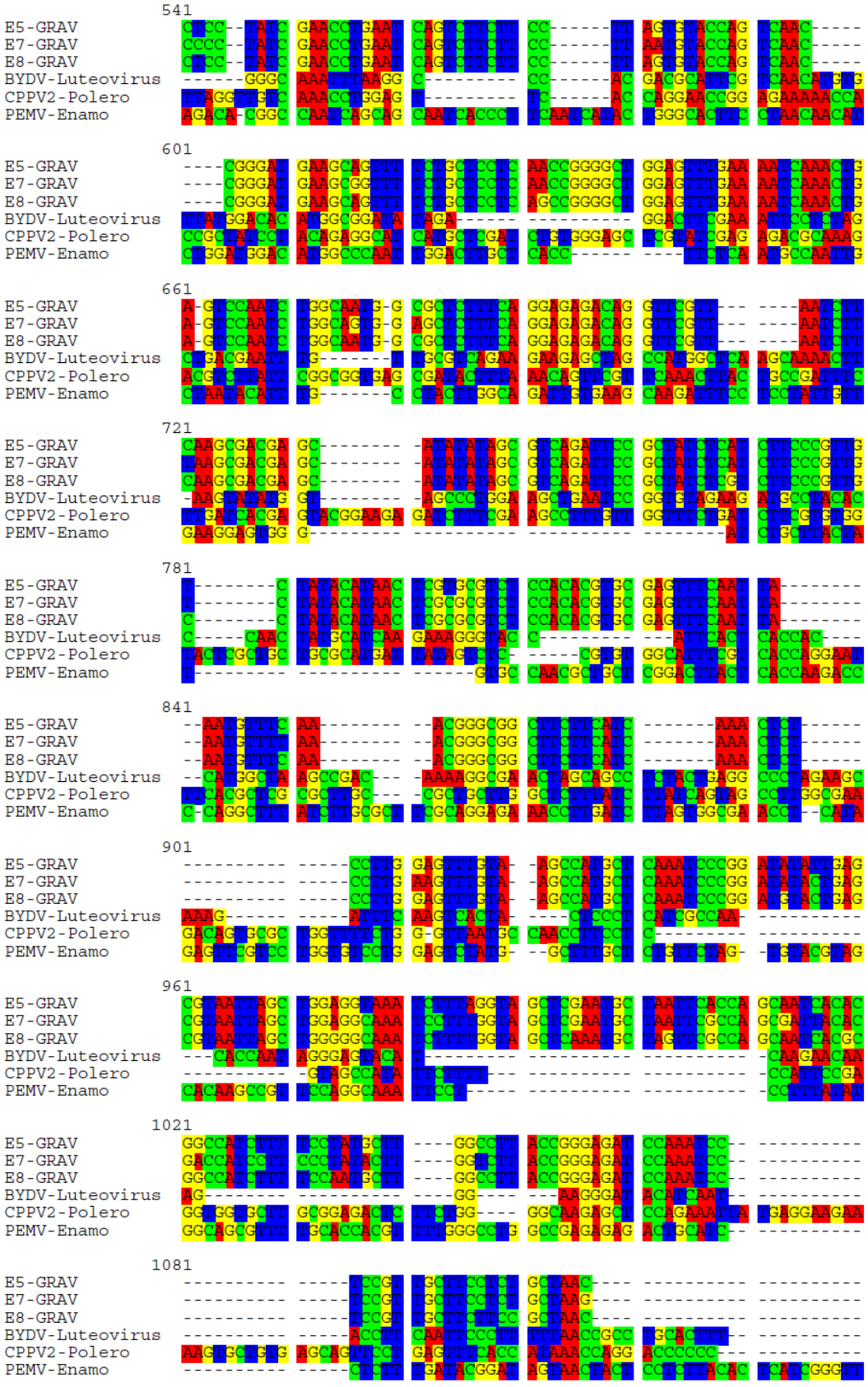


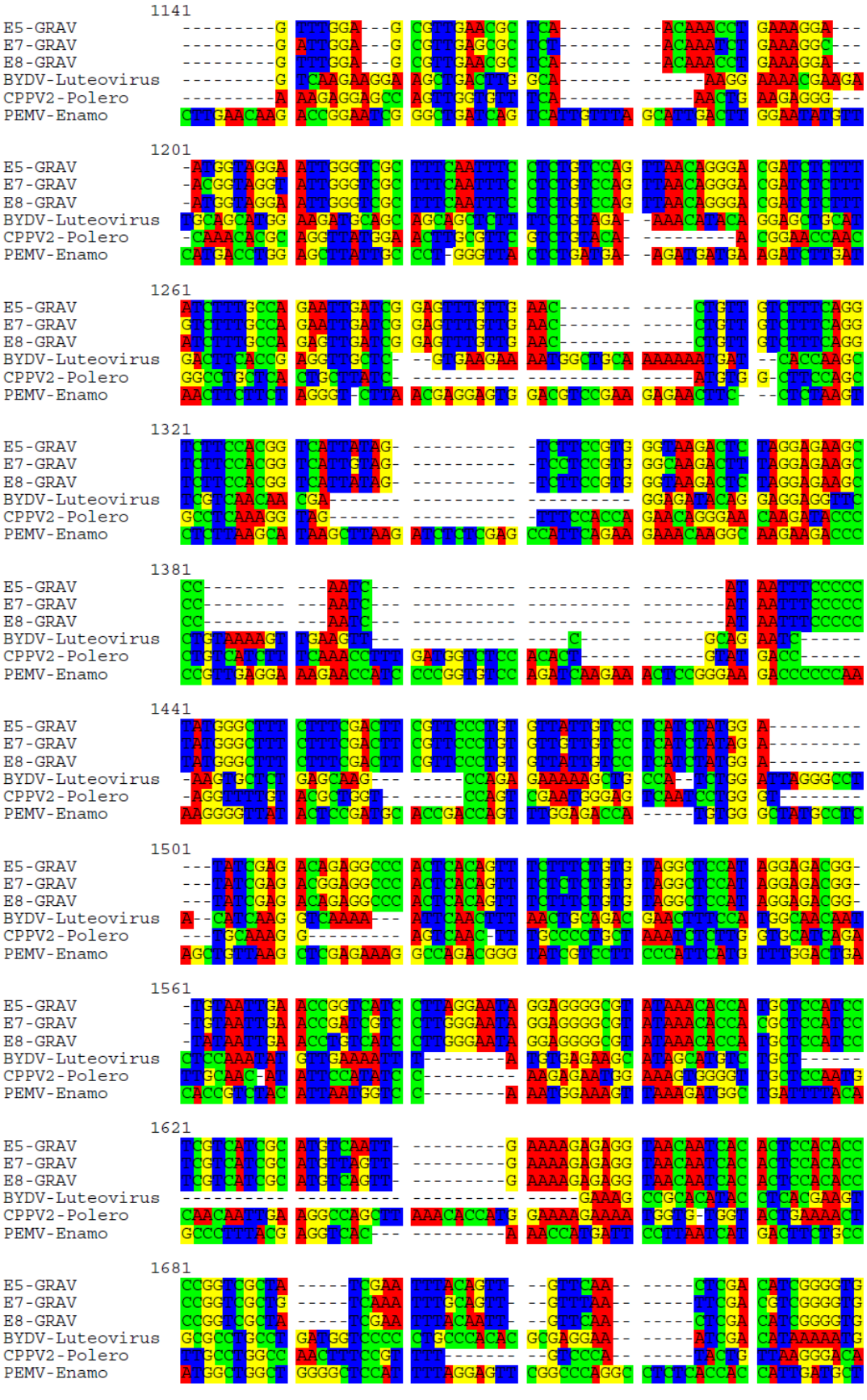
Appendix IV: Aligned GRAV-CP sequences

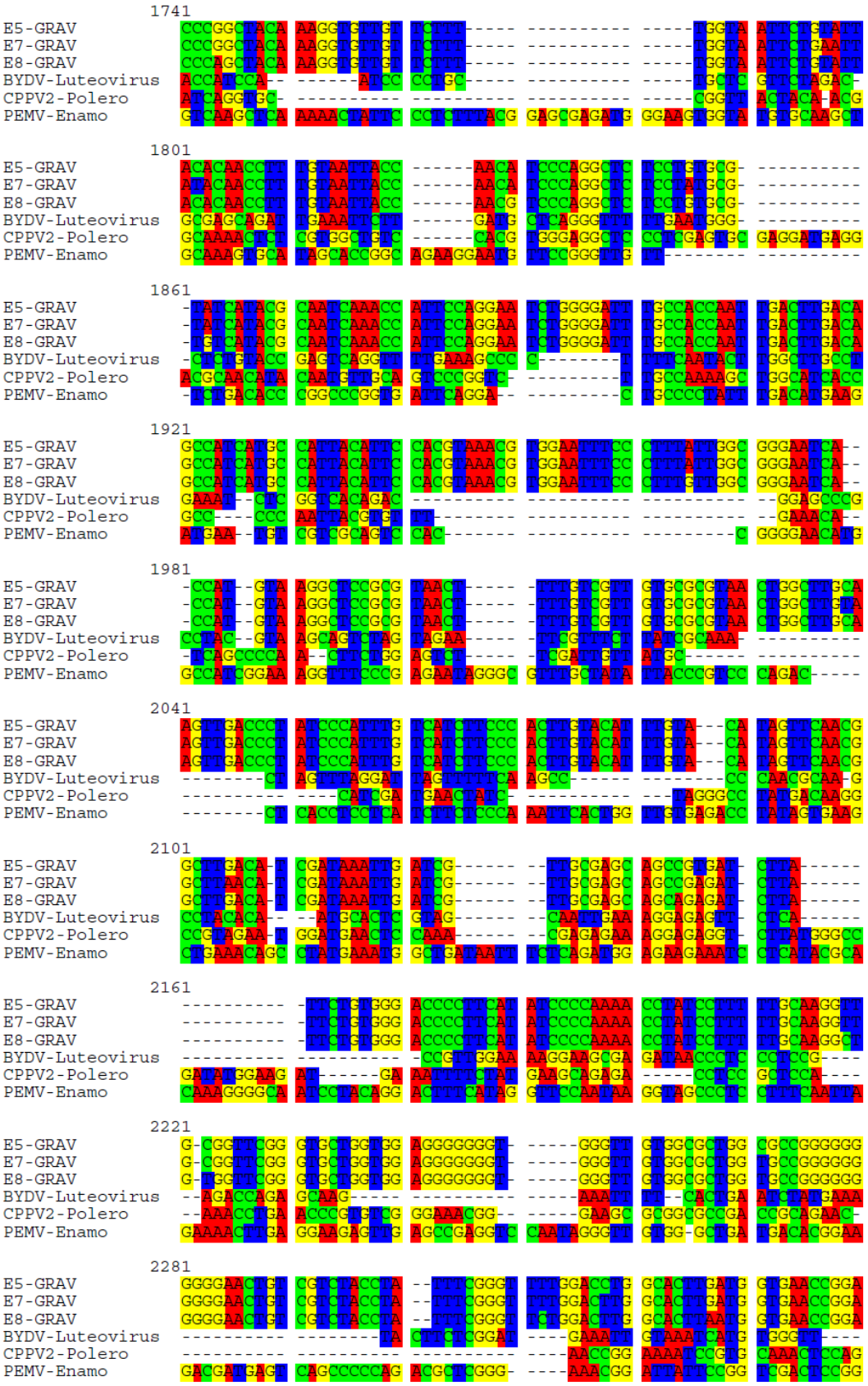


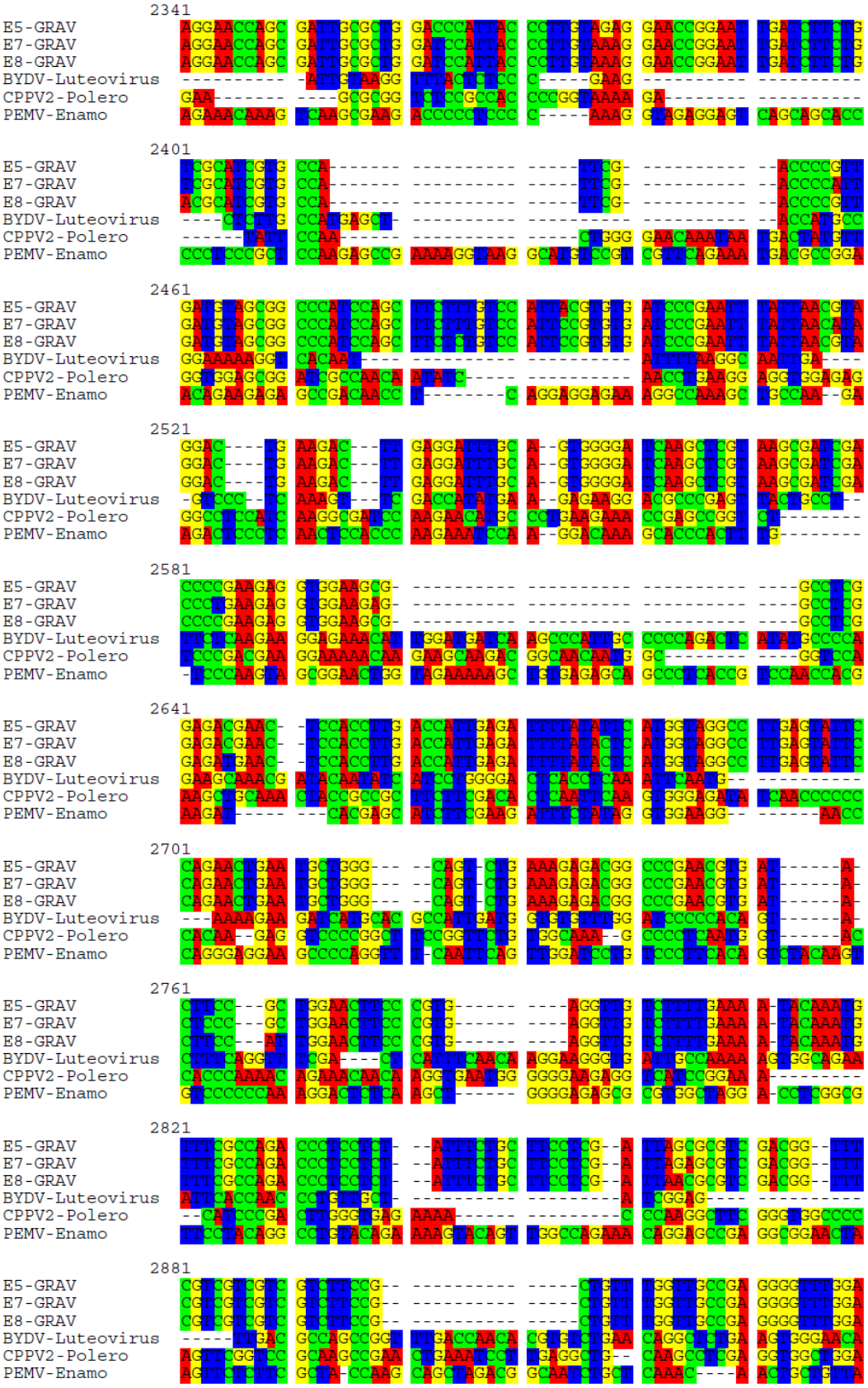
Appendix V: GRAV sequences aligned with assigned representatives from the Luteoviridae family.

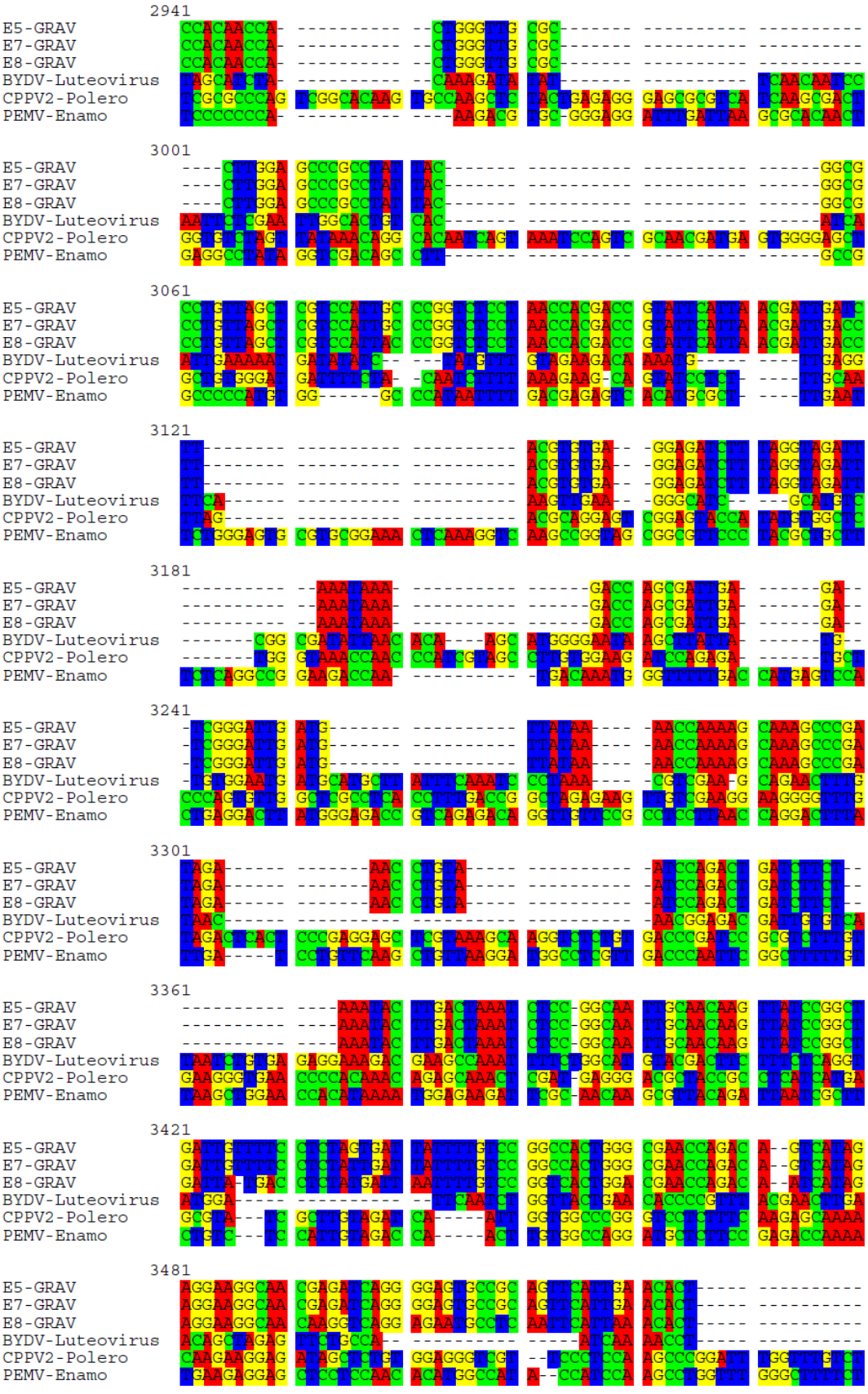


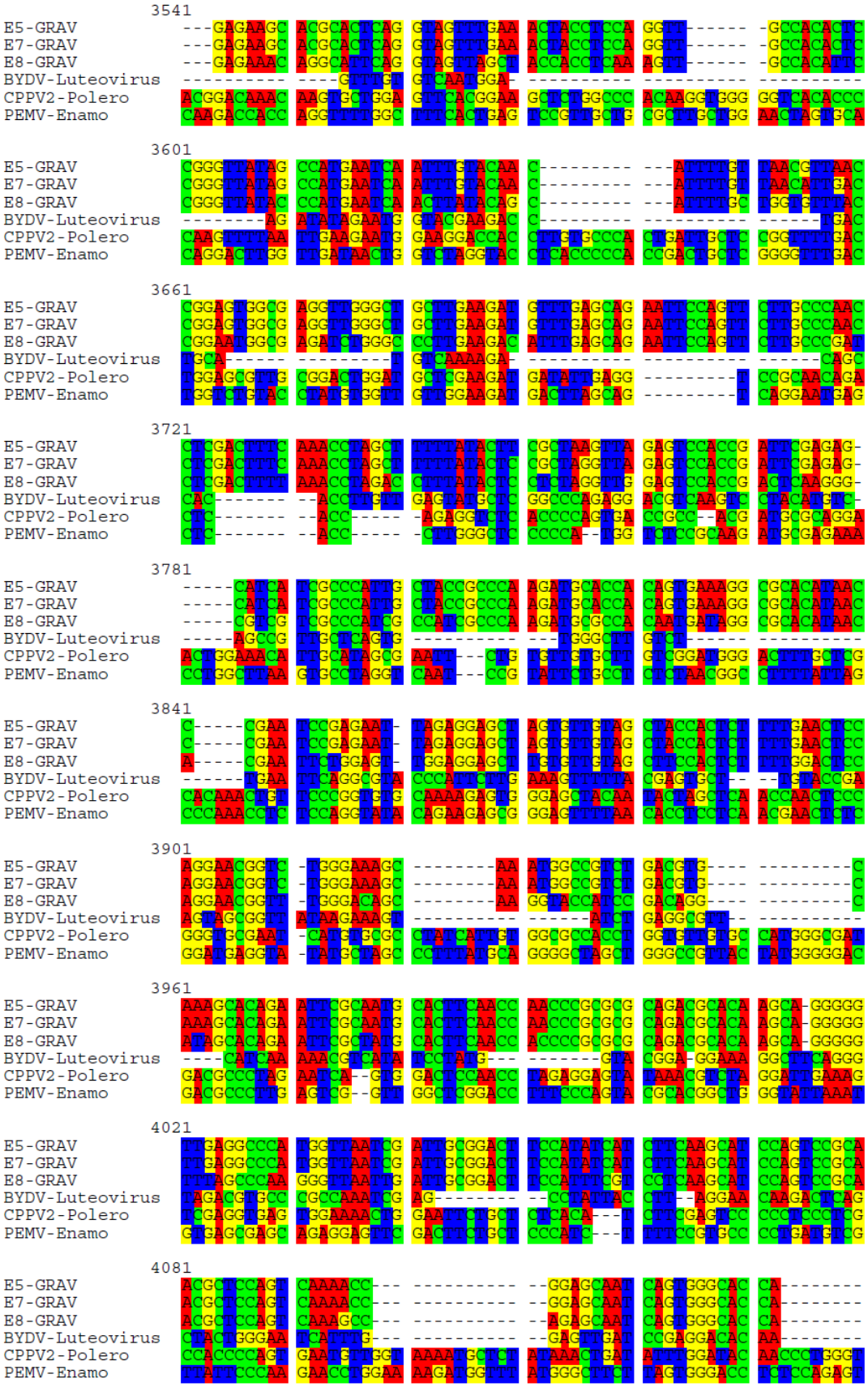

541
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
601
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
661
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
721
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
781
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
841
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
901
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
961
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
1021
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
1081
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo

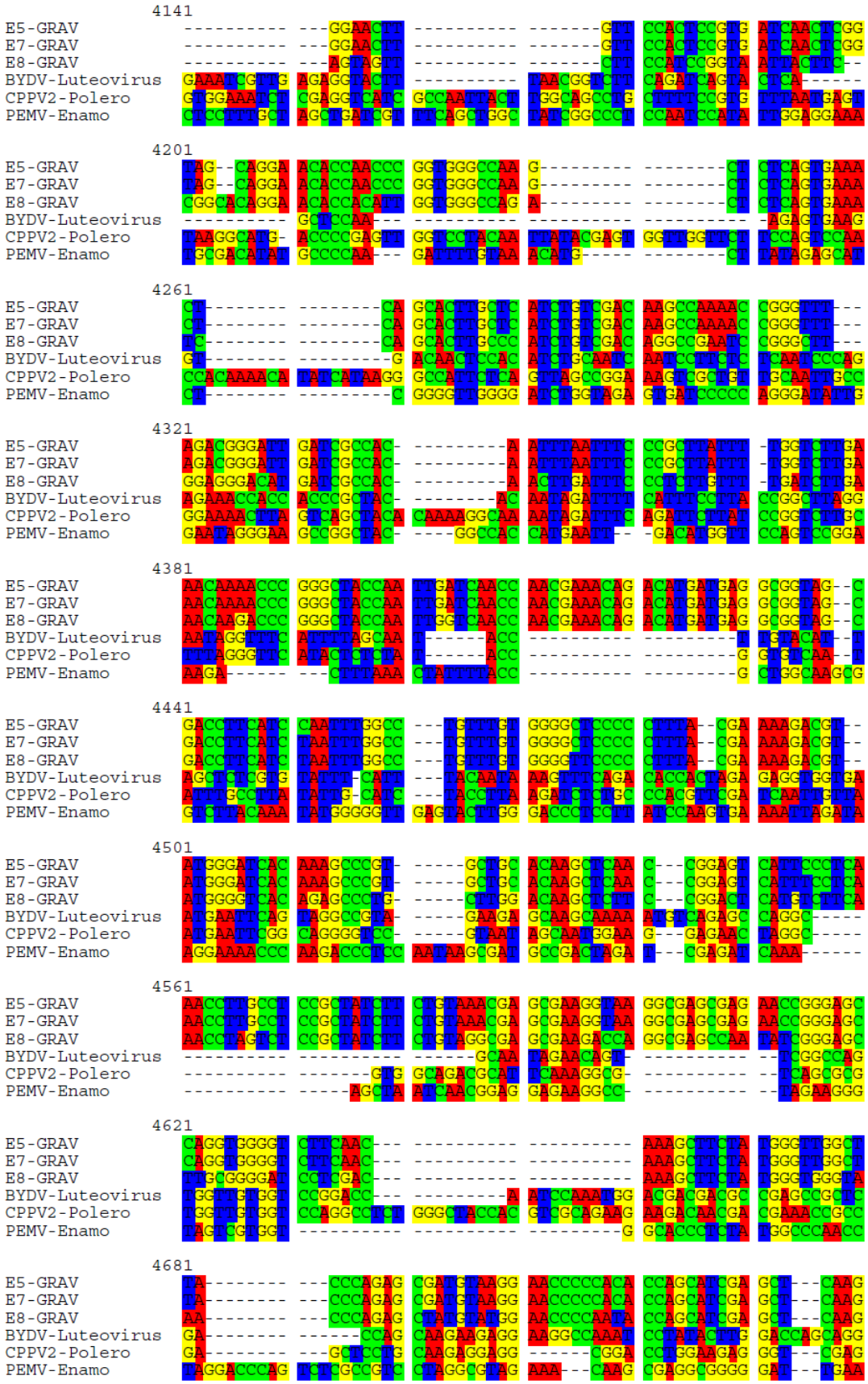

1141
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
1201
1261
1321
1381
1441
1501
1561
1621
1681

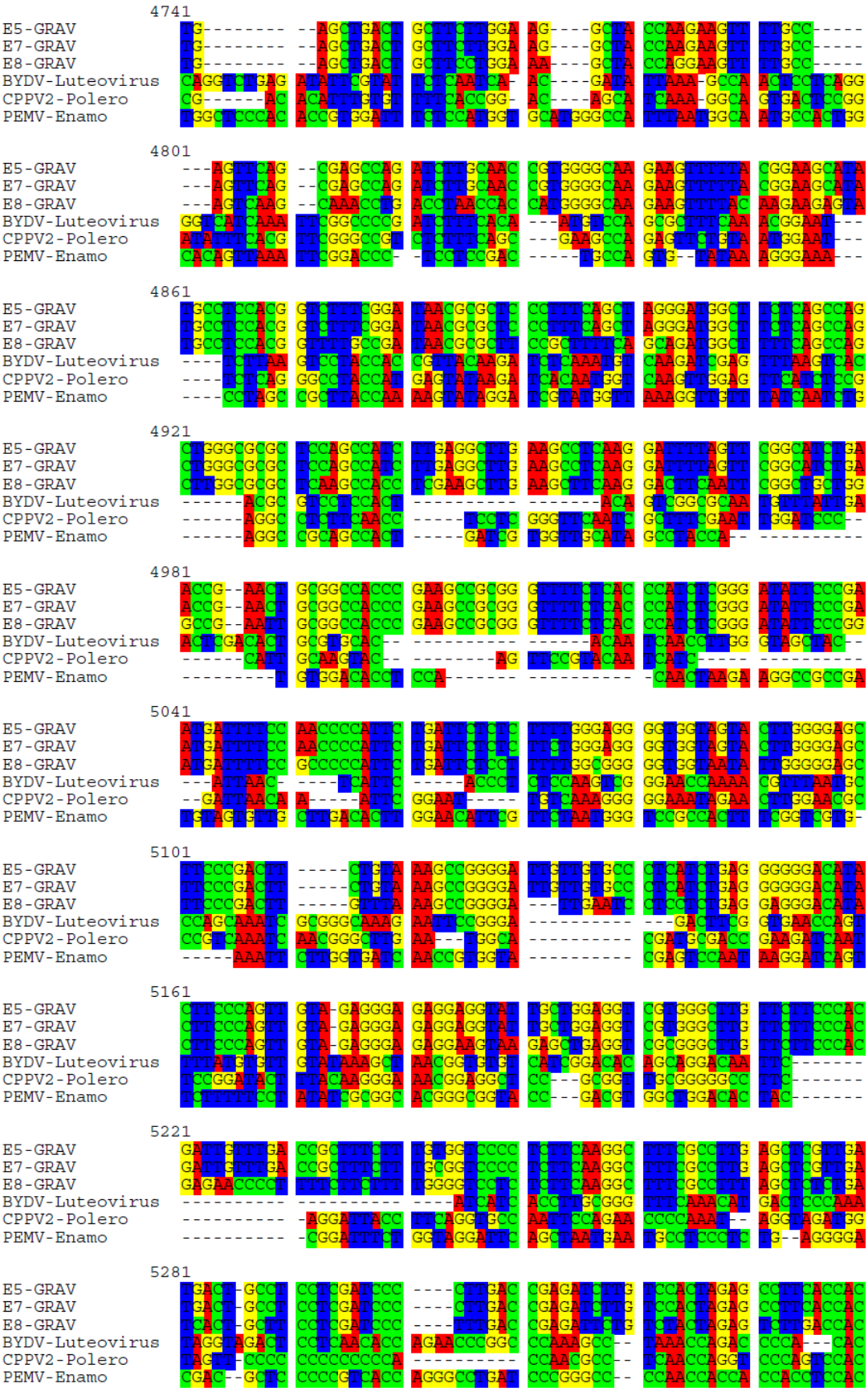

1741
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
1801
1861
1921
1981
2041
2101
2161
2221
2281

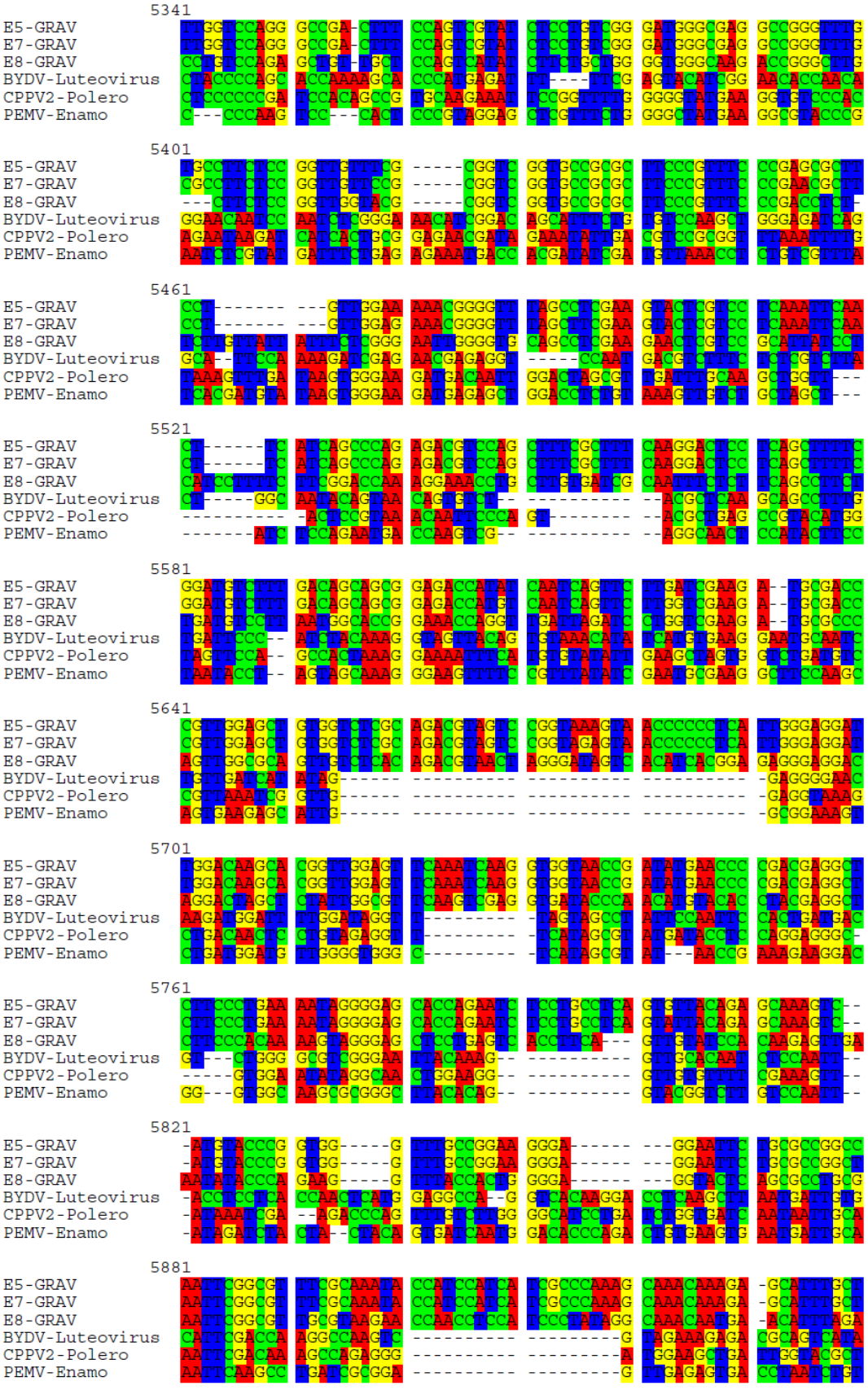

2341
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
2401
2461
2521
2581
2641
2701
2761
2821
2881

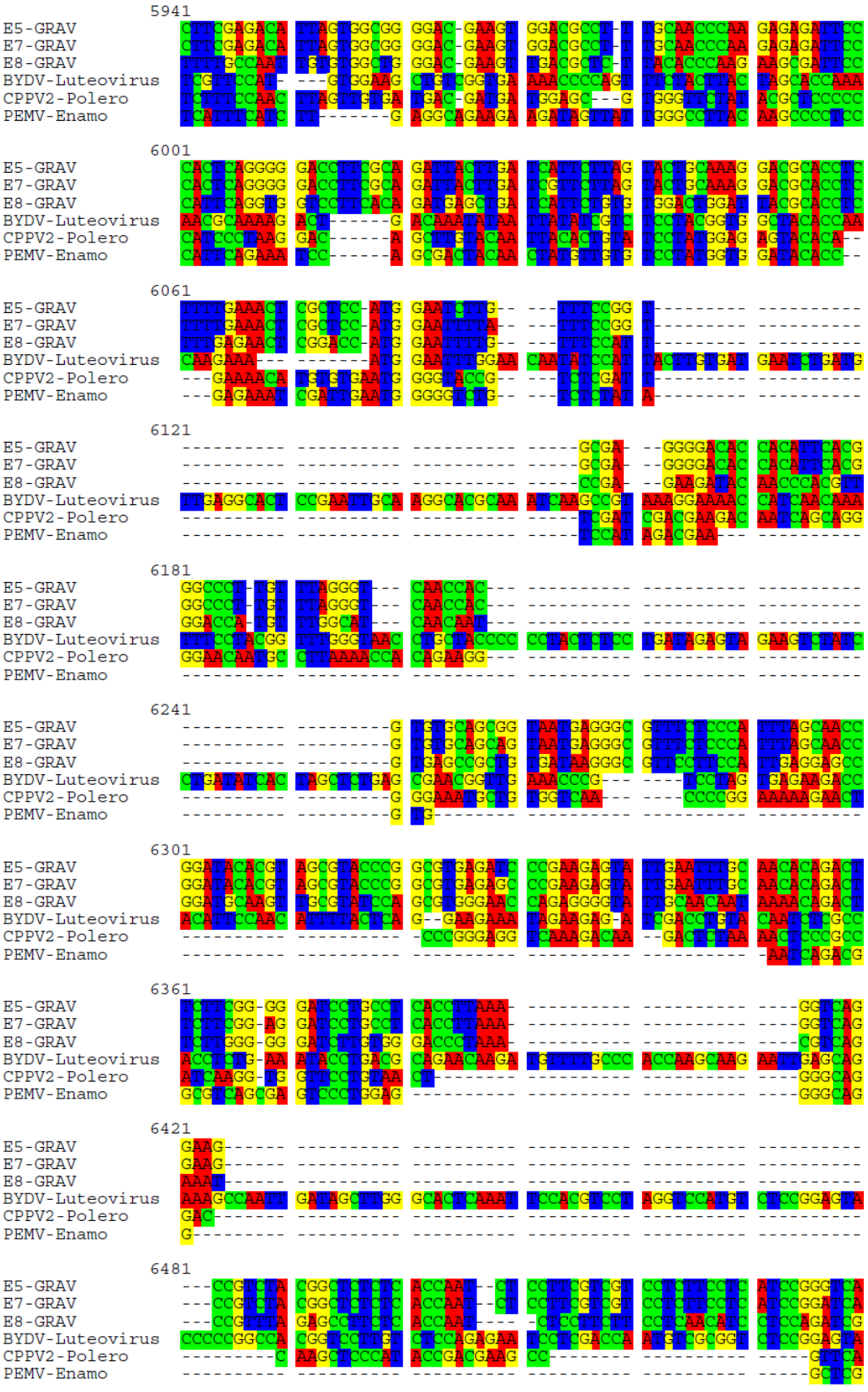

2941
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
3001
3061
3121
3181
3241
3301
3361
3421
3481

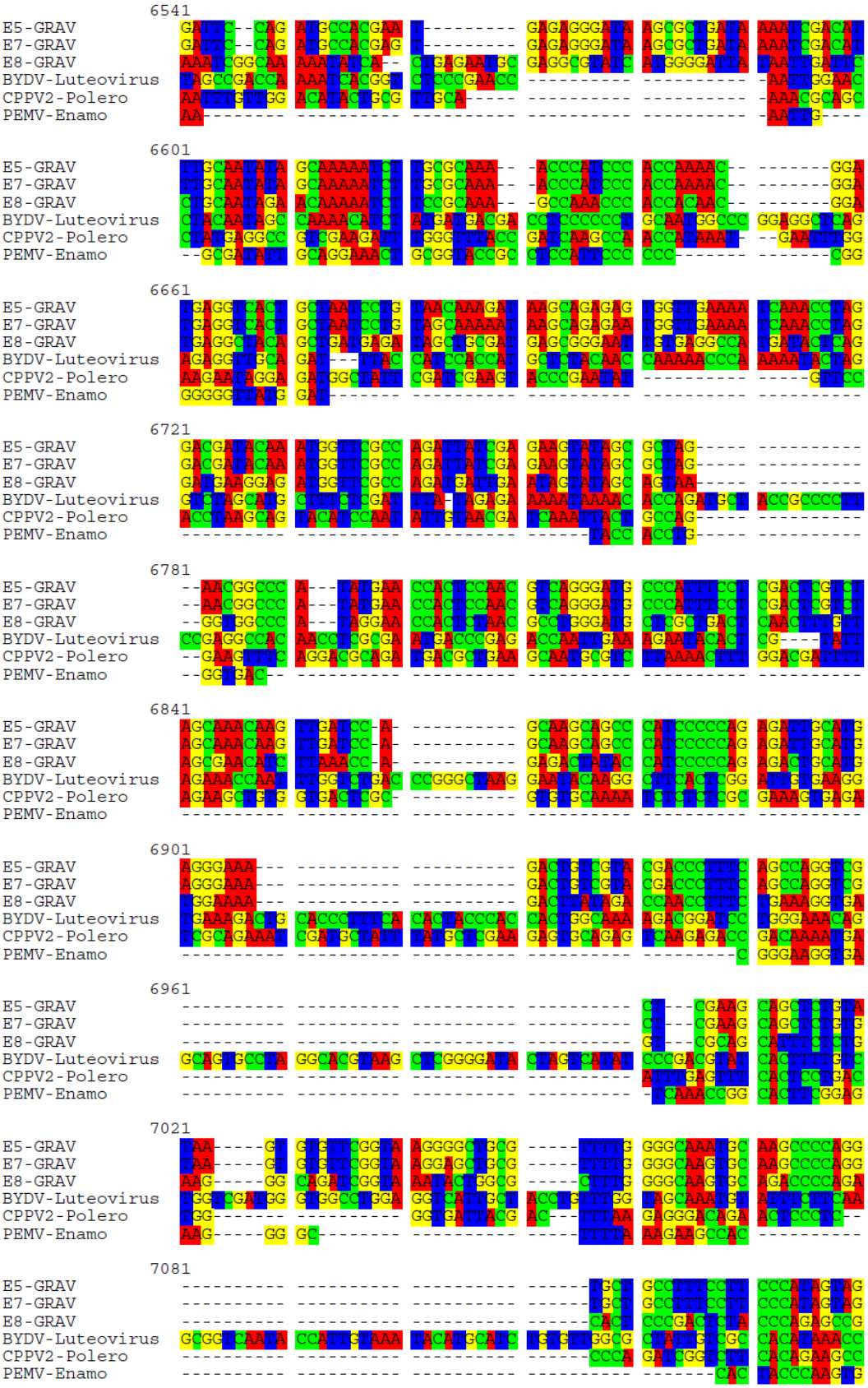

3541
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
3601
3661
3721
3781
3841
3901
3961
4021
4081

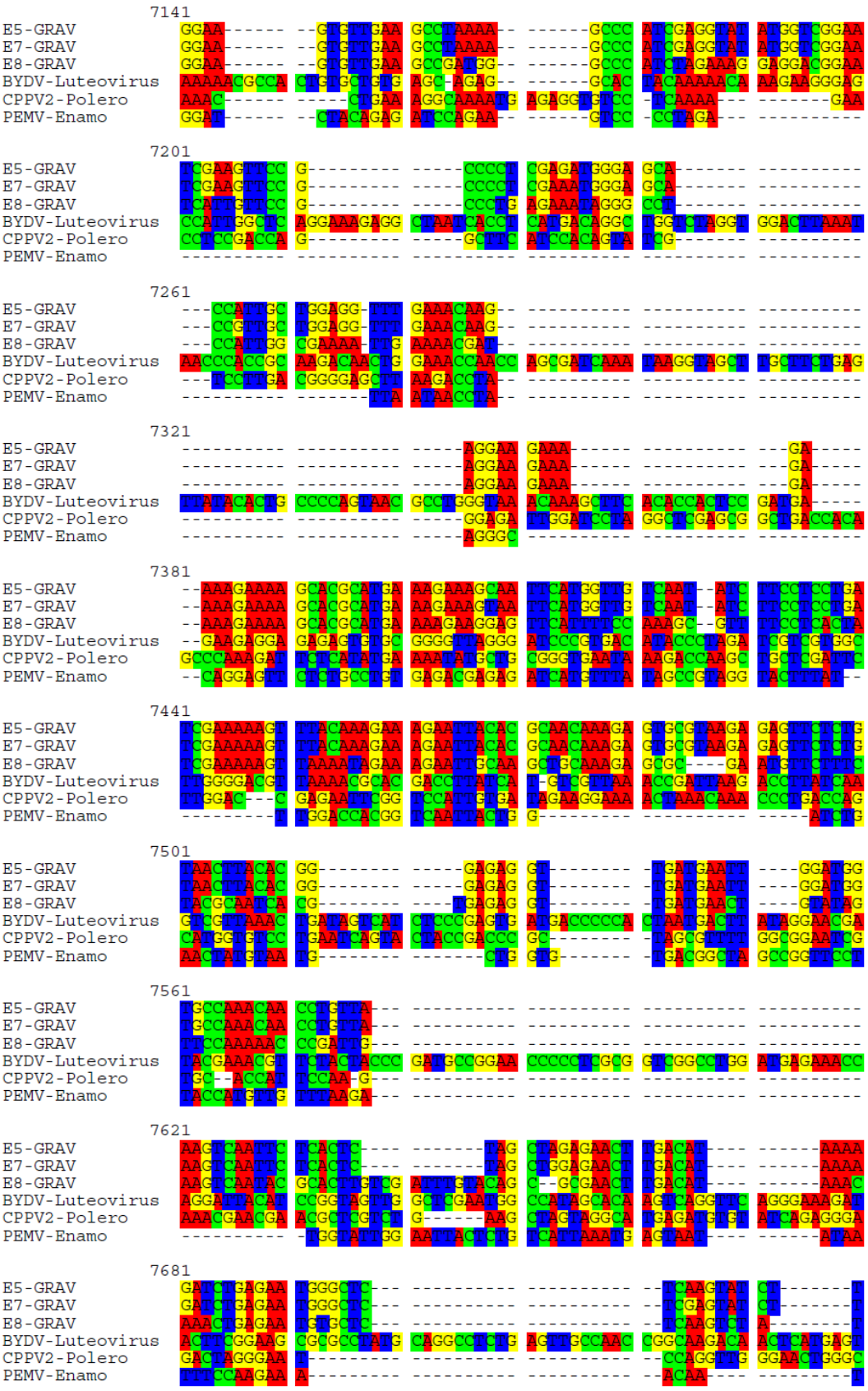

4141
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
4201
4261
4321
4381
4441
4501
4561
4621
4681

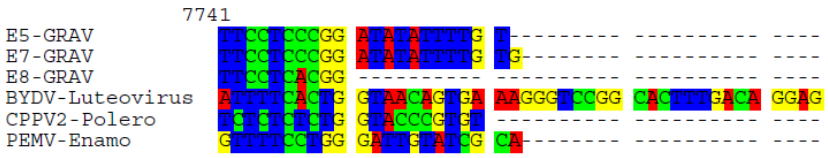

4741
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
4801
4861
4921
4981
5041
5101
5161
5221
5281


5341
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
5401
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
5461
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
5521
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
5581
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
5641
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
5701
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
5761
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
5821
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
5881
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo


5941
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
6001
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
6061
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
6121
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
6181
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
6241
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
6301
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
6361
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
6421
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
6481
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo


6541
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
6601
6661
6721
6781
6841
6901
6961
7021
7081


7141
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo
7201
7261
7321
7381
7441
7501
7561
7621
7681


7741
E5-GRAV
E7-GRAV
E8-GRAV
BYDV-Luteovirus
CPPV2-Polero
PEMV-Enamo

Appendix VI: RNA quantities

Sample	Nucleic Acid	Unit	260/280	260/230	Sample Type
1	711.3	ng/µl	2.11	2.23	RNA
2	944.8	ng/µl	2.12	2.18	RNA
3	526.6	ng/µl	2.05	1.95	RNA
4	442.3	ng/µl	1.84	1.17	RNA
5	550.9	ng/µl	2.02	1.85	RNA
6	243.6	ng/µl	1.99	1.37	RNA
7	470	ng/µl	2.05	2.14	RNA
8	221.8	ng/µl	1.78	1.03	RNA
9	343.9	ng/µl	1.92	0.75	RNA
9	333.2	ng/µl	1.94	0.74	RNA
10	1010.2	ng/µl	2.14	2.14	RNA
11	2512	ng/µl	2.09	2.14	RNA
12	1015.2	ng/µl	2.04	1.69	RNA
13	5921	ng/µl	2.1	2.12	RNA
14	4322.1	ng/µl	2.12	2.06	RNA
15	1114.6	ng/µl	2.04	1.83	RNA
17	6364.1	ng/µl	2.03	2.1	RNA
18	14059.7	ng/µl	2	2.04	RNA
19	2542.9	ng/µl	2.1	2.12	RNA

Appendix VII: Locations of the surveyed groundnut fields

COUNTY	ALTITUDE	LONGTUDE	LATITUDE	SEASON
Busia	1296	E034.31286	N00.31286	Long rain
Busia	1296	E034.31286	N00.31286	Long rain
Busia	1296	E034.31286	N00.31286	Long rain
Busia	1257	E034.23760	N00.29552	Long rain
Busia	1257	E034.23760	N00.29552	Long rain
Busia	1257	E034.23760	N00.29552	Long rain
Busia	1257	E034.23760	N00.29552	Long rain
Busia	1264	E034.23464	N00.29288	Long rain
Busia	1264	E034.23464	N00.29288	Long rain
Busia	1264	E034.23464	N00.29288	Long rain
Busia	1264	E034.23464	N00.29288	Long rain
Busia	1310	E034.28268	N00.31942	Long rain
Busia	1310	E034.28268	N00.31942	Long rain
Busia	1310	E034.28268	N00.31942	Long rain
Busia	1284	E034.32163	N00.32496	Long rain
Busia	1284	E034.32163	N00.32496	Long rain
Busia	1284	E034.32163	N00.32496	Long rain
Busia	1280	E034.32230	N00.32128	Long rain
Busia	1280	E034.32230	N00.32128	Long rain
Busia	1280	E034.32230	N00.32128	Long rain
Busia	1185	E034.19242	N00.40588	Long rain
Busia	1193	E034.20306	N00.41242	Long rain
Busia	1193	E034.20306	N00.41242	Long rain
Busia	1193	E034.20306	N00.41242	Long rain
Busia	1193	E034.20306	N00.41242	Long rain
Busia	1193	E034.20306	N00.41242	Long rain
Busia	1191	E034.20351	N00.41551	Long rain
Busia	1191	E034.20351	N00.41551	Long rain
Busia	1191	E034.20351	N00.41551	Long rain
Busia	1191	E034.20351	N00.41551	Long rain
Busia	1191	E034.20351	N00.41551	Long rain
Busia	1191	E034.20351	N00.41551	Long rain
Busia	1193	E034.20446	N00.41642	Long rain
Busia	1193	E034.20446	N00.41642	Long rain
Busia	1193	E034.20446	N00.41642	Long rain
Busia	1193	E034.20446	N00.41642	Long rain
Busia	1193	E034.20446	N00.41642	Long rain
Busia	1172	E034.13031	N00.41994	Long rain
Busia	1179	E034.10503	N00.41319	Long rain
Busia	1179	E034.10503	N00.41319	Long rain
Busia	1179	E034.10503	N00.41319	Long rain
Busia	1175	E034.10410	N00.40845	Long rain

Busia	1175	E034.10410	N00.40845	Long rain
Busia	1175	E034.10410	N00.40845	Long rain
Busia	1175	E034.10410	N00.40845	Long rain
Busia	1175	E034.10410	N00.40845	Long rain
Busia	1200	E034.11015	N00.43571	Long rain
Busia	1200	E034.11015	N00.43571	Long rain
Busia	1200	E034.11015	N00.43571	Long rain
Siaya	1340	E034.35617	N00.07417	Long rain
Siaya	1340	E034.35617	N00.07417	Long rain
Siaya	1340	E034.35617	N00.07417	Long rain
Siaya	1340	E034.35617	N00.07417	Long rain
Siaya	1367	E034.34445	N00.05377	Long rain
Siaya	1367	E034.34445	N00.05377	Long rain
Siaya	1367	E034.34445	N00.05377	Long rain
Siaya	1290	E034.43913	S00.04043	Long rain
Siaya	1290	E034.43913	S00.04043	Long rain
Busia	1164	E034.31875	S00.23808	Long rain
Busia	1164	E034.31875	S00.23808	Long rain
Busia	1164	E034.31875	S00.23808	Long rain
Siaya	1176	E034.32233	S00.23762	Long rain
Siaya	1176	E034.32233	S00.23762	Long rain
Siaya	1176	E034.32233	S00.23762	Long rain
Siaya	1191	E034.32590	S00.23710	Long rain
Siaya	1191	E034.32590	S00.23710	Long rain
Siaya	1191	E034.32590	S00.23710	Long rain
Siaya	1182	E034.32560	S00.23590	Long rain
Siaya	1182	E034.32560	S00.23590	Long rain
Siaya	1182	E034.32560	S00.23590	Long rain
Siaya	1291	E034.27107	N00.10636	Long rain
Siaya	1291	E034.27107	N00.10636	Long rain
Siaya	1291	E034.27107	N00.10636	Long rain
Siaya	1291	E034.27107	N00.10636	Long rain
Siaya	1282	E034.28773	N00.10851	Long rain
Siaya	1282	E034.28773	N00.10851	Long rain
Siaya	1282	E034.28773	N00.10851	Long rain
Siaya	1282	E034.28773	N00.10851	Long rain
Siaya	1167	E034.18727	S00.03375	Long rain
Siaya	1167	E034.18727	S00.03375	Long rain
Bungoma	1476	E034.54118	N00.61600	Long rain
Bungoma	1476	E034.54118	N00.61600	Long rain
Bungoma	1476	E034.54118	N00.61600	Long rain
Bungoma	1465	E034.54144	N00.61531	Long rain
Bungoma	1465	E034.54144	N00.61531	Long rain
Bungoma	1465	E034.54144	N00.61531	Long rain

Bungoma	1465	E034.54144	N00.61531	Long rain
Bungoma	1404	E034.46411	N00.69718	Long rain
Bungoma	1404	E034.46411	N00.69718	Long rain
Bungoma	1404	E034.46411	N00.69718	Long rain
Bungoma	1404	E034.46411	N00.69718	Long rain
Bungoma	1350	E034.44774	N00.68771	Long rain
Bungoma	1350	E034.44774	N00.68771	Long rain
Bungoma	1350	E034.44774	N00.68771	Long rain
Bungoma	1350	E034.44774	N00.68771	Long rain
Busia	1261	E034.34343	N00.64906	Long rain
Busia	1261	E034.34343	N00.64906	Long rain
Busia	1261	E034.34343	N00.64906	Long rain
Busia	1299	E034.33460	N00.66420	Long rain
Busia	1299	E034.33460	N00.66420	Long rain
Busia	1446	E034.36942	N00.69250	Long rain
Busia	1446	E034.36942	N00.69250	Long rain
Busia	1446	E034.36942	N00.69250	Long rain
Busia	1446	E034.36942	N00.69250	Long rain
Busia	1441	E034.36904	N00.69389	Long rain
Busia	1441	E034.36904	N00.69389	Long rain
Busia	1441	E034.36904	N00.69389	Long rain
Busia	1441	E034.36904	N00.69389	Long rain
Busia	1447	E034.39956	N00.67659	Long rain
Busia	1447	E034.39956	N00.67659	Long rain
Busia	1447	E034.39956	N00.67659	Long rain
Busia	1447	E034.39956	N00.67659	Long rain
Busia	1446	E034.39912	N00.67713	Long rain
Busia	1446	E034.39912	N00.67713	Long rain
Busia	1446	E034.39912	N00.67713	Long rain
Busia	1446	E034.39912	N00.67713	Long rain
Busia	1416	E034.38803	N00.74552	Long rain
Busia	1416	E034.38803	N00.74552	Long rain
Busia	1416	E034.38803	N00.74552	Long rain
Busia	1416	E034.38803	N00.74552	Long rain
Bungoma	1535	E034.76328	N00.61117	Long rain
Bungoma	1535	E034.76328	N00.61117	Long rain
Bungoma	1535	E034.76328	N00.61117	Long rain
Bungoma	1591	E034.79065	N00.62522	Long rain
Bungoma	1591	E034.79065	N00.62522	Long rain
Bungoma	1591	E034.79065	N00.62522	Long rain
Bungoma	1591	E034.79065	N00.62522	Long rain
Bungoma	1591	E034.79065	N00.62522	Long rain
Bungoma	1560	E034.80379	N00.62807	Long rain
Bungoma	1560	E034.80379	N00.62807	Long rain

Bungoma	1560	E034.80379	N00.62807	Long rain
Bungoma	1560	E034.80379	N00.62807	Long rain
Bungoma	1560	E034.80379	N00.62807	Long rain
Bungoma	1628	E034.78019	N00.63291	Long rain
Bungoma	1628	E034.78019	N00.63291	Long rain
Bungoma	1628	E034.78019	N00.63291	Long rain
Bungoma	1628	E034.78019	N00.63291	Long rain
Bungoma	1275	E034.40377	N00.59442	Long rain
Bungoma	1275	E034.40377	N00.59442	Long rain
Bungoma	1275	E034.40377	N00.59442	Long rain
Bungoma	1275	E034.40377	N00.59442	Long rain
Bungoma	1275	E034.40377	N00.59442	Long rain
Bungoma	1344	E034.44003	N00.68440	Long rain
Bungoma	1344	E034.44003	N00.68440	Long rain
Bungoma	1344	E034.44003	N00.68440	Long rain
Bungoma	1374	E034.44807	N00.70103	Long rain
Bungoma	1374	E034.44807	N00.70103	Long rain
Bungoma	1397	E034.44944	N00.70175	Long rain
Bungoma	1397	E034.44944	N00.70175	Long rain
Bungoma	1397	E034.44944	N00.70175	Long rain
Kakamega	1520	E034.78662	N00.14787	Long rain
Kakamega	1520	E034.78662	N00.14787	Long rain
Kakamega	1530	E034.66257	N00.05551	Long rain
Kakamega	1530	E034.66257	N00.05551	Long rain
Kakamega	1530	E034.66257	N00.05551	Long rain
Kakamega	1519	E034.66122	N00.05523	Long rain
Kakamega	1519	E034.66122	N00.05523	Long rain
Kakamega	1519	E034.66122	N00.05523	Long rain
Kakamega	1513	E034.66036	N00.05441	Long rain
Kakamega	1513	E034.66036	N00.05441	Long rain
Kakamega	1513	E034.66036	N00.05441	Long rain
Kakamega	1513	E034.66036	N00.05441	Long rain
Kakamega	1558	E034.74823	N00.00325	Long rain
Kakamega	1558	E034.74823	N00.00325	Long rain
Kakamega	1558	E034.74823	N00.00325	Long rain
Kakamega	1558	E034.74823	N00.00325	Long rain
Kakamega	1558	E034.74823	N00.00325	Long rain
Kakamega	1600	E034.81551	N00.01565	Long rain
Kakamega	1600	E034.81551	N00.01565	Long rain
Kakamega	1600	E034.81551	N00.01565	Long rain
Kakamega	1600	E034.81551	N00.01565	Long rain
Kakamega	1600	E034.81551	N00.01565	Long rain
Kakamega	1589	E034.81600	N00.01505	Long rain
Kakamega	1589	E034.81600	N00.01505	Long rain

Kakamega	1589	E034.81600	N00.01505	Long rain
Kakamega	1589	E034.81600	N00.01505	Long rain
Kakamega	1589	E034.81600	N00.01505	Long rain
Kakamega	1684	E034.82533	N00.03115	Long rain
Kakamega	1684	E034.82533	N00.03115	Long rain
Kakamega	1684	E034.82533	N00.03115	Long rain
Homabay	1313	E034.57562	S0059917	Long rain
Homabay	1313	E034.57562	S0059917	Long rain
Homabay	1313	E034.57562	S0059917	Long rain
Homabay	1313	E034.57562	S0059917	Long rain
Homabay	1313	E034.57562	S0059917	Long rain
Homabay	1338	E034.58366	S00.60474	Long rain
Homabay	1338	E034.58366	S00.60474	Long rain
Homabay	1338	E034.58366	S00.60474	Long rain
Homabay	1339	E034.58286	S00.60896	Long rain
Homabay	1339	E034.58286	S00.60896	Long rain
Homabay	1339	E034.58286	S00.60896	Long rain
Homabay	1339	E034.58286	S00.60896	Long rain
Homabay	1343	E034.58385	S00.61199	Long rain
Homabay	1343	E034.58385	S00.61199	Long rain
Homabay	1343	E034.58385	S00.61199	Long rain
Homabay	1343	E034.58385	S00.61199	Long rain
Homabay	1329	E034.12975	S00.70017	Long rain
Homabay	1329	E034.12975	S00.70017	Long rain
Homabay	1329	E034.12975	S00.70017	Long rain
Homabay	1339	E034.12822	S00.70061	Long rain
Homabay	1339	E034.12822	S00.70061	Long rain
Homabay	1339	E034.12822	S00.70061	Long rain
Homabay	1356	E034.14912	S00.68648	Long rain
Homabay	1356	E034.14912	S00.68648	Long rain
Homabay	1325	E034.17109	S00.68328	Long rain
Homabay	1325	E034.17109	S00.68328	Long rain
Homabay	1325	E034.17109	S00.68328	Long rain
Homabay	1454	E034.63848	S00.62621	Long rain
Homabay	1454	E034.63848	S00.62621	Long rain
Homabay	1454	E034.63848	S00.62621	Long rain
Homabay	1454	E034.63848	S00.62621	Long rain
Homabay	1454	E034.63848	S00.62621	Long rain
Homabay	1465	E034.64264	S00.62588	Long rain
Homabay	1465	E034.64264	S00.62588	Long rain
Homabay	1465	E034.64264	S00.62588	Long rain
Homabay	1465	E034.64264	S00.62588	Long rain
Homabay	1473	E034.64319	S00.62610	Long rain
Homabay	1473	E034.64319	S00.62610	Long rain

Homabay	1473	E034.64319	S00.62610	Long rain
Homabay	1473	E034.64319	S00.62610	Long rain
Homabay	1468	E034.63749	S00.62680	Long rain
Homabay	1468	E034.63749	S00.62680	Long rain
Homabay	1468	E034.63749	S00.62680	Long rain
Homabay	1468	E034.63749	S00.62680	Long rain
Homabay	1449	E034.63797	S00.62541	Long rain
Homabay	1449	E034.63797	S00.62541	Long rain
Homabay	1449	E034.63797	S00.62541	Long rain
Homabay	1449	E034.63797	S00.62541	Long rain
Homabay	1449	E034.63797	S00.62541	Long rain
Homabay	1455	E034.63761	S00.62468	Long rain
Homabay	1455	E034.63761	S00.62468	Long rain
Homabay	1455	E034.63761	S00.62468	Long rain
Homabay	1408	E034.63527	S00.62322	Long rain
Homabay	1408	E034.63527	S00.62322	Long rain
Homabay	1408	E034.63527	S00.62322	Long rain
Homabay	1408	E034.63527	S00.62322	Long rain
Homabay	1433	E034.63490	S00.62302	Long rain
Homabay	1433	E034.63490	S00.62302	Long rain
Homabay	1433	E034.63490	S00.62302	Long rain
Homabay	1433	E034.63490	S00.62302	Long rain
Homabay	1448	E034.63409	S00.62329	Long rain
Homabay	1448	E034.63409	S00.62329	Long rain
Homabay	1448	E034.63409	S00.62329	Long rain
Homabay	1448	E034.63409	S00.62329	Long rain
Homabay	1438	E034.63251	S00.62313	Long rain
Homabay	1438	E034.63251	S00.62313	Long rain
Homabay	1438	E034.63251	S00.62313	Long rain
Homabay	1438	E034.63251	S00.62313	Long rain
Homabay	1438	E034.63251	S00.62313	Long rain
Homabay	1449	E034.62985	S00.62512	Long rain
Homabay	1449	E034.62985	S00.62512	Long rain
Homabay	1449	E034.62985	S00.62512	Long rain
Homabay	1449	E034.62985	S00.62512	Long rain
Homabay	1351	E034.58273	S00.61512	Short rain
Kakamega	1552	E034.72606	N00.12995	Short rain
Homabay	1351	E034.58273	S00.61512	Short rain
Homabay	1351	E034.58273	S00.61512	Short rain
Busia	1225	E034.33674	N00.65445	Short rain
Busia	1320	E034.35423	N00.73924	Short rain
Homabay	1351	E034.58273	S00.61512	Short rain
Busia	1389	E034.38952	N00.71382	Short rain
Busia	1416	E034.37709	N00.69690	Short rain

Busia	1410	E034.37346	N00.69641	Short rain
Busia	1407	E034.39107	N00.71532	Short rain
Busia	1407	E034.39107	N00.71532	Short rain
Busia	1221	E034.17486	N00.36270	Short rain
Busia	1229	E034.17559	N00.36128	Short rain
Busia	1416	E034.37709	N00.69690	Short rain
Kakamega	1552	E034.72606	N00.12995	Short rain
Kakamega	1518	E034.78664	N00.14788	Short rain
Busia	1438	E034.37261	N00.69267	Short rain
Kakamega	1597	E034.70835	N00.07816	Short rain
Kakamega	1552	E034.72606	N00.12995	Short rain
Kakamega	1518	E034.78664	N00.14788	Short rain
Busia	1181	E034.33303	N00.62172	Short rain
Busia	1462	E034.39536	N00.67814	Short rain
Busia	1462	E034.39536	N00.67814	Short rain
Busia	1336	E034.35728	N00.74348	Short rain
Busia	1189	E034.33092	N00.62213	Short rain
Homabay	1336	E034.54420	S00.57751	Short rain
Busia	1228	E034.33865	N00.65436	Short rain
Busia	1343	E034.33363	N00.63637	Short rain
Busia	1218	E034.33158	N00.63656	Short rain
Siaya	1267	E034.32716	S00.06093	Short rain
Siaya	1259	E034.32898	S00.05704	Short rain
Busia	1441	E034.39369	N00.67833	Short rain
Busia	1458	E034.39904	N00.68400	Short rain
Busia	1455	E034.39881	N00.67807	Short rain
Siaya	1267	E034.32716	S00.06093	Short rain
Siaya	1267	E034.32716	S00.06093	Short rain
Busia	1467	E034.39631	N00.67923	Short rain
Busia	1469	E034.39760	N00.67960	Short rain
Busia	1469	E034.39760	N00.67960	Short rain
Busia	1390	E034.39028	N00.71010	Short rain
Kakamega	1592	E034.75635	N00.11998	Short rain
Kakamega	1592	E034.75635	N00.11998	Short rain
Busia	1229	E034.17955	N00.36007	Short rain
Busia	1382	E034.38951	N00.71284	Short rain
Busia	1379	E034.38913	N00.71270	Short rain
Busia	1390	E034.39028	N00.71010	Short rain
Busia	1395	E034.39230	N00.71068	Short rain
Busia	1395	E034.39230	N00.71068	Short rain
Busia	1336	E034.35728	N00.74348	Short rain
Busia	1379	E034.38913	N00.71270	Short rain
Busia	1385	E034.38935	N00.71435	Short rain
Busia	1440	E034.37812	N00.69597	Short rain

Busia	1430	E034.37445	N00.69515	Short rain
Busia	1385	E034.38935	N00.71435	Short rain
Busia	1382	E034.38951	N00.71284	Short rain
Busia	1430	E034.37445	N00.69515	Short rain
Busia	1395	E034.39273	N00.71085	Short rain
Busia	1407	E034.39782	N00.70042	Short rain
Busia	1410	E034.39635	N00.70070	Short rain
Busia	1361	E034.36237	N00.73834	Short rain
Busia	1395	E034.39273	N00.71085	Short rain
Busia	1364	E034.36440	N00.74005	Short rain
Busia	1363	E034.36406	N00.74013	Short rain
Busia	1389	E034.38952	N00.71382	Short rain
Busia	1306	E034.27793	N00.31820	Short rain
Busia	1234	E034.15780	N00.32863	Short rain
Bungoma	1307	E034.37951	N00.61641	Short rain
Bungoma	1307	E034.37951	N00.61641	Short rain
Busia	1234	E034.15808	N00.32931	Short rain
Busia	1234	E034.15808	N00.32931	Short rain
Busia	1234	E034.15808	N00.32931	Short rain
Bungoma	1431	E034.47524	N00.71529	Short rain
Bungoma	1431	E034.47524	N00.71529	Short rain
Busia	1306	E034.27793	N00.31820	Short rain
Bungoma	1324	E034.38408	N00.61557	Short rain
Bungoma	1432	E034.47522	N00.71455	Short rain
Bungoma	1432	E034.47522	N00.71455	Short rain
Bungoma	1427	E034.47611	N00.71328	Short rain
Siaya	1274	E034.34568	S00.08004	Short rain
Siaya	1274	E034.34568	S00.08004	Short rain
Siaya	1274	E034.34568	S00.08004	Short rain
Siaya	1274	E034.34568	S00.08004	Short rain
Bungoma	1431	E034.47524	N00.71529	Short rain
Busia	1205	E034.16887	N00.35417	Short rain
Busia	1211	E034.17033	N00.35276	Short rain
Kakamega	1544	E034.66602	N00.06374	Short rain
Busia	1234	E034.15780	N00.32863	Short rain
Busia	1205	E034.16887	N00.35417	Short rain
Busia	1201	E034.17838	N00.36920	Short rain
Bungoma	1324	E034.38408	N00.61557	Short rain
Bungoma	1324	E034.38408	N00.61557	Short rain
Bungoma	1437	E034.47462	N00.47153	Short rain
Busia	1277	E034.15487	N00.31860	Short rain
Bungoma	1427	E034.47611	N00.71328	Short rain
Siaya	1197	E034.34181	S00.32585	Short rain
Busia	1289	E034.33158	N00.33419	Short rain

Busia	1190	E034.16438	N00.36786	Short rain
Siaya	1303	E034.34433	S00.08766	Short rain
Busia	1356	E034.35058	N00.71592	Short rain
Siaya	1303	E034.34433	S00.08766	Short rain
Busia	1277	E034.15487	N00.31860	Short rain
Busia	1199	E034.16658	N00.36056	Short rain
Busia	1199	E034.16658	N00.36056	Short rain
Busia	1286	E034.27865	N00.31569	Short rain
Bungoma	1437	E034.47462	N00.47153	Short rain
Bungoma	1437	E034.47462	N00.47153	Short rain
Bungoma	1441	E034.47380	N00.71567	Short rain
Bungoma	1441	E034.47380	N00.71567	Short rain
Bungoma	1441	E034.47380	N00.71567	Short rain
Busia	1312	E034.28217	N00.31770	Short rain
Busia	1312	E034.28217	N00.31770	Short rain
Siaya	1303	E034.34433	S00.08766	Short rain
Kakamega	1539	E034.69537	N00.06794	Short rain
Bungoma	1427	E034.47611	N00.71328	Short rain
Bungoma	1436	E034.47055	N00.71428	Short rain
Bungoma	1436	E034.47055	N00.71428	Short rain
Bungoma	1436	E034.47055	N00.71428	Short rain
Busia	1182	E034.16068	N00.36627	Short rain
Siaya	1303	E034.34433	S00.08766	Short rain
Kakamega	1539	E034.69537	N00.06794	Short rain
Bungoma	1284	E034.39766	N00.59157	Short rain
Busia	1202	E034.13985	N00.40976	Short rain
Kakamega	1540	E034.66439	N00.06210	Short rain
Kakamega	1540	E034.66439	N00.06210	Short rain
Bungoma	1284	E034.39766	N00.59157	Short rain
Kakamega	1525	E034.66335	N00.05916	Short rain
Bungoma	1284	E034.39766	N00.59157	Short rain
Kakamega	1525	E034.66335	N00.05916	Short rain
Busia	1271	E034.23274	N00.30274	Short rain
Siaya	1223	E034.33007	S00.33214	Short rain
Siaya	1223	E034.33007	S00.33214	Short rain
Busia	1286	E034.27865	N00.31569	Short rain
Bungoma	1284	E034.39766	N00.59157	Short rain
Busia	1274	E034.23130	N00.30292	Short rain
Siaya	1197	E034.34181	S00.32585	Short rain
Siaya	1197	E034.34181	S00.32585	Short rain
Bungoma	1284	E034.39766	N00.59157	Short rain
Kakamega	1525	E034.66335	N00.05916	Short rain
Busia	1274	E034.23130	N00.30292	Short rain
Kakamega	1498	E034.66174	N00.05348	Short rain

Kakamega	1498	E034.66174	N00.05348	Short rain
Kakamega	1498	E034.66068	N00.05534	Short rain
Kakamega	1498	E034.66068	N00.05534	Short rain
Kakamega	1498	E034.66068	N00.05534	Short rain
Busia	1274	E034.23130	N00.30292	Short rain
Bungoma	1271	E034.40218	N00.59605	Short rain
Siaya	1223	E034.33007	S00.33214	Short rain
Busia	1230	E034.18511	N00.35133	Short rain
Bungoma	1271	E034.40218	N00.59605	Short rain
Bungoma	1271	E034.40218	N00.59605	Short rain
Busia	1195	E034.14187	N00.40381	Short rain
Bungoma	1271	E034.40218	N00.59605	Short rain
Busia	1304	E034.28094	N00.31598	Short rain
Busia	1311	E034.27917	N00.31702	Short rain
Busia	1311	E034.27917	N00.31702	Short rain
Homabay	1336	E034.54420	S00.57751	Short rain
Bungoma	1271	E034.40218	N00.59605	Short rain
Busia	1277	E034.23585	N00.30755	Short rain
Busia	1195	E034.14187	N00.40381	Short rain
Siaya	1336	E034.35868	S00.07524	Short rain
Siaya	1336	E034.35868	S00.07524	Short rain
Siaya	1336	E034.35868	S00.07524	Short rain
Siaya	1336	E034.35868	S00.07524	Short rain
Busia	1286	E034.27865	N00.31569	Short rain
Siaya	1301	E034.34427	S00.08644	Short rain
Siaya	1301	E034.34427	S00.08644	Short rain
Siaya	1301	E034.34427	S00.08644	Short rain
Bungoma	1283	E034.39667	N00.59429	Short rain
Bungoma	1283	E034.39667	N00.59429	Short rain
Bungoma	1283	E034.39667	N00.59429	Short rain
Bungoma	1283	E034.39667	N00.59429	Short rain
Bungoma	1283	E034.39667	N00.59429	Short rain
Busia	1285	E034.33087	N00.32232	Short rain
Busia	1183	E034.10686	N00.41479	Short rain
Busia	1183	E034.10686	N00.41479	Short rain
Homabay	1345	E034.58086	S00.60979	Short rain
Busia	1285	E034.33087	N00.32232	Short rain
Homabay	1374	E034.54052	S00.57785	Short rain
Homabay	1327	E034.58371	S00.60816	Short rain
Homabay	1327	E034.58371	S00.60816	Short rain
Homabay	1327	E034.58371	S00.60816	Short rain
Homabay	1327	E034.58371	S00.60816	Short rain
Homabay	1327	E034.58371	S00.60816	Short rain
Homabay	1345	E034.58086	S00.60979	Short rain

Homabay	1337	E034.54174	S00.57543	Short rain
Busia	1186	E034.10749	N00.41389	Short rain
Homabay	1337	E034.54174	S00.57543	Short rain
Homabay	1337	E034.54174	S00.57543	Short rain
Busia	1289	E034.32445	N00.32220	Short rain
Busia	1183	E034.12789	N00.47977	Short rain
Busia	1186	E034.10749	N00.41389	Short rain
Busia	1289	E034.33158	N00.33419	Short rain
Homabay	1374	E034.54052	S00.57785	Short rain
Homabay	1362	E034.54165	S00.57765	Short rain
Busia	1203	E034.10750	N00.42081	Short rain
Homabay	1362	E034.54165	S00.57765	Short rain
Siaya	1303	E034.34433	S00.08766	Short rain
Homabay	1374	E034.54052	S00.57785	Short rain
Homabay	1345	E034.58086	S00.60979	Short rain
Homabay	1336	E034.54420	S00.57751	Short rain
Homabay	1352	E034.58208	S00.61374	Short rain
Homabay	1352	E034.58208	S00.61374	Short rain
Homabay	1352	E034.58208	S00.61374	Short rain
Homabay	1352	E034.58208	S00.61374	Short rain
Homabay	1352	E034.58208	S00.61374	Short rain
Homabay	1345	E034.58086	S00.60979	Short rain
Busia	1320	E034.35423	N00.73924	Short rain
Busia	1284	E034.33161	N00.32302	Short rain
Busia	1284	E034.33161	N00.32302	Short rain
Busia	1183	E034.12789	N00.47977	Short rain
Busia	1320	E034.35423	N00.73924	Short rain
Homabay	1345	E034.58086	S00.60979	Short rain
Homabay	1336	E034.58243	S00.60857	Short rain
Homabay	1336	E034.58243	S00.60857	Short rain
Homabay	1336	E034.58243	S00.60857	Short rain
Homabay	1336	E034.58243	S00.60857	Short rain
Homabay	1336	E034.58243	S00.60857	Short rain
Homabay	1336	E034.58243	S00.60857	Short rain
Busia	1320	E034.35423	N00.73924	Short rain
Busia	1320	E034.35423	N00.73924	Short rain
Homabay	1362	E034.54165	S00.57765	Short rain
Homabay	1329	E034.45344	S00.69450	Short rain
Homabay	1325	E034.53304	S00.69321	Short rain
Kakamega	1553	E034.71852	N00.12469	Short rain
Busia	1277	E034.32221	N00.32168	Short rain
Homabay	1362	E034.54165	S00.57765	Short rain
Homabay	1329	E034.45344	S00.69450	Short rain
Homabay	1325	E034.53304	S00.69321	Short rain

Homabay	1325	E034.53304	S00.69321	Short rain
Homabay	1370	E034.53998	S00.57697	Short rain
Homabay	1370	E034.53998	S00.57697	Short rain
Homabay	1329	E034.45344	S00.69450	Short rain
Homabay	1334	E034.53652	S00.69514	Short rain
Homabay	1334	E034.53652	S00.69514	Short rain
Busia	1206	E034.17614	N00.36898	Short rain
Homabay	1329	E034.53646	S00.69228	Short rain
Homabay	1329	E034.53646	S00.69228	Short rain
Busia	1182	E034.10469	N00.41272	Short rain
Homabay	1370	E034.53998	S00.57697	Short rain
Homabay	1334	E034.53652	S00.69514	Short rain
Homabay	1329	E034.53646	S00.69228	Short rain
Homabay	1334	E034.53652	S00.69514	Short rain
Homabay	1325	E034.53304	S00.69321	Short rain
Busia	1182	E034.10469	N00.41272	Short rain
Busia	1257	E034.32043	N00.32471	Short rain
Busia	1277	E034.32221	N00.32168	Short rain
Busia	1185	E034.10552	N00.41298	Short rain
Bungoma	1481	E034.53045	N00.60687	Short rain
Bungoma	1514	E034.533119	N00.61361	Short rain
Bungoma	1479	E034.52184	N00.61094	Short rain
Bungoma	1490	E034.526390	N00.617222	Short rain
Bungoma	1509	E034.58068	N00.62845	Short rain
Bungoma	1557	E034.59395	N00.66004	Short rain
Bungoma	1515	E034.60737	N00.66895	Short rain
Bungoma	1538	E034.61226	N00.67426	Short rain
Bungoma	1747	E034.72624	N00.82073	Short rain
Bungoma	1935	E034.72564	N00.85590	Short rain
Kakamega	1469	E034.62708	N00.21789	Short rain
Kakamega	1469	E034.62708	N00.21789	Short rain
Kakamega	1469	E034.62708	N00.21789	Short rain
Kakamega	1469	E034.52639	N00.21789	Short rain
Kakamega	1490	E034.62357	N00.22022	Short rain
Kakamega	1490	E034.62357	N00.22022	Short rain
Kakamega	1490	E034.62357	N00.22022	Short rain
Kakamega	1490	E034.62357	N00.22022	Short rain
Kakamega	1490	E034.62357	N00.22022	Short rain
Kakamega	1490	E034.62357	N00.22022	Short rain
Kakamega	1490	E034.62357	N00.22022	Short rain
Kakamega	1490	E034.62357	N00.22022	Short rain
Kakamega	1490	E034.62357	N00.22022	Short rain
Kakamega	1490	E034.62357	N00.22022	Short rain
Kakamega	1490	E034.62357	N00.22022	Short rain

Appendix VII: Publications from this study

International Journal of Genetics and Genomics
2019; 7(4): 98-102
http://www.sciencepublishinggroup.com/j/ijgg
doi: 10.11648/j.ijgg.20190704.12
ISSN: 2376-7340 (Print); ISSN: 2376-7359 (Online)

Incidence of Groundnut Rosette Disease (GRD) and Genetic Diversity of Groundnut Rosette Assistor Virus (GRAV) in Western Kenya

Benard Mukoye[1, 3, *], Millicent Florence Owuor Ndonga[1], Hassan Karakacha Were[2]

[1]Department of Biological Sciences, School of Natural Sciences, Masinde Muliro University of Science and Technology (MMUST), Kakamega, Kenya

[2]Department of Agriculture and Land Use Management (ALUM), School of Agriculture, Veterinary Science and Technology, Masinde Muliro University of Science and Technology (MMUST), Kakamega, Kenya

[3]Department of Biosafety and Phytosanitary Services, Kenya Plant Health Inspectorate Service (KEPHIS), Nairobi, Kenya

Email address:
btemukoye@gmail.com (B. Mukoye), mndonga@mmust.ac.ke (M. F. O. Ndonga), hwere@mmust.ac.ke (H. K. Were)

[*]Corresponding author

To cite this article:
Benard Mukoye, Millicent Florence Owuor Ndonga, Hassan Karakacha Were. Incidence of Groundnut Rosette Disease (GRD) and Genetic Diversity of Groundnut Rosette Assistor Virus (GRAV) in Western Kenya. *International Journal of Genetics and Genomics*. Vol. 7, No. 4, 2019, pp. 98-102. doi: 10.11648/j.ijgg.20190704.12

Received: September 11, 2019; **Accepted**: October 4, 2019; **Published**: October 16, 2019

Abstract: This study determined the incidence of groundnut rosette disease (GRD) and genetic diversity of groundnut rosette assistor virus (GRAV, genus *Luteovirus*) in western Kenya. The diseases is a major constraint of groundnuts in Sub-Saharan Africa (SSA) causing up to 100% yield losses in severe cases. Among the GRD associated viruses, GRAV plays a crucial role in vector transmission of the other viruses. Therefore understanding the genetics of GRAV across SSA could enhance development of resistance to the disease. In Kenya, groundnuts are mainly grown in western region, however, the yields are poor mainly due to GRD. Information on occurrence and distribution of GRD in western Kenya was not documented and little was known about the characteristics of associated viruses. Two diagnostic surveys were conducted in six counties; Bungoma, Busia, Homabay, Kakamega, Siaya and Vihiga. Symptomatic and asymptomatic groundnut were collected in RNA*later*® solution for laboratory analysis. Total RNA was extracted from the leaf samples using RNeasy Mini Kit (Qiagen) according to the manufacturers' protocol and used for double stranded cDNA synthesis using the SuperScript II kit. The cDNA was column-purified with the DNA Clean & ConcentratorTM-5 – DNA kit. The samples were then processed with the transposon-based chemistry library preparation kit (Nextera XT, Illumina) following manufacturer's instructions. The fragment sizes structure of the DNA libraries was assessed using the Agilent 2100 Bioanalyzer. The indexed denatured DNA libraries were sequenced (200-bp paired-end sequencing) on the Illumina MiSeq platform (Illumina). Reads quality check was done using FastQC. Trimmed reads were used for de novo assembly and contigs aligned to the viral genomes database using CLC Genomics Workbench 10.1.2. The assembled contigs were subjected to a BLASTn search against the GenBank database. Phylogenetic analyses and comparisons were performed using the MEGA X. Average incidence was 53% and 41% in the short and long rain seasons, respectively. Chlorotic rosette was the dominant symptom followed by Green rosette and Mosaic. The GRAV coat protein (GRAV-CP) gene sequences revealed 97-100% identity with GeneBank isolates showing very slight variations across SSA. The study concludes that GRD incidence is high in western Kenya and that GRAV is highly conserved across SSA. The study recommends an urgent need to curb GRD, possibly through the exploitation of pathogen derived resistance (PDR) with GRAV as the suitable candidate.

Keywords: Incidence, GRAV, Kenya, Diversity

1. Introduction

Groundnuts, (*Arachis hypogaea L.),* is the fifth most important annual oilseed and food legume crop. It is grown in diverse environments throughout the semi-arid and sub-tropical regions, in nearly 100 countries, in the six continents of the world [1]. Groundnut production is of great value in terms of income and nutrition for smallholder farmers in East Africa [2, 3]. Resource poor smallholder farmers grow nearly 75 - 80% of the world's groundnuts in developing countries obtaining yields of 500-800kg/ha, as opposed to the potential yield of >2.5t/ha [4]. In western Kenya, an average of 600 – 700 kg/ha is achieved which is less than 30-50% of the potential yield [2]. Low yields are mainly attributed to poor quality seeds, drought, poor agronomic practices, numerous pests and diseases caused by numerous pathogenic viruses, fungi, bacteria and nematodes [5, 3]. Among the viral diseases, Groundnut rosette disease (GRD) is the most devastating in Sub-Saharan Africa (SSA) causing an estimated annual loss of US$156 million every year [6]. The disease is caused by association between Groundnut rosette assistor virus (GRAV), Groundnut rosette *umbravirus* (GRV) and a Satellite-RNA (Sat-RNA) of GRV [7]. To be transmitted by aphids, GRV and Sat-RNA are packaged within the GRAV coat protein [8].

Groundnut rosette assistor virus (GRAV) belongs to the family *Luteoviridae* [9]. The GRAV virion are isometric shaped with 28nm diameter non-enveloped particles of polyhedral symmetry. It has a single stranded positive sense RNA non-segmented genome of 6900 nt that encodes both structural and non-structural proteins [10]. It is suggested that GRAV encodes six open reading frames (ORFs) just like other *luteoviruses*. The GRAV virions are composed of 24.5kDa single coat protein (CP) subunits. This virus is anti-genetically related to *Potato leaf roll virus, Beet western yellow virus* and *Bean/pea leaf roll virus* [11].

Both chlorotic and green rosette symptoms occur throughout the SSA, and sometimes occur in the same field [12]. A less common third symptom variant, called mosaic rosette, resulting from mixed infection by the Sat-RNA causing chlorotic and green mottled variant, has been reported from East Africa [11, 6]. Infected groundnut leaves may also show symptoms other than the typical chlorotic or green rosette [8].

In Eastern Uganda, green rosette symptoms predominate [13]. This is in contrast with [14], who reported that chlorotic rosette symptoms of GRD have been the predominant form throughout SSA and western Kenya. The dynamics of the GRD virus symptomatology, therefore, needs constant monitoring. For example, in Nigeria, a there was shift from green to chlorotic rosette over a period of about 20 years. The shift could be due to changes in the genome sequences of GRD associated agents or other factors [13].

Survey conducted by [14] showed that GRD incidence ranged between 40% in areas of western Kenya surveyed in the groundnut growing seasons of 1997-1998 and Sat-RNA shared 89-95% nucleotide identity with those from Malawi and Nigeria. Since then, no other survey has been conducted to ascertain the current status of GRD in the region. In addition, no genomic sequences of any of the GRD associated viruses existed in the GeneBank from western Kenya. With the dynamics of the disease, this hinders proper diagnosis of GRD and development of management strategies. This study determined the incidence of GRD and assed the sequence diversity of GRAV of isolates from western Kenya.

2. Materials and Methods

2.1. Field Survey

The GRD diagnostic survey was conducted in all the major groundnut growing areas in western Kenya during the short rains (October to December 2016) and long rains season (May to July 2017). The following Counties were covered: Bungoma, Busia, Homabay, Kakamega, Siaya and Vihiga. Sampling of groundnut farms was done by stopping at regular predetermined intervals, of 3-8 km along motorable roads that traverses each sampling area. The survey were conducted, by walking through groundnut fields, and visually inspecting groundnut crops for symptomatic leaves. Disease incidence was calculated according to [15], as the percentage of plants showing GRD virus symptoms, to the total number of plants observed in the field. GRD viral incidence was scored using a rating scale according to [15], where: low incidence = 1-20%; moderate incidence = 21-49% and high incidence = 50-100%. The types of GRD symptoms observed were recorded. The collected data on GRD virus incidence and severity, was subjected to analysis of variance (ANOVA), using Statistical Analysis System (SAS) program version 9.3.1 software. Pairwise comparisons of means was done using Least Significance Differences (LSD) for multiple-means comparison method at $P \leq 0.05$ confidence level.

Symptomatic and asymptomatic leaves were collected in 10ml falcon tubes containing RNA*later*® RNA Stabilization Solution and put in a cool box. The samples were kept in the fridge and used for molecular studies. Geographical Positioning System (GPS) (entrex venture HC GARMIN™), was used to record the latitude, longitude and altitude of the sampled farms.

2.2. RNA Extraction, Sequencing and Sequence Analyses

Total RNA was extracted from the leaf samples using RNeasy Mini Kit (Qiagen) according to the manufacturers' protocol and used for double stranded cDNA synthesis using the SuperScript II (Thermo Fisher Scientific, Waltham, USA) kit. The cDNA was column-purified with the DNA Clean & ConcentratorTM-5 – DNA kit (Zymo Research, Irvine, USA). The samples were then processed with the transposon-based chemistry library preparation kit (Nextera XT, Illumina) following manufacturer's instructions. The fragment sizes structure of the DNA libraries was assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies,

Santa Clara, USA). The indexed denatured DNA libraries were sequenced (200-bp paired-end sequencing) on the Illumina MiSeq platform (Illumina).

Reads quality check was done using FastQC (version 0.11.5). Reads were then trimmed to remove poor quality sequences. Trimmed reads [16] were used for de novo assembly and contigs aligned to the viral genomes database (ftp://ftp.ncbi.nih.gov/genomes/Viruses/all.fna.tar.gz/, downloaded on October 2017) using CLC Genomics Workbench 10.1.2. The assembled contigs were subjected to a BLASTn search against the GenBank database [17]. Complete and partial GRV Sat-RNA sequences used for comparison and phylogenetic analyses were retrieved from GenBank (http://www.ncbi.nlm.nih.gov/). Phylogenetic analyses and comparisons were performed using the MEGA X [18].

3. Results

3.1. Incidence of GRD

A total of 526 farms were sampled in six (6) counties (253 in long rain and 273 in short rain). The main symptoms observed in all Counties in order of abundance, starting from the most common, were chlorotic rosette, green rosette and mosaic. Generally, GRD incidence was high during the short rain season (53%) than the long rain season (41%) in all Counties. High mean GRD incidence was recorded in Kakamega in the short rain season (68.92%) while moderate incidence was in Bungoma (30.89%) during the long rain season. There was a significant difference in GRD incidence among the counties (p=0.011). Siaya County had the overall lowest incidence which was significantly different from that of Kakamega but did not vary significantly from that of Bungoma, Busia, Vihiga and Homabay counties (Table 1).

Table 1. Mean GRD incidence (%) per County.

County	Season	N	Mean (%)	Std. Error of Mean (+/-)
Bungoma	Long rain	45	30.89	4.534
	Short rain	47	66.51	4.295
Busia	Long rain	74	43.36	3.526
	Short rain	108	46.56	2.728
Homabay	Long rain	73	48.60	3.919
	Short rain	55	48.22	4.025
Kakamega	Long rain	30	43.47	5.283
	Short rain	17	94.12	4.779
Siaya	Long rain	31	33.94	4.820
	Short rain	26	43.23	6.645
Vihiga	Short rain	20	47.50	6.412
	Long rain	253	41.51	1.962
	Short rain	273	53.04	1.909

3.2. Diversity of GRAV

Four GRAV coat protein (CP) gene sequences were assembled (600 nt). The four were compared with GRAV CP gene sequences from Malawi, Nigeria and Ghana available in the GeneBank. The comparison revealed 97-100% identity with the Kenyan isolates. Isolates GRAV-5 and GRAV-19 each had 100% identity with M16GCP (AF195824.1) and 99% with M8GCP (AF195502.1) then 98% with the other Malawian, Nigerian and Ghanaian Isolates. Isolate GRAV-22 had 99% identity with isolates M16GCP and M8GCP then 98% with the other Malawian, Nigerian and Ghanaian Isolates. Isolate GRAV-12 had 100% identity with M16GCP and 99% with M8GCP from Malawi, then 98% with the rest of Malawian, Ghanaian and Nigerian isolates except N29GCP (AF195828.1) and N15GCP (AF195825.1) that showed 97% identity. In phylogenetic tree, all Kenyan isolates clustered together with isolate M16GCP. In general all western Kenya isolates exhibited closest identity and grouped together with some Malawian isolates, M16GCP and M8GCP than the rest of Malawian, Nigerian and Ghanaian isolates (Figure 1).

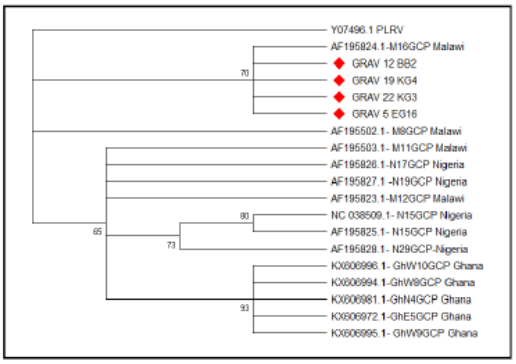


Figure 1. Phylogenetic tree of the 600nt western Kenya GRAV CP and GeneBank isolates.

The evolutionary history was inferred by using the Maximum Likelihood method based on the Kimura 2-parameter model. The tree is rooted on of a distantly related *Luteovirus* (Potato leaf roll virus – Y07496.1 PLRV). Bootstrap confidence values (500 replications) are shown.

4. Discussion

Groundnut rosette is the most prevalent disease of groundnuts in western Kenya. The disease was recorded in every County that was surveyed with incidences of up to 100% at some farm levels. The short rain season recorded higher incidence (53%) than the long rains (41%). This could be attributed to the high vector pressure during the short rains as compared to the long rains season when the aphid pressure is low as a result of heavy rains that wash the insects away. A study by [12], found that periods of long rains negatively affected GRD progression as aphid vector pressure was low. [19], reported a positive correlation between potato disease incidence and aphid numbers. This further supports the implication that virus disease incidence variations between the seasons contributed to by differences in vector pressure. Incidence increased with increase in severity due to early infection leading to intensification of the viruses as the plant

grows and build-up of inoculum for vectors to spread to nearby plants. Groundnut rosette is a polycyclic disease whereby diseased plants from previous cropping season serves as inoculum sources for initiating subsequent disease spread [8]. In western Kenya, groundnuts are grown in two cropping seasons (long rains and short rains) and due to limitation in land to practice shift cultivation, the same piece of land is continuously used to grow the same or related host crops in the subsequent cropping season. Therefore, GRD infected groundnuts and possibly hosts of any of the GRD associated viruses remaining from the long rains season serves as immediate sources of the GRD agents beginning the disease cycle at early stages of crop development in the short rains cropping season. Such initial infections that occur at early stages of plant growth enhance repeated cycles of infections thus increasing the severity of the disease in the groundnut fields [6].

All major GRD symptoms were observed in the surveyed region with chlorotic rosette being most prevalent followed by green rosette. This supports the findings of [14], who reported chlorotic rosette to be the most prevalent GRD symptoms in the region. The high prevalence of the chlorotic rosette could also be attributed to its higher transmission efficiency compared to green rosette. This observation concurs with that of [20], who reported minimum acquisition feeding periods of 4 h and 8 h for chlorotic and green rosette respectively and the median latent periods of 26.4 h, 38.4 h respectively, for chlorotic and green rosette. The mosaic symptom has not been previously reported but was distributed in most of the surveyed region. This suggests that there is evolution of new variants of Sat-RNA in western Kenya that might be causing these new symptoms or the mosaic was due to another causal agent. A total of 10 variants of Sat-RNA have been reported to be associated with the various GRD symptoms [21]. A mixture of either variants, especially the chlorotic and green rosette and/or the mild ones, are likely to induce the mosaic symptoms [8]. It is therefore possible that some of these variants occur in western Kenya in mixed infections, thus causing the varied symptom observed, especially the mosaic. Apart from the typical rosette symptoms, other symptoms including severe leaf curling and bunching were observed. This suggests that there is wider variability in expression of GRD and could be due to more severe variants of associated viruses or other agents. It is worth noting that from the Next generation Sequences (NGS) of this study, other than GRD associated viruses, other viruses were detected (data not shown) and could be the reason for some of the new symptoms observed on groundnuts [22].

The four GRAV CP gene sequences from western Kenya clustered together and had 97 – 100% identity with those from Malawi, Ghana and Nigeria implying that there was no much difference among the western Kenya GRAV CP gene isolates. Kenyan GRAV CP isolates exhibited closest identities with Malawian isolates than Nigerian and Ghanaian isolates. This findings concurs with [14] and [23], who observed closer identity between sequences from the same geographical region as compared to those from separate geographical regions. In the study, [14] found that Kenyan isolates of GRAV CP gene shared 98% nucleotide identity with Malawian isolates as compared to 96-97% with those from Nigeria. [23], observed that Ghanaian GRAV CP gene sequence isolates had 98-99% nucleotide identity as compared to 97-99% with Malawian isolates. Such differences due to geographical distances could be as a result of differences in environmental conditions that bring about variations in evolution of the viruses. All western Kenya GRAV CP isolates were closest to Malawian isolates M16GCP and M8GCP (99-100%) than the other isolates from Malawi, Nigeria and Ghana. A similar observation was noted by [14], where two of the Kenya isolates in the study (K1 and K2), specifically from western Kenya were closest to M16GCP and M8GCP than with the rest of her isolates from other regions in Kenya. This could imply that the GRAV CP gene from western Kenya have not evolved for at least the last 20 years. However variation could exist in GRAV from other regions in Kenya. In general all GRAV CP gene sequences both in this study and those in GeneBank shared 97-100% nucleotide identity. This implies that GRAV CP gene is highly conserved across the wide geographical region in Sub-Saharan Africa. It can thus be targeted as a suitable candidate for development of pathogen-derived resistance (PDR) through genetic engineering that can be used across Sub-Saharan Africa [9, 23].

5. Conclusion

Groundnut Rosette (GRD) is still the major disease of groundnuts and is present whenever groundnuts are grown in western Kenya. Chlorotic rosette is the most prevalent form of symptom on groundnuts in western Kenya. The mosaic rosette is an emerging symptom in groundnuts and could be due to dual infection by Sat-RNA variants or other agents. The GRAV CP gene is less diverse even with wide geographical distance.

The four GRAV sequences were deposited in GeneBank with accession numbers LC480460 (GRAV 12), LC480461 (GRAV 22), LC480462 (GRAV 19) and LC480463 (GRAV 5).

Acknowledgements

This work was funded by The royal society of UK, The International Foundation for Science, The National Research Fund (NRF-Kenya) and the ILRI BecA hub. We are grateful to Dr. Wellington Ekaya and the entire BecA capacity building team for allowing this work to be done at the BecA labs.

References

[1] Kumar, P. L. & Waliyar, F., (Ed). (2007). Diagnosis and detection of viruses infecting ICRISAT mandate crops: *Methods Manual*. Patancheru 502 324, Andhra Pradesh, India; *International Crops Research Institute for the Semi-Arid Tropics*. 133pp.

[2] Kidula, N., Okoko, N., Bravo-Ureta, B. E., Thuo, M. & Wasilwa, L. (2010). A preliminary analysis of yield differences in groundnuts between research and non-research farmers in Kenya. In paper presented at the 12th KARI biennial scientific conference, 8-12 November 2010, Naiobi Kenya.

[3] Okello, D. K., Birima, M. & Deom, C. M., (2010). Overview of groundnut research in Uganda: Post, present and future. *Afr. J. Biotechnol.* 9: 6448-6459.

[4] Kayondo, S. I., Rubaihayo, P. R., Ntare, B. R., Gibson, P. I., Edema, R., Ozimati, A. & Okello, D. K. (2014). Genetics of resistance to groundnut rosette virus disease: *African crop science journal,* 22: 21-29. ISSN: 1021-9730/2014.

[5] Mutegi, C. K. (2010). The extend of aflatoxin and aspergillus section flavi, penicillium spp. and Rhizopus spp. contamination of peanuts from households in Western Kenya and the causative factors of contamination. *PhD dissertation, University of Kwazulu-Natal, Pietermaritzburg.* South Africa.

[6] Waliyar, F., Kumar, P. L., Ntare, B. R., Monyo, E., Nigam, S. N., Reddy, A. S., Osiru, M. & Diallo, A. T. (2007). A Century of Research on Groundnut Rosette Disease and its Management. *Information Bulletin* no. 75. Patancheru 502 324, Andhra Pradesh, India. *International Crops Research Institute for the Semi-Arid Tropics,* 40 pp. ISBN 978-92-9066-501-4.

[7] Taliansky, M. E. & Robinson, D. J. (2003). Molecular Biology of umbraviruses*: Phantom warriors. J. Gen. Virol.* 84: 1951-1960.

[8] Naidu, R. A., Robinson, D. J., & Kimmins, F. M. (1998a). Detection of each of the causal agents of groundnut rosette disease in plants and vector aphids by RT-PCR. *J. Virol. Methods* 76: 9-18.

[9] Deom, C. M., Naidu, R. A., Chiyembekeza, A. J., Ntare, B. R. & Subrahmanyam, P. (2000). Sequence diversity with the three agents of groundnut rosette disease. *Phytopathol.* 90: 214-219. doi: 10.1094/PHYTO.2000.90.3.214

[10] Murant, A. F., & Kumar, I. K. (1990). Different variants of the satellite RNA of groundnuts rosette virus are responsible for the chlorotic and green forms of groundnut rosette disease. *Ann. Appl. Biol.* 117: 85-92.

[11] Scott, K. P., Farmer, M. J., Robinson, D. J., Torrence, L. & Murant, A. F. (1996). Comparison of the coat protein of groundnut rosette assistor virus with those of other *luteovirus*. *Ann. Appl. Biol.* 128: 77-83.

[12] Mugisa, I. O., Karungi, J., Akello, B., Ochwo-Ssemakula, M. K. N., Biruma, M., Okello, D. K. & Otim, G. (2016). Determinants of groundnut rosette virus disease occurrence in Uganda. *Elseviercropprotectionjournal.*http://dx.doi.org/10.1016/j.cropro.2015.10.019.

[13] Okello, D. V, Akello, L. B, Tukamuhabwa, P., Odongo, T. L, Ochwo-Ssemakula, M., Adriko, J. & Deom, C. M. (2014). Groundnut rosette disease symptom types, distribution and management of the disease in Uganda. *African journal of plant science.* 8: 153-163.

[14] Wangai, A. W., Pappu, S. S., Pappu, H. R., Okoko, N., Deom, C. M. & Naidu, R. A. (2001). Distribution and characteristics of groundnut rosette disease in Kenya. *Plant Disease,* 85: 470-474.

[15] Reddy, D. V. R. (1991). Groundnut viruses and virus diseases; Distribution, identification and control. *Rev. Plant Pathol.* 70: 665-678.

[16] Haas B. J, Papanicolaou A, Yassour M, Grabherr M, Blood P. D, Bowden J, Couger M. B, Eccles D, Li B, Lieber M, MacManes M. D, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey C. N, Henschel R, LeDuc R. D, Friedman N, Regev A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* 8: 1494–1512. https://doi.org/10.1038/nprot.2013.084.

[17] Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* 215: 403– 410. https://doi.org/10.1016/S0022-2836(05)80360-2.

[18] Kumar, S., Stecher, G., Li M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* 35: 1547-1549.

[19] Were, H. K., Kabira, J. N., Kinyua, Z. M., Olubayo, F. M., Karinga, J. K., Aura, J., Lees, A. K., Cowan, G. H. & Torrance, L. (2013). Occurrence and Distribution of Potato Pests and Diseases in Kenya. *Potato Research* 56: 325–342.

[20] Misari, S. M., Abraham, J. M., Demski J. W., Ansa, O. A., Kuhn, C. W, Casper, R. & Breyel, E. (1988a). Aphid transmission of the viruses causing chlorotic and green rosette diseases of peanut in Nigeria. *Plant Disease* 72: 250-253.

[21] Blok, V. C., Ziegler, A., Robinson, D. J. & Murant, A. F. (1994). Sequences of 10 variants of the satellite- like RNA -3 of groundnut rosette virus. *Virology* 202: 25-32.

[22] Mukoye, B., Mangeni, B. C., Sue, J., Ndonga, M. F. O., & Were, H. K. (2018). Next Generation Sequencing as a tool in modern pest diagnosis. A case study of groundnuts (*Arachis hypogaea*) as a potential host of new viruses in western Kenya. Conference proceedings: The 2nd Phytosanitary Conference, 4th – 8th June, 2018 at KEPHIS-Nairobi, Kenya.

[23] Appiah, A. S., Sossah, L. F., Tegg, S. R., Offei, K. S. & Wilson, R. C. (2017). Assessing sequence diversity of goundnut rosette disease agents and the distribution of *groundnut rosette assistor virus* in major groundnut-producing regions of Ghana. *Trop. Plant Pathol.* Doi: 10.1007/s40858-017-0140-x.

Vol. 12(1), pp. 22-29, January-June 2020
DOI: 10.5897/IJGMB2019.0178
Article Number: 81BC2FC63204
ISSN 2006-9863
Copyright © 2020
Author(s) retain the copyright of this article
http://www.academicjournals.org/IJGMB

ACADEMIC JOURNALS
expand your knowledge

International Journal of Genetics and Molecular Biology

Full Length Research Paper

Distribution of groundnut rosette disease and sequence diversity of groundnut rosette virus associated satellite RNA (Sat-RNA) in Western Kenya

Benard Mukoye[1,2*], Anthony Simiyu Mabele[1], Millicent Florence Owuor Ndonga[1], Bonphace Collins Mangeni[1] and Hassan Karakacha Were[1]

[1]Masinde Muliro University of Science and Technology (MMUST), Kakamega, Kenya.
[2]Kenya Plant Health Inspectorate Service (KEPHIS), Nairobi, Kenya.

Received 18 June, 2019; Accepted 29 October, 2019.

Production of groundnuts (*Arachis hypogaea* L.) in Western Kenya is mainly constrained by groundnut rosette disease (GRD) which cause up to 100% yield loss. This disease expresses different symptoms as a result of variations in the groundnut rosette virus (GRV) associated satellite-ribonucleic acid (GRV Sat-RNA). Over the past 20 years, no work had been done to document the status of the disease in Kenya. Additionally, no sequences of any of the GRD associated viruses were available in the GeneBank from Kenya. This study determined the distribution of GRD and the genetic diversity of GRV Sat-RNA. Sampling was done in main groundnut growing areas of Western Kenya during the long and short rain seasons in 2016/2017. Total RNA was extracted from the leafy samples collected using RNeasy Mini Kit (Qiagen) according to the manufacturers' protocol and used for double stranded cDNA synthesis using the SuperScript II kit. DNA libraries were sequenced on the Illumina MiSeq platform (Illumina). Reads were used for *de novo* assembly and contigs aligned to the viral genomes database using CLC Genomics Workbench 10.1.2. The assembled contigs were subjected to a BLASTn search against the GenBank database. Average GRD incidence was 53 and 41% in the short and long rain seasons, respectively. Chlorotic rosette was the dominant symptom followed by green rosette and mosaic. Nucleotide sequences of Sat-RNA revealed identities of 88 to 100% between the Kenyan isolates and those from Malawi, Nigeria and Ghana. All Kenya isolates clustered closest with green rosette variants of Malawi except one which clustered with chlorotic/yellow blotch variants. Rosette is widely distributed in Western Kenya and occurs wherever groundnuts are grown. The variations of GRD symptoms in Western Kenya could be due to the existence of different variants of Sat-RNA or other agents.

Key words: Groundnuts, satellite-ribonucleic acid (Sat-RNA), diversity, Western Kenya.

INTRODUCTION

Groundnut (*Arachis hypogaea L.)* is the fifth most important annual oilseed and food legume crop. It is grown in diverse environments throughout the semi-arid and sub-tropical regions, in nearly 100 countries, in the six continents of the world (Kumar and Waliyar, 2007). Groundnut production is of great value in terms of income and nutrition for smallholder farmers in East Africa (Kidula et al., 2010; Okello et al., 2010). Resource poor smallholder farmers grow nearly 75 to 80% of the world's groundnuts in developing countries obtaining yields of

500 to 800 kg/ha, as opposed to the potential yield of >2.5 t/ha (Kayondo et al., 2014). In Western Kenya, an average of 600 to 700 kg/ha is achieved which is less than 30 to 50% of the potential yield (Kidula et al., 2010). Low yields are mainly attributed to poor quality seeds, drought, poor agronomic practices, numerous pests and diseases caused by numerous pathogenic viruses, fungi, bacteria and nematodes (Mutegi, 2010; Okello et al., 2010). Among the viral diseases, groundnut rosette disease (GRD) is the most devastating in sub-Saharan Africa (SSA) causing an estimated annual loss of US$156 million every year (Waliyar et al., 2007). The disease is caused by association between groundnut rosette assistor virus (GRAV), groundnut rosette virus (GRV) and a Satellite-RNA (Sat-RNA) of GRV (Taliansky and Robinson, 2003). Variations in Sat-RNA have been shown to be responsible for different rosette symptoms (chlorotic, green and mosaic rosette) (Taliansky and Robinson, 1997; Olorunju et al., 2001; Kayondo et al., 2014). Both chlorotic and green rosette symptoms occur throughout the SSA, and sometimes occur in the same field (Mugisa et al., 2016). A less common third symptom variant, called mosaic rosette, resulting from mixed infection by the Sat-RNA causing chlorotic and green mottled variant, has been reported from East Africa (Scott et al., 1996; Waliyar et al., 2007). Infected groundnut leaves may also show symptoms other than the typical chlorotic or green rosette (Naidu et al., 1999).

In Eastern Uganda, green rosette symptoms predominate (Okello et al., 2014). This is in contrast with the findings of Wangai et al. (2001), who reported that chlorotic rosette symptoms of GRD have been the predominant form throughout SSA and Western Kenya. The dynamics of the GRD virus symptomatology, therefore, needs constant monitoring. For example, in Nigeria, a there was shift from green to chlorotic rosette over a period of about 20 years. The shift could be due to changes in the genome sequences of GRD associated agents or other factors (Okello et al., 2014).

Survey done by Wangai et al. (2001) showed that GRD incidence ranged between 40% in areas of Western Kenya surveyed in the groundnut growing seasons of 1997-1998 and Sat-RNA shared 89 to 95% nucleotide identity with those from Malawi and Nigeria. Since then, no other survey has been conducted to ascertain the current status of GRD in the region. In addition, no genomic sequences of any of the GRD associated viruses existed in the GeneBank from Western Kenya. With the dynamics of the disease, this hinders proper diagnosis of GRD and development of management strategies. This study determined the distribution of GRD and assessed the sequence diversity of Sat-RNA of isolates from Western Kenya.

MATERIALS AND METHODS

Field survey

The GRD diagnostic sampling was conducted in all the major groundnut growing areas in Western Kenya during the short rains (October to December 2016) and long rains season (May to July, 2017). The following counties were covered: Bungoma, Busia, Homabay, Kakamega, Siaya and Vihiga. Sampling of groundnut farms was done by stopping at regular predetermined intervals, of 3 to 8 km along motorable roads that traverses each sampling area. The survey was conducted, by walking through groundnut field and visually inspecting groundnut crops for symptomatic leaves. Disease incidence was calculated according to Reddy (1991), as the percentage of plants showing GRD virus symptoms, to the total number of plants observed in the field. GRD viral incidence was scored using a rating scale according to Reddy (1991) where: low incidence = 1-20%; moderate incidence = 21-49%, and high incidence = 50-100%. The types of GRD symptoms observed were recorded. The collected data on GRD virus incidence and severity, was subjected to analysis of variance (ANOVA), using Statistical Analysis System (SAS) program version 9.3.1 software (SAS Institute, 2013). Pairwise comparisons of means were done using Least Significance Differences (LSD) for multiple-means comparison method at $P \leq 0.05$ confidence level.

Symptomatic and asymptomatic leaves were collected in 10 ml falcon tubes containing RNA*later*® RNA Stabilization Solution and put in a cool box. The samples were kept in the fridge and used for molecular studies. Geographical Positioning System (GPS) (entrex venture HC GARMIN™), was used to record the latitude, longitude and altitude of the sampled farms.

RNA extraction, sequencing and sequence analyses

Total RNA was extracted from the leaf samples using RNeasy Mini Kit (Qiagen) according to the manufacturers' protocol and used for double-stranded cDNA synthesis using the SuperScript II (Thermo Fisher Scientific, Waltham, USA) kit. The cDNA was column-purified with the DNA Clean and ConcentratorTM-5-DNA kit (Zymo Research, Irvine, USA). The samples were then processed with the transposon-based chemistry library preparation kit (Nextera XT, Illumina) following manufacturer's instructions. The fragment sizes structure of the DNA libraries was assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, USA). The indexed denatured DNA libraries were sequenced (200-bp paired-end sequencing) on the Illumina MiSeq platform (Illumina).

Reads quality check was done using FastQC (version 0.11.5). Reads were then trimmed to remove poor quality sequences using Trimmomatic (V 0.36) software (Bolger et al., 2014). Trimmed reads (Haas et al., 2013) were used for *de novo* assembly and contigs aligned to the viral genomes database (ftp://ftp.ncbi.nih.gov/genomes/Viruses/all.fna.tar.gz/, downloaded on October 2017) using CLC Genomics Workbench 10.1.2. The assembled contigs were subjected to a BLASTn search against the GenBank database (Altschul et al., 1990). Complete and partial GRV Sat-RNA sequences used for comparison and phylogenetic analyses were retrieved from GenBank (http://www.ncbi.nlm.nih.gov/). Phylogenetic analyses and comparisons were performed using the MEGA X (Kumar et al., 2018). Tobacco bushy top virus - KU997687.1 TBTV was used as an outgroup.

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

Author(s) agree that this article remain permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Table 1. Mean GRD incidence (%) per County.

County	Season	N	*Mean (%)	Std. error of mean (+/-)
Bungoma	Long rain	45	30.89[b]	4.534
	Short rain	47	66.51[b]	4.295
Busia	Long rain	74	43.36[b]	3.526
	Short rain	108	46.56[b]	2.728
Homabay	Long rain	73	48.60[b]	3.919
	Short rain	55	48.22[b]	4.025
Kakamega	Long rain	30	43.47[b]	5.283
	Short rain	17	94.12[a]	4.779
Siaya	Long rain	31	33.94[b]	4.820
	Short rain	26	43.23[b]	6.645
Vihiga	Short rain	20	47.50[b]	6.412
Total	Long rain	253	41.51	1.962
	Short rain	273	53.04	1.909

*Means with the same letter within the column are not significantly different.

Table 2: Percentage frequency of the occurrence of the three main GRD symptoms in western Kenya.

Symptom	Percent
Chlorotic rosette (CR)	58.6
Green rosette (GR)	27.4
Mosaic (MS)	14.1

RESULTS

A total of 526 farms were sampled in six (6) counties (253 in long rain and 273 in short rain). Generally, GRD incidence was high during the short rain season (53%) than the long rain season (41%) in all counties. High mean GRD incidence was recorded in Kakamega in the short rain season (68.92%) while moderate incidence was in Bungoma (30.89%) during the long rain season. There was a significant difference in GRD incidence among the counties (p=0.011, df=521, F=3.322). Siaya County had the overall lowest incidence which was significantly different from that of Kakamega but did not vary significantly from that of Bungoma, Busia, Vihiga and Homabay counties (Table 1).

Generally, GRD infected plants were dwarf with increased tillering although some were tall but expressed other major symptoms associated with GRD. The main symptoms observed in all counties in order of abundance, starting from the most common, were chlorotic rosette, green rosette and mosaic. Chlorotic rosette was recorded in 58.6% of the surveyed farms, green rosette in 27.4% while mosaic was observed 14.1% of farms (Table 2). Other symptoms observed were leaf rolling, upward leaf curling and severe leaf bunching (Figure 1). The distribution of the major symptoms is shown in Figure 2.

Diversity of GRV Sat-RNA

Six complete genomes of GRV Sat-RNA were assembled. The sequences varied slightly in the number of nucleotides (nt) ranging between 896 and 901 nt (Table 3).

The six Sat-RNAs from Kenya were then compared with those from the GeneBank. In the phylogenetic tree, all Kenyan isolates formed two distinct clusters together with Malawian isolates. Isolates E7 and E8 clustered with M11S, isolates BUG1-21, BG3-18 and KG8-1 clustered together with M16S while isolate EG16-5 is grouped with

Figure 1. Some of the virus-like symptoms observed in the surveyed fields. A: Dwarfed plant with green rosette; B: Severe chlorosis (yellow) on young leaves and dwarfing; C: Severe young leaf rolling and bunching on a dwarfed plant; D: Mosaic mostly on young leaves.

M24S. All Nigerian isolates clustered together is similar to Ghanaian isolates. Sequence identities between 88 and 100% of the Kenyan isolates and those from Malawi, Nigeria and Ghana were revealed. Very close identities between 92 and 100% were observed between the Kenyan isolates and those from Malawi, followed by Nigerian isolates (90-93%) and least with Ghanaian isolates (86-89%). Isolate BUG1-21 had 100, 99 and 98% identities with M16S, M12S, and M11S, respectively, which are all green rosette variants and 94% with M24S (chlorotic variant). While the other Western Kenya isolates (KG8-1, BUG1-21, BG3-18, E7 and E8) had 92 to 95% identity with Malawian isolate M24S (chlorotic rosette variant), isolate EG16-5 (Kakamega) showed the closest identity (97%) with this isolate. The same isolate EG16-5 was the only that clustered together with M24S, all chlorotic isolates (Z29702.1, Z29703.1) and yellow blotch (Z29710.1, Z29711.1). Isolates E7 and E8 were closest to Malawian isolate M11S with 97 and 99% identity, respectively. Isolates BG3-18 and KG8-1 were closest to Malawian isolates M16S displaying 97% identity (Figure 3).

DISCUSSION

Groundnut rosette disease is the most prevalent disease of groundnuts in Western Kenya. The disease was recorded in every county that was surveyed with incidences of up to 100%. The short rain season recorded higher incidence (53%) than the long rains (41%). This can be attributed to the high vector pressure during the short rains as compared to the long rains season when the aphid pressure is low as a result of heavy rains that wash the insects away. This concurs with the findings by Mugisa et al. (2016) that periods of long rains negatively affected GRD progression when aphid pressure is low.

All major GRD symptoms were observed in the surveyed region with chlorotic rosette being the most prevalent followed by green rosette. This concurs with Wangai et al. (2001) who reported chlorotic rosette to be the most prevalent GRD symptom in the region. The high prevalence of the chlorotic rosette could also be attributed to its higher transmission efficiency compared to green rosette. In a study, Misari et al. (1988a), reported

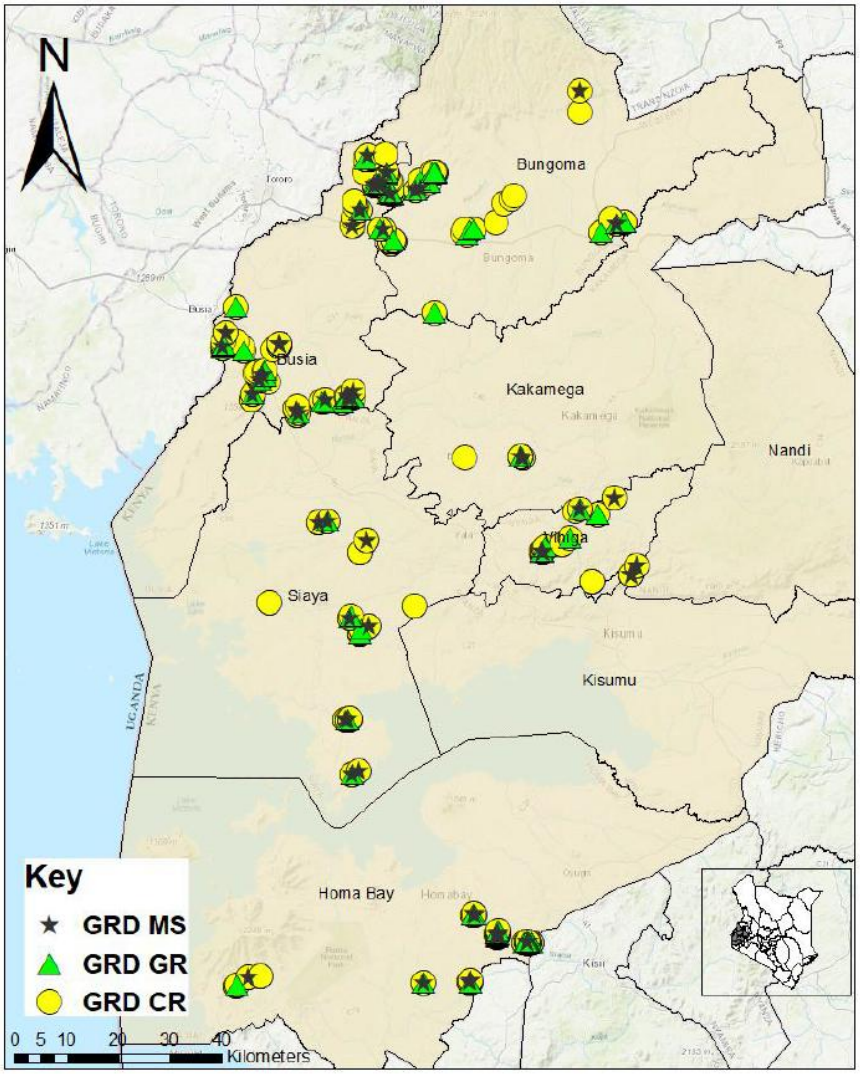


Figure 2. A map of Western Kenya showing the distribution of GRD symptoms in the surveyed counties. **MS**-Mosaic, GR-Green Rosette, CR-Chlorotic Rosette.

minimum acquisition feeding periods of 4 and 8 h for chlorotic and green rosette, respectively and the median latent periods of 26.4 and 38.4 h, respectively, for chlorotic and green rosette. In this study, a new symptom, the mosaic, which had not been previously reported in Western Kenya, was observed in most of the

Table 3. Description of the Sat-RNA sequences assembled.

Sample ID	Sat-RNA ID	Sequence length (nt)	County of origin
EG16	EG16-5	901	Kakamega
E7	E7	896	Siaya
E8	E8	897	Busia
BUG1	BUG1-21	901	Busia
KG8	KG8-1	898	Kakamega
BG3	BG3-18	901	Bungoma

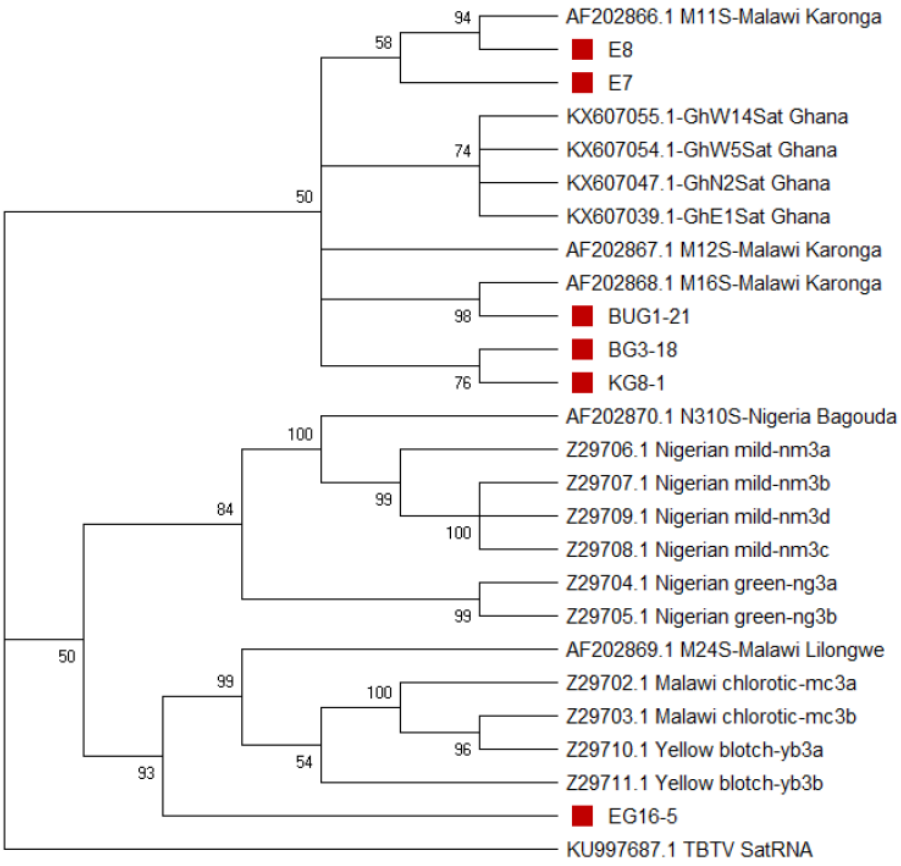


Figure 3. Phylogenetic tree of Western Kenya Sat-RNA and GeneBank isolates. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model (Tamura and Nei, 1993). The tree is rooted on Sat-RNA of a distantly related *Umbravirus* (Tobacco bushy top virus - KU997687.1 TBTV). Bootstrap confidence values (500 replications) are shown.

surveyed counties. This suggests that there is evolution of new variants of Sat-RNA that might be causing these new symptoms. A total of 10 variants of Sat-RNA have been reported as being associated with the various GRD symptoms (Blok et al., 1994). A mixture of either variants, especially the chlorotic and green rosette and/or the mild ones, are likely to induce the mosaic symptoms (Naidu et al., 1998). It is therefore possible that the variants of sat-RNA reported in this study occur in Western Kenya in mixed infections, thus causing the mosaic observed. It is worth noting that from the Next Generation Sequences (NGS) used in this study, order than GRV Sat-RNA, other viruses were detected (data not shown) and could be the reason for some of the new symptoms observed on groundnuts (Mukoye et al., 2018).

The Western Kenya Sat-RNAs sequences showed close identity (92-100%) to Malawian isolates than those from Ghana and Nigeria (88-93%). This implies that the genetic diversity of the Sat-RNA become more varied with wide geographical distance. Kenya and Malawi are located in Eastern Africa while Ghana and Nigeria are in West Africa thus having a wider geographical separation than Malawi. This finding is in line with Wangai et al. (2001) who observed a closer sequence relationship between Kenyan Sat-RNA isolates and those from Malawi. However, this study has reported sequence identity of up to 100% with Malawian isolates as opposed to 95% reported by Wangai et al. (2001). This suggests that more variants of Sat-RNA exist in Western Kenya that are contributing to the diverse symptoms expressed by GRD. Since this study used NGS which has been demonstrated to be more reliable in detection of new or poorly characterized viruses (Rott et al., 2017), it has revealed new variants of Sat-RNA in Western Kenya. Besides, there were variations among the Western Kenya Sat-RNA isolates similar to Malawian isolates where they formed distinct clusters in the phylogenetic tree. The isolate EG16-5 was the most distinct and clustered together with chlorotic and yellow blotch Sat-RNA variants. This suggests that this isolate is related to the chlorotic rosette symptom that was prevalent in the surveyed areas.

This study concludes that GRD is still the major viral disease of groundnuts in Western Kenya and occurs wherever groundnuts are grown in the region. The disease expresses varied symptoms with chlorotic rosette being the most prevalent form. The observed variations in the symptoms were due to the presence of diverse variants of the symptom inducing agent, Sat-RNA. The use of NGS has revealed that new variants of Sat-RNA exist in Western Kenya.

The six Kenyan Sat-RNAs have been deposited in the GeneBank with accession numbers LC469779, LC472299, LC472300, LC472301, LC472302 and LC472303.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

ACKNOWLEDGEMENTS

This work was funded by Royal Society of UK, International Foundation for Science, National Research Fund (NRF-Kenya) and ILRI BecA hub. They are grateful to Dr. Wellington Ekaya and the entire BecA capacity building team for allowing this work to be done at the BecA labs.

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. Journal of Molecular Biology 215:403-410.

Blok VC, Ziegler A, Robinson DJ, Murant AF (1994). Sequences of 10 variants of the satelitte- like RNA -3 of groundnut rosette virus. Virology 202:25-32.

Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics, btu170.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, LeDuc RD, Friedman N, Regev A (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature Protocols 8:1494-1512.

Kayondo SI, Rubaihayo PR, Ntare BR, Gibson PI, Edema R, Ozimati A, Okello DK (2014). Genetics of resistance to groundnut rosette virus disease: African Crop Science Journal 22:21-29.

Kidula N, Okoko N, Bravo-Ureta BE, Thuo M, Wasilwa L (2010). A preliminary analysis of yield differences in groundnuts between research and non-research farmers in Kenya. In paper presented at the 12th KARI biennial scientific conference, 8-12 November 2010, Naiobi Kenya.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular Biology and Evolution 35:1547-1549.

Kumar PL, Waliyar F (2007). Diagnosis and detection of viruses infecting ICRISAT mandate crops: Methods Manual. Patancheru 502 324, Andhra Pradesh, India; International Crops Research Institute for the Semi-Arid Tropics.133 p.

Misari SM, Abraham JM, Demski JW, Ansa OA, Kuhn CW, Casper R, Breyel E (1988a). Aphid transmission of the viruses causing chlorotic and green rosette diseases of peanut in Nigeria. Plant Disease 72:250-253.

Mugisa IO, Karungi J, Akello B, Ochwo-Ssemakula MKN, Biruma M, Okello DK, Otim G (2016). Determinants of groundnut rosette virus disease occurrence in Uganda. Elsevierciopprotectionjournal.http://dx.doi.org/10.1016/j.cropro.2015.10.019.

Mukoye B, Mangeni BC, Sue J, Ndong'a MFO, Were HK (2018). Next generation sequencing as a tool in modern pest risk analysis: a case study of groundnuts (*Arachis hypogaea*) as a potential host of new viruses in Western Kenya. Kenya Plant Health Inspectorate Service 2nd phytosanitary conference, 4th – 8th June, 2018, Nairobi – Kenya.

Mutegi CK (2010). The extend of aflatoxin and aspergillus section flavi, penicillium spp. and Rhizopus spp. contamination of peanuts from households in Western Kenya and the causative factors of contamination. PhD dissertation, University of Kwazulu-Natal,Pietermaritzburg . South Africa.

Naidu RA, Bottenberg H, Subrahmanyam P, Kimminns FM, Robinson DJ, Thresh JM (1998). Epidemiology of groundnut rosette virus disease: Current status and future research needs. Annals of Applied Biology 132:525-548.

Naidu RA, Kimmins FM, Deom CM, Subrahimanyam P, Chiyeubekeza AJ, van der Merwe PJA (1999). Groundnut rosette: Avirus disease affecting groundnut production in Sub-Saharan Africa. Plant Disease 83:700-709.

Okello DK, Birima M, Deom CM (2010). Overview of groundnut

research in Uganda: Post, present and future. African Journal of Biotechnology 9:6448-6459.
Okello DV, Akello LB, Tukamuhabwa P, Odongo TL, Ochwo-Ssemakula M, Adriko J, Deom CM (2014). Groundnut rosette disease symptom types, distribution and management of the disease in Uganda. African Journal of Plant Science 8:153-163.
Olorunju PE, Ntare BR, Pande S, Reddy SV (2001). Additional sources of resistance to groundnut rosette disease in groundnut germplasm and breeding lines. Annals of Applied Biology 159:259-268.
Reddy DVR (1991). Groundnut viruses and virus diseases; Distribution, identification and control. Review of Plant Pathology 70:665-678.
Rott M, Xiang Y, Boyes I, Belton M, Saeed H, Kesanakurti P, Hayes S, Lawrence T, Birch C, Bhagwat B, Rast H (2017). Application of Next Generation Sequencing for diagnostic testing of tree fruit viruses and viroids. Plant Disease 101:1489-1499.
SAS Institute (2013). SAS/STAT 9.3.1: User guide. SAS Publishers, India.
Scott KP, Farmer MJ, Robinson DJ, Torrence L, Murant AF (1996). Comparison of the coat protein of groundnut rosette assistor virus with those of other *luteovirus*. Annals of Applied Biology 128:77-83.
Taliansky ME, Robinson DJ (1997). Trans-acting untranslated elements of groundnut rosette virus satelitte RNA are involved in symptom production. Journal of General Virology 78:1277-1285.
Taliansky ME, Robinson DJ (2003). Molecular Biology of umbraviruses: Phantom warriors. Journal of General Virology 84:1951-1960.
Tamura K, Nei M (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution 10:512-526.
Waliyar F, Kumar PL, Ntare BR, Monyo E, Nigam SN, Reddy AS, Osiru M, Diallo AT (2007). A Century of Research on Groundnut Rosette Disease and its Management. Information Bulletin no.75.Patancheru 502 324, Andhra Pradesh, India. International Crops Research Institute for the Semi-Arid Tropics 40 p. ISBN 978-92-9066-501-4.
Wangai AW, Pappu SS, Pappu HR, Okoko N, Deom CM, Naidu RA (2001). Distribution and characteristics of groundnut rosette disease in Kenya. Plant Disease 85:470-474.