MOLECULAR CHARACTERIZATION OF VIRUSES INFECTING COMMON BEAN (*Phaseolus vulgaris* L.) AND REACTION OF BEAN GENOTYPES TO VIRUS INFECTION

RABSON MPUNDU MULENGA

(B.Sc. Biological and Chemical Sciences, University of Zambia, M.Sc. Crop Protection, University of Nairobi)

A THESIS SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS

FOR THE AWARD OF THE DEGREE OF

DOCTOR OF PHILOSOPHY IN PLANT PATHOLOGY

DEPARTMENT OF PLANT SCIENCE AND CROP PROTECTION

FACULTY OF AGRICULTURE

UNIVERSITY OF NAIROBI

2022

DECLARATION

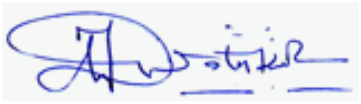
This thesis is my original work and has not been submitted for award of a degree in any other University

Rabson Mpundu Mulenga

Signature  Date: 10th July 2022

This dissertation has been submitted with our approval as University supervisors

Prof. Douglas W. Miano
Department of Plant Science and Crop Protection,
University of Nairobi

Signature  Date: 10th July, 2022

Dr. Felister M. Nzuve
Department of Plant Science and Crop Protection,
University of Nairobi

Signature  Date: 14th July, 2022

Dr. Olufemi J. Alabi
Department of Plant Pathology and Microbiology
Texas A&M AgriLife Research and Extension Center
WeslacoTexas A&M University

Signature.  Date: 12th July, 2022

Dr. Evans Kaimoyo
Department of Biological Sciences
University of Zambia

Signature  Date: 13th July 2022

DECLARATION OF ORIGINALITY

Name of Student	Rabson Mpundu Mulenga
Registration Number	A80/52585/2018
College	College of Agriculture and Veterinary Sciences
Faculty	Faculty of Agriculture
Department	Department of Plant Science and Crop Protection
Course Name	Doctor of Philosophy in Plant Pathology
Title of Work	Molecular characterization of viruses infecting common bean (*Phaseolus vulgaris* L.) and reaction of bean genotypes to virus infection

Declaration

1. I understand what plagiarism is and I am aware of the University's policy.
2. I declare that this thesis is my original work and has not been submitted for examination, award of a degree or publication. Where other people's work or my own work has been used, this has properly been acknowledged and referenced in accordance with the University of Nairobi's requirements.
3. I have not sought or used the services of any professional agencies to produce this work.
4. I have not allowed and shall not allow anyone to copy my work with the intention of passing it off as his or her own work.
5. I understand that any false claim in respect of this work shall result in disciplinary action, in accordance with University Plagiarism Policy.

Signature  Date: 10th July 2022

DEDICATION

I dedicate this work to my wife Grace Lombe Mulenga and our children Kabwe, Samuel, Precious and Janina for their encouragement and support during the challenging times of my study. The mere questioning of how my day was by Kabwe, Samuel and Precious and the muttering of words of welcome from the one-year-old Janina gave me sobering and refreshing moments every time I returned home after a long absence. May the Almighty God allocate long fruitful lives to all of you. Finally, to my mother Alufonshina Musonda for providing the foundation for my education and to all my family members and church mates for prayers and moral support.

ACKNOWLEDGEMENTS

I wish to sincerely thank my supervisors Prof. Douglas W. Miano, Dr. Felister M. Nzuve, Dr. Evans Kaimoyo and Dr. Olufemi J. Alabi for their professional guidance, expertise, and support during my studies. I also want to thank my Mentor Dr. Juliet Akello for instilling in me confidence in the face of challenges during my research work, financial support and for all the valuable guidance provided to me. Let me also extend my deepest thanks to Dr. Patrick C. Chikoti for facilitating unhindered access to laboratory facilities.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFLP	Amplified fragment length polymorphism
AMV	Alfalfa mosaic virus
BCMV	Bean common mosaic virus
BCMNV	Bean common mosaic necrosis virus
BDMV	Bean dwarf mosaic virus
BYMV	Bean yellow mosaic virus
CABMV	Cowpea aphid-borne mosaic virus
CalGMV	Calopogonium golden mosaic virus
BGYMV	Bean golden yellow mosaic virus
CMV	Cucumber mosaic virus
CMoMV	Carrot mottle mimic virus
CMoV	Carrot mottle virus
CP	Capsid protein
CPMMV	Cowpea mild mottle virus
CPPV3	Cowpea polerovirus 3
DAS-ELISA	Double sandwich enzyme linked immunosorbent assay
DNA	Deoxynucleic acid
ETBTV	Ethiopian tobacco bushy top virus
FAOSTAT	Food and Agriculture Organization Corporate Statistical database
gRNA	Genomic ribonucleic acid
HTS	High-thoughput sequencing
ICTV	International Committee on Taxonomy of Viruses
LAC	Latin American Countries
NTR	Nontranslated region

OPMV	Opium poppy mosaic virus
ORF	Open reading frame
PCR	Polymerase chain reaction
PG	Pathogroup
PeMoV	Peanut mottle virus
PEMV-2	Pea enation mosaic virus 2
PSbMV	Pea seed-borne mosaic virus
PvEV-1	Phaseolus vulgaris endornavirus 1
PvEV-2	Phaseolus vulgaris endornavirus 2
RNA	Ribonucleic acid
RAPD	Random amplified polymorphic DNA
RFLP	Restriction fragment length polymorphism
RT-PCR	Reverse transcription polymerase chain reaction
SBMV	Southern bean mosaic virus
SNP	Single nucleotide polymorphism
SSA	sub-Saharan Africa
sgRNA	Subgenomic RNA
SSR	Simple-sequence repeats
ssRNA	Single stranded RNA
satRNA	Satellite RNA
TBTV	Tobacco bushy top virus
TAS-ELISA	Triple antibody sandwich enzyme-linked immunosorbent assay
TNA	Total nucleic acid
UTR	Untranslated region
ZSA	Zambia Statistical Agency

GENERAL ABSTRACT

Virus disease symptoms are frequent in common bean (*Phaseolus vulgaris* L.) fields in Zambia, but information is scanty about the identities, distribution and genetic diversities of causal pathogens. This study was conducted to survey for common bean virus-like diseases in farmers' fields in Zambia, conduct molecular characterization of viruses identified and screen Zambian common bean cultivars for resistance to bean common mosaic necrosis virus (BCMNV) and bean common mosaic virus (BCMV). To address these knowledge gaps, surveys were conducted from March to May of 2018 in 128 common bean fields across six provinces of Zambia located in agro-ecological zones (AEZs) II and III. In total, 640 leaf tissue samples (symptomatic = 585; non-symptomatic = 55) were collected for virus diagnoses. From these, a subset of 223 samples that were selected based on symptom diversity and disease severity were subsumed into nine composite samples and subjected to total nucleic acid (TNA) extractions. Each of the nine TNA samples were diagnosed by high throughput sequencing (HTS) and the generated sequence reads were bioinformatically analyzed. Subsequently, the 640 samples were screened for the HTS-detected viruses using reverse transcription polymerase chain reaction (RT-PCR) and the results were validated by Sanger sequencing. Analysis of the combined HTS data obtained from composite samples produced nine distinct viruses belonging to five genera (*Potyvirus*, *Cucumovirus*, *Endornavirus*, *Sobemovirus* and *Umbravirus*). Screening of the 640 samples showed that 67% (429/640) of the samples were positive for at least one of the nine viruses either as single (65.1%; 417/640) or mixed (~1.9%; 12/640) infections. Southern bean mosaic virus (SBMV) was the most frequently detected virus accounting for ~29.4% (188/640) of the samples, followed by phaseolus vulgaris endornavirus1 (PvEV-1) ~9.2% (59/640). The remaining seven viruses, BCMV, BCMNV, cowpea aphidborne mosaic virus (CABMV), cucumber mosaic virus (CMV), Ethiopian tobacco bushy top virus (ETBTV), peanut mottle virus

(PeMoV) and PvEV-2, occurred at incidences of 0.3% (2/640) for CABMV to 7.7% (49/640) for BCMNV and PvEV-2. Across the AEZ, there was more virus diversity in AEZ II (5 to 8 viruses) than in AEZ III (3 to 5 viruses).

De novo assembly of the HTS generated sequence reads resulted in 24 virus-aligned sequences (BCMV=3, BCMNV=3, CABMV=2, ETBTV=2, PeMoV=2, CMV=1, PvEV-1=1, PvEV-2=1, SBMV=9). In pairwise comparison, ETBTV, SBMV and PeMoV sequences shared 88 to 99.4% nucleotide (nt) identities with corresponding sequences of respective global viruses whereas sequences of viruses CMV, BCMNV, BCMV, PvEV-1 and PvEV-2 shared 94 to 100% nt identities with corresponding global sequences. Further analyses revealed that the three BCMV sequences are putative recombinants whereas PeMoV isolate CP-com-1 and SBMV sequence Mse-3 are putative mutants.

Phylogenetic analyses of the 24 virus sequences and global sequence homologs showed that virus sequences from this study clustered severally on the phylogenetic trees. For example, BCMNV and BCMV formed clusters with global isolates of known BCMNV and BCMV pathogroups (PGs). Thus, to establish PGs of BCMNV and BCMV from this study, the two viruses in the bean samples were assayed in standard differential common bean cultivars. Four PGs (I, III, VIa and VIb) were identified with the occurrence of PGs I and III in Zambia being reported for the first time in this study. The identified PGs were used to screen 14 common bean cultivars for resistance to BCMNV and BCMV. Two released varieties Lwangeni and Lunga that carry the resistance gene *bc-3* were resistant to all four PGs whereas those bearing resistance gene *bc-1*^{2} were susceptible to viruses in PGs VIa and VIb. Therefore, farmers are encouraged to plant the two varieties especially in BCMNV and BCMV hotspots. The results from this study will be used to design diagnostic tools for detecting vommon bean viruses in Zambia.

CHAPTER ONE

INTRODUCTION

1.1 Background information

Common bean (*Phaseolus vulgaris* L.) is a vital source of macro- and micro-nutrients in low-income countries (Katungi *et al.,* 2009; Celmeli *et al.,* 2018; Rawal and Navarro, 2019). It is a low input crop predominantly grown by small-scale farmers usually on less than one hectare of land in most sub-Saharan African (SSA) countries (Ampofo and Massamo, 1998). The small-cultivated areas in turn often produce enough for consumption with excess sold for income generation.

In the developing world and in Zambia in particular, low yields caused by abiotic and biotic factors are common in farmers' fields. These factors coupled with little or no fertilizer inputs substantially reduce productivity (Katungi *et al.,* 2009). The abiotic factors include poor soils, drought, leached soils, and high soil salinity (Beebe *et al.,* 2014). The main biotic factors are arthropod pests and various fungal, bacterial, and viral diseases (Wortmann *et al.,* 1998; Hillocks *et al.,* 2006; Buruchara *et al.,* 2010; Akhavan *et al.,* 2013). Insect pests and fungal/bacterial pathogens can attack the crop at the vegetative growth stage as well as in storage. These can be countered by treatment with chemical and biological control agents. However, viral pathogens affect the crop mainly at the vegetative growth stage and they lack therapeutics once the plant becomes infected.

1.2 Viruses infecting common bean

Common bean-infecting viruses have been a subject of several reviews with the most recent by Worrall *et al.* (2015). In Africa, several viruses have been reported from the crop, including bean common mosaic necrosis virus (BCMNV; genus *Potyvirus*), bean common mosaic virus

(BCMV; genus *Potyvirus*, Vetten and Allen, 1991; Njau and Lyimo, 2000), southern bean mosaic virus (SBMV; genus *Sobemovirus*) (Mwaipopo *et al.,* 2018), Phaseolus vulgaris endornavirus1 (PvEV-1; genus *Alphaendornavirus*, Nordenstedt *et al.,* 2017), Phaseolus vulgaris endornavirus2 (PvEV-2; genus *Alphaendornavirus*, Nordenstedt *et al.,* 2017), cowpea aphid-borne mosaic virus (CABMV; genus *Potyvirus*, Mwaipopo *et al.,* 2018; Wainaina *et al.,* 2019), cucumber mosaic virus (CMV; genus *Cucumovirus*, Mutuku *et al.,* 2018), bean yellow mosaic virus (BYMV; genus *Potyvirus*, Vetten and Allen, 1991) and alfalfa mosaic virus (AMV; genus *Alfamovirus*, Spence and Walkey, 1995) peanut mottle virus (PeMoV; genus *Potyvirus*, Vetten and Allen, 1991; Spence and Walkey, 1995; Mwaipopo *et al.,* 2018).

1.3 Statement of the problem

The production of common bean in Africa is beset by low yields of <1 t/ha compared with 1 to ≥2 t/ha reported for South America, Europe and North America (FAOSTAT, 2018). Although country specific yield data varies from a mean of 0.5 to 2.6 t/ha across the African continent, mean yield data for the last decade has consistently remained below 1 t/ha. The suboptimal yields are partly attributable to diseases caused by viruses. Viral symptoms are frequent in bean fields Africa-wide and are usually attributed to BCMNV and BCMV; the two extensively studied viruses (Worrall *et al.,* 2015). However, studies have shown that BCMNV and BCMV occur in less than 40% of viral infections in common bean fields in Africa with 60% caused by unknown viruses (Segundo *et al.,* 2008; Mwaipopo *et al.,* 2018). Bean-infecting viruses reported in previous studies in Zambia include CMV, PeMoV, BCMNV and BCMV. However, the geographical distributions of these viruses are unknown, making it difficult to manage the diseases they induce across the country. The knowledge gap also makes it difficult to institute interventions aimed at limiting the negative impact of disease infections on yield. Further, there

is no genome sequence data for common bean-infecting viruses in Zambia, making it difficult to develop robust molecular diagnostic tools that can be deployed for their routine diagnosis. Typically, all the reported viruses from Zambia were detected using immunological assays. However, such detection methods exclude viruses for which there are no antisera or where certain viruses possess similar coat protein epitopes leading to possible cross-reactivities of the detection antibodies (Kushwaha *et al.,* 2010). Further, serological assays seldom distinguish between strains and genetic variants. Thus, the characterization of common bean viruses in Zambia remain incomplete and thus knowledge of the virome of bean viruses circulating in farmers's fields in the country is desirable. This study was conceived to plug the critical knowledge gap using high-throughput sequencing technology to mine novel and known viruses across Zambian common bean fields.

Resistance to BCMNV and BCMV is regulated by the dominant *I* gene and six recessive genes; *bc-u,* allelic genes (*bc-1, bc-1*^{2}, *bc-2, bc-2*^{2}) and *bc-3*. The distribution of these genes in commonly grown Zambian bean cultivars is not known. Even for cultivars that may carry the resistance genes (R-genes), their continued effectiveness against genetic variants of both BCMNV and BCMV needs to be evaluated. Further, whereas the stated R-genes confer resistance to BCMNV and BCMV, their reactions to other common bean-infecting viruses is unknown. Therefore, even in cases where BCMNV and BCMV are managed through resistant cultivars, other viruses may potentially remain pathogenic and continue to cause suboptimal yields. Thus, this study is needed to provide vital information useful in managing common bean viral diseases.

1.4 Justification

Understanding the geographical distribution of virus diseases in common bean farmers' fields, identifying their causal pathogens and their molecular diversity is the first logical step to developing disease management strategies. A study of geographical distribution of common bean-infecting viruses is needed to provide information on their cold and hotspots. Moreover, molecular characterization of viruses infecting common bean is useful in developing diagnostic tools that can be used in disease diagnosis. Bean common mosaic necrosis virus and Bean common mosaic virus are still the two most important viruses of common bean in SSA. Screening of Zambian common bean cultivars for resistance to BCMNV and BCMV is useful for identifying cultivars with desirable resistance to the two viruses and also to help in identifying sources of genetic registance that will be included in future breeding efforts. Information generated in this study will be widely used by plant pathologists, breeders and farmers in managing diseases of virus aetiology.

1.5 Study objectives

The broad objective of this study was to determine the virome of Zambian common bean fields and generate information useful for management of viral diseases.

Specific objectives of the study were:

i. To determine distribution of common bean-infecting virus diseases in farmers' fields in Zambia.
ii. To characterize identified viruses using molecular techniques.
iii. To evaluate Zambian common bean cultivars for resistance to bean common mosaic necrosis virus and bean common mosaic virus.

1.6 Hypotheses

i. Virus diseases and their causal pathogens are not widely distributed in farmers' fields in Zambia.

ii. Diversity of virus species is similar across different agro-ecological zones.

iii. Sources of genetic resistance to bean common mosaic necrosis virus and bean common mosaic virus do not exist within the gene pool of common bean cultivars in Zambian.

CHAPTER TWO

LITERATURE REVIEW

2.1 History of common bean crop

Common bean (*Phaseolus vulgaris* L.) traces its origins to the wild bean species in the Mesoamerican and Andean neotropic ecologies (Singh *et al.,* 1991; Chacón *et al.,* 2005; Bitocchi *et al.,* 2012; Bitocchi *et al.,* 2013; Bellucci *et al.,* 2014; Schmutz *et al.,* 2014). In recent studies, the multi-location origin of common bean was supported by evidence provided in a combined multi-loci sequence data and morphological traits analyses (Rendón-Anaya *et al.,* 2017). It is thought that its original wild progenitor may have been repeatedly selected during the domestication period (Rendón-Anaya *et al.,* 2017).

The genus *Phaseolus* L. (Family: *Fabaceae*) comprises a large number of cultivated and wild species described as bush, dwarf or climbing plants according to the growth habit (OECD, 2016). A complete description of the floral characteristic features of *P. vulgaris* is given by Wortmann (2006) and it showed that the crop shares many features of the family. However, the coiled and single or twin-turn termini of the keel of the flower are exclusive to *P. vulgaris* (Gentry, 1969). To date, five species, *P. dumosus* Macfady, *P. coccineus*, *P. vulgaris*, *P. lunatus* and *P. acutifolius* are grown globally (Bellucci *et al.,* 2014).

Phaseolus vulgaris is classified as either Mesoamerican (small seeded) or Andean (big seeded) Diversity Panel (ADP) (Mamidi *et al.,* 2011). Most of the Mesoamerican gene pool is cultivated in North America with the ADP prevalent across Europe, South America and parts of Africa (Cichy *et al.,* 2015).

2.2 Importance of common bean

Dry bean (common bean) is a reliable dietary protein and starch source for most people in Latin American countries (LAC), the Caribean, sub-Saharan Africa (SSA) and Asia (Rawal and Navarro, 2019). It is also a good source of needed micronutrients among them zinc, selenium, iodine and iron (Broughton *et al.,* 2003). Over 300 million people in LAC and SSA consume common bean (Petry *et al.,* 2015). In SSA alone, more than 200 million people depend on common bean as a primary staple (Schmutz *et al.,* 2014). In Africa, Eastern Africa has the highest per capita consumption of common bean (Petry *et al.,* 2015; Rawal and Navarro, 2019). In Zambia, like many African countries, data on common bean consumption levels are underestimated due to consumption data compilation challenges (Hotz and McClafferty, 2007; Blair *et al.,* 2010). However, consumption levels are higher in wealthier households than lower income households in Zambia (Hichaambwa *et al.,* 2009).

2.3 Production of common bean in Zambia

Over the past decade (2008-2018), global production of common bean increased from an average of 22 to 27 million tons at a calculated mean yield of 1.1 t/ha (Table 2.1). Africa-wide production of common bean increased from an average of 4 to ~7 million tons at a calculated yield of 0.9 t/ha in 2018 (FAOSTAT, 2018) (Table 2.1). In Africa, the largest producers of common bean are Eastern African countries led by Tanzania (Mwaipopo *et al.,* 2018; Rawal and Navarro, 2019). Farmers within each country mainly produce for subsistence (Petry *et al.,* 2015; Rawal and Navarro, 2019) with marginal surplus to sell. Country specific data also show progressive increase in bean production in SSA during the last decade (Table 2.1). In Zambia, common bean cultivation increased from 59,586 ha to 83,635 ha but averaged nearly 46,000 tons during the last decade (2008-2017) according to the Zambia Statistical Agency (ZSA, 2017),. However, average

yields of 0.5 t/ha are common in Zambia (Hamazakaza *et al.*, 2014) and are nearly two-fold lower than the Africa-wide average yield of 0.9 t/ha (FAOSTAT, 2018) and four-fold below the experimental yield potential (2 t/ha) for most cultivars. Such low yields could be caused by a composite of biotic and abiotic constraints.

Table 2.1. Global production of common bean (dry beans) between 2008 and 2018.

		Yield (t/ha)		**Total Production (Tonnes)**	
Region	**Sub-region**	**Year (2008)**	**(2018)**	**Year (2008)**	**(2018)**
Africa		0.75*	0.87*	4,197,706*	7,117,815*
	Eastern Africa	0.77	0.9	3,044,478	5,007,411
	Northern Africa	-	5.1	115,057	139,424
	Southern Africa	1.1	1.0	63,198	77,089
	Western Africa	0.66	0.96	400,88	745,195
America		1.0*	1.0*	7,498,284*	7,475,729*
	North America	2.0	2.4	1,425,490	1,449,220
	South America	0.95	1.0	4,212,782	3,831,158
	Central America	0.8	0.8	1,676,145	1,893,889
	Caribbean	0.8	0.5	183,867	301,462
ASIA		0.7*	0.6*	9,710,726*	12,361,630*
	Central ASIA	0.7	1.1	72,742	532,151
	Eastern ASIA	1.4	1.5	2,134,267	1,720,238
	Southern ASIA	0.4	0.4	3,430,377	6,556,363
	Western ASIA	1.7	2.5	191,43	250,092
Europe		1.7*	1.8*	393,293*	381,267*
	Eastern Europe	1.8	1.9	246,407	328,709
	Southern Europe	1.7	1.5	133,009	52,558
	Western Europe	2.6	-	8,669	-
Total(mean)		(1.0*)	(~1.1*)	21,800,009	27,336,441

*=numbers represented are global and include regions not presented in the table. Source of data: FAOSTAT 2018.

2.4 Constraints to production of common bean

Common bean is produced in diverse cropping systems covering both monomodal and bimodal rainfall patterns (Wortmann *et al.,* 1998). Yields are suboptimal in many African countries and this is attributed to biotic and abiotic stresses. Abiotic stresses include high soil salinity, high temperatures, drought, and severely leached soils mostly in high rainfall ecologies where soil acidity is common (Graham and Vance, 2003; Beebe *et al.,* 2014). Biotic stresses that cause significant yield losses include arthropod pests and several diseases caused by fungal, bacterial and viral pathogens (Wortmann *et al.,* 1998). Several arthropod pests of common bean have been documented in SSA attacking various parts of the bean plant and causing differing levels of damages (Wortmann *et al.,* 1998; Buruchara *et al.,* 2010). Singh and Schwartz (2010) published a catalogue of diseases of common bean and their yield reducing potential. Notably, in severe cases, plant viruses can induce yield losses of up to 100% (Hagedorn and Inglis, 1986). Viruses occupy a special place as a constraint to production in that whereas diseases caused by other pathogen types can be controlled by chemical treatment, viral diseases lack such therapeutics.

2.5 Historical perspectives of viruses infecting common bean

The dispersal of the two Mesoamerican and Andean gene pools of common bean across the world in the 16th century introduced common bean to many different cropping systems. Fortunately, common bean is adaptable to many cropping systems (Raatz *et al.,* 2019) but it is also vulnerable to many diseases, among them several viruses (Morales, 2006). Viruses have been associated with common bean diseases since the end of the 19th century (Morales, 2006). Zaumeyer and Thomas (1957) detailed several viruses associated with various symptoms in a monographic study of bean diseases. They include alfafa dwarf mosaic virus (ADMV), curly top virus, pod mottle virus, southern bean mosaic virus (virus 4), common bean mosaic virus (virus

1), yellow bean mosaic virus (virus 2) and others. Although the nomenclature for some of the viruses has changed over time, they are globally distributed (Drijfhout *et al.,* 1978; Vetten *et al.,* 1992; Flores-Este´vez *et al.,* 2003; Verhoeven *et al.,* 2003; Morales, 2006; Pasev and Kostova, 2015; Worrall *et al.,* 2015; Johary *et al.,* 2016; Mwaipopo *et al.,* 2018).

Among plant viruses frequently detected in infected common bean plants, BCMNV and BCMV are the two major yield limiting and consequently ecomincally impactful viruses (Worrall *et al.,* 2015). In Africa, most reports on common bean viruses are from East Africa particularly Kenya and Tanzania. Due to the destructive nature and economic importance of BCMNV (formerly serotype A) and BCMV (formerly serotype B), earlier studies focused on the two viruses to the neglect of other viruses. The earliest report of serotypes of BCMV in Kenya was in 1973 and 1936 in Tanzania (Kulkarni and Muguga, 1973). A series of surveys by Vetten and Allen (1991) in many sub-Saharan African countries formed the basis for country specific future studies. For example, Omunyin *et al.* (1995) and Mangeni *et al.* (2014) studied the distribution of BCMNV and BCMV in different localities of Kenya whereas several studies were conducted in Tanzania (Vetten and Allen, 1991; Spence and Walkey, 1995; Myers *et al.,* 2000; Njau and Lyimo, 2000; Mwaipopo *et al.,* 2018) and Uganda (Sengooba *et al.,* 1997). In Zambia, apart from the limited studies by Vetten and Allen (1991) and Spence and Walkey (1995), there has been no comprehensive study of common bean-infecting viruses and allusion to viruses of common bean in Zambia is limited to BCMNV and BCMV.

Other than BCMNV and BCMV, literature is replete with other bean-infecting viruses (Vetten and Allen, 1991; Sengooba *et al.,* 1997; Shahraeen *et al.,* 2005; Mutuku *et al.,* 2018; Mwaipopo *et al.,* 2018; Wainaina *et al.,* 2019). Morales (2006) documented some of the viruses infecting

common bean globally with their detections based mostly on traditional diagnostic assays. In a more recent comprehensive survey that utilized advanced diagnostic tools, Mwaipopo *et al.* (2018) reported 15 viruses belonging to 11 genera in Tanzania. In Kenya, Mutuku *et al.* (2018) also reported presence of CMV, PvEV-1, PvEV-2, and BCMNV based on metagenomics analysis of collected samples. With the increased application of more sensitive culture dependent (PCR and RT-PCR) and independent (HTS) detection tools, the virome landscape of most common bean farming systems is just evolving.

In Zambia, limited information exists on the presence and spread of CMV, PeMoV, BCMNV and BCMV with the detections of these viruses having been achieved through studies by Vetten and Allen (1991) and Ndomba (2002). Further sequence data on the reported viruses (CMV, PeMoV, BCMNV and BCMV) is unavailable. Based on recent detections of a diverse repertoire of viruses not previously known to exist within the African bean cropping systems (Mwaipopo *et al.*, 2018), it is quite probable that similar viruses could be present in farmers' fields in Zambia. The lack of such vital information therefore, presents a knowledge gap that could be resolved by conducting a detailed study of bean-infecting viruses in Zambia.

2.6 Viruses infecting common bean

Common bean is vulnerable to many viruses that cause various symptoms and yield losses. Viruses frequently reported in common bean fields include; BCMV, BCMNV, cowpea mild mottle virus (CPMMV), CMV, CABMV, PeMoV and SBMV (Table 2.1). The BYMV is infrequent whereas bean golden yellow mosaic virus (BGYMV: genus *Begomovirus*), Calopogonium golden mosaic virus (CalGMV: genus *Begomovirus*) and squash yellow mild mottle virus (SYMMoV: genus *Begomovirus*) are to-date localized to continental South America

(Karkashian *et al.,* 2011). More recently, two cryptic non-pathogenic viruses PvEV-1 and PvEV-2 were reported from Tanzania and Kenya, further adding to the complex of viruses occuring in common bean fields (Nordenstedt *et al.,* 2017; Mutuku *et al.,* 2018; Mwaipopo *et al.,* 2018). Apart from BGYMV and CalGMV, the rest have been documented across the African continent (Mwaipopo *et al.,* 2018). Despite this extensive knowledge of the viruses of common bean, new information is still emerging from other studies using more sensitive detection methods that show the involvement of viruses from other genera (Table 2.2).

2.6.1 Potyviruses

The family *Potyviridae* (genus: *Potyvirus*) includes many plant-infecting viruses causing diseases across a vast array of plant hosts among them common bean (Wylie *et al.,* 2017). Potyviruses that infect common bean include BCMNV, BCMV, CABMV, BYMV and PeMoV. These viruses have successfully adapted to their hosts over time and are ably transmitted through two main pathways; insect vectors and seed.

Like other potyviruses, virions of common bean-infecting viruses of the family *Potyviridae* are filamentous, measuring 680-900nm and 11-20nm in length and diameter, respectively (Fang *et al.,* 1995). They are encapsidated in 30-47 kilodaltons (kDa) capsid protein arranged in helical pattern around the genomic RNA. The organization of common bean infecting potyvirid genome (Figure 2.1) comprises the typical 10 viral proteins starting with the protein viral P1 to the capsid protein (CP) and may possess 5' and 3' Nontranslated regions (NTRs) whose lengths are variable among the genomes. The 3' terminus is flanked by a polyadenylate tract (Wylie *et al.,* 2017). The RNA is single-stranded, positive sense with variable lengths ranging from 9-10 kb (Worrall *et al.,* 2015).

Table 2.2. Viruses infecting common bean globally

Virus	Genus (Family)	Distribution	Reference
Bean common mosaic necrosis virus (BCMNV)	*Potyvirus* (*Potyviridae*)	Global	Worrall *et al.*, 2015
(BCMV Serotype A)			Drijfhout *et al.*, 1978; Vetten *et al.*, 1992
Bean common mosaic virus (BCMV)	*Potyvirus* (*Potyviridae)*	Global	Berger *et al.*, 1997
(BCMV Serotype B)			Drijfhout *et al.*, 1978; Vetten *et al.*, 1992
Cowpea mild mottle virus	*Carlavirus* (*Betaflexiviridae*)	Global	
(CPMMV)			Zanardo and Carvalho, 2017
Cucumber mosaic virus (CMV)	*Cucumovirus* (*Bromoviridae*)	Tanzania, Iran, Kenya, Zambia	Mwaipopo *et al.*, 2018; Mutuku *et al.*, 2018
		USA, Uganda	Zitter and Murphy, 2009, Morales, 2006
Cowpea aphid-borne mosaic virus (CABMV)	*Potyvirus* (*Potyviridae*)	Kenya, Tanzania and Zambia	Mwaipopo *et al.*, 2018; Mutuku *et al.*, 2018
Southern bean mosaic virus (SBMV)	*Sobemovirus* (Unassigned)	Africa, Europe, ASIA, USA	Verhoeven *et al.*, 2003, Lamptey and Hamilton, 1974; Givord, 1981
Bean golden yellow mosaic virus (BGYMV)	*Begomovirus* (*Geminiviridae*)	Mexico, Central America	Morales, 2006
Phaseolus vulgaris endornavirus1 (PvEV-1)	*Endornavirus*	Mexico, Central	Wakarchuk and Hamilton, 1990

	(*Endornaviridae*)	America, Carribeans	
			Okada *et al.,* 2013; Nordenstedt *et al.,* 2017
Phaseolus vulgaris endornavirus2 (PvEV-2)	*Endornavirus* (*Endornaviridae*) America	Mexico, Central	Wakarchuk and Hamilton, 1990
Bean golden mosaic virus (BGMV)	*Begomovirus* (*Geminiviridae*)	South America	Morales, 2006
Ethiopian tobacco bushy top virus (ETBTV)	*Umbravirus* (*Tombusviridae*)	**Tanzania	Mwaipopo *et al.,* 2018
Peanut mottle virus (PeMoV)	*Potyvirus* (*Potyviridae*)		Vetten and Allen, 1991
Bean yellow mosaic virus (BYMV)	*Begomovirus* (*Geminiviridae*)	Kenya, Iran	Fauquet *et al.,* 2005 (ICTV-Report)
Bean dwarf mosaic virus (BDMV)	*Begomovirus* (*Geminiviridae*)		Morales *et al.,* 1990; Morales, 2006

Source: Much of the information was sourced from Morales, 2006. But information on the distribution of viruses was also sourced from other literature like Abraham *et al.,* 2014. ** A short sequence fragment that resembled four umbraviruses including ETBTV was detected by high-throughput sequencing (HTS) from common bean pooled samples making it difficult to associate it with any of the four viruses (Mwaipopo *et al.,* 2018).

Between the 5' NTR and 3' NTR lies a large region that codes for a major non-functional polyprotein that can be proteolytically cleaved into nine functional proteins (Adams *et al.*, 2005; Oana *et al.*, 2009; Ivanov *et al.*, 2014; Wylie *et al.*, 2017). Additionally, a very interesting protein found in all common bean-infecting potyviruses, the Pretty Interesting Potyvirus-ORF (PIPO) results from the -1/+2 frame shift that occurs at a GA_6 conserved site within the P3 gene (Chung *et al.*, 2008; Olspert *et al.*, 2015). The PIPO together with the other nine proteolytically self-cleaved proteins perform different important viral functions reviewed in Revers and Garcia (2015).

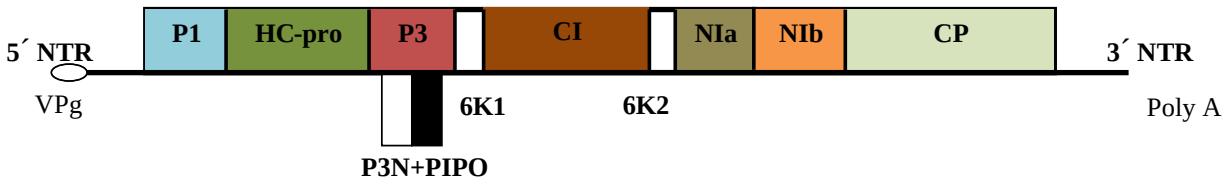


Figure 2.1. A general genomic map (not drawn to scale) of members of the family *Potyviridae*. VPg=Viral protein genome-linked represented by an unshaded ellipse, HC-pro= Helper component protein, P3=Viral protein 3, P3N+PIPO =P3N-pretty interesting potyvirus-ORF; PIPO protein represented by vertical stripes, 6K1=6 kilodalton protein, CI=cylindrical inclusion (CI), 6K2=6 kilodalton 2, NIa=nuclear inclusion-a, NIb=nuclear inclusion-b and CP=Capsid protein. The 3´ NTR (non-translated region) terminus is polyadenylated (Revers and Garcia, 2015).

2.6.2 Sobemoviruses

The genus *Sobemovirus* comprise viruses that infect diverse host plants but specific viruses have narrow host range. For example, SBMV only infects common bean and southern cowpea mosaic virus (SCPMV) is restricted to cowpea (*Vigna unguiculata* (L.) Walp.) (Hacker and Fowler, 2000).

The genus name *Sobemovirus* is derived from SBMV (Tremaine and Hamilton, 1983). Like all other members of the genus, the genome of SBMV is positive sense single stranded RNA (+ssRNA) and is encapsidated in icosahedral capsid measuring 30 nm in diameter (Hull, 1995). Southern bean mosaic virus genome (Figure 2.2) is polycistronic measuring 4.2 kb and consisting of a single coat protein (approximately 30kDa), genomic RNA (gRNA) and subgenomic RNA (sgRNA) molecules (Hull, 1995; Sõmera *et al.*, 2015). The long polyprotein gRNA is proteolytically cleaved by the N-terminal protease (also called sobemovirus peptidase) resulting in four ORFs that code for different functional viral proteins (Figure 2.2) (Sõmera *et al.*, 2015).

Figure 2.2. General structural organization of southern bean mosaic virus (SBMV) genome. 5´ NTR= 5´ nontranslated region, VPg=virus genome-linked protein, 3´ NTR=3´ terminus is not polyadenylated. ORF1=Replicase, ORF2a=RNA dependent RNA-polymerase (RdRP), ORF2b= movement protein and ORF3=coat protein (CP), adapted from Sõmera *et al.* (2015)

Wu *et al.* (1987) proposed the genome organization of SBMV and elucidated its four overlapping coding components, which are similar for all sobemoviruses except for *Imperata yellow mottle virus* infecting *Imperata cylindrical* and maize (Sérémé *et al.*, 2008). The genome structure recognized by the International Committee on Taxonomy of Viruses (ICTV) for all sobemoviruses sequenced to date comprises 5' terminal that is linked to the VPg and a 3' terminal lacking the A-tail and four overlapping protein-ORFs that are

conserved across the members of the genus (Sivakumaran and Hacker, 1998; Sõmera *et al.*, 2015). The genomic RNA (gRNA) acts as mRNA during the translation of the long polypeptide resulting into four ORFs.

The first protein (P1; ORF1) is involved in systemic silencing (Sarmiento *et al.*, 2007; Lacombe *et al.*, 2010) and virus movement (Sivakumaran and Hacker, 1998; Chowdhury and Savithri, 2011). The P2b results from -1 ribosomal slippage at the P2a 3' terminal continuing to form P2b, which is the RNA-dependent RNA polymerase (RdRp) of the SBMV. The ORF3 (P3) encodes the coat protein (CP) of the genome and is translated, unlike the preceding ORFs, from the sgRNA (Dwyer *et al.*, 2003). The CP is associated with short and long distance movements. Previous studies of common bean viruses did not report occurrence of SBMV in Zambia presumably due to inadequacies of the detection assays used.

2.6.3 Cucumoviruses

Cucumber mosaic virus (CMV: genus *Cucumovirus*) is a cosmopolitan virus infecting a vast majority of agriculturally important plants both monocots and dicots (Edwardson and Christie, 1991). Among the crops infected is common bean (Morales, 2006; Jacquemond, 2012). Several virus-detection studies in common bean fields have revealed global occurrence of CMV (Vetten and Allen, 1991; Shahraeen *et al.*, 2005; Njau *et al.*, 2006; Garcı´a-Arenal *et al.*, 2008; Azizi and Shams-bakhsh, 2014; Mutuku *et al.*, 2018).

The genome of CMV is tripartite comprising three (+)ssRNA molecules so called RNA1, RNA2 and RNA3 based on their decreasing sequence size (Figure 2.3). Each RNA molecule is encapsidated in 180 subunit capsid proteins and measure 29 nm in diameter (Jacquemond, 2012). RNA1 is the largest genomic component of CMV. It is monocistronic coding for 1a gene (Ali Rezaian *et al.*, 1985; Revathy and Bhat, 2017). RNA2 is bicistronic and encodes two genes, 2a and 2b. The 2a component is processed from the gRNA whereas 2b (Ding *et*

al., 1994) results from a +1 frameshift, generating a short ORF (Jacquemond, 2012). RNA3 is the smallest of the three RNA molecules and it is bicistronic comprising two genes, 3a and 3b that are separated by an intercistronic region. Like 2a, 3a is expressed from gRNA whereas 2b and 3b are products of subgenomic RNA (sgRNA4A) and RNA4 expressions, respectively (Jacquemond, 2012).

The genes 1a and 2a are important in virus replication as they harbour the viral components of the replicase of CMV (Palukaitis *et al.*, 1992). The 5' terminal of 1a gene possesses a putative methyltransferase whereas the 3' terminal has a helicase motif (Gorbalenya *et al.*, 1989; Rozanov *et al.*, 1992; Revathy and Bhat, 2017). Gene 2b has a triple role acting as post transcription gene silencing (PTGS) suppressor (Li *et al.*, 1999), a movement protein (MP) in systemic movement (Wang *et al.*, 2004) and vectors-mediated transmission (Ziebel *et al.*, 2011). The functions of the two gene components (3a and 3b) of the RNA3 molecule are elucidated in other studies (Shintaku *et al.*, 1992; Suzuki *et al.*, 1995).

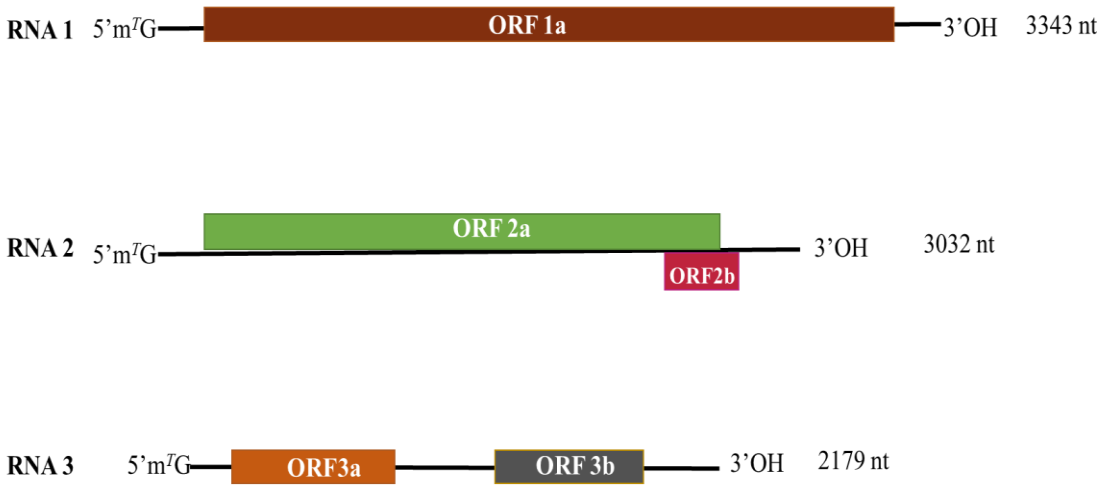


Figure 2.3. Genome map for cucumber mosaic virus (CMV) depicting three RNA molecules, their sizes and open reading frames (ORFs). M^7G means 7-methyl guanosine. nt means nucleotide. Adapted from Jacquemond (2012).

2.6.4 Umbraviruses

The genus *Umbravirus* (family *Tombusviridae*) comprises viruses that are distinct in that they lack structural proteins and therefore do not form particles (Taliansky *et al.*, 2000; Taliansky and Robinson, 2003). Currently, there are 11 recognized species in this genus, including Ethiopian tobacco bushy top virus (ETBTV) (https://talk.ictvonline.org/ictv-reports/ictv_online_report/). Since umbraviruses do not encode a CP, they rely on helper viruses, mostly from the genera *Polerovirus*, *Luteovirus* and *Enamovirus* for transcapsidation during vector-mediated transmission (Taliansky and Robinson, 2003). In nature, individual umbraviruses may have limited host range but at experimental level, this may be wider (Abraham *et al.*, 2014; Tang *et al.*, 2015).

In the pathology of umbraviruses, the involvement of satellite RNA (satRNA) molecules is critical in pathogenesis (Taliansk *et al.*, 2000; Mo *et al.*, 2011; Abraham *et al.*, 2014; Tang *et al.*, 2015). Thus for successful transmission to occur, the triad combination of an umbravirus, a satRNA and helper virus in some of the umbraviruses should be acquired and co-transmitted by vectors (Abraham *et al.*, 2014), but subsequent invasion of susceptible host cells leading to disease symptom development is not dependent on the helper virus (Taliansky and Robinson, 2003).

The annotation of the virus sequence reveals a common genome map for all members of the genus *Umbravirus* comprising one linear, (+)ssRNA (Gibbs *et al.*, 1996; Mo *et al.*, 2003; Abraham *et al.*, 2014; Tang *et al.*, 2015). The sizes of genomes sequenced so far are between 4-4.2 kb and consist of a short 5' UTR, three gRNA encoded ORFs and one sgRNA expressed ORF (Figure 2.4). The 3' UTR is not polyadenylated and varies in length. As earlier indicated, the genome does not possess structural proteins; the CP. Open reading frame 2 is separated from ORF3 by an intergenic region varying in length from 109-nt in

groundnut rosette virus (GRV) (Taliansky *et al.,* 1996) to 182 in ETBTV (Abraham *et al.,* 2014). The ORF1 and ORF2 presumably undergo co-translation through -1frameshift and hence together perform RdRp function of the virus (Koonin and Dolja, 1993). The ORF3 is associated with two major functions, RNA stabilization and systemic translocation of the virus (Ryabov *et al.,* 1999a; Ryabov *et al.,* 2001). The ORF4 is known to perform cell-to-cell movement in umbraviruses. This was deduced from experiments in which GRV encoded ORF4 performed the function of MPs of CMV and potato virus X (PVX) regardless of the presence or absence of their respective MPs (Ryabov *et al.,* 1998; Ryabov *et al.,* 1999b).

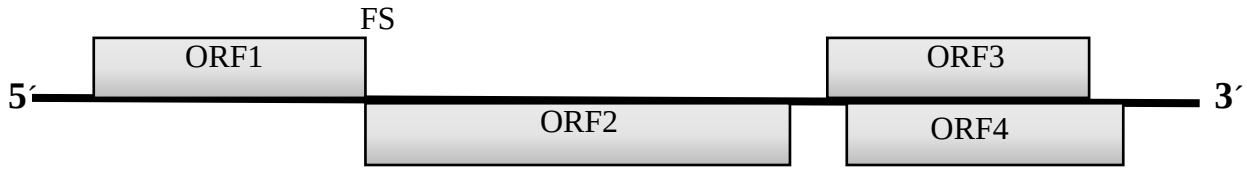


Figure 2.4. A typical genomic organization of umbravirus. The blocks represent different open reading frames (ORFs). The −1 frameshift event (FS) occurs in ORFs 1 and 2. The ORF3 functions as long distance movement protein. The ORF4 product facilitates short distance virus movement. Adapted from Tang *et al.* (2015)

2.6.5 Alphaendornaviruses

Endornaviruses (family *Endornaviridae*) are a group of viruses that infect plant (Nordenstedt *et al.,* 2017) and non plant hosts. The plants include rice (*Oryza sativum*), bell paper (*Capsicum annuum*), avocado (*Persea americana*) fungi (*Phytophthora spp, Helicobasidium mompa, Gremmeniella abietina*). Legumes (family *Fabaceae*) that include common bean are susceptible to infection by alphaendornaviruses too (Okada *et al.,* 2013; Nordenstedt *et al.,*

2017). The virus can persist in common bean for a long time and vertical rather than horizontal transmission is possible (Nordenstedt *et al.,* 2017).

Whole genomes of endornaviruses isolated from common bean (*Phaseolus vulgaris* L.) were first sequenced from cultivar Black Turtle Soup (BTS) by Okada *et al.* (2013) and the viruses were provisionally named as Phaseolus vulgaris endornavirus 1 (PvEV-1) and Phaseolus vulgaris endornavirus 2 (PvEV-2). This nomenclature was accepted by the International Committee on Taxonomy of Viruses (ICTV) (Valverde *et al.,* 2019). The PvEV-1 and PvEV-2 are single-stranded RNA, however, *invivo* they occur as double-stranded replicative RNA (ds-RNA) intermediates (Roossinck, 2015); the stable form that is easily isolated. Genome-wide analysis revealed that genomes of the distinct virus code for a polyprotein that is proteolytically cleaved at recognizable sites to produce functional proteins (Okada *et al.,* 2013; Nordenstedt *et al.,* 2017) (Figure 2.5). Typical of members of the family *Endornaviridae,* genomes of PvEV-1 and PvEV-2 are devoid of the capsid gene and thus they do not exist as true virons (Valverde *et al.,* 2019).

Recently, PvEV-1 and PvEV-2 were discovered by illumina sequencing of nucleic acid isolated from bean leaf samples collected from farmer grown common bean fields in Tanzania and Kenya (Mutuku *et al.,* 2018; Mwaipopo *et al.,* 2018). Interest in the study of these viruses is not instigated by any known adverse effect the viruses cause on production or yield and current scientific knowledge is devoid of insights into the impact of the two viruses on crop growth and yield (Nordenstedt *et al.,* 2017; Mwaipopo *et al.,* 2018). Regardless, replicative viruses in a plant, even if they cause virtually no disease symptoms, lead to yield reduction of some measure (Morales, 2006). Thus, regardless of the lack of information available on the impact of endornaviruses infecting common bean, it is plausible to assume that they affect plant growth to some extent.

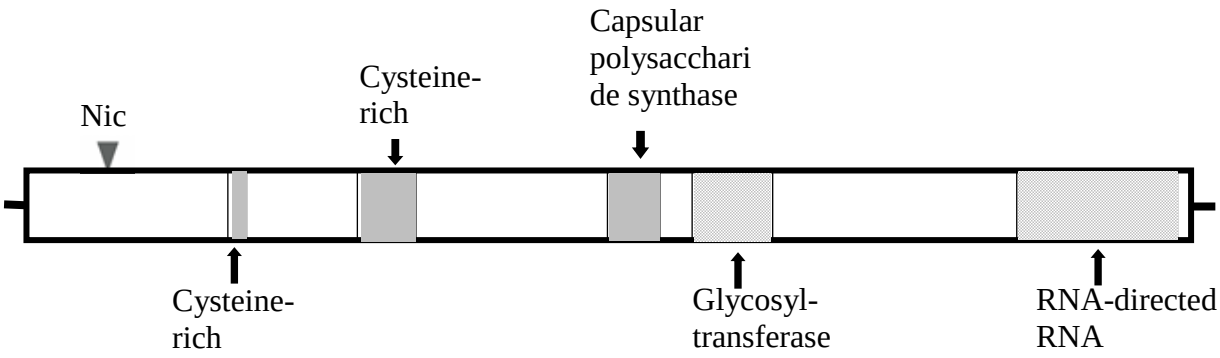


Figure 2.5. Genome organization of Phaseolus vulgaris alphaendornavirus (PvEV). The different functional proteins of the genome are shown as boxes (Okada *et al.*, 2013).

2.7 Transmission of viruses and disease spread

Natural transmission of plant viruses has always been a topical issue in plant virus disease epidemiology. Three major transmission pathways are recognized: true seeds, arthropod vectors, and vegetative plant materials (Dickinson, 2003). Common bean-infecting viruses are transmitted through true seed and arthropod vectors and therefore, these two will be the focus of this section.

True seeds are very important biological propagative materials. Unlike animal viruses whose long distance dispersal is conveyed by infected individuals, plants are sessile but the viruses infecting them can exploit seed dispersal as a conduit to invade new areas that may be near or far from the infection-loci (Dickinson, 2003; Sastry, 2013; Revers and Garcia, 2015). Hence, even at a low rate, seed transmission can result in significant disease outbreaks (Revers and Garcia, 2015). Viruses are transmitted at different efficiencies largely because of differing abilities of viruses to enter the suspensor, which is the precursor of the developing embryo (Revers and Gracia, 2015). Apart from potyviruses, other notable seed transmitted common bean infecting viruses are CMV, SBMV and the endonarviruses PvEV-1 and PvEV-2. There

is no evidence yet of seed transmissibility of ETBTV and other umbraviruses (Ryabov *et al.*, 2012).

Most viruses are transmitted from plant to plant through arthropod vectors and this route form a major pathway for virus spread (Hull, 2014). There are several modes of arthopod vector transmission of plant viruses: non-persistent non-circulative; semipersistent non-circulative; circulative persistent non-propagative; and circulative persistent propagative (Andret-Link and Fuchs, 2005). In a few cases, a specific arthropod vector may even transmit several viruses via multiple transmission modes (Andret-Link and Fuchs, 2005) as is the case with the aphid (*Myzus persicae* (Sulzer) (Pinheiro *et al.*, 2019).

Aphids vector the majority of bean-infecting viruses and those infecting other crops (Nault, 1997; Ng and Falk, 2006). They transmit CMV (Wamonje *et al.*, 2020), SBMV (Sõmera *et al.*, 2015 with references therein), BCMV and BCMNV (Worrall *et al.*, 2015; Wamonje *et al.*, 2020), ETBTV (Abraham *et al.*, 2014), CABMV and PeMoV among other viruses.

The combination of aphid and seed transmissibility of most of the viruses, coupled with the fact that in Zambia farmers share uncertified seeds for planting (Hamazakaza *et al.*, 2014) make high prevalence of viruses in most cropping systems highly probable. Additionally, common bean is usually intercropped with maize or other crops in many African farming practices (Assefa *et al.*, 2016; Nassary *et al.*, 2020). Whereas the production benefits of such practices make them attractive to farmers (Mucheru-Muna *et al.*, 2010), there is a likelihood that they may influence vector population dynamics, their feeding behavior, and ultimately disease dispersal (Boudreau, 2013).

2.8 Relative importance of common bean infecting viruses

A disease is considered important if it is widely distributed or frequently occurs on an important crop, affects growth, yield and quality attributes of the crop, and results in damaging economic loss. Based on this, some diseases are more important than others. Globally, many common bean infecting viruses have been detected in several studies; they include BCMNV, BYGMV, CABMV, BCMV, SBMV, BGMV, AMV, bean leafroll virus (BLRV), CMV and bean pod mottle virus (BPMV) (Flores-Estévez *et al.,* 2003; Morales, 2006; Azizi and Shams-bakhsh, 2014; Mwaipopo *et al.,* 2017, 2018). However, the level of crop damage and extent of distribution differs from one virus to another across countries. For example, BGMV, which is important in Brazil, Argentina and Bolivia (Singh and Schwartz, 2010) has not been reported in Africa. SBMV that was reported in few plants in about eight countries in Africa as early as 1974 (Lamptey and Hamilton, 1974; Givord, 1981) has not been reported in the major bean growing areas in East Africa until recently when Mwaipopo *et al*. (2018) reported its occurrence in common bean fields in Tanzania for the first time. Additionally, few reports exist of CMV detected from common bean samples in Ethiopia (Spence and Walkey, 1995), Kenya (Mutuku *et al*., 2018), Tanzania (Mwaipopo *et al.,* 2017; 2018), Zambia and Zimbabwe (Vetten and Allen, 1991). BYMV was only reported in Kenya and PeMoV has been reported in Zambia and Tanzania previously (Vetten and Allen, 1991; Mwaipopo *et al.,* 2018).

Conversely, BCMNV and BCMV are widely distributed infecting common bean worldwide (Worrall *et al.,* 2015) albeit in varying proportions. The two viruses are the most prevalent and considered the most destructive causing yield losses of between 35 to 100% in the most susceptible varieties (Damayanti *et al.,* 2008; Li et al., 2014). Particularly BCMNV poses real

challenge in areas where it is endemic especially as it relates to resistance breeding involving the introgression of the *I* gene in desired bean cultivars.

Further, BCMNV and BCMV exist as a complex of molecularly diverse strains some of which can break resistance in known resistant genotypes (Feng *et al.,* 2015). Drijfhout (1978) demonstrated reactive differentiation when several strains of the BCMNV and BCMV were screened across 10 common bean differential cultivars. The importance of these two viruses has led to many screening studies in search of genetic resistance to BCMNV and BCMV in a global effort to remedy the impact of these two viruses on crop health (Omunyin, 1984; Morales, 1989; Miklas *et al.,* 2000). Further, insights into genetic variability of isolates or pathotypes of BCMNV and BCMV has been a subject of intensive study (Larsen *et al.,* 2011; Feng *et al.,* 2014, Feng *et al.,* 2015; Pasev and Kostova, 2015). In Zambia, Spence and Walkey (1995) reported two BCMNV/BCMV pathotypes (IVb and VIa) from just 14 samples whereas Kaitisha (2003) indicated the presence of other strains of both viruses in the country. Evidently, BCMNV and BCMV and their pathotypes constitute a major pathological challenge to bean production in Zambia and hence is a just focus in this study (Kaitisha 2003).

2.9 Genetic variability and pathotypes of BCMNV and BCMV

RNA viruses employ several replication strategies that maintain the fidelity of their nucleic acid compositions. However, the replication enzyme involved, RNA-dependent RNA polymerase (RdRp), is a low fidelity enzyme that lacks proofreading mechanism (Domingo and Holland, 1997; Mandary *et al.,* 2019). Hence, occasional misincorporation of nucleotide bases occurs during replication and when this happens, it leads to the evolution of heterogeneous population of viruses (Smith *et al.,* 1997). In a way, this could help a virus adapt to new environments, evade host defense mechanisms, and exploit new transmission

pathways. Virus evolution may occur through recombination, frame shift mutation, re-assortment of segmented virus genomes, gene duplication, deletion or point mutation (Hull, 2002).

Recombination is a common phenomenon among RNA viruses and it happens when two distinct viruses exchange segments of sequence information between nucleotide strands during replication. Based on extensive sequence analysis, Revers *et al*. (1996) confirmed high frequency of the occurrence of recombinants among potyviruses. This common occurrence is the likely cause of evolutionary rise in strain diversity of many RNA viruses (Cervera *et al.*, 1993). Specifically, recombination has been associated with strain diversity among isolates of BCMNV and BCMV (Larsen *et al.*, 2005). A good example is that of a single natural recombination event between BCMNV isolates NL-3 and RU1 that resulted in the recombinant strain NL-3K (Strausbaugh *et al.*, 2003; Larsen *et al.*, 2005). The "new" strain induced more severe symptoms in some bean differential host cultivars (cvs) than the parental isolates (Strausbaugh *et al.*, 2003; Larsen *et al.*, 2005). Other BCMV recombinants such as isolates RU1M and 1755a resulted from recombination events between isolates RU1-OR and NL-1/US1, respectively, with an unknown minor parent (Feng *et al.*, 2014; Feng *et al.*, 2015).

Molecular variants can also arise from genome reassortment. This process is common in viruses with segmented genomes such as those in the family *Tospoviridae* (Hanley and Weaver, 2008). Variants that arise through reassortment increase or decrease in their frequencies depending on their fitness. The fittest variants increase in frequency (positive selection) in a population whereas the less fit ones decrease (negative selection). Variants that undergo positive selection drive high diversity whereas those that undergo negative selection lead to low population diversity (Garcia-Arenal *et al.*, 2001).

Variants could be serologically indistinguishable from the parent virus but may show differences in their pathogenicity and other biological properties. Host response differences between variants of a particular virus population have been extensively exploited for classifying pathotypes especially among BCMNV and BCMV isolates (Drijfhout, 1978). Drijfhout (1978) conducted the pioneering work on pathotyping of BCMV that resulted in identification of seven pathogroups (PGs). The groups were determined by profiling their reaction with 10 common bean differential cvs. The BCMNV strains were grouped into PGs III (strain NL-8) and VI (strains TN-1, NL-3, NL-3K, NL-5) (Drijfhout, 1978; Worrall *et al.*, 2015). BCMV has the most diverse strains and hence more PGs (I, II, IV, V and VII) (Drijfhout, 1978). Recently, a recombinant strain of BCMV (isolate 1755a) was assigned to a new PG (VIII) because of its novel interactions with resistance alleles (Feng *et al.,* 2015). Thus, regardless of the information currently available viruses are continuously evolving leading to novel strains that in some cases exhibit different pathological profiles.

2.10 Management of common bean virus diseases

Management of common bean viruses or any other viruses is achieved by several approaches. Phytosanitation, use of insecticides and breeding for resistance are some of the common methods of management (Worrall *et al.,* 2015). Phytosanitation involves eradication of sources of virus inoculum via removal or roguing of infected plants. This is more effective if combined with virus indexing of seed lots in order to reduce the occurrence of primary infection. Large insect vector populations can be controlled with insecticides and thus limit the spread of viral diseases. In common bean fields, insecticides can be used to successfully manage aphid vectored viruses with circulative modes of transmission due to the time-lag between virus acquisition and transmission (Westwood and Stevens, 2010). However, for BCMNV and BCMV, acquisition time is very short and the viruses are nonpersistent and

noncirculative, making it possible for the vectors to transmit the viruses before lethal chemical toxicity levels are attained (Thottappilly, 1992; Westwood and Stevens, 2010). Further, there is a limited range of environmentally friendly insecticides that can be applied, most chemicals are deleterious to natural enemies, and they are often cost-prohibitive to subsistence farmers (Westwood and Stevens, 2010).

Breeding for genetic resistance is the sure way to attain durable resistance (Haley *et al.,* 1994; Morales and Kornegay, 1996). In common bean, this is based on a single dorminant gene and a series of recessive allelic resistance genes that confer resistance to both BCMNV and BCMV in their different combinations (Drijfhout, 1978). By itself, the dominant *I* gene is capable of conferring resistance to all BCMV pathotypes. However, in the presence of necrotic strains of BCMNV, the plants possessing dominant *I* genes die due to hypersensitive response to infection (Drijfhout, 1978; Worrall *et al.,* 2015). The remedy is to mask the deleterious effect of interactions between *I* gene and BCMNV in *I +bc-3* combinations resulting in broad spectrum resistance (Kelly *et al.,* 1995).

2.11 Resistance of common bean to viruses

Genetic interactions between common bean cultivars and BCMV was extensively covered by Drijfhout (1978). In his study, Drijfhout studied the strains of BCMV and identified the associated interactive resistance genes. Generally, resistance to BCMV and BCMNV is attained by pyramiding the dominant *I* gene with any of the six epistatic genes (Haley *et al.,* 1994; Mukeshimana *et al.,* 2005). In this way, desired resistance to all BCMNV/BCMV pathotypes is attained while masking the hypersensitive reaction in *I* gene-carrying cultivars.

Although initial breeding efforts appeared to have contained the problem of resistance to BCMNV and BCMV, there are still new virus strains that overcome resistance in common bean. For instance, a recombinant of BCMV named RU1-OR was found to overcome

resistance in cvs carrying *bc-u* and *bc-1*^{2} or *bc-2*^{2} resistance genes (Feng *et al.,* 2014); implying that recombinants may alter the known reactions of some bean cultivars hence the need to update scientific knowledge with new studies.

2.12 Diagnosis of common bean viruses

2.12.1 Enzyme-Linked Immunosorbent Assay

The enzyme-linked immunosorbent assay (ELISA) developed by Clark and Adams (1977), is a vastly useful detection tool in virus diagnostics. The ELISA method is premised on the reaction of antibodies with specific antigenic determinants (or epitopes) of pathogens. ELISA has wide applications as a tool for detecting pathogens in medical and agricultural settings. In agriculture and specifically plant virology, application of ELISA assays led to the detection and quantification of various viral agents in plants of diverse crops using monoclonal antibodies (MAbs), polyclonal antibodies (PAbs), or their combinations. ELISA offers three advantages; it is cheaper, easy to use and scalable to large sample sizes (Boonham *et al.,* 2014).

Whereas monoclonal antibodies (MAbs) are raised to be specific to a particular protein epitope of the pathogen of interest and hence can only identify the antigen for which they were produced, polyclonal antibodies (PAbs) consist of a collection of different antibodies in the antisera (Boonham *et al.,* 2014). For common bean viruses that exist as many strains such as BCMV and BCMNV, both MAbs and PAbs have been used to discriminate among virus strains from different plant parts (Arli-Sokmen *et al.,* 2016). To increase sensitivity and accuracy of the method, different formats of ELISA assays have been developed including double-antibody sandwich (DAS)-ELISA and triple-antibody sandwich (TAS)-ELISA (Thomas *et al.,* 1986; Harrison *et al.,* 2002).

Although ELISA methods have become integral parts of routine detection procedures, their application is limited by three major disadvantages. ELISA requires high quality antisera that is often only produced in specialized laboratories, is often inapplicable for the identification of novel viruses, may cross react with a novel virus and lacks adequate resolution for closely related viruses thereby confounding strain typing (Boonham *et al.,* 2014). The advent of polymerase chain reaction (PCR) has spurred the identification of different forms of very similar or different pathogens in uniplex PCR for single infections or multiplex PCR for mixed infections (Alabi *et al.,* 2008; Panno *et al.,* 2012; Aloyce *et al.,* 2013).

2.12.2 Polymerase chain reaction

The PCR is a more sensitive method than ELISA in that it is able to discriminate between molecularly similar viruses that ordinarily can not be differentiated in ELISA. It employs oligonucleotides designed to bracket and amplify segments of the genomes of target pathogens. Since its advent, PCR assays have been widely used in many studies of viruses causing plant diseases. In the case of RNA viruses, a variant of PCR, reverse transcription (RT)-PCR converts RNA to complementary DNA (cDNA) that can be used in PCR (Petrović *et al.,* 2010). Although PCR and RT-PCR are sensitive and have been used widely, their use for the detection of novel pathogens is challenging. This limits the application of the technique in detecting previously uncharacterized pathogens for which no sequence data exist for designing primers. One way of circumventing this limitation is to use degenerate primers designed specific to sequences of members of a particular genus. This has been used in detecting several existing and new begomoviruses (Deng *et al.,* 1994; Wyatt and Brown, 1996; Pita *et al.,* 2001) and potyviruses (Chen *et al.,* 2001; Ha, 2007). Unfortunately, novel viruses that are highly divergent from members of specific genera or potential members of a putative novel genus may escape detection with degenerate primers. Such limitations can be circumvented with the use of high-throughput sequencing (HTS) method otherwise called

next generation sequencing (NGS). ELISA and PCR/ RT-PCR require a *priori* knowledge of the target pathogen but the relatively lower cost and versatility of the two methods make them useful in many routine detections (Kushwaha *et al.,* 2010). However, for generic identification of novel pathogens, HTS is more useful (Pecman *et al.,* 2017).

2.12.3 High-throughput sequencing

The HTS is a very sensitive technology that exploits the power of metagenomics for pathogen diagnosis. It offers a culture independent detection of pathogens of interest, captures broader spectrum of mutations than Sanger sequencing, can extract genetic information in large amounts, and is faster (Helmy *et al.,* 2016). Through the use of the HTS many previously unknown viruses infecting common beans were reported in Tanzania (Mwaipopo *et al.,* 2018). Nordenstedt *et al*. (2017) used HTS to detect the presence of PvEV-1 and PvEV-2 in Nicaragua and Tanzania. Zongoma *et al.* (2017) reported three plant-pathogenic viroids infecting grapevines, grapevine yellow speckle viroid 1 (GYSVd-1), GYSVd-2, and *Hop stunt viroid*, from samples collected from *Vitis spp* in Nigeria. Mutuku *et al*. (2018) and Wainaina *et al*. (2019) sequenced BCMNV, CABMV and CMV from common bean leaf tissue samples from Kenya.

In spite of these achievements, there are numerous challenges facing sequencing technologies in developing countries. Helmy *et al*. (2016) itemized them as high cost of sequencing, complex bioinformatics coupled with a large amount of sequencing analysis tools from which researchers have to find best options for their data and cost of library preparation and sequencing reagents.

2.12.4 Biological characterization of BCMNV and BCMV

Regardless of the sophistication of sequence-based detection tools, they cannot describe the differences in phenotypic reactions induced by different isolates of viruses infecting common

bean. The phenotypic reactions resulting from interactions of the common bean differential cultivars with different PGs of BCMNV and BCMV have been reviewed in section 2.11.

2.13 Molecular markers and resistance genes in common bean

Molecular markers are useful in plant identification and genetic improvements. The useful heritable feature of markers is that they are known to evolve together with gene loci that condition a target phenotypic expression (Jiang, 2013). This feature was successfully exploited in identifying markers for marker-assisted selection (MAS) breeding (Haley *et al.,* 1994; Melotto *et al.,* 1996) in conferring broad resistance to different BCMV and BCMNV strains in bean plants (Kelly *et al.,* 1995; Mukeshimana *et al.,* 2005; Pasev *et al.,* 2014). A particular advantage of using MAS is that it enables faster turnaround time of the breeding process.

In common beans, except the dominant *I* gene, the molecular sequences of all the six known allelic recessive genes are known (Worrall *et al.,* 2015). Markers exist in many forms, among them restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), simple sequence repeats (SSR), variable number tandem repeats (VNTR) and single nucleotide polymorphism (SNP). The PCR based marker, RAPD, can easily be confirmed using marker specific oligonucleotides (Haley *et al.,* 1994; Melotto *et al.,* 1996). Understanding resistance gene pool within a population of common bean genotypes is critical to resistance breeding. The availability of RAPD oligonucleotides makes it possible to combine molecular and bioassay approaches in confirming markers within a select genotype of different plants.

CHAPTER THREE

SURVEY AND METAGENOMICS ANALYSIS OF VIRUSES IN COMMON BEAN (*Phaseolus vulgaris* L.) FIELDS IN ZAMBIA

3.1 Abstract

The production of common bean (*Phaseolus vulgaris* L.) is adversely affected by virus-like diseases globally, but little is known about the occurrence, distribution, and diversity of common bean-infecting viruses in Zambia. Consequently, field surveys were conducted during 2018 season in 128 fields across six provinces of Zambia and 640 common bean leaf tissue samples were collected with (n=585) or without (n=55) symptoms. The prevalence of symptomatic fields was 100%, but incidence of symptomatic plants ranged from 32% to 67.5%. Metagenomic analyses of nine composite samples revealed the occurrence of isolates of *Bean common mosaic necrosis virus*, *Bean common mosaic virus*, *Cowpea aphid-borne mosaic virus*, *Peanut mottle virus*, Southern bean mosaic virus (SBMV), *Cucumber mosaic virus*, Phaseolus vulgaris alphaendornavirus 1 (PvEV-1), PvEV-2 and *Ethiopian tobacco bushy top virus*. Screening of the survey samples by RT-PCR for the viruses detected by high throughput sequencing revealed the prevalence of single (65.2% or 417/640) over mixed (1.9% or 12/640) infections in the samples. SBMV was the most frequently detected virus, occurring in ~29.4% (188/640) of the samples and at a prevalence rate of 58.6% (75/128) across fields. The results showed that diverse virus species are present in Zambian common bean fields and the information will be useful for the management of common bean viral diseases.

3.2 Introduction

Common bean (*Phaseoulus vulgaris* L.) is an inexpensive substitute for animal protein sources which are cost prohibitive in many African countries. Nearly 28 million tons of dry beans was produced globally in 2018, of which Africa accounted for 22.8% (FAOSTAT, 2019). Common bean is consumed by over 300 million people South America, the Carribeans and sub-Saharan Africa (Petry *et al.,* 2015), including in Zambia where it is commonly consumed by low income households (Hichaambwa *et al.,* 2009). Over the past decade, harvested area for dry beans in Zambia almost doubled, from about 42,500 ha in 2003/2004 to over 71,000 ha in 2011 (ZSA, 2017). However, yields over the same period remained relatively low at ≤0.55 tons/ha versus potential yields of 1.5-2.5 tons/ha in experimental fields (Hamazakaza *et al.,* 2014). The low yields are attributed to several factors, especially pests and diseases.

Diseases caused by viruses limit the production of common bean in Africa and globally (Morales, 2006; Beebe *et al.,* 2013). The viruses include peanut mottle virus (PeMoV), cowpea aphid-borne mosaic virus (CABMV), bean common mosaic virus (BCMV), bean common mosaic necrosis virus (BCMNV); alfalfa mosaic virus (AMV; genus *Alfamovirus*); cucumber mosaic virus (CMV; genus *Cucumovirus*); and southern bean mosaic virus (SBMV; genus *Sobemovirus*) (Spence and Walkey, 1995; Sengooba *et al.,* 1997; Njau *et al.,* 2006; Mutuku *et al.,* 2018; Mwaipopo *et al.,* 2018; Wainaina *et al.,* 2019). Further, two double stranded RNA (dsRNA) endonaviruses, i.e. Phaseolus vulgaris alphaendornavirus-1 (PvEV-1) and PvEV-2, were recently reported in Kenya and Tanzania (Nordenstedt *et al.,* 2017; Wainaina *et al.,* 2017; Mutuku *et al.,* 2018; Mwaipopo *et al.,* 2018). A recent study using highthroughput sequencing (HTS) also documented viruses belonging to five additional genera (*Umbravirus, Crinivirus, Carlavirus, Caulimovirus and Cytorhabdovirus*) from

common bean samples in Tanzania, revealing a more complex virome of common bean fields in Africa than previously reported (Mwaipopo *et al.,* 2018).

However, symptoms described for common bean viral diseases are difficult to attribute exclusively to specific virus species. Thus, distinguishing these viruses based on foliar symptoms is unreliable; perhaps with the exception of BCMNV that induces a diagnostic "black root" or necrotic symptoms in infected *I*-gene bearing bean plants (Drijfhout, 1978; Morales, 1989; Flores-Estévez *et al.,* 2003). But similar necrotic symptoms were documented in bean plants infected with some temperature-dependent strains of BCMV at >30°C (Feng *et al.,* 2014), further confounding the etiology of common bean viral diseases.

In studies conducted nearly three decades ago, the occurrences of five common bean viruses (BCMV, BCMNV, CABMV, CMV and PeMoV) in Zambia were reported based on screening of very limited number of samples (n = 14 to 91) using immunological assays (ELISA) and/or bean differential bioassays (Vetten and Allen, 1991; Spence and Walkey, 1995; Ndomba, 2002). In addition to limited sensitivity, neither ELISA nor differential cultivars are useful for interrogating the complex molecular characteristics of these virus species. Due to the limitations in the resolution power of serological and biological assays, it is possible that previously uncharacterized common bean infecting viruses are present in Zambia (Spence and Walkey, 1995). Previous reports also lack information on the country-wide prevalence and field incidences of common bean infecting viruses in Zambia. To fill these knowledge gaps, a study was conducted with the following aims: (i) to determine the incidence and distribution of virus diseases in common bean fields in Zambia, (ii) to identity the causal or associated viruses, and (iii) to determine the genetic diversities of the most prevalent viruses.

3.3 Materials and Methods

3.3.1 Survey areas, field assessment of incidence and severity of virus-like symptoms

Disease surveys were conducted from March to May 2018 in six of the ten provinces of Zambia. The country is divided into mainly three agroecological zones (AEZ) based on the amount of rainfall received. The six provinces (Luapula, Northern, Northwestern and Muchinga) are located in AEZ III (rainfall >1200 mm) while Central and Eastern are in AEZ II (rainfall 800-1000 mm). The selection of the surveyed provinces was based on the data obtained from the Zambia Statistics Agency (ZSA) that showed that the six provinces accounted for 93% of common bean production in Zambia (ZSA Crop focused, 2017). The surveyed common bean fields were targeted along feeder roads or main roads and a minimum of 7 km between field interval was maintained in areas with high common bean field intensity while the interval was increased to 30 km in areas with low field intensity. The numbers of surveyed fields per province were Luapula (25), Northern (21), Muchinga (20), Central (20), Eastern (20), and Northwestern (22), thus totaling 128 field locations across the six provinces.

A province was considered as a single sampling domain and the sample size determination per sampling domain was as recommended by Sseruwagi *et al*. (2004). At each field, 30 plants were visually inspected along two diagonal transects (15 plants/transect). The overall disease incidence per province was calculated as previously described (Mwaipopo *et al.*, 2018). Brifely, disease incidence was calculated as a percentage of symptomatic plants out of the total number of plants assessed within that province. Disease symptom severity was assessed on a scale of 1 to 5 in which 1 = no virus symptoms, 2 = mild symptoms on trifoliate leaves, 3= moderate symptoms on trifoliate leaves, 4 = severe and widespread symptoms on most or all of the trifoliate leaves and 5= very severe widespread symptoms on all fully formed trifoliate leaves that may also include stunted growth and leaf deformation. The

overall mean severity for each province was calculated by finding the average of all scores (2 to 5). Coordinates for all the surveyed fields were taken using a handheld global positioning system device (Garmin International Inc.USA).

3.3.2 Distribution of common bean cultivars in surveyed provinces

Data on the distribution of common bean cultivars encountered during the survey was also collected. Information obtained included source of seed planted (open market, registered seed supplier or recycled) for three seasons (past two planting seasons and the current season). The data collected was analyzed to decipher common bean cultivar preferences in each province and across the six provinces.

3.3.3 Sample collection and isolation of total nucleic acid

Leaf tissue samples were collected from four to eight plants that represented the spectrum of observed foliar symptoms in each field. Both symptomatic and symptomless plants were sampled in the surveyed 128 fields resulting in a total of 640 samples (symptomatic = 585; non-symptomatic = 55) collected during the surveys. Each sample was preserved dry on silica gel and transported to the Plant Pathology laboratory, Mount Makulu Central Research Station, Chilanga, Zambia. Total nucleic acid (TNA) was isolated from 50 mg of each sample using a CTAB (cetyltrimethylammonium bromide) method (Chang *et al.,* 1993) with modification. In the modified protocol, L-Spermidine was not included in the buffer. Hence the composition of the CTAB isolation buffer used was 2% CTAB, 2% PVP-58, 100 mM Tris-HCL (pH 8.0), 25 mM EDTA, 2 M NaCl, 2-mercaptoethanol (0.2%). The rest of the TNA isolation procedure remained the same. The recovered pellets were re-suspended in 30 µl nuclease-free water, quantified using NanoDrop ONE spectrophotometer (Thermo Fisher Scientific, CA) and the quality of TNA checked on 1.2% agarose gel that was prestained with 0.5 mg/ml ethidium bromide.

3.3.4 High-throughput sequencing

To enable initial diagnosis of the samples by high throughput sequencing (HTS), subsets of 30 symptomatic samples were randomly selected per province and 250 ng/µL total nucleic acid aliquots per sample within each subset were pooled and thoroughly mixed into one composite sample to reduce analytical costs. Care was taken to ensure that each sample subset encompassed the spectrum of foliar symptoms encountered across the surveyed fields within each province. Thus, six composite samples were derived from Central (CP-1), Eastern (EP-1), Luapula (LP-1), Muchinga (MP-1), Northern (NP-1), and Northwestern (NWP-1) provinces. Three additional composite samples were made similarly from two research fields in Northern (Msa-3; composite of 12 samples) and Eastern (Mse-3; composite of 17 samples) provinces and a 5-ha commercial field in Central province (Com-1; composite of 14). This resulted in a total of nine composite total nucleic acid samples that were advanced for HTS analysis. An aliquot of 80 µL total nucleic acids per composite sample was mixed with 20 µL RNAStable solution (Biomatrica, Japan) and the sample mixture was vacuum dried. The nine dried nucleic acid samples were shipped for sequencing at Inqaba Biotechnical Industries Limited (Pretoria, South Africa) where they were rehydrated by adding 40 µL DEPC treated water and incubated for 10-15 minutes at room temperature. Following quality analysis, samples with ≥5.0 RNA Integrity Number (RIN) were advanced for complementary DNA library preparations using the ScriptSeq™ Complete Kit (plant leaf) (Illumina, CA, USA). Each library was sequenced on the MiSeq platform (Illumina, USA) in 2 x 150 bp mode. The library specific raw Illumina reads were trimmed for adapters using the CLC genomics workbench 6.5 (https://digitalinsights.qiagen.com). The trimmed reads were mapped to the *Phaseolus vulgaris* genome (GenBank accession no. KI548526) using Bowtie2 v2.2.8 (Langmead and Steven, 2012) to subtract the host-specific sequences. The remaining nonhost reads were imported into Geneious R11.1.2 (https://www.geneious.com) and

assembled *de novo*, with default settings. The resulting contigs were sorted by length and used for BLASTn query of the National Center for Biotechnology Information (NCBI) databases. Separately, the trimmed reads were mapped to reference virus sequences obtained from GenBank using the "map to reference genome" tool in Geneious. Subsequently, the mapped or *de novo* assembled virus-like sequences were aligned in MAFFT (Katoh *et al.*, 2002) to obtain a consensus agent-specific genome sequence. The virus genomes were checked for open reading frames (ORF) using Geneious R11.1.2 and the coding regions were translated using the Expasy-translate tool (https://web.expasy.org/translate/). Comparisons were made with exemplar virus isolates and sequences with intact gene encoding regions plus complete 5'- and 3'- untranslated regions (UTRs) were labeled as complete genomes while those that possess intact coding regions and incomplete 5′-UTR and/or 3′-UTR were designated as near complete genomes.

3.3.5 Screening of field-collected samples for HTS detected viruses

Virus-specific primers were designed based on the HTS-derived data (Table 3.1) and used to screen the 640 survey samples. Complementary DNA (cDNA) was synthesized from approximately 0.2 µg of total nucleic acid per sample with the Improm-II^{RT} Reverse Transcription System (Promega, Madison, WI), according to the manufacturer's recommendation. Each virus-specific PCR reaction assay consisted of 1X DreamTaq Buffer, 0.2 mM dNTP, 1U of Dream Taq DNA polymerase, 32 ng cDNA, and 0.2 mM of each forward and reverse primer. The PCR cycling conditions were initial denaturation at 94°C for 2 min, followed by 35 cycles of 94°C for 45 seconds, primer pair specific annealing temperature for 30s, 72°C for 1-3 min (Table 3.1) and final extension at 72°C for 7 min. Known virus-specific positive and negative controls were included in the assays. The PCR products were loaded unto 1.2% agarose gel prestained with 0.5 µg/mL ethidium bromide

alongside a 100 bp or 1 kb DNA Ladder (NEB, UK), electrophoresed in 1X Tris-acetate-EDTA buffer and visualized under UV light using the Gel Doc XR System (Bio-Rad Laboratories, Hercules, CA).

3.3.6 Sanger sequencing and sequence analyses

To confirm the PCR results, virus-specific DNA bands were selected from each province such that they represented all HTS detected viruses in the pooled samples. The selected amplicons were gel eluted with Zymoclean gel DNA recovery kit (Zymoclean, Irvine, CA, USA), according to the manufacturer's protocol. The eluates were then directly sequenced in both orientations at the Virology Laboratory, Department of Infectious Diseases, School of Veterinary Science, University of Zambia. The raw sequences were trimmed for oligonucleotides and a contig was generated for each amplicon in BioEdit program (Hall, 1999) followed by BLASTn analysis (Altschul *et al.*, 1990) to ascertain virus identity and obtain corresponding fragments of published virus genomes to be included in downstream analyses. Multiple sequence alignments were performed for each virus-specific sequence datasets with the program MUSCLE (Edgar 2004) in MEGA7 (Kumar *et al.,* 2016) and the alignment file was used to estimate the best-fit evolutionary substitution model using MEGA7. Phylogenetic reconstruction was achieved by the Maximum Composite Likelihood method using the model-of-best-fit, with 1,000 bootstrap replications. Pairwise comparisons were performed using the sequence demarcation tool (SDT v.1.2) as per Muhire *et al*. (2014).

3.3.7 Field Data analysis

The frequency of occurrence of each of the 17 common bean cultivars (cvs) encountered during the survey were compiled. Occurrence of each of the virus symptom and its visually estimated severity levels across the six provinces and for each of the 17 common bean

Table 3.1. Primers used in reverse transcription polymerase chain reaction (RT-PCR) for the amplification of viruses from survey samples.

Target virus	Prime Name	Primer sequences 5'–3' direction	Annealing temperature (°C)	Target region	Product size (bp)	Reference
CABMV	CABMV-F	GATGGAGATATGGAGTTGTC	55	HC-Pro, P3, CI	2621	This study
	CABMV-R	CTAAGCCCTCTCATTCTAAG		…….		
PvEV-1	EV1-F	AATAGCTCTCAAGACTCAAGCC	55	gp1	769	This study
	EV1-R	TGGGCAGTCCTTGTTATCATC		…….		
PvEV-2	EV2-F	GTTAATGCACGACACGGATATTG	55	gp1	788	This study
	EV2-R	GTGTCTTGTTTGGCACGTAATC		……		
BCMV	BCMV-F	AAGGATGTGAATGCTGGTTCTA	55	CP	536	This study
	BCMV-R	CTCTTGCTCGATCCGATGTT		…….		
PeMoV	PeMoV-cpF	CATTTYACAGCCGCTCCTATT	52	NIb, CP, UTR	2100	This study
	PeMoV-cpR	AGCGATATACACACCAACACTC		…….		
SBMV	SBMV-1F	AGCTGGATTTCCTACCTTTGTG	50	UTR, MP, P2a	873	This study
	SBMV-1R	GGCGTCATCTCCGTTTATCTT		…….		
CMV	CMV-ep2aF	GAGCTGGGTGACTCTGTTAAT	50	RNA2	559	This study
	CMV-ep2aR	GGAAATCACACCACCACTTTG		……		
ETBTV	ETMse-1F	GCCGAGGTAAATGGCATGATA	52	RdRP, REP	988	This study
	ETMse-1R	TGTTGTAAGCCGGAGCAATC		…….		
BCMNV	BCMNVFcpF2	GCTGGGGCCGATGAGAG	55	CP	711	Mwaipopo *et al.*, 2018
	BCM-NVcommonR	GTCCCKTGCAGTGTGCCT		…….		

CABMV = *Cowpea aphid-borne mosaic virus*, PvEV-1 = *Phaseolus vulgaris endornavirus* 1, PvEV-2 = *Phaseolus vulgaris endornavirus* 2, BCMV = *Bean common mosaic virus*, PeMoV = *Peanut mottle virus*, SBMV = *Southern bean mosaic virus*, CMV = *Cucumber mosaic virus*, ETBTV = *Ethiopian tobacco bushy top virus* and BCMNV = *Bean common mosaic necrosis virus*. HC-Pro=Helper Component protein, CP = Coat protein, UTR = Untraslated region, MP = Movement protein, RdRP = RNA-dependent RNA polymerase, REP = Replicase, CI = Cylindrical inclusions, P3 = potyviral protein 3, RNA2 = RNA segment 2, Nib = Nuclear inclusion protein b, gp1 = Glycoprotein 1, P2a = Protease 2a

cultivars were compiled and analyzed using the Chi-square statistics program implemented in Statistical Package for Social Sciences (SPSS) version 22 (SPSS Inc, Chicago, IL). Bar and pie charts for frequencies of sources of seed for planting and distribution of seed varieties were plotted in Microsoft excel 2016 (Microsoft, Redmond, WA). To determine the effect of source of seed (recycled, open market and registered seed suppliers) for planting on the incidence and severity of disease in the farmers' fields, a T-test was conducted in R-Statistics, V4.11.2021-08-10.

3.4 Results

3.4.1 Assessment of virus-like disease symptoms in common bean fields in Zambia

For each sampling domain, a minimum of 20 farmers' fields, ranging from 0.125 to 1.0 ha in size, were selected for visual assessment of foliar virus symptoms and sample collection. Diverse virus-like disease symptoms were observed in farmer's fields across the six surveyed provinces distributed into two AEZs. They included mosaic patterns, upward and downward leaf curl, vein clearing, severe leaf deformation and rolling, vein banding and rugosity (Fig. 3.1 A-F). Of the 640 leaf tissue samples, 585 were symptomatic while 55 were symptomless. The most prevalent symptoms were mosaic and rugosity, occuring in 76.6% (98/128) of the surveyed common bean fields. Incidence of virus-like symptoms ranged from 32.5% in Central Province to 67.5% in Eastern Province, averaging 46.5% for all the 128 fields (Table 3.2). Across the two AEZs, disease incidence was higher in AEZ II (50%) than in AEZ III (~44.8%). Incidence of viral diseases on common bean plants varied significantly ($P<0.0001$) across the AEZs. However, there was no significant difference in disease incidence across the common bean fields within the same AEZ. Seventeen common bean cultivars were encountered during the survey and the frequencies of their cultivation across fields varied greatly, with a predominance of Kabulangeti, Solwezi and two other varieties (White and

Yellow) (Table 3.3). The 17 cultivars showed varying degrees of disease incidence, ranging from 20% for Kifukuma to 100% for Lanyati (Table 3.3). Disease symptom severity among the cultivars also varied significantly (χ^2 =399.40, df = 74, P<0.001), ranging from 2.3 to 3.2 for Kifukuma and Mkalasonga, respectively (Table 3.3). An investigation of the source of seed for planting among the farmers revealed that the seeds were sourced open market, recycled and registered seed suppliers. Farmers predominantly planted seed from open markets (Fig 3.2 A and B) except for Northern Province where most of the farmers (61.9% or 13/21) propagated recycled seeds. However, there was no significant difference in terms of incidence (P-value=0.09827) and severity (P-value=0.58) between open market and recycled seed source. Data for the registered seed suppliers was not included in the analysis because there were only four fields planted with seed from registered seed suppliers compared with 34 for recycled seed source and 50 for open market, respectively. Notably, open market seeds are usually not certified by the Seed Control and Certification Institute (SCCI) of Zambia.

3.4.2 Detection of viruses in composite samples by HTS

High-throughput sequencing of the nine composite samples resulted in the detection of nine distinct virus species belonging to five genera. The viruses are: BCMNV, BCMV, CABMV and PeMoV (genus *Potyvirus*), PvEV-1 and PvEV-2 (genus *Endornavirus*), SBMV (genus *Sobemovirus*), ETBTV (genus *Umbravirus*), and CMV (genus *Cucumovirus*). SBMV was detected in eight (EP-1, MP-1, CP-1, Msa-3, Mse-3, NP-1, LP-1 and NWP-1) of the nine composite samples, PvEV-1 and PvEV-2 in seven samples (EP-1, MP-1, CP-1, Msa-3, Mse-3, NP-1 and LP-1), BCMNV in four samples (EP-1, MP-1, CP-1 and Mse-3), each of CMV and CABMV was detected in three samples (Mse-3, EP-1 and LP-1) and (EP-1, Com-1 and Mse-3), respectively. ETBTV was detected in samples EP-1 and Com-1, whereas PeMoV and BCMV were detected in one sample each i.e, Com-1 from Central Province for PeMoV

Table 3.2 Incidence of virus-like symptomatic common bean plants and the distribution of viruses in farmer's fields in six provinces of Zambia

Province	No. Fields	No. Samples	Field Incidence of virus-like symptoms (%)	Number of samples that were positive in RT-PCR per province									No. of samples with co-infections per province		
				BCMV	BCMNV	PvEV-1	PvEV-2	ETBTV	CMV	PeMoV	CABMV	SBMV	PeMoV ETBTV	*BCMNV+ BCMV	PvEV-1 BCMNV
Luapula	25	131	53.9	0	0	11	5	3	18	0	0	44	0	0	0
Northern	21	97	46.4	0	0	30	21	0	0	0	0	22	0	0	0
Muchinga	20	98	41.7	0	5	6	7	0	0	0	0	27	0	0	0
Central**	20	90	32.5	0	7	4	0	8	0	8	0	6	2	0	0
Eastern**	20	121	67.5	16	35	5	16	7	10	0	2	20	0	8	2
Nwestern	22	103	37	0	2	3	0	0	0	0	0	69	0	0	0
Total (Means)	128	640	(46.5)	16	49	59	49	18	28	8	2	188	2	8	2

**=Provinces marked are located in agro-ecological zone II (AEZII; rainfall >800 mm) and the unmarked are in AEZIII (rainfall >1000 mm).
*=All the 8 BCMNV+BCMV mixed infections were recorded from one field in Eastern Province. CABMV = *Cowpea aphid-borne mosaic virus*; PvEV-1 = *Phaseolus vulgaris endornavirus* 1; PvEV-2 = Phaseolus vulgaris endornavirus 2; BCMV = *Bean common mosaic virus*; PeMoV = *Peanut mottle virus*; SBMV = *Southern bean mosaic virus*; CMV = *Cucumber mosaicvirus*; ETBTV = *Ethiopian tobacco bushy top virus* and BCMNV = *Bean common mosaic necrosis virus.* Incidence (%) of infected plants was calculated as number of plants with virus-like symptoms out of 30 plants inspected), multiplied by 100. Symptom severity was only considered for mosaic symptoms leading to leaf deformation

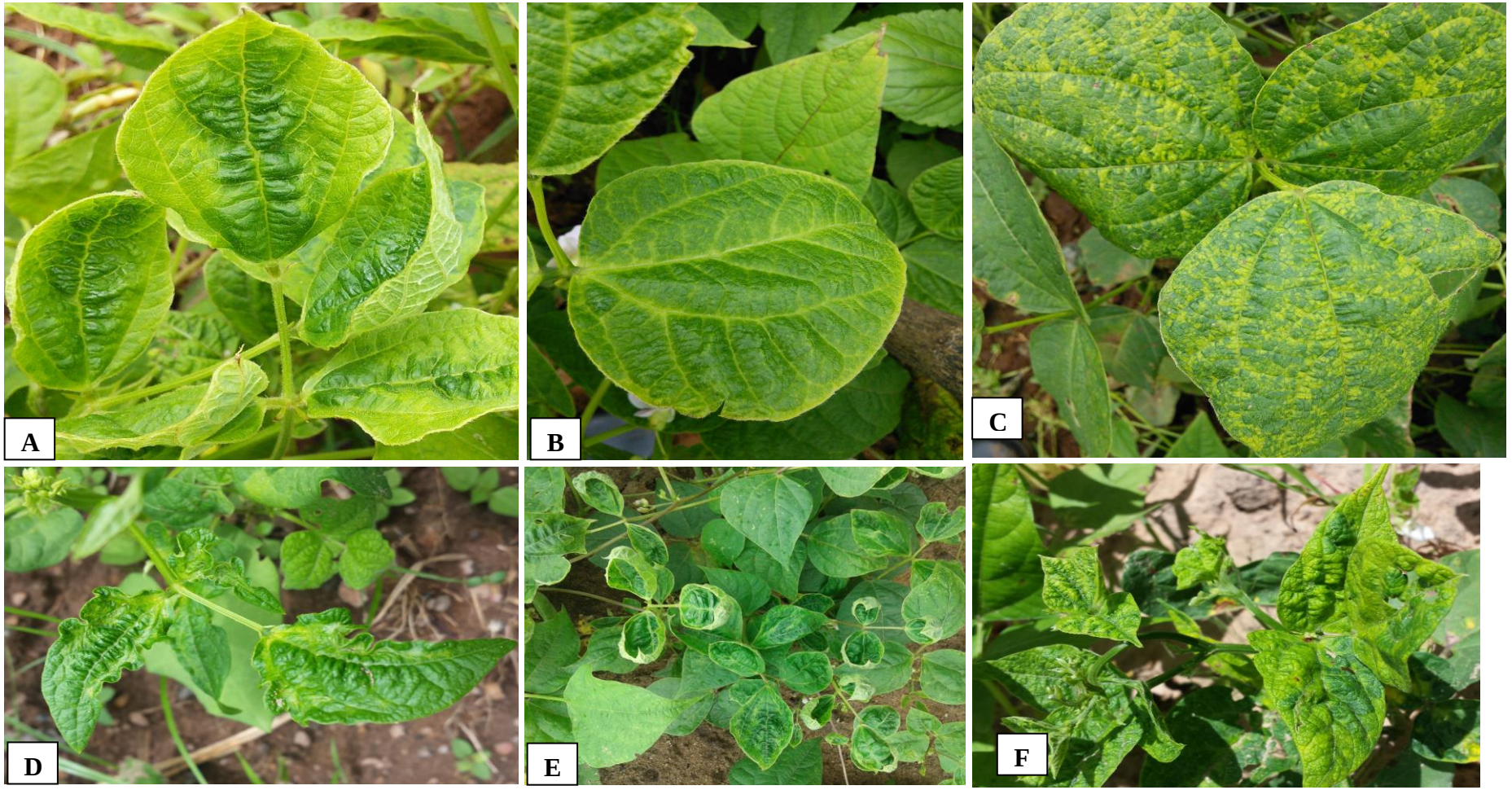
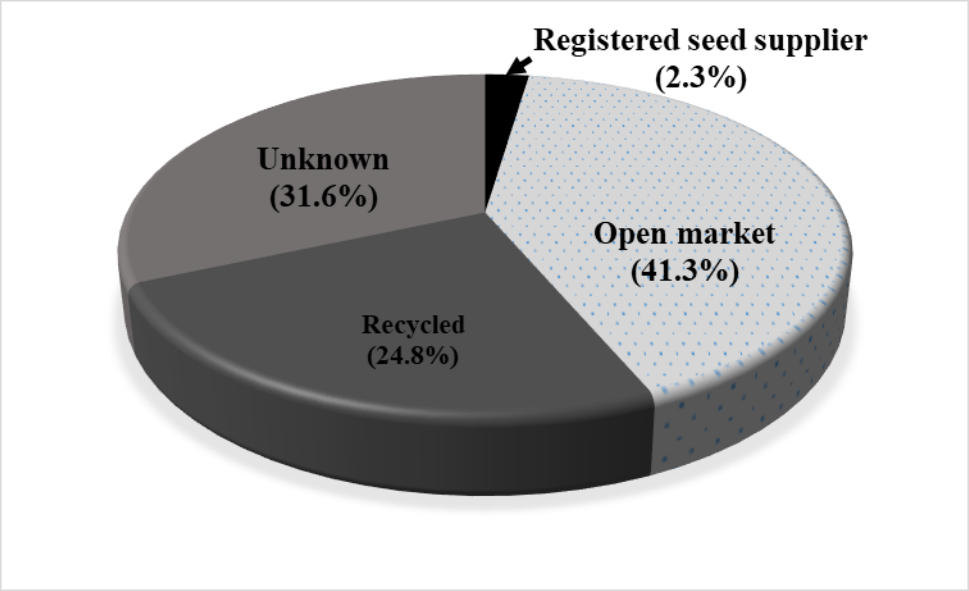
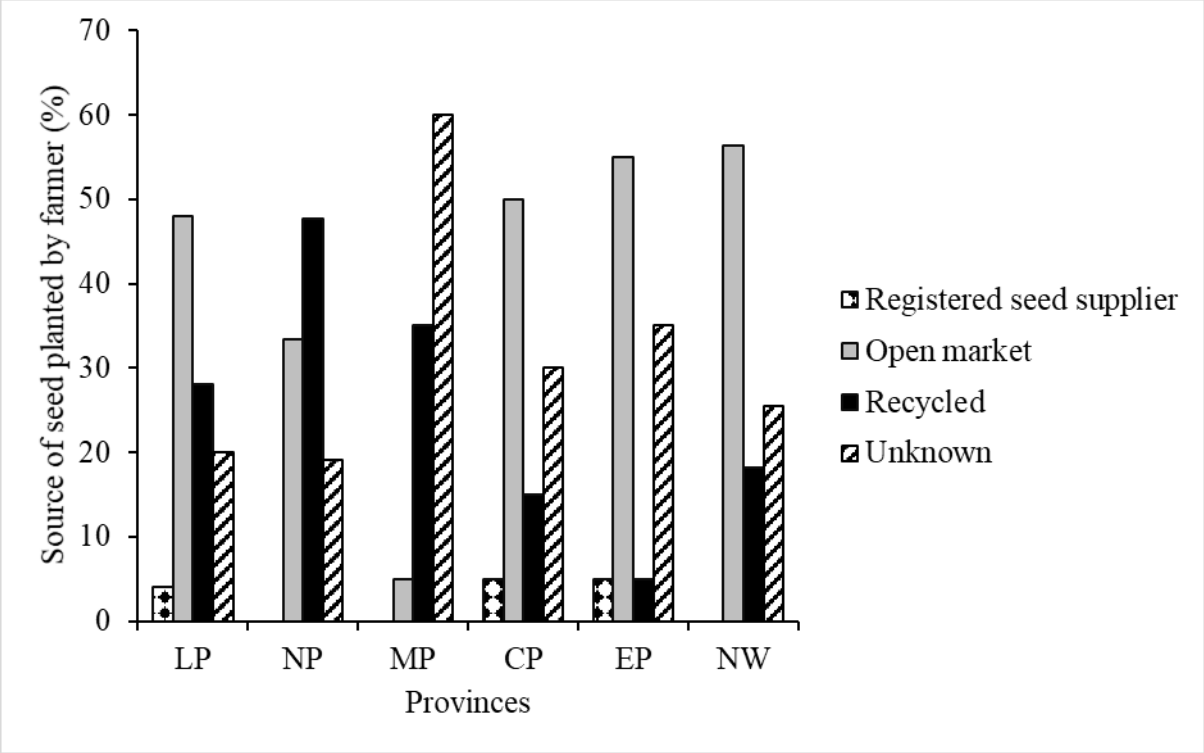


Figure 3.1. Diversity of common bean foliar viral disease symptoms on naturally infected plants in farmers' fields in Zambia.
A, Typical bean common mosaic leaf symptoms; **B**, vein clearing; **C**, mottling; **D**. Vein banding; **E**. upward leaf curling; **F**, bright patchy yellowing and mosaic.



A



B

Figure 3.2A. Country-wide prevalence of common bean seed sources for planting by farmers. **B.** Sources of common bean seeds planted by farmer in different provinces of Zambia. LP = Luapula province, NP = Northern Province, MP = Muchinga Province, CP = Central Province, EP = Eastern Province and NW = Northwestern Province. About 87.5% (112/128) respondents were reachable and 12.5% (16/128) were not reachable. Unknown means that the farmer was either unreachable or that the farmer did not disclose the source of the planted seed.

Table 3.3. Distribution of common bean cultivars encountered during the survey and assessments of their levels of disease incidence, severity and detected viruses.

Occurrence of common bean cultivars in provinces										
Cultivars	LP	NP	MP	CP	EP	NWP	% of (No. of fields in which the cultivar occurred)	Disease Incidence (%)	Severity	Occurrence of RT-PCR detected viruses in samples (No. of positive samples/total samples collected for each cultivar)
Kampemba[b]	+	-	-	-	-	-	3.3 (3)	66.6	2.67	SBMV (7/16)
Lukupa[a]	+	-	-	-	-	-	1.1 (1)	60	2.65	PvEV-1 (3/8), PvEV-2 (1/8)
Lusaka[b]	+	+	+	-	-	-	3.3 (3)	67.8	2.72	BCMNV (2/12), SBMV(7/12)
Kabulangeti[a]	+	+	+	+	+	-	27.8 (25)	51.8	2.56	SBMV (27/88), BCMNV (5/88) ETBTV (1/88), PvEV-1 (8/88), PvEV-2 (6/88)
White**	+	+	+	+	+	-	26.7 (24)	40.2	2.59	BCMNV (1/73), PvEV-1 (10/73), PvEV-2 (8/73), SBMV(5/73), ETBTV (1/73)
Solwezi[b]	+	-	-	+	+	+	13.3 (12)	45	2.56	SBMV (27/72), BCMNV (1/72)
Inzenga[b]	-	-	-	-	-	+	1.1 (1)	73.3	2.82	SBMV (2/7), PvEV-1 (3/7)
Red bean**	-	-	-	-	+	-	1.1 (1)	86.7	2.47	BCMNV (4/10), BCMV (3/10)
Chipupuwe[b]	-	-	-	-	-	+	1.1 (1)	83.3	3.04	SBMV (1/6)
Kifukuma[b]	-	-	-	-	-	+	1.1 (1)	20	2.33	SBMV (2/6)
Serenje[b]	-	-	-	-	-	+	1.1 (1)	50	2.47	SBMV (5/8)
Kansenga[b]	-	+	-	-	-	-	1.1 (1)	26.7	2.25	SBMV (1/8)
Lanyati[b]	-	-	-	-	+	-	1.1 (1)	100	2.97	BCMNV (2/10), BCMV (2/10)
Mangasa[b]	-	-	-	-	-	+	2.2 (2)	41.7	2.48	SBMV (8/11)
Kadyalima[b]	-	-	-	-	+	-	2.2 (2)	55	2.76	BCMNV (3/16), BCMV (1/16)
Mkalasonga[b]	-	-	-	-	+	-	1.1 (1)	83.3	3.2	BCMNV (7/10), BCMV (2/10)
Yellow**	-	+	+	+	-	+	11.1 (10)	40.4	2.41	BCMNV (1/43), SBMV (5/43), PvEV-1 (5/43), PvEV-2 (1/43)
Mean (Total)							(90)	58.3	2.64	
P value ($p \leq 0.05$)							…..	<0.001	<0.001	

+ =cultivar present in at least once field within the province; – = cultivar absent in all fields within the province. LP=Luapula province, NP=Northern Province, MP=Muchinga Province, CP=Central Province, EP=Eastern Province, and NWP=Northwetern Province. a = improved cultivar, b = landraces. Identification of cultivar names was based on farmers' knowledge. ** = Names of these cultivars are not known but farmers identified them by seed coat colour.

and EP-1 from Eastern Province for BCMV. Further results on the molecular characterization of these viruses are provided in Chapter 4.

3.4.3 Incidence and geographical distribution of the HTS-detected viruses

DNA bands of the expected sizes were obtained for each of the HTS-detected viruses using published or newly designed primers (Table 3.1). Of the 640 samples, ~65.2% (417/640) gave virus-specific DNA bands for only one virus, indicating that they were singly infected, ~1.9% (12/640) had mixed infections of different virus combinations, while the remaining 33% (211/640) tested negative. SBMV was the most predominant virus, occurring in ~29.4% (188/640) of the samples, followed by PvEV-1 (9.2% or 59/640), BCMNV and PvEV-2 (7.7% or 49/640 each), CMV (4.4% or 28/640), ETBTV (2.8% or 18/640), BCMV (2.5% or 16/640), PeMoV (1.3% or 8/640), and CABMV (0.3% or 2/640). Among the few mixed-infected samples, 66.7% (8/12) had BCMNV+BCMV, 16.7% (2/12) had PvEV-1+BCMNV, and 16.7% (2/12) had PeMoV+ETBTV (Table 3.2). All the eight mixed infections of BCMNV + BCMV were from the same field. Whereas the majority (98.1% or 421/429) of the virus positive samples were symptomatic, eight samples (1.9% or 8/429) were symptomless.

In terms of virus prevalence, SBMV was more widely distributed with its occurrence in 58.6% (75/128) of fields spread across all the six provinces, PvEV-1 occurred in 35.2% (45/128) of the fields across six provinces, PvEV-2 in 18.0% (23/128) of the fields across four provinces, BCMNV in 18.8% (24/128) of the fields across four provinces, ETBTV in 5.5% (7/128) of fields across three provinces, CMV in 7% (9/128) of the fields across two provinces, BCMV in 3.9% (5/128) of fields in Eastern province, PeMoV in 0.8% (1/128) of fields in Central province, and CABMV in 0.8% (1/128) of fields in Eastern province (Table 3.2). In terms of geographical distribution of the detected viruses, eight of the nine viruses

occurred in Eastern province, followed by Luapula and Central provinces with five viruses each and Muchinga Province with four viruses, Northern and Northwestern provinces with three viruses each (Fig. 3.3). Thus, overall there was more virus diversity in samples collected from AEZ II than those from AEZ III.

3.4.4 DNA fingerprinting of amplified viruses and phylogenetic analysis

Representative virus-specific DNA bands (BCMV = 13), BCMNV = 8, SBMV = 8, CMV = 3, CABMV = 3, PeMoV = 2, ETBTV = 1, PvEV-1 = 2, and PvEV-2 = 2) were eluted from the gel and Sanger sequenced in both orientations and comparatively analysed with corresponding sequences of global isolates of each virus.

The BCMV (MW534341-53) and BCMNV (MW534354-61) nt partial sequences clustered into their respective virus-specific clades on the phylogenetic tree (Fig. 3.4A); the sequence datasets for each virus shared 96 to 100% nt identity with each other and the same levels of nt identities with global isolates of the corresponding viruses. A significant number (20/22) of isolates of both viruses from this study were also closely related to isolates from East Africa (Kenya and Tanzania) suggesting wider dispersal of the BCMV and BCMNV viruses.

The partial SBMV nt sequences were 84 to 99% identical to each other and 82 to 99% identical to corresponding global sequences of the virus. They also segregated into two main phylogroups with isolates Mse-13 and Mse-17 (MN653953-54) clustering separately from the rest of the SBMV isolates from Zambia (Fig. 3.4B), thus indicating the presence of genetically diverse isolates of SBMV in common bean fields in Zambia. The three partial CMV RNA2 sequences obtained in this study (MW534362-64) shared 100% nt identity with each other and 97% nt identity with the isolates from Kenya (MH567343, MH567348 and MH567353). The sequences from Zambia also clustered tightly with the CMV isolates from Kenya within the subgroup IA clade (Fig. 3.4C), suggesting their common ancestry. The four

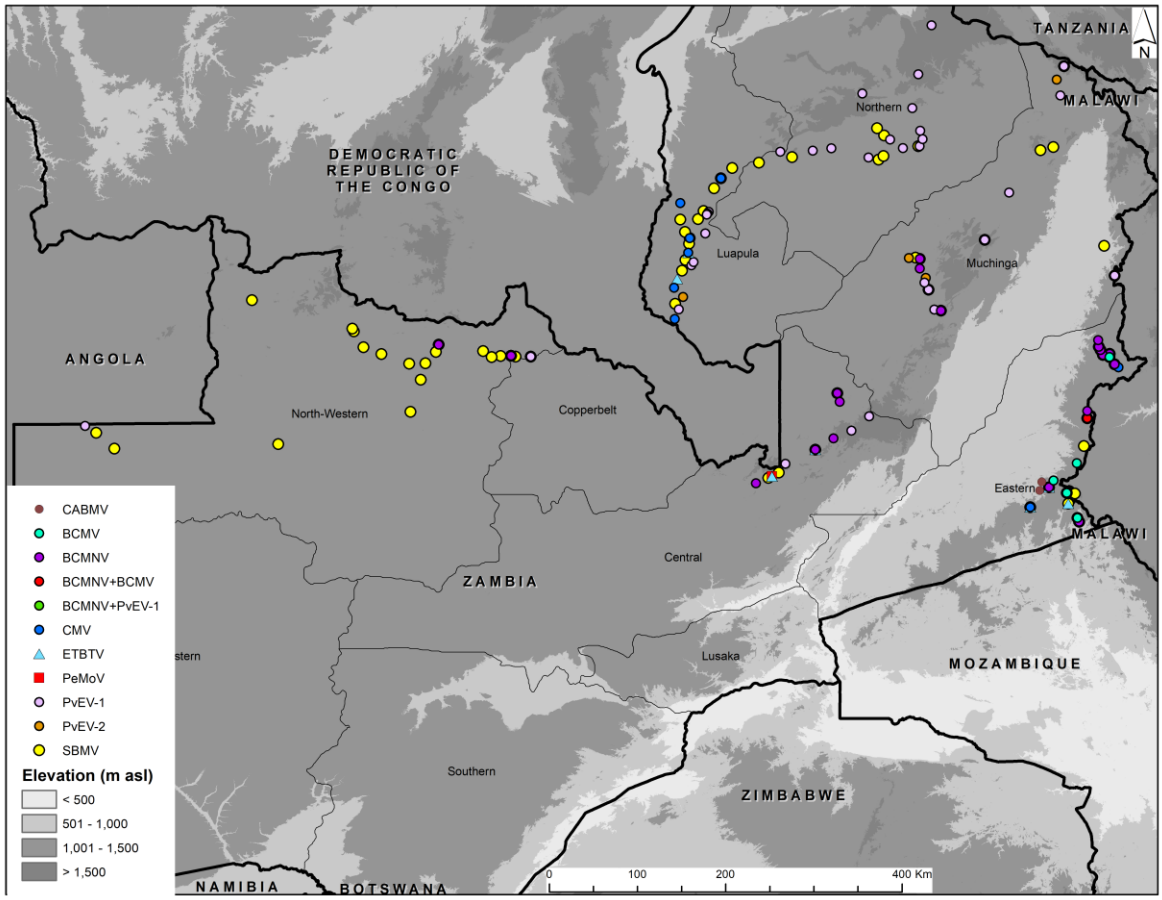


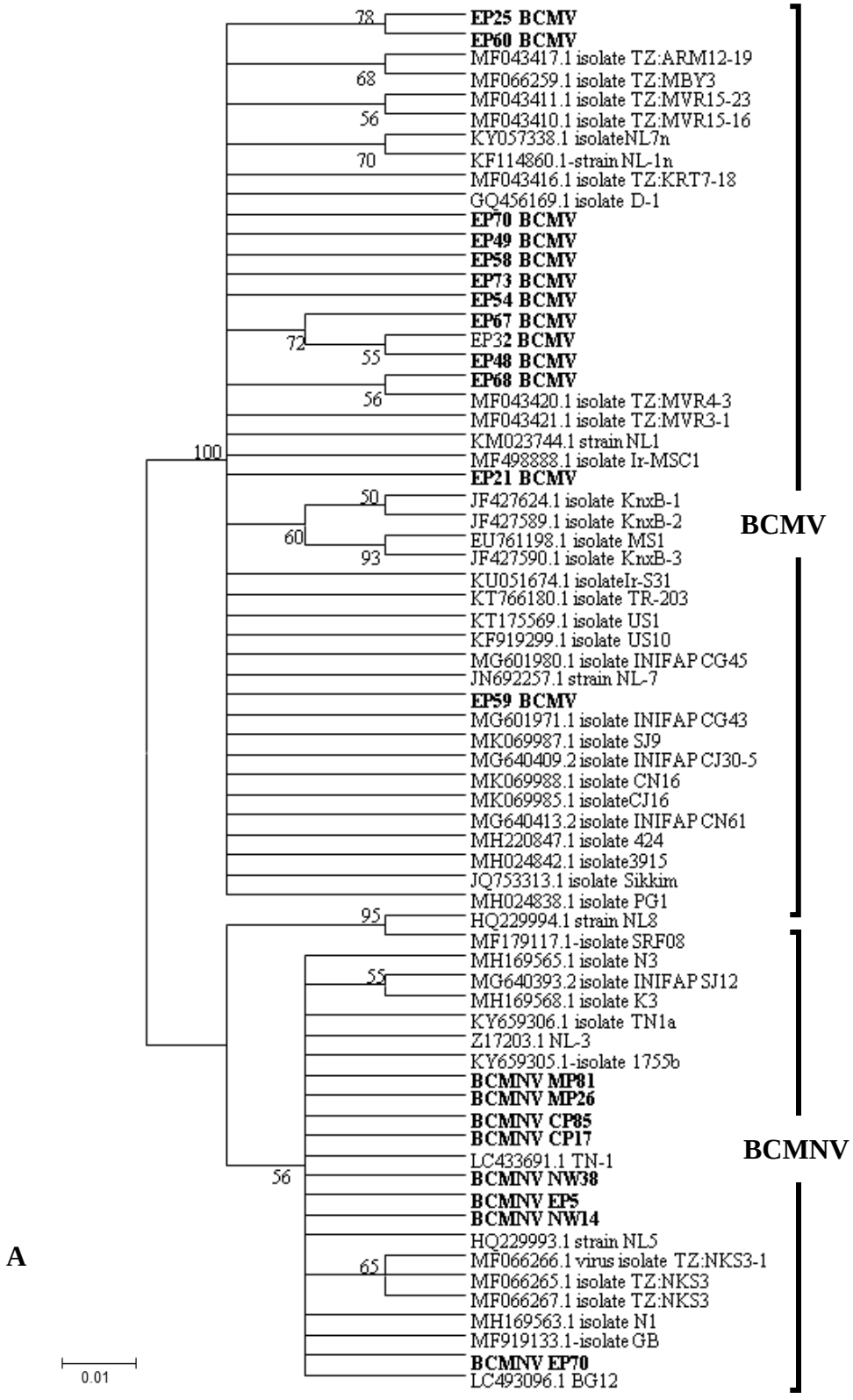
Figure 3.3. Country map of Zambia depicting distribution of common bean-infecting viruses in the six surveyed provinces of Luapula, Northern, Muchinga, Central, Eastern, and Northwestern. The four provinces (Luapua, Northern, Muchinga and Nortwestern) and the two provinces (Central and Eastern) are located in Agroecological zones (AEZs) III (mean annual rainfall >1000 mm) and II (mean annual rainfall ≤800 mm), respectively.

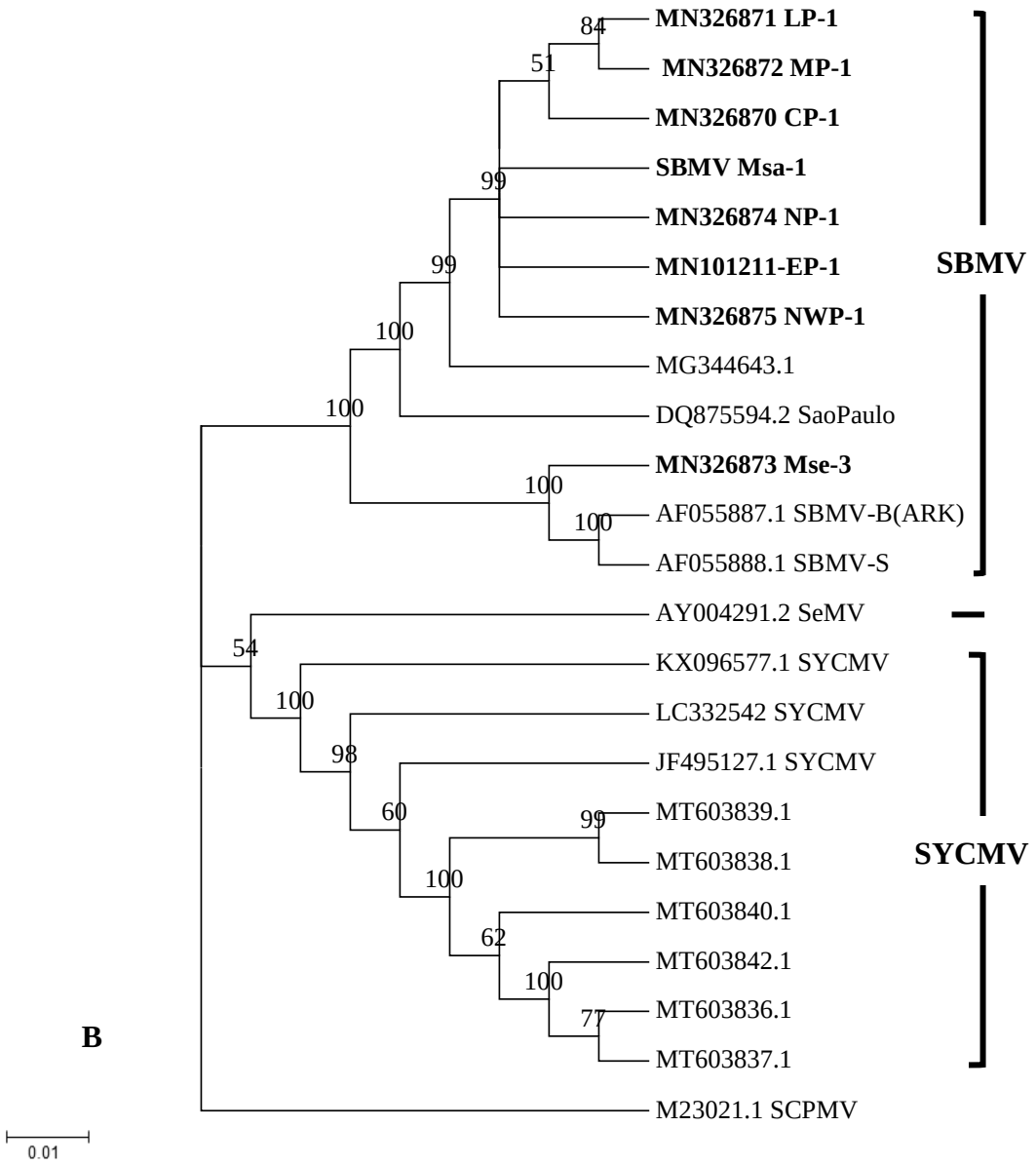
partial PeMoV genome sequences shared 89 to 100% nt sequence identities with each other and 97 to 100% nt identities with global sequences of the virus based on pairwise comparisons. The isolates CP-414-1 and CP415-56 (MT900843-44) shared 98% nt identity with each other and 89 to 91% nt identities with global sequences of exemplar viruses. On the phylogenetic tree, the two PeMoV sequences clustered separately indicating genetic diversity among the PeMoV viruses in Zambia. The three CABMV partial sequences obtained in this study (CABMV-ZM, CAB-4 and CAMB-WP12) shared 99 to 100% nt identities with each other and 82 to 97% with global sequences of CABMV based on pairwise comparisons. They

shared very close phylogenetic relationship and clustered tightly with isolate CABMV-Z (AF348210) from Zimbabwe (Figure 3.4E). Sequences of two isolates of ETBTV (MT225090-91) were obtained in this study. They shared 99% nt identity with each other, 97% nt identity with ETBTV isolate Malawi (LC494673) from neighboring Malawi, and 88% nt identity with the isolate 18-2 (KJ918748) from Ethiopia. The ETBTV isolates were also more phylogenetically related with isolate Malawi than with isolate 18-2 (Figure 3.4F). Two partial sequences of two distinct viruses were obtained from two virus-specific DNA fragments. The one Sanger-derived sequence specific to Phaseolus vulgaris endornavirus 1 isolate EV1 (MW534365) shared maximum 99.2% nt identities with exemplar global viruses based on pairwise comparison. The other partial sequence was specific to Phaseolus vulgaris endornavirus 2 isolate EV2 (MT534366) and shared maximum 100% nt identity with homologous sequences of global isolates of the virus.

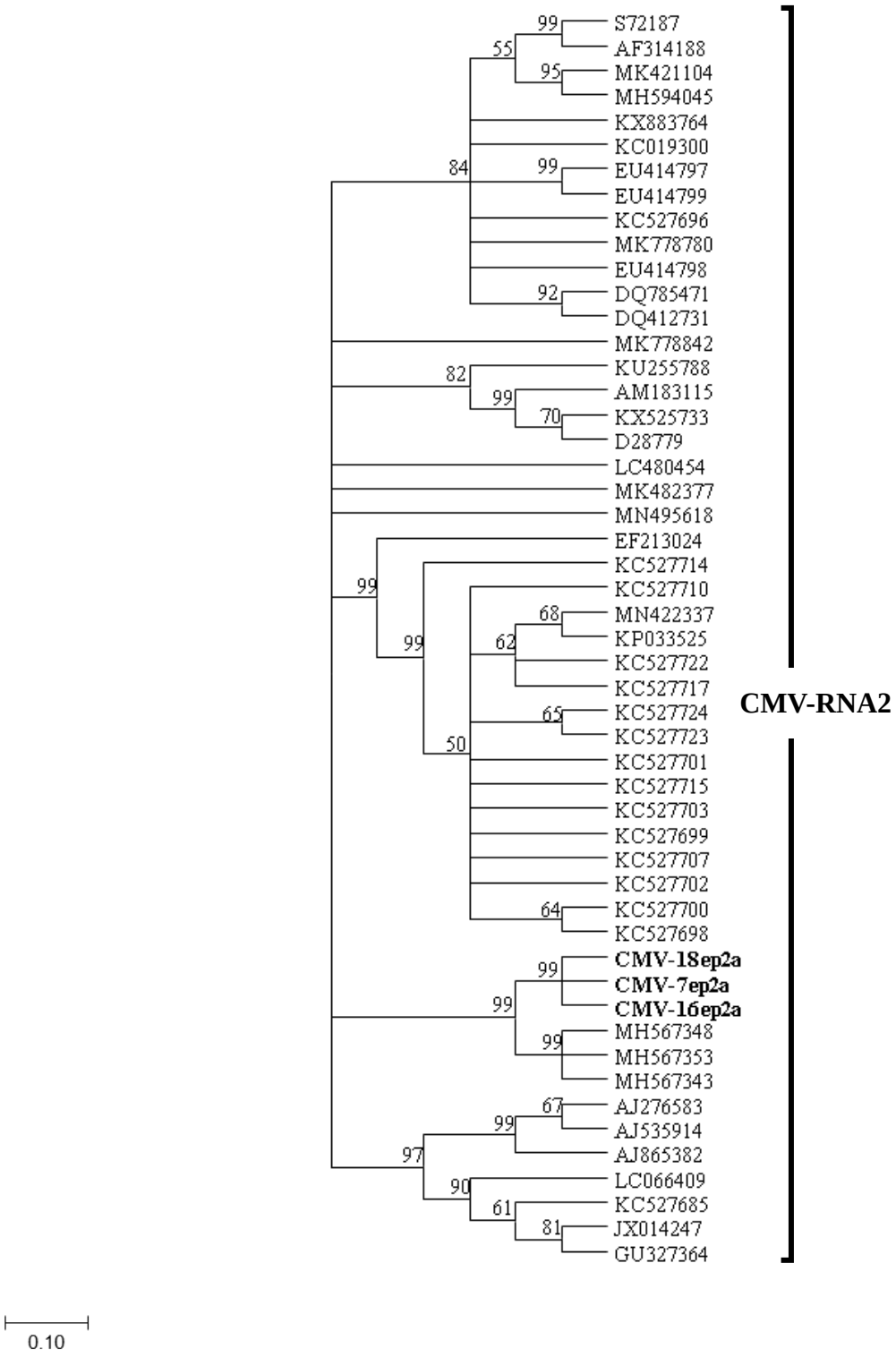
3.5 Discussion

In this study, nine common bean-infecting viruses belonging to five distinct genera were detected across 128 famer's fields in six provinces of Zambia using a combination of HTS, RT-PCR and Sanger sequencing. The viruses BCMNV, BCMV, CABMV and PeMoV (genus *Potyvirus*), CMV (genus *Cucumovirus*), ETBTV (genus *Umbravirus*), PvEV-1 and PvEV-2 (genus *Alphaendornavirus*), and SBMV (genus *Sobemovirus*) occurred mainly as single infections in the analyzed samples but a small proportion of mixed infected plants were also identified. In previous studies, four potyviruses (BCMNV, BCMV, CABMV and PeMoV) and a cucumovirus (CMV) were reported in Zambia using serological and biological assays (Vetten and Allen 1991; Kannaiyan and Hachiwa 1993; Spence and Walkey 1995).

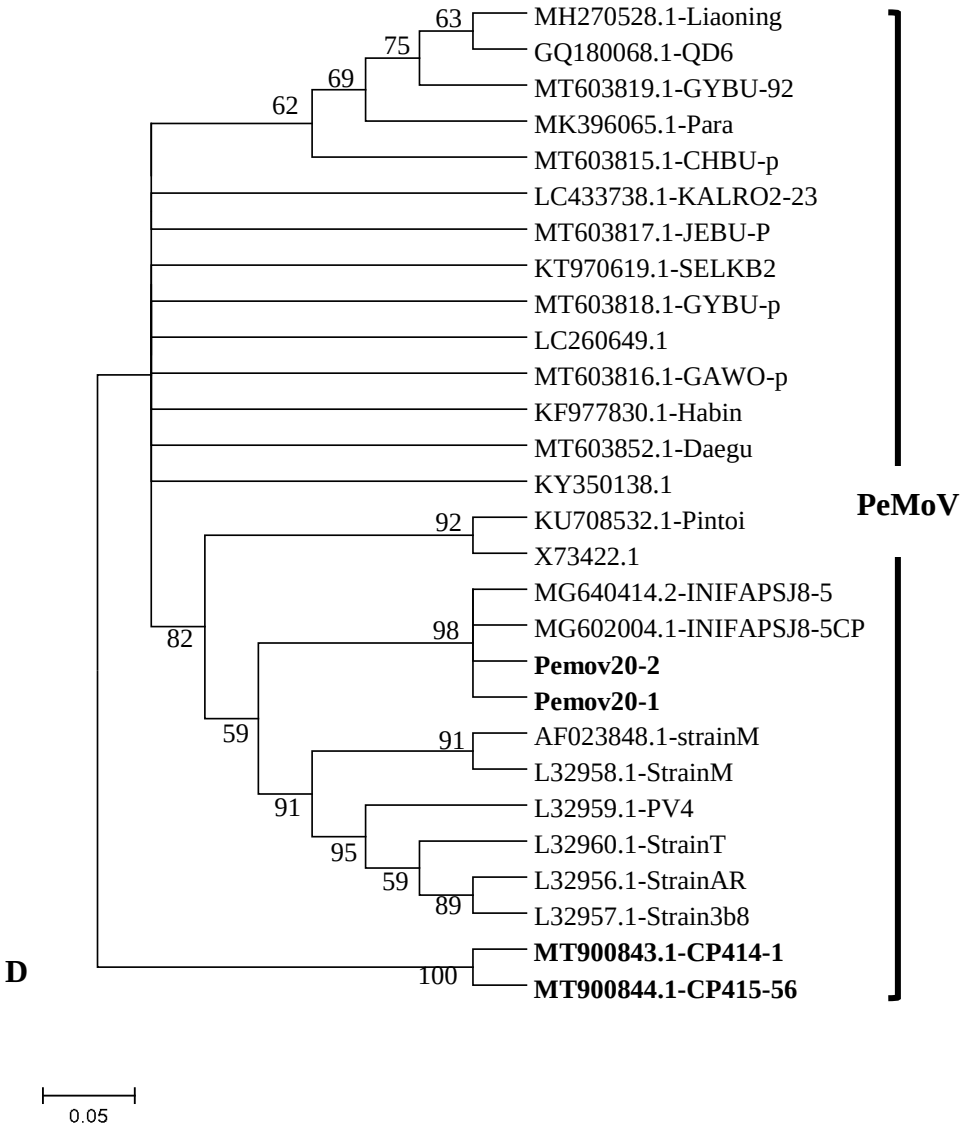

78
EP25 BCMV
EP60 BCMV
MF043417.1 isolate TZ:ARM12-19
68
MF066259.1 isolate TZ:MBY3
MF043411.1 isolate TZ:MVR15-23
56
MF043410.1 isolate TZ:MVR15-16
KY057338.1 isolateNL7n
70
KF114860.1-strain NL-1n
MF043416.1 isolate TZ:KRT7-18
GQ456169.1 isolate D-1
EP70 BCMV
EP49 BCMV
EP58 BCMV
EP73 BCMV
EP54 BCMV
EP67 BCMV
72
EP32 BCMV
55
EP48 BCMV
EP68 BCMV
56
MF043420.1 isolate TZ:MVR4-3
MF043421.1 isolate TZ:MVR3-1
KM023744.1 strain NL1
100
MF498888.1 isolate Ir-MSC1
EP21 BCMV
50
JF427624.1 isolate KnxB-1
JF427589.1 isolate KnxB-2
60
EU761198.1 isolate MS1
93
JF427590.1 isolate KnxB-3
KU051674.1 isolateIr-S31
KT766180.1 isolate TR-203
KT175569.1 isolate US1
KF919299.1 isolate US10
MG601980.1 isolate INIFAP CG45
JN692257.1 strain NL-7
EP59 BCMV
MG601971.1 isolate INIFAP CG43
MK069987.1 isolate SJ9
MG640409.2 isolate INIFAP CJ30-5
MK069988.1 isolate CN16
MK069985.1 isolateCJ16
MG640413.2 isolate INIFAP CN61
MH220847.1 isolate 424
MH024842.1 isolate3915
JQ753313.1 isolate Sikkim
MH024838.1 isolate PG1
95
HQ229994.1 strain NL8
MF179117.1-isolate SRF08
MH169565.1 isolate N3
55
MG640393.2 isolate INIFAP SJ12
MH169568.1 isolate K3
KY659306.1 isolate TN1a
Z17203.1 NL-3
KY659305.1-isolate 1755b
BCMNV MP81
BCMNV MP26
BCMNV CP85
BCMNV CP17
LC433691.1 TN-1
56
BCMNV NW38
BCMNV EP5
BCMNV NW14
HQ229993.1 strain NL5
MF066266.1 virus isolate TZ:NKS3-1
65
MF066265.1 isolate TZ:NKS3
MF066267.1 isolate TZ:NKS3
MH169563.1 isolate N1
MF919133.1-isolate GB
BCMNV EP70
LC493096.1 BG12
0.01
BCMV
BCMNV
A

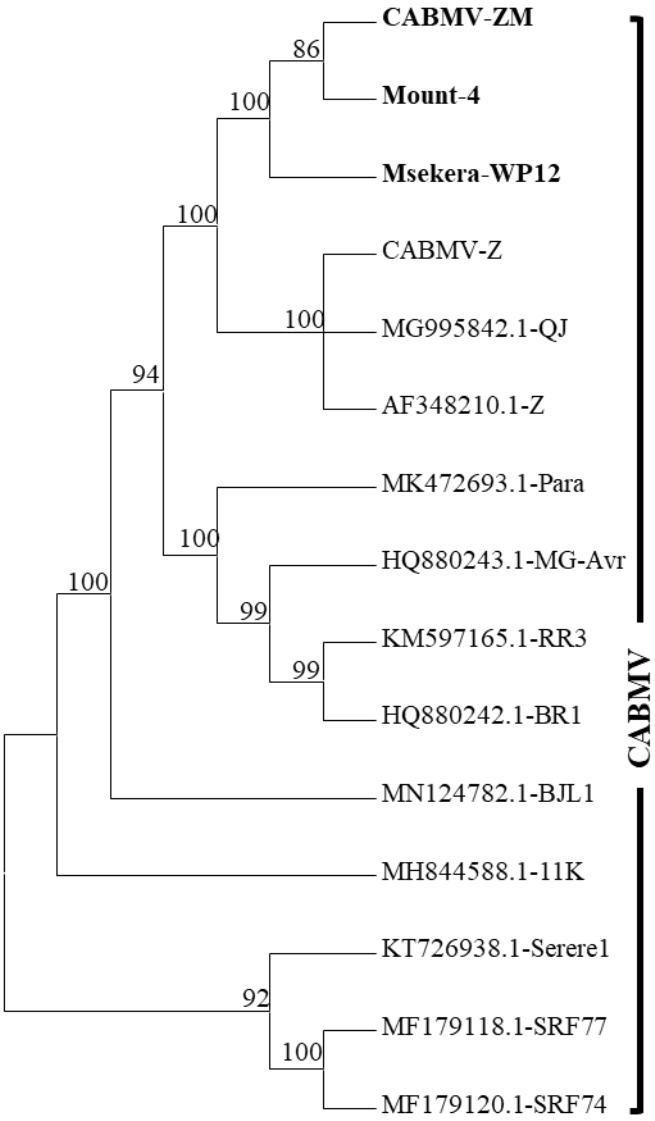

MN326871 LP-1
MN326872 MP-1
MN326870 CP-1
SBMV Msa-1
MN326874 NP-1
MN101211-EP-1
MN326875 NWP-1
MG344643.1
DQ875594.2 SaoPaulo
MN326873 Mse-3
AF055887.1 SBMV-B(ARK)
AF055888.1 SBMV-S
AY004291.2 SeMV
KX096577.1 SYCMV
LC332542 SYCMV
JF495127.1 SYCMV
MT603839.1
MT603838.1
MT603840.1
MT603842.1
MT603836.1
MT603837.1
M23021.1 SCPMV
SBMV
SYCMV
84
51
99
99
100
100
100
100
54
100
98
60
99
100
62
100
77
0.01

B

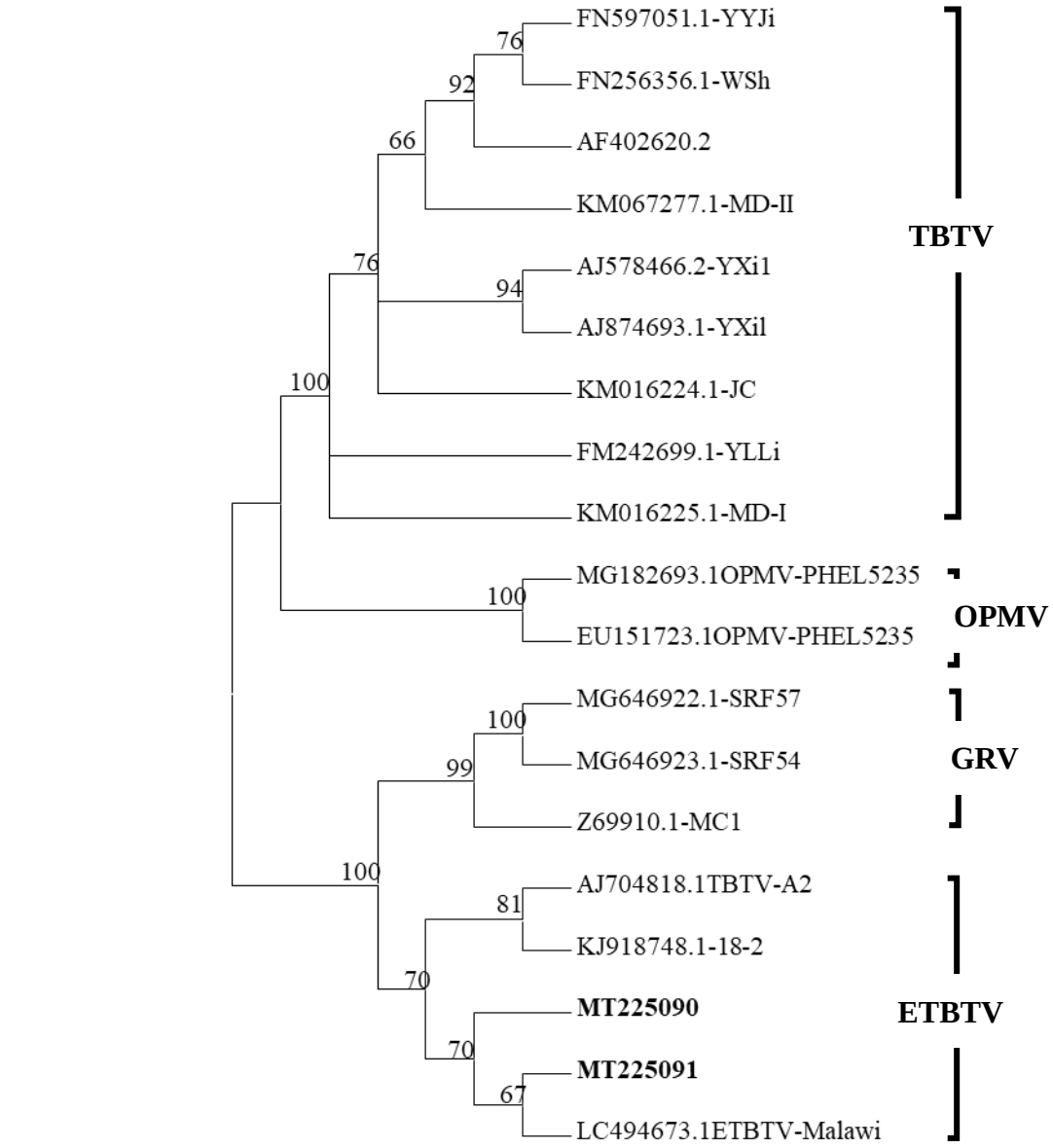

S72187
AF314188
MK421104
MH594045
KX883764
KC019300
EU414797
EU414799
KC527696
MK778780
EU414798
DQ785471
DQ412731
MK778842
KU255788
AM183115
KX525733
D28779
LC480454
MK482377
MN495618
EF213024
KC527714
KC527710
MN422337
KP033525
KC527722
KC527717
KC527724
KC527723
KC527701
KC527715
KC527703
KC527699
KC527707
KC527702
KC527700
KC527698
CMV-18ep2a
CMV-7ep2a
CMV-16ep2a
MH567348
MH567353
MH567343
AJ276583
AJ535914
AJ865382
LC066409
KC527685
JX014247
GU327364
CMV-RNA2
0.10

C


63
75
69
62
MH270528.1-Liaoning
GQ180068.1-QD6
MT603819.1-GYBU-92
MK396065.1-Para
MT603815.1-CHBU-p
LC433738.1-KALRO2-23
MT603817.1-JEBU-P
KT970619.1-SELKB2
MT603818.1-GYBU-p
LC260649.1
MT603816.1-GAWO-p
KF977830.1-Habin
MT603852.1-Daegu
KY350138.1
92
KU708532.1-Pintoi
X73422.1
MG640414.2-INIFAPSJ8-5
98
82
MG602004.1-INIFAPSJ8-5CP
Pemov20-2
Pemov20-1
59
91
AF023848.1-strainM
L32958.1-StrainM
91
L32959.1-PV4
95
L32960.1-StrainT
59
L32956.1-StrainAR
89
L32957.1-Strain3b8
MT900843.1-CP414-1
100
MT900844.1-CP415-56
PeMoV
D
0.05


CABMV-ZM
Mount-4
Msekera-WP12
CABMV-Z
MG995842.1-QJ
AF348210.1-Z
MK472693.1-Para
HQ880243.1-MG-Avr
KM597165.1-RR3
HQ880242.1-BR1
MN124782.1-BJL1
MH844588.1-11K
KT726938.1-Serere1
MF179118.1-SRF77
MF179120.1-SRF74
86
100
100
100
94
100
100
99
99
92
100
CABMV
0.01

E


FN597051.1-YYJi
FN256356.1-WSh
AF402620.2
KM067277.1-MD-II
AJ578466.2-YXi1
AJ874693.1-YXil
KM016224.1-JC
FM242699.1-YLLi
KM016225.1-MD-I
MG182693.1OPMV-PHEL5235
EU151723.1OPMV-PHEL5235
MG646922.1-SRF57
MG646923.1-SRF54
Z69910.1-MC1
AJ704818.1TBTV-A2
KJ918748.1-18-2
MT225090
MT225091
LC494673.1ETBTV-Malawi
76
92
66
76
94
100
100
99
100
81
70
70
67
TBTV
OPMV
GRV
ETBTV

F

0.10

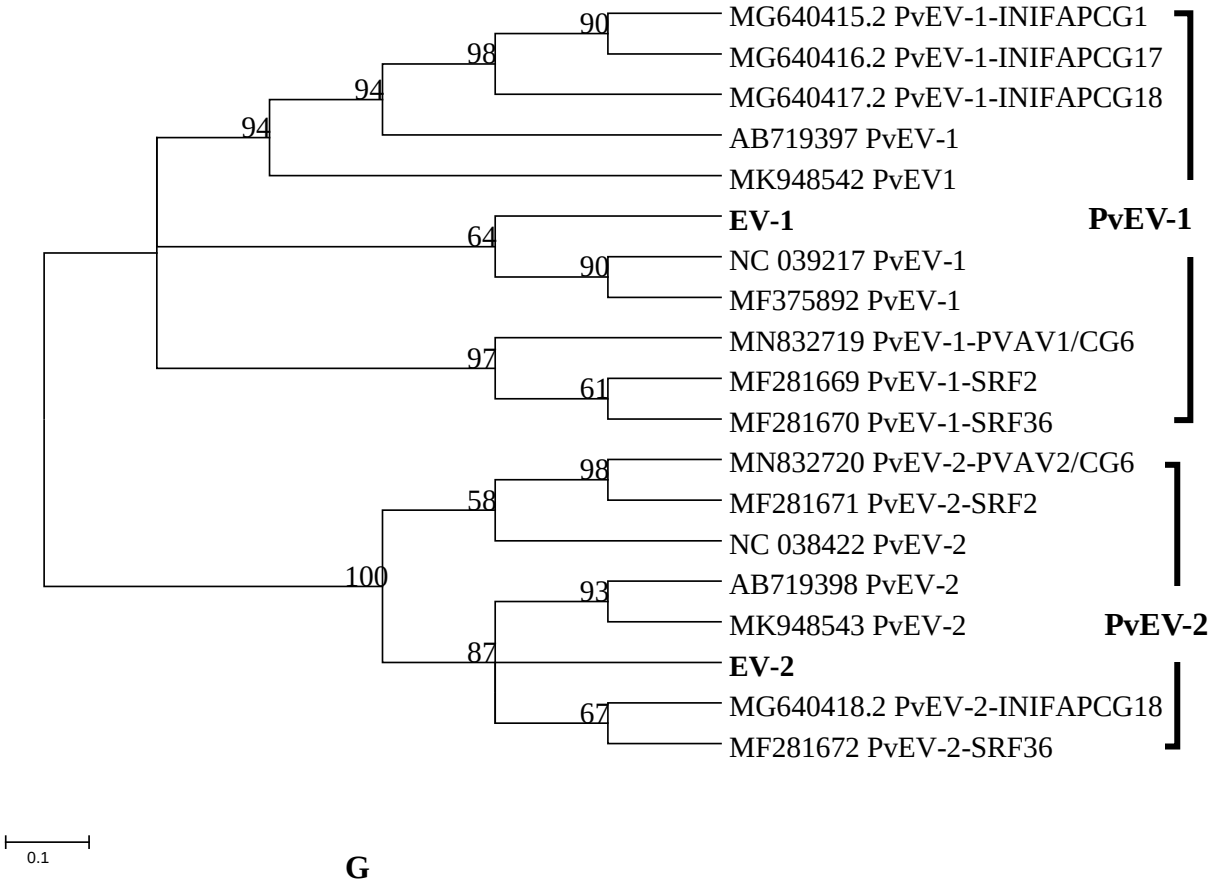


Figure 3.4 Unrooted cladograms of virus sequences detected from the survey samples and corresponding sequenes of isolates retrieved from GenBank. Phylogenetic trees reconstructed using the Maximum Composite Likelihood Method based on best-fit model the Kimura 2 (K+2) (A, B, C and E), Tamura 3 (T+3) (D, F and G) models. BCMNV = *Bean common mosaic necrosis virus*; BCMV = *Bean common mosaic virus*; CABMV = *Cowpea aphid-borne mosaic virus*; CMV = *Cucumber mosaic virus*; ETBTV = *Ethiopian tobacco bushy top virus*; PeMoV = *Peanut mottle virus*; PvEV-1 and 2= Phaseolus vulgaris endornavirus 1 and 2; SBMV = *Southern bean mosaic virus*; TBTV=*Tobacco bushy top virus*; SeMV=*Sesame mosaic virus*; SYCMV=*Soya bean yellow common mosaic virus*; SCPMV=*Southern cowpea mosaic virus*; GRV=*Groundnut rosette virus*; OPMV=*Opium poppy mosaic virus*. Isolates from this study are shown in bold font.

However, the definitive identities of the genetic variants of these viruses were not determined until now since neither serology nor bioassays have the resolution power required for such a task. In addition to presenting sequence-based evidence for the occurrences of genetic variants of five previously reported viruses of common bean in Zambia for the first time, the results of this study also confirm the occurrences of ETBTV, PvEV-1, PvEV-2, and SBMV in Zambia. Additionally, this study represents the first documentation of the countrywide distribution of all nine viruses in common bean fields in Zambia as illustrated in Table 3.1 and Figure 3.1.

The high incidences of virus-like symptoms (67.5%) coupled with the detection of 8/9 viruses in Eastern province and their low incidences coupled with the detection of predominantly SBMV in Northwestern province (37%) (Table 3.2) denote the two extremes of hot and cold spots for common bean viruses in Zambia. The higher disease incidence and virus diversities recorded in Eastern province could be due to greater proximities of the surveyed common bean farms in this province to the land border shared between Zambia and Malawi (Fig. 3.1) which may have facilitated the transboundary movement of viruliferous vectors. In addition, farmers in Eastern province grew more of the susceptible common bean landraces (Table 3.3), which may have also contributed to its status as a hotspot for viruses. The Eastern province also has the most complex 'crop culture', relative to other provinces, in that mixed cropping of several crops such as tobacco, legumes, maize, cucurbits are common. Although farmers typically implement mixed cropping as insurance against crop failure, the crops often grown in mixed cropping with common beans are hosts to the same set of viral pathogens and their vectors hence they serve as continuous sources of virus inoculum to each other. In contrast, mixed cropping in Northwestern province involve common bean and sweet potato; two crops that are seldom reported to share similar viruses.

The results from this study indicate that most Zambian common bean farmers grow locally purchased uncertified seeds, a practice that is common in Africa and especially in Eastern and Southern Africa (David and Sperling, 1999; Katungi *et al.,* 2009; McGuire and Sperling, 2016; Wilkus *et al.,* 2018). Katungi *et al*. (2009) also observed that transactions of seeds over long distances are limited, hence common bean seed cultivars grown in certain localities have limited geographical distribution. This implies that local spread of common bean viruses is influenced more by involvement and abundance of insect vectors in the transmission of common bean-infecting viruses rather than movement of seed. Thus, management of viral diseases would require a careful selection of crops in rotation with common bean that limit exponential multiplication of vectors.

The detection and widespread distribution of SBMV, PvEV-1 and PvEV-2 in Zambia was a surprising finding considering that the viruses were not identified in previous reports (Vetten and Allen, 1991). It is highly improbable that the three viruses were recently introduced in Zambia through seed or other alternative transmission pathways. Thus, a more plausible explanation for the failure to detect SBMV, PvEV-1 and PvEV-2 in previous reports could be due to the limitation in the detection methods used, the lack of symptoms induced in infected plants (as in the case of PvEV-1/2) to warrant further studies and the biased focused study of BCMNV and BCMV as the most important viruses infecting common bean in East, Central and Southern Africa (Sengooba *et al.,* 1997; Myers *et al.,* 2000; Njau and Lymo, 2000; Mangeni *et al.,* 2014). This view is supported by a recent review of viruses infecting common bean in Tanzania (Mwaipopo et al., 2017). Further, such asymmetrical studies excluded other potentially important viruses. Although SBMV, PvEV-1 and PvEV-2 have been reported in several common bean growing areas globally and from a few countries in Africa, there is a lack of information on their occurrence in countries that share land borders with Zambia, except for Tanzania (Mwaipopo *et al.,* 2018). Thus, the genome sequence data generated in

this study will be useful in future studies to decipher the evolutionary histories of the three viruses from Zambia and other African countries. Previous studies revealed that SBMV, PvEV-1 and PvEV-2 share similar transmission pathways; vertically through contaminated seeds and pollen (Lecoq *et al.,* 2011; Nordenstedt *et al.,* 2017) and horizontally through insect vectors (Tremaine and Hamilton 1983; Nordenstedt *et al.,* 2017). With such diverse transmission routes, it is possible that a single introduction in one area could quickly spread to susceptible hosts in other regions (Johansen *et al.,* 1994). Of the six provinces surveyed, five share long and usually poorly regulated borders with one or two of the eight countries neighboring Zambia. Thus, the widespread distribution of SBMV, PvEV-1 and PvEV-2 has epidemiological implications beyond the geographical boundaries of Zambia. For example, the occurrence of SBMV at several sites in Northwestern province bordering both Angola and the Democratic Republic of Congo (DRC), poses a real likelihood of its spread to both countries. Hence, this calls for vigilance and implementation of necessary phytosanitary regulations that could minimize the spread of SBMV to other countries in the region.

The relatively higher incidences of BCMNV (7.7%) and its occurrence in four of the six surveyed provinces compared to BCMV, which occurred only in Eastern province with a country-wide incidence of 2.5% is consistent with the findings from other studies. For example, BCMNV is reported to be endemic in Central and East Africa where it is often found in greater occurrences in many countries in both regions (Worrall *et al.,* 2015). In a recent study, BCMV was reported in Tanzania only in two of the six surveyed agroecological regions compared to BCMNV that was present in samples from five regions (Mwaipopo *et al.,* 2018). BCMNV was also the only potyvirus detected in a metagenomic study of viruses infecting common bean in Kenya (Mutuku *et al.,* 2018). In Zambia, a preponderance of serotype A (BCMNV) over serotype B (BCMV) was reported in a limited number of samples collected from Eastern province of Zambia (Vetten and Allen 1991). Nevertheless, the

observed restricted distribution of BCMV to Eastern province in this study was unexpected because both BCMNV and BCMV have similar transmission pathways being through seed dispersal, aphid vector species and human conveyance for long distance spread of the two diseases. Incidentally, most transmissibility studies of BCMNV and BCMV were conducted prior to reclassification of what was referred to as BCMV into BCMNV and BCMV, thus currently literature is not clear on transmission characteristics of the two closely related viral agents (Worrall *et al.,* 2015). Regardless, it is reasonable to hypothesize that the transmission efficiencies in seed and aphid vectors for the two viruses may be varied and this could be the reason for the observed differences in the distribution of BCMNV and BCMV in this study.

Generally, data in this study does not show any specific trends in the incidence and severity of virus symptoms between improved and landraces although most improved common bean cultivars in Zambia are deemed resistant to BCMNV and BCMV (Muimui, 2015). Over the years, there has been intensive efforts for resistance breeding. The selection processes for particular preferred straits may inadvertently erode the gene pool thereby reducing genetic variability and hereditary resistance that can prove costly in the event of seismic shifts in disease epidemics (Luquet *et al.,* 2012). The focus on selecting common bean cultivars with resistance genes to BCMNV and BCMV may have inadvertently eroded hereditary resistance to SBMV and other viruses reported in this study resulting in progenies that are susceptible to SBMV. Hence, SBMV is now widespread and may become the most important disease of common bean in Zambia.

This study presents comprehensive data on the distribution of common bean-infecting viruses in Zambia. It confirmed the occurrence of previously reported viruses BCMNV, BCMV, CABMV, CMV and PeMoV and revealed the occurrence of ETBTV, SBMV, PvEV-1 and PvEV-2 reported for the first time in Zambia. The virus distribution map generated provides details of the hotspots of detected viruses, which could be useful for breeding and

management purposes. It is hoped that this information will also help virologists from other countries bordering Zambia to expand their studies to include viruses other than BCMNV and BCMV. The diagnostic assays designed in this study are already being used by the national legumes programmes in Zambia to monitor viruses in common bean fields.

CHAPTER FOUR

MOLECULAR CHARACTERISATION OF VIRUSES INFECTING COMMON BEAN IN ZAMBIA

4.1 Abstract

Common bean-infecting viruses and their molecular diversities occurring in farmers' fields are not well studied in Zambia. Thus, this study aimed to characterize viruses by metagenomics analysis of field collected samples. Of the 640 leaf tissue samples that were collected during the survey for common bean viruses in six provinces of Zambia, 223 samples selected based on symptom diversity were subsumed into nine composite samples. The composite samples were subjected to Illumina sequencing resulting in 22.9 M high-throughput sequencing (HTS) raw paired-end reads. About 748,897 contigs (150-300 nt in length) were further assembled into 24 complete and near complete genomes of nine viruses belonging to five genera. These include bean common mosaic necrosis virus (BCMNV; genus *Potyvirus*), bean common mosaic virus (BCMV; *Potyvirus*), cowpea aphid-borne mosaic virus (CABMV; *Potyvirus*), peanut mottle virus (PeMoV; *Potyvirus*), Phaseolus vulgaris alphaendornavirus 1 (PvEV-1; *Endornavirus*), Phaseolus vulgaris alphaendornavirus 2 (PvEV-2; *Endornavirus*), Ethiopian tobacco bushy top virus (ETBTV; genus *Umbravirus*), cucumber mosaic virus (CMV; genus *Cucumovirus*), and southern bean mosaic virus (SBMV; genus *Sobemovirus*). Pairwise analyses of the sequences of the different viruses detected in this study revealed 88 to ~100% nucleotide (nt) sequence identities with global isolates of corresponding viruses. Hence, results from this study indicate considerable variabilities among viruses circulating in common bean fields in Zambia. Information from this study will be useful in developing diagnostic tools for detecting common bean-infecting viruses.

4.2 Introduction

There are at least 30 viruses that have been characterized from naturally infected common bean plants (Loebenstein and Thottappilly, 2004) worldwide, among them being members of the genera *Potyvirus*, *Endornavirus*, *Umbravirus*, *Cucumovirus* and *Sobemovirus*. The geographical distribution and frequency of occurrence is not very well known for some viruses such as Ethiopian tobacco bushy top virus (ETBTV; genus *Umbravirus*) whereas others such as cucumber mosaic virus (CMV; genus *Cucumovirus*) (Palukaitis and García-Arenal, 2003; Jacquemond, 2012), bean common mosaic virus (BCMV) and bean common mosaic necrosis virus (BCMNV) (genus *Potyvirus*) (Worrall *et al.,* 2015), and southern bean mosaic virus (SBMV; genus *Sobemovirus*) (Tamm and Truve, 2000) are widespread.

The genus *Potyvirus* has the largest number of economically important plant viruses capable of causing significant yield losses in different crops including common bean (Revers and Garcia, 2015). Several potyviruses have been recovered from infected common bean plants among them BCMV, BCMNV, CABMV and PeMoV (Worrall *et al.*, 2015; Mutuku *et al.,* 2018; Mwaipopo *et al.,* 2018). Typical genome organization of potyvirids depicts positive sense single stranded RNA (+ssRNA) measuring 9,300 to 10,800 nucleotides (Nigram *et al.*, 2019). Their genome organization, the functional proteins and the 5' and 3'´-untranslated regions (UTRs) are well studied (Siaw *et al.,* 1985; Lain *et al.*, 1988; Riechmann et al., 1990; Revers and Garcial, 2015). The genomes of potyviruses show considerable plasticity, a consequence of mutation and recombination (Revers *et al.,* 1996). This has given rise to diversity among the species that also reflect differences in their biological properties. Thus, the availability of genome sequences of field isolates of these viruses may be useful to further elucidate their evolutionary histories and molecular epidemiology.

Umbraviruses have been documented infecting legumes in several studies. The characteristic umbravirus genome is (+)ssRNA, polycistronic and comprises four overlapping ORFs that code for different functions, but they lack the capsid protein (CP) (Ryabov *et al.,* 2012; Abraham *et al.,* 2014). The short and long-distance proteins are overlapping and an intercistronic region, that varies in length among different umbraviruses, separates ORF2 and ORF3 (Taliansky *et al.,* 1996; Abraham *et al.,* 2014). Due to the absence of CP, insect vector transmission of umbraviruses requires the involvement of a helper virus, usually of the genus *Polerovirus* or *Luteovirus* (Murant, 1993; Taliansky *et al.,* 2000). However, after infection, umbraviruses are capable of autonomous replication *in planta*, although a satellite RNA is required in the development of disease symptoms (Murant *et al.,* 1988; Murant and Kumar, 1990; Abraham *et al.,* 2014). The natural host range of most umbraviruses is restricted to particular families. For example, ETBTV was reported to be restricted to solanaceous species and repeated attempts to infect members of the family *Fabaceae* with the virus was unsuccessful (Abraham *et al.,* 2014). However, a recent metagenomic analysis of field collected common bean samples revealed short contigs that matched to several umbraviruses including suspected ETBTV (Mwaipopo *et al.,* 2018), suggesting a possible expansion of the natural host range to include common bean. Thus, the status of common bean as a natural host of umbraviruses requires further investigation.

Another virus that is known to infect common bean is Southern bean mosaic virus (SBMV; genus *Sobemovirus*). The SBMV is restricted to common bean as its natural host (Tremaine and Hamilton, 1983). Like other members of the genus *Sobemovirus*, the SBMV genome is approximately 4.2 kb, (+)ssRNA and is encapsidated in isometric virions measuring 30 nm in diameter (Tremaine and Hamilton, 1983). The genome organization depicts four overlapping ORFs (ORF1, Replicase; ORF2a, RNA-dependent RNA polymerase (RdRP); ORF2b, movement protein [MP]; ORF3, CP). Marked differences in genome organization exist

among members of the genus *Sobemovirus* and sequence identities of the different genes vary considerably (Tamm and Truve, 2000). Thus, it is important to conduct further characterization of field isolates of SBMV from Zambia and anlyze them comparatively with sequences of global isolates of the virus.

Cucumber mosaic virus (CMV; genus *Cucumovirus*) is cosmopolitan, seedborne or vectored by over 75 aphids (Palukaitis *et al.*, 1992) and highly genetically diverse (Jacquemond, 2012; Kim *et al.,* 2014). In Africa, CMV occurs in various crops including common bean in single or mixed infection with other viral agents (Vetten and Allen, 1991; Mutuku *et al.,* 2018; Mwaipopo *et al.,* 2018). Its tripartite (+)RNA genome has been well elucidated (Roossinck, 2001). Based on serological profiles, amino acid sequences of the CP regions and hybridization characteristics, CMV is categorized into subgroups I and II (Roossinck *et al.*, 1999). Subgroup I is further divided into IA and IB based on the sequence variations of the 3' UTR of RNA3. Subgroups IA and II are globally distributed but IB is largely found in Asia (Roossinck, 2001). However, recent studies of CMV infecting common bean in Kenya detected subgroup IB isolates in Kirinyaga (Mutuku *et al.,* 2018). Considering that sequence isolates of CMV reported in chapter 3 were based on sequence fragments from RNA 2 segment, it was necessary to study all three components of the tripartite genome of CMV from Zambia in order to understand their relatedness with global isolates. Thus, the objective of this study was to conduct molecular characterization of representative viruses isolated from infected common bean plants from Zambia. The results of this study will be used to design molecular diagnostic tools that can be used in detecting viruses in seed distributed to farmers in Zambia and in monitoring possible emergency of molecular strains of the viruses.

4.3 Materials and Methods

4.3.1 Selection of isolates for virus characterization

In total, 640 symptomatic leaf samples of common bean plants were collected from farmers' fields across the six provinces of Zambia of which 223 representative samples were subsumed into nine samples as described in chapter 3. Metagenomics analysis of the samples and subsequent bioinformatics were performed as described in Chapter 3. Sequences of the different virus sequences obtained from HTS generated data files for individualcomposite samples (CP-1, EP-1, LP-1, MP-1, NP-1, NWP-1, Mse-3, Msa-3 and Com-1) (chapter three) were manually inspected for completeness of open reading frames (ORFs) relative to their closest reference global isolates. Sequences of respective viruses that did not have missing sequence fragments (gaps), possessed complete ORFs, complete or near complete 5´- and 3´-untranslated regions (UTR) were selected from each of the nine HTS data files. Thus, a total of 24 sequences representing various viruses (Table 4.1) were selected and subjected to various bioinformatics analyses (Section 3.3.4).

4.3.2 Characterization and phylogenetic analysis of HTS-generated viral sequences

Each of the 24 viral sequences were BLASTn queried against the NCBI to obtain corresponding global isolates of different viruses and separate data files were generated that included the HTS generated isolate (s) from this study and corresponding global virus species recovered from the NCBI. The nucleotide sequences in the data files were used to analyze the viruses for various genomic properties of the isolates from Zambia compared with global sequences. In analysis, pairwise comparisons were achieved using the Sequence Demarcation Tool v1.2 (SDTv1.2) (Muhire *et al.,* 2014). Further, similarity analysis was performed in SimPlot v3.5.1 program (Lole *et al.,* 1999). To determine putative recombination, alignment

Table 4.1. High throughput sequencing detected viruses and number of viruses with complete and near complete sequences for each composite sample

Province	Pooled sample	No. of raw reads	Virus aligned pair-end Reads	% of virus-specific reads	Virus detected	No. of near complete and complete genomes obtained from each composite sample
Eastern	EP-1	3,360,675.00	216,844.00	6.45	SBMV	1
		…..	…..	…..	BCMNV-NL3	1
		…..	…..	…..	BCMV	3
		…..	…..	…..	PvEV-1	1
		…..	…..	…..	CABMV	1
	Mse-1	4,186,537.00	188,160.00	4.49	SBMV	1
		…..	…..	…..	BCMNV	1
		…..	…..	…..	CMV	1
		…..	…..	…..	ETBTV	2
		…..	…..	…..	CABMV	1
Central	CP-1	4,061,565.00	143,632.00	3.54	SBMV	1
		…..	…..	…..	CABMV	1

		…..	…..	…..	PeMoV	1
	Com-1	960,007.00	1,132.00	0.12	PeMoV	1
Luapula	LP-1	3,055,744.00	337,250.00	11.04	SBMV	1
		…..	…..	…..	PvEV-2	1
Northern	NP-1	3,274,091.00	55,042.00	1.68	SBMV	1
	Msa-3	902,200.00	914.00	0.1	SBMV	1
Muchinga	MP-1	2,407,078.00	62,724.00	2.61	SBMV	1
		…..	…..	…..	BCMNV-TN1a	1
Northwestern	NWP-1	766,491.00	82,314.00	10.74	SBMV	1

files were imported into SplitsTree v4.5.1 to generate reticulate (non-tree like) phylogenetic networks. Hudson and Bryant (2006) defined reticulate network as a phylogenetic tree in which evolutionary histories are depicted by reticulate events such as hybridization, horizontal gene transfer or recombination. Thus, when the topology of a phylogenetic tree includes nodes with more than two parents, putative recombination events are inferred. Based on this definition, sequence datasets that resulted in the detection of reticulate phylogenetic networks in Splits Tree v4.5.1 program were imported into Recombination Detection Program version 4.99 (RDP4) (Martin et al., 2015) and subjected to recombination scanning with seven recombination detection methods. Recombination was considered authentic if recombination signal was detected by at least four recombination detection methods with a statistical *P*-value <0.05. The results from the two recombination detection methods were used to detect recombination in the difference genomic sequences for the different viruses obtained in this study. Where necessary, synonymous and non-synonymous substitutions were estimated using the Synonymous and Non-synonymous Analytical Program (SNAP) v2.1.1 (http://www.hiv.lanl.gov). The SNAP calculates synonymous and non-synonymous substitution rates based on a set of codon-aligned nucleotide sequences (Alicai *et al.*, 2016). Phylogenetic analyses were performed in MEGA7 for each sequence in order to infer evolutionary histories of the viruses included in this study.

4.4 Results

4.4.1 Sequence analyses and pairwise comparisons

In total 24 complete and near complete genome sequences of several viruses (BCMV = 3, BCMNV = 3, SBMV = 8, CMV = 1, CABMV = 3 PeMoV = 2, ETBTV = 2, PvEV-1 = 1, and PvEV-2 = 1) were recovered from the nine composite samples. The largest number of sequences were recovered from two samples (EP-1 and Mse-3) from Eastern Province (Table 4.1).

4.4.1.1 *Ethiopian tobacco bushy top virus*

The full lengths of the two ETBTV sequences from Zambia were determined to be 4,239 nt (CP414-1) and 4,238 nt (EP417) respectively, in contrast to ETBTV isolate 18-2 from Ethiopia (Accession no.KJ918748) which is 4,236 nt. At their 5´ end, the two genomes possess short untranslated regions (UTRs) that are 10 nt long beginning with 'GGG' whereas the 3´ termini end with 'CCC' similar to isolate 18-2. Both sequences have similar genome organisation with open reading frames starting and ending at similar positions. A 182 nt intergenic region spans nt positions 2,547 to 2,729 like isolate 18-2. The lengths of their 3´ UTRs differ slightly (CP414-1; 639 nt and EP417; 638 nt) and thus are 2 to 3-nt longer than that of isolate 18-2. The predicted ORFs of the ETBTV sequences from Zambia showed all the four cistrons; the RNA-dependent RNA-polymerase (RdRP), replication associated protein (RAP), long distance movement protein, and short distance movement protein that are arranged in a fashion characteristic of members of the genus *Umbravirus*. Consistent with all members of the genus *Umbravirus*, the RdRP, RAP, and ORF3 are derived from genomic RNAs (gRNAs) whereas ORF4 is derived from subgenomic RNA (sgRNA). The genomes lack the structural coat protein, a feature common to all umbraviruses and the reason they require a helper virus for vector transmission.

The lengths of the four ORFs are comparable to those of isolate 18-2 from Ethiopia. ORF1 is 876 nt and includes the canonical shifty heptanucleotide sequence AAAUUUU found just before the TAG stop codon. ORF2 encodes the RdRP and is translated via a -1 frameshift from ORF1, thus encoding a combined 844 aa ORF1/ORF2 protein. The overlapping gRNA and sgRNA are 258 and 262 aa and they are associated with long distance and cell-to-cell movement, respectively. Recently, ETBTV was isolated from infected *N. tabacum* in Malawi (Udagawa *et al*., 2020). The full length of the ETBTV isolate from Malawi (LC494673) is 4,186 nt in contrast to ETBTV isolates 18-2 (4236 nt), CP414-1 (4239 nt) and EP417 (4,238

nt), possesses 8-nt at its 5′ UTR beginning with ‘GG’ and 641-nt at the 3′ terminus ending with ‘CCC’. The ORFs of the ETBTV isolate Malawi are comparable in length to both ETBTV isolate 18-2 from Ethiopia, and the two sequences from Zambia. However, isolate Malawi possesses a rather shorter intergenic region (130-nt) spanning nt positions 2545 to 2675.

BlastN search of the sequence in the GenBank gave the highest score (96%) with ETBTV isolate Malawi (LC494673). The ETBTV sequences EP-1 (OK670701) and CP414-1 (MT225089) shared 99% nt identity between them and 80 to 97.5% nt identities with global sequences of other umbraviruses. Genome-wide pairwise comparison using the SDTv1.2 (Muhire *et al*., 2014) showed the complete genomes of ETBTV from Zambia shared 97.5% nt identity with isolate Malawi and 88% nt identity with ETBTV isolate 18-2 and they are distantly related to GRV isolate SRF54 (MG646923) at 73.6%. Gene-specific pairwise nt/aa comparisons of the different ORFs of the isolate from Zambia with ETBTV isolates Malawi, 18-2 and GRV isolates SRF54 and SRF57 showed variations across the four ORFs especially with the GRV isolates (Table 4.2). Nucleotide and aa percentage identities between the isolate from Zambia and isolates Malawi and isolate 18-2 ranged from 96 to 98.8% and 80 to 91%, respectively, and lower identities (21 to 78.5%) with GRV isolates SRF54 and SRF57. The highest (98 to 98.8%) nt/aa identities between the ETBTV sequences from Zambia and their closest homolog (LC494673) were recorded in ORF4 and the lowest (95 to 96.6%) was recorded in ORF1.

Phylogenetic analysis of 32 complete genomes of several members of the genus *Umbravirus* (2 from Zambia and 30 retrieved from GenBank) resulted in several clusters. The ETBTV sequences from Zambia tightly grouped with ETBTV isolate Malawi and loosely with ETBTV isolate 18-2 from Ethiopia. The GRV isolates SRF54 and SRF57 formed a distinct

group (Fig 4.1), further supporting that ETBTV from Zambia belongs to the same species as ETBTV isolates 18-2 and Malawi. These observations were further supported by SimPlot scanning of the entire genome (Fig 4.2).

Table 4.2. Comparisons of nucleotide (nt) and amino acid (aa) sequences of the four open reading frames (ORFs) of Ethiopian tobacco bushy top virus (ETBTV) sequence CP414-1 and sequences of groundnut rosette virus isolates SRF54 and SRF57 and global ETBTV isolates.

	RdRP (ORF2)	REP (ORF1)	MP (ORF3)	MP (ORF4)
Isolate	nt/aa	nt/aa	nt/aa	nt/aa
ETBTV-Malawi	97.3/97.8	96.6/95.2	98.6/96.2	98.3/98.8
ETBTV 18-2	86.9/90	85.4/84.2	89.9/89.1	91/80.5
SRF54	78.5/74.5	64.6/55.9	76.2/19.8	77.4/19.8
SRF57	78.4/74.5	66/55.8	76.3/19.4	77.6/21.6

An interesting feature of umbraviruses is their lack of a coat protein cistron, making them reliant on the capsid proteins of helper viruses for genome encapsidation and insect vector transmission (Murant *et al.*, 1993). Usually, the helper viruses for umbraviruses are from the family *Tombusviridae* or *Solemoviridae*. Surprisingly, no luteovirid sequence was detected in the Illumina reads obtained from composite samples Com-1 and EP-1. To further investigate the association of a helper virus with the common bean ETBTV isolates from Zambia, the individual samples in the composite sample Com-1 (n = 14) and EP-1 (n=30) and six additional samples from ETBTV-positive field samples were screened by RT-PCR for presence of a polerovirus or luteovirus using a pair of generic primers that was modified from Robertson *et al.* (1991) and Chomič *et al.* (2010). The expected 245 bp DNA fragment was successfully amplified from five (Com-1=1, EP-1=1 and three others) of the 50 samples and the two randomly selected amplicons were Sanger sequenced. The sequences (MT900845-46) obtained from both amplicons shared 100% nt identity with each other and 96% with

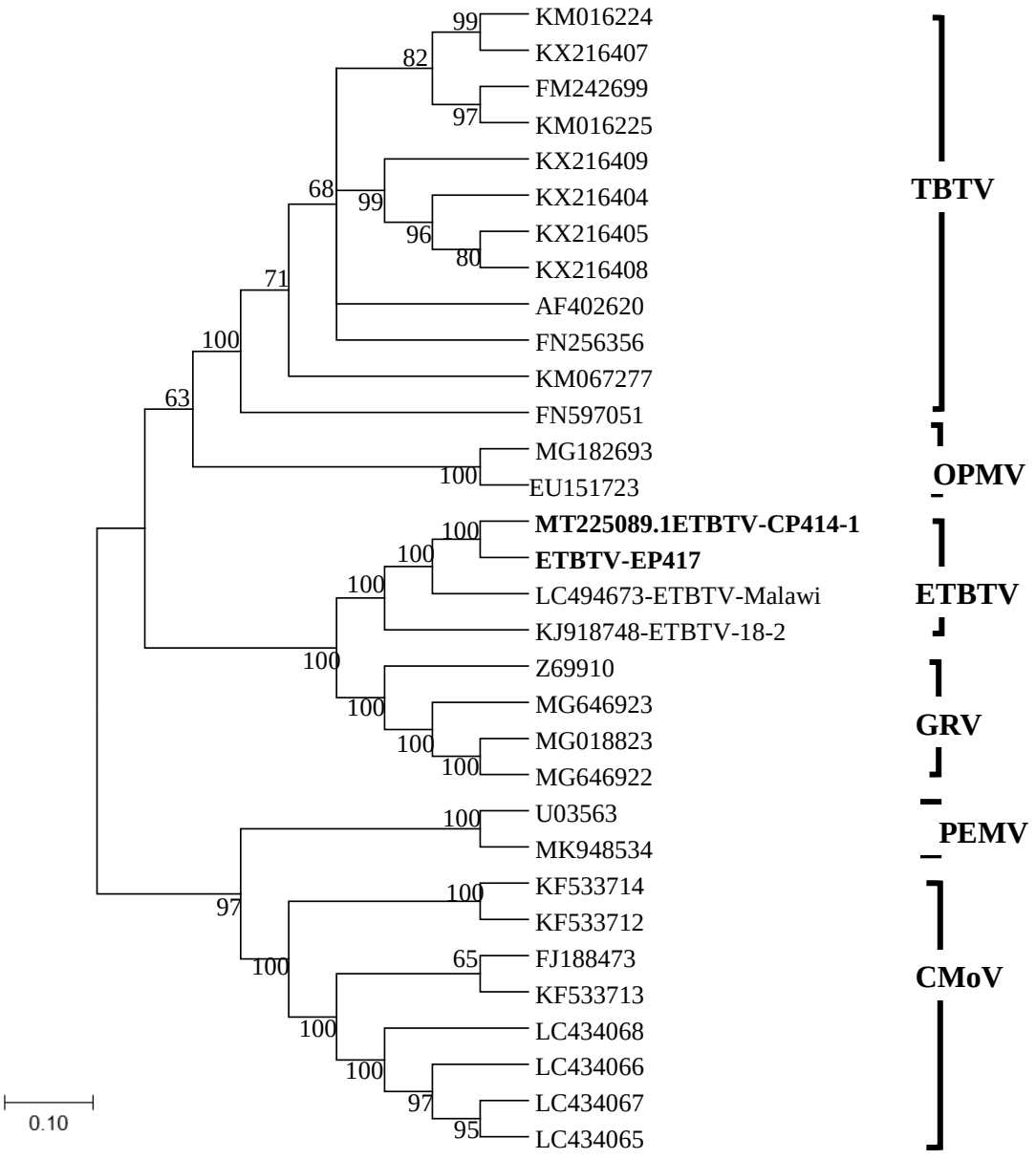


Figure 4.1. Maximun Likelihood phylogenetic analysis of 32 complete genome sequences of umbraviruses. ETBTV = *Ethiopian tobacco bushy top virus*, GRV = *Groundnut rosette virus*, TBTV = *Tobacco bushy top virus*, OPMV = *Opium poppy mosaic virus*, CMoV = *Carrot mottle virus*, PEMV-2 = *Pea enation mosaic virus* 2.

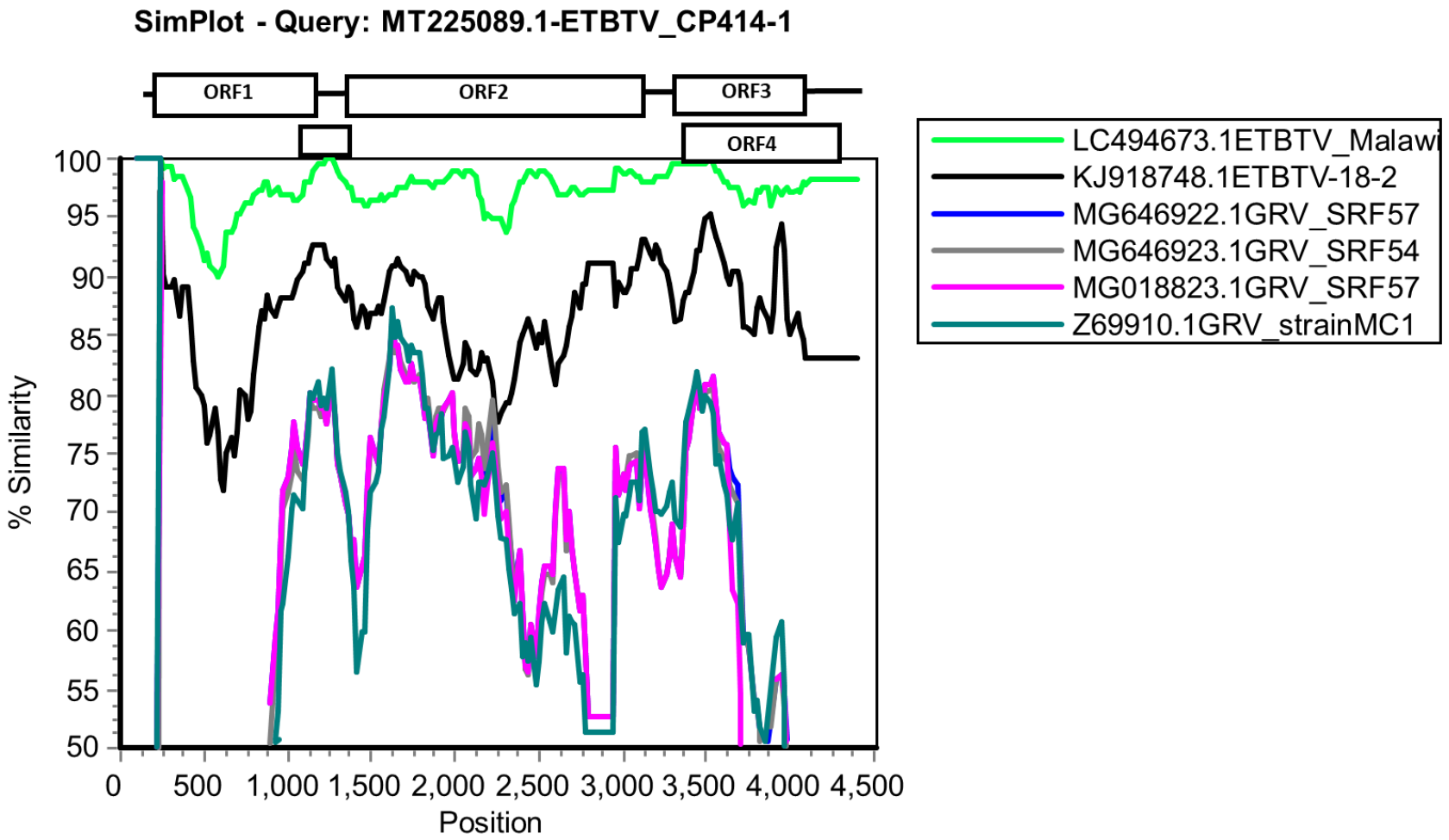


Figure 4.2. Similarity analysis of complete genome sequences of isolates of *Ethiopian tobacco bushy top virus* isolates (ETBTV: CP414-1, Malawi and 18-2) and *groundnut rosette virus* (GRV: SRF54 and SRF57) were performed using SimPlot.

cowpea polerovirus1 (KX599163). To further verify the identity of the putative helper virus, total nucleic acid for one of the samples was independently Illumina sequenced and a genome-length (5,902 nt) near complete sequence of isolate Com-54 (OK670702) was assembled. BLASTn querying of the recovered sequence and pairwise comparison with the exemplar virus CPPV1 (KY364846) produced similar results albeit with a lower (80.6%) amino acid identity. Since no other polerovirus, enamovirus and luteovirus sequence was detected in the ETBTV-positive samples, CPPV 1-Pv (Tentative name) was assumed the putative helper virus for the common bean-infecting ETBTV isolates from Zambia.

Recently, a 549 nt partial sequence of the RdRP gene of tobacco bushy top virus (TBTV) was recovered from tobacco plants with bushy top symptoms in Zimbabwe and deposited in GenBank. Classification of tombusvirids is reliably inferred from the structural relationships of their viral RdRPs (White, 2020). Thus, to understand its relatedness with ETBTV isolate from Zambia, the corresponding partial fragments were extracted from ETBTV sequences and used in pairwise comparison and phylogenetic analyses. In pairwise comparison, TBTV-A2 shared 94.6% and 88.1% with ETBTV isolates 18-2 (KJ918748) and CP414-1 (MT225089), respectively. Phylogenetic analysis of 18 partial sequences (2 from this study and 17 from GenBank) showed that TBTV-A2 (AJ704818) clustered in the same group as ETBTV-Malawi (LC494673), 18-2, CP414-1 and EP-1 (Fig 4.3).

4.4.1.2 *Ethiopian tobacco bushy top virus Satellite RNA (ETBTV-satRNA)*

Three sequences were recovered via HTS from composite samples Com-1 and Mse-3 collected from Central and Eastern provinces. The length of each sequence was determined to be 521 bp and their 5′- and 3′- ends shared identical ‘GGG’ and ‘CCC’ sequences with ETBTV isolates CP414-1 from Zambia and 18-2 from Ethiopia, respectively.

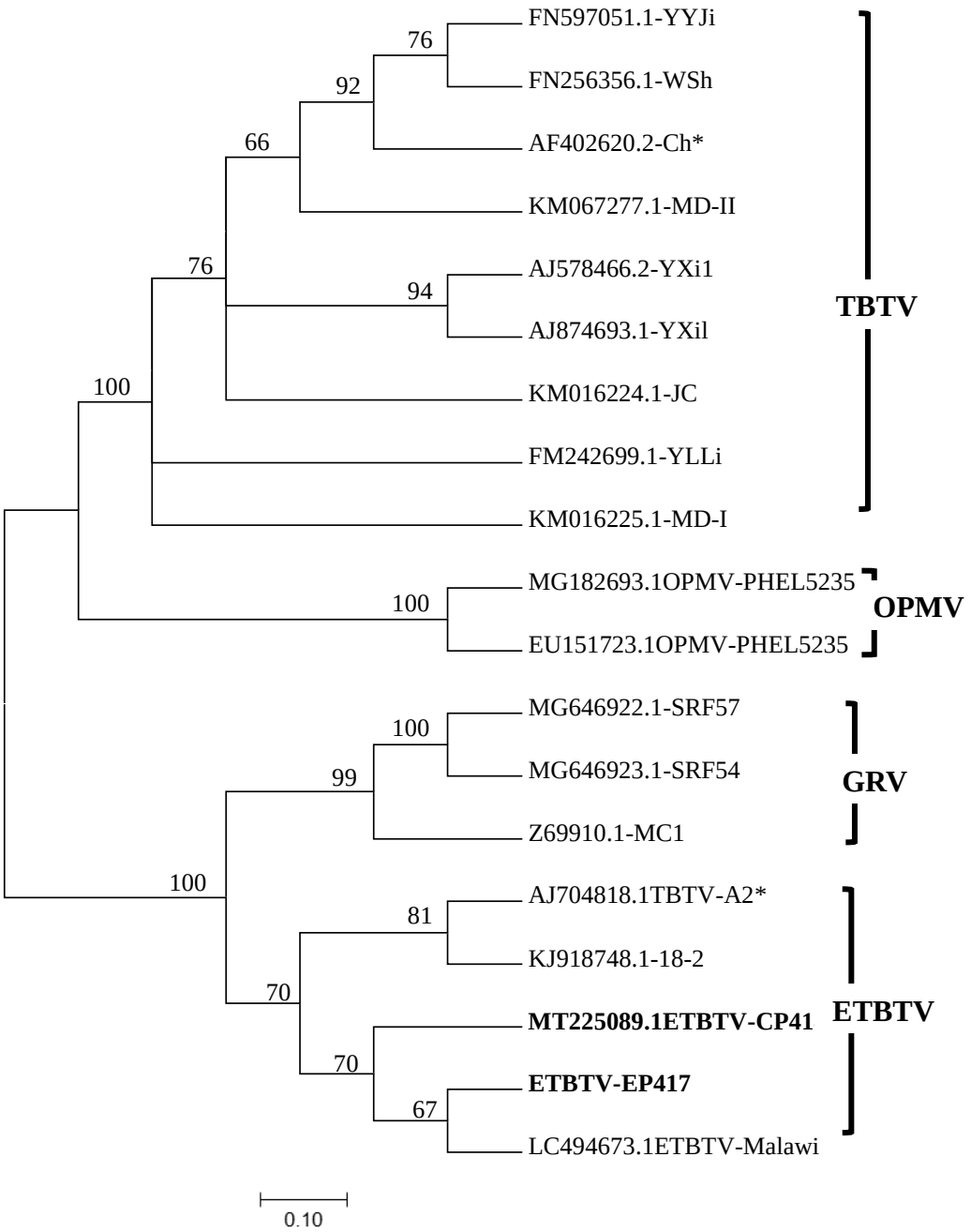


Figure 4.3. Phylogenetic analysis of 19 RNA dependent RNA polymerase partial sequences of umbraviruses. ETBTV = *Ethiopian tobacco bushy top virus*, GRV = *Groundnut rosette virus*, TBTV = *Tobacco bushy top virus*, OPMV = *Opium poppy mosaic virus*. * = The two viruses were originally thought to belong to the same virus species.

BLASTn query of the three RNA sequences (CP417-3: MT225092; CP418-4: MT225093; CP419-4: MT225094) revealed highest nt identity (95.4%) with ETBTV-satRNA isolate 18-2 (KJ918747), thus confirming the molecular nature of the sequences. Similar to satRNA isolate 18-2, no putative ORF was predicted in the ETBTV satRNA sequences derived in this study in line with the working definition of satellites (Walker *et al.,* 2021). In pairwise comparison, the three satRNA sequences shared 98 to 99.6% nt identity with each other and 93 to 95.4% nt identity with ETBTV-satRNA isolate 18-2. Much lower percent nucleotide identity was obtained when the sequences were compared with the GRV-associated satRNA sequences (Data not shown).

4.4.1.3 *Southern bean mosaic virus*

The eight assembled HTS-derived SBMV sequences obtained in this study ranged in length from 4,059 to 4,111 nt typical of members of the genus *Sobemovirus*. The lengths of their 5'- and 3'- UTRs ranged from 77 to 82 and 108 to 121 nt but they are presumably missing 11-14 nt and 9-21 nt sequences in their 5' and 3' UTRs, respectively, based on the the length of reference genomes retrieved from GenBank. The genomes are (+)ssRNA viruses, polycistronic with the expected typical sobemoviral four overlapping ORFs (MP, P2a, P2ab and CP). The nucleotide sequence lengths of each of the coding regions are equal in seven of the eight genomes but the MP and P2a genes for isolate Mse-3 (MN326873) are longer by three and 12 nucleotides, respectively. Additionally, the genome sequences of isolates MP-1, NP-1, LP-1, CP-1, Msa-3, EP-1 and NWP-1 have a 166 bp conserved region spanning nt 2644 to 2809 (Fig 4.3) that is also present and homologous in length to SBMV isolate TZ:SBMV pooled RNA reads_21-24HXH_15_CONTIG174:2017 (MG344643) from Tanzania (Fig 4.4). This 166 bp conserved region is absent in the genome sequence of isolates Mse-3 (this study), isolates SBMV B(ARK) (AF055887) and SBMV-S (AF055888) to which it is more closely related (Fig 4.4).

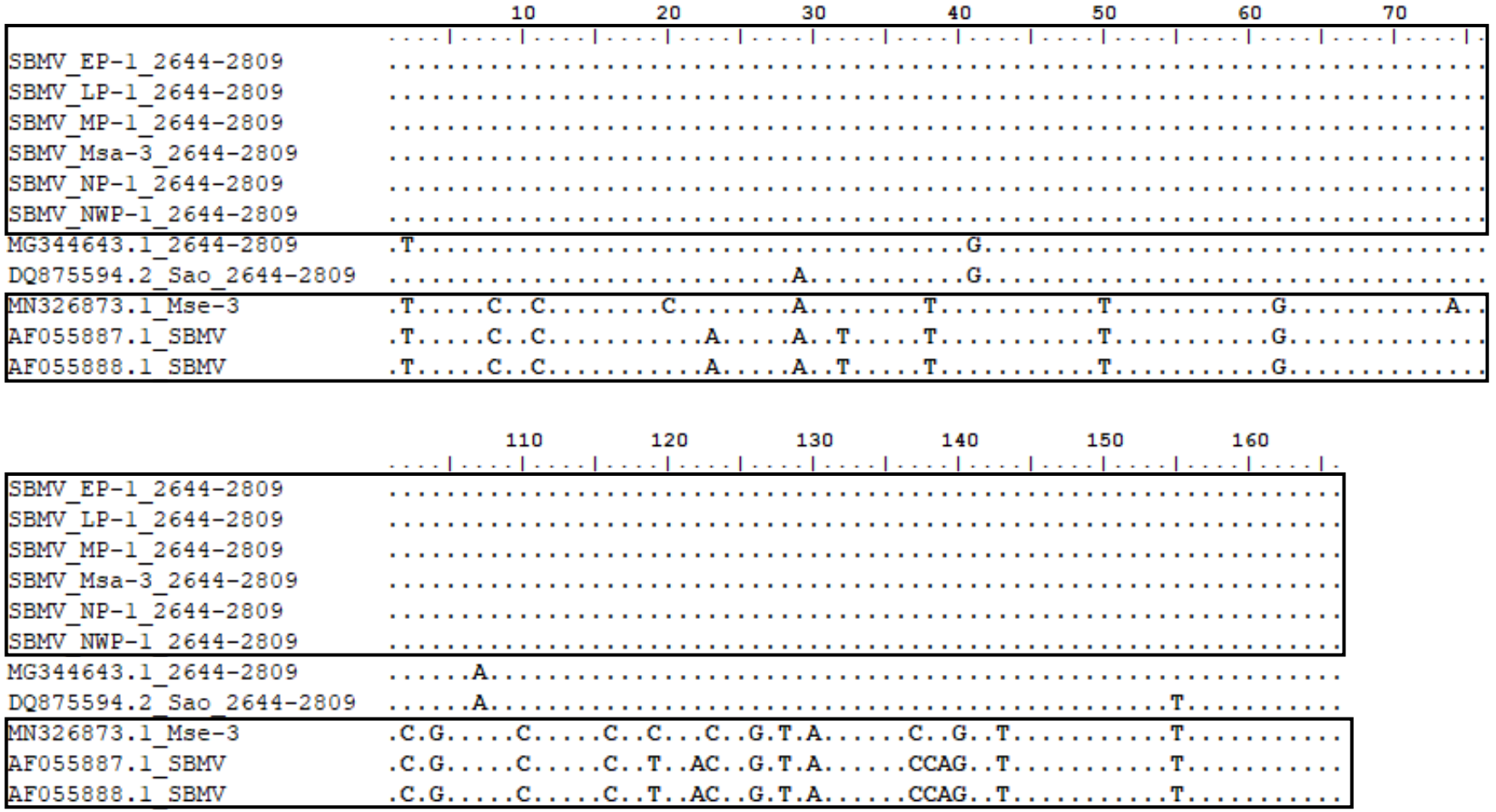
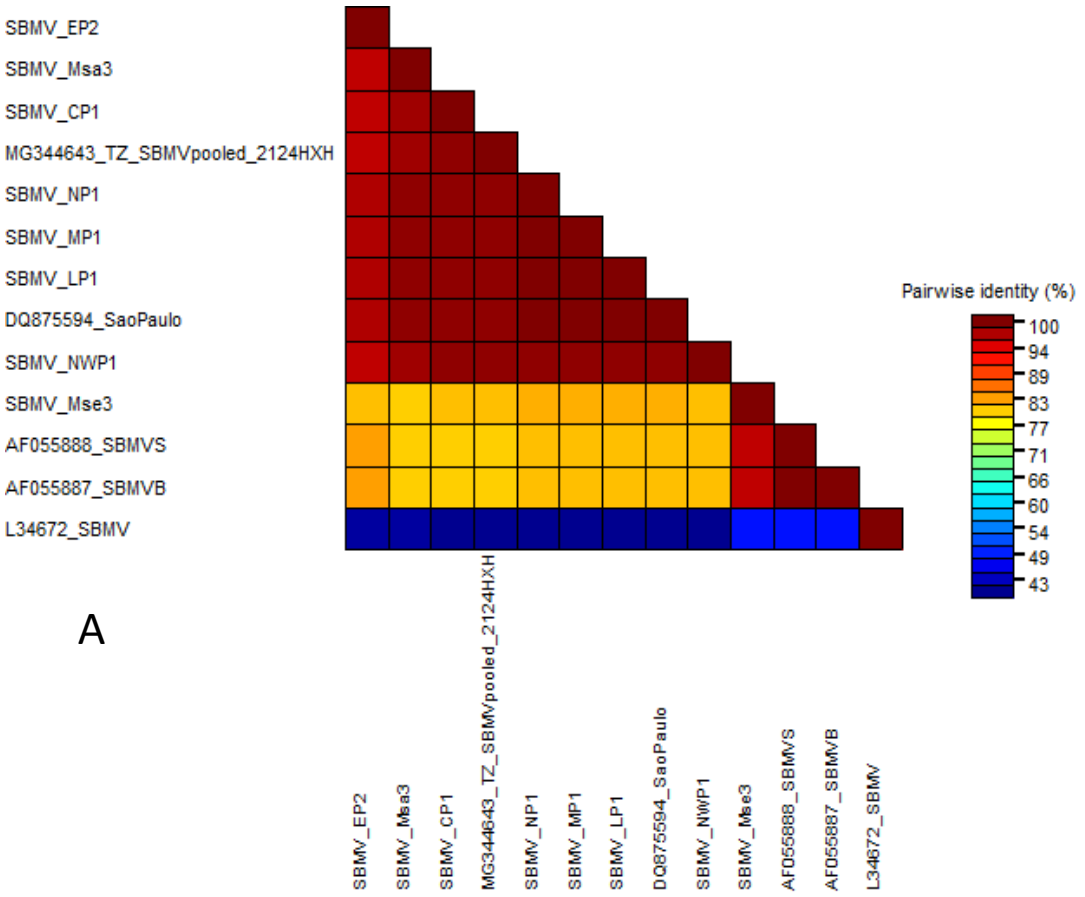


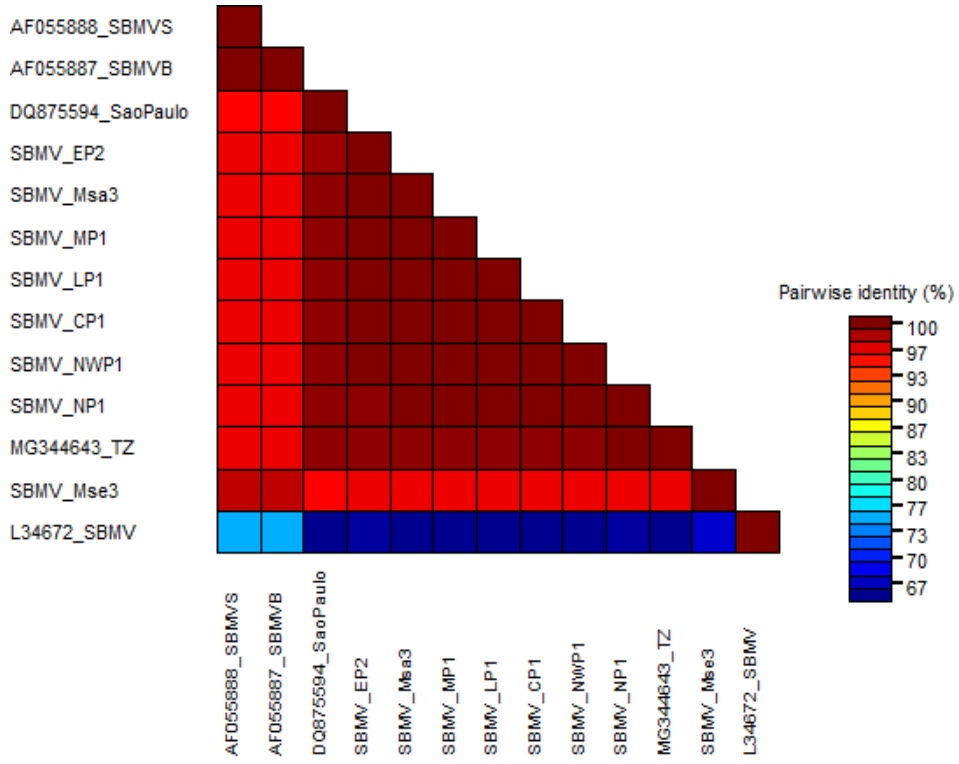
Figure 4.4. A 166 bp conserved region is common in the genomes of Southern bean mosaic virus (SBMV) isolates from Zambia, unlike in isolate TZ:SBMV pooled RNA reads_21-24HXH_15_CONTIG174:2017 from Tanzania and Sao Paulo from Brazil. The region starts at nucleotide 2,644 to 2,809 within the P2ab gene for isolates EP-1, LP-1, MP-1, Msa-3, NP-1 and NWP-1. This conserved sequence fragment is lacking in the genomes of SBMV isolate Mse-3 (this study), SBMV-B (AF055887), and SBMV-S (AF055888).

In pairwise comparison, the eight genome length SBMV sequences derived in this study shared 89 to 100% identity with each other and 85 to 99.4% with SBMV isolates retrieved from GenBank. Isolate Mse-3 (MN326873) was most identical (97.2%) to SBMV-B(ARK) (AF055887) but distantly related (89%) to TZ:SBMV (MG344643). The converse was true for the other seven isolates from Zambia in that they were more closely related (98 to 99.4%) to TZ:SBMV (MG344643) than to SBMV-B(ARK) with which they shared 98 to 99.4% identity. Further analysis revealed that isolates from this study shared 80-100% MP, 96-100% P2a 94-100% P2ab, and 93-100% CP aa sequence identities amongst themselves (Fig 4.5a-d). Specifically, isolate SBMV-Mse-3 was closely related (96-99.2% aa) in the MP, P2a and CP genes to isolates SBMV-B(ARK) and SBMV-S than to the other seven isolates from Zambia. However, Mse-3 was closely related (~95% aa) in the P2ab to the other seven isolates from Zambia than to SBMV B(ARK) and SBMV-S (78% aa). These results obtained from the SDT1.2 analysis mirror those obtained from the SimPlot scanning (Fig 4.6).

Considering that the MP (ORF1) and Px (ORFx) proteins are the most variable regions of the sobemovirus genomes (Sõmera *et al.,* 2015), both genes were compared across the isolates included in this study. Generally, MP gene is highly conserved for isolates Sao Paulo and Tz-21-22HXH and all the isolates from Zambia except Mse-3 (Fig 4.7A) further confirming the uniqueness of Mse-3 among the SBMVs in this study. Similarly, the PX (ORFx) is conserved for most of the isolates except for EP-1, Msa-3 and Mse-3 (Fig 4.7B). To further determine the observed relatedness, 23 sobemovirus sequences (8 from this study and 15 from GenBank) were included in phylogenetic analysis. The isolates from Zambia formed two distinct clusters, I and II. Cluster I included seven isolates (MP-1, NP-1, LP-1, CP-1, Msa-3, EP-1 and NWP-1) closely related to isolates Sao Paulo and Tz_21-21HXH from Brazil and Tanzania, respectively, whereas cluster II comprised Mse-3 closely related to isolates SBMV B(ARK) and SBMV S (Fig 4.8).


SBMV_EP2
SBMV_Msa3
SBMV_CP1
MG344643_TZ_SBMVpooled_2124HXH
SBMV_NP1
SBMV_MP1
SBMV_LP1
DQ875594_SaoPaulo
SBMV_NWP1
SBMV_Mse3
AF055888_SBMVS
AF055887_SBMVB
L34672_SBMV
Pairwise identity (%)
100
94
89
83
77
71
66
60
54
49
43

A


AF055888_SBMVS
AF055887_SBMVB
DQ875594_SaoPaulo
SBMV_EP2
SBMV_Msa3
SBMV_MP1
SBMV_LP1
SBMV_CP1
SBMV_NWP1
SBMV_NP1
MG344643_TZ
SBMV_Mse3
L34672_SBMV
Pairwise identity (%)
100
97
93
90
87
83
80
77
73
70
67

B

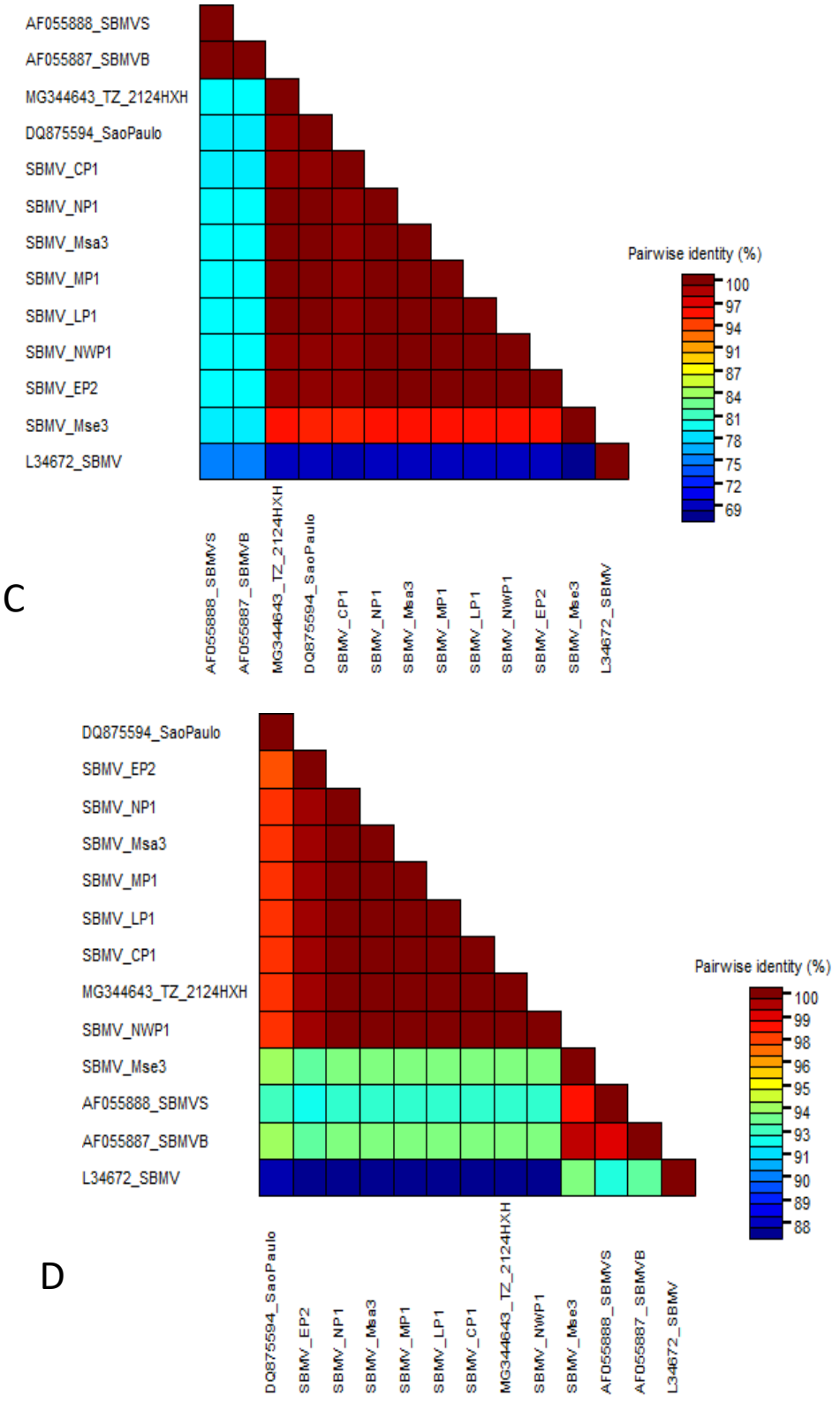


Figure 4.5. Colour coded pairwise identities of amino acid sequences of different genes of Southern bean mosaic virus (SBMV) sequences from this study and from GenBank. A = MP, B = P2a, C = P2ab, and D = CP.

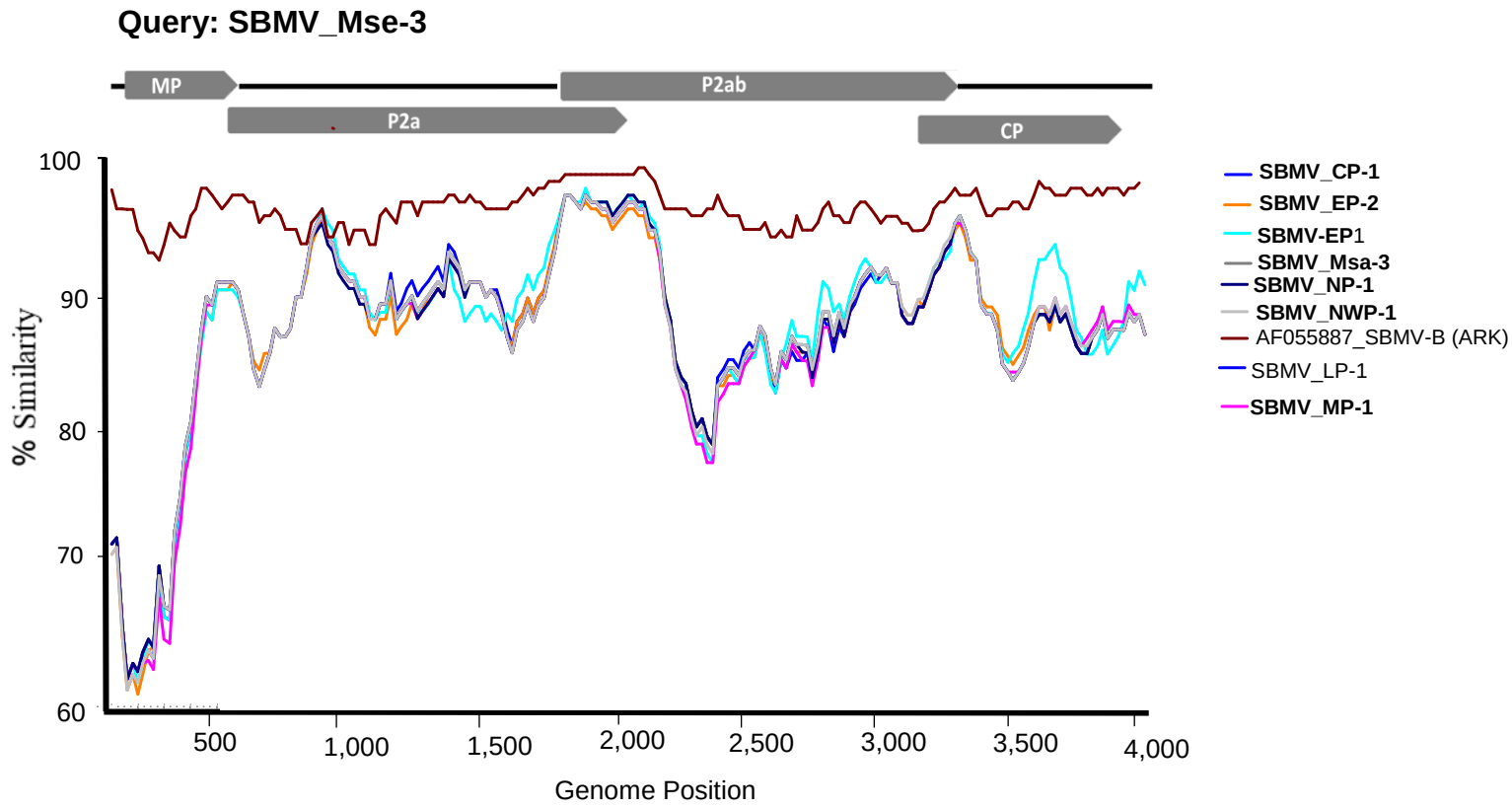


Figure 4.6. Similarity analysis of complete genome sequences of Southern bean mosaic virus (SBMV) isolates from Zambia were performed using SimPlot with Mse-3 as the query sequence. Sequence identity between Mse-3 and the rest of the sequences was lowest in the movement protein (MP) region except for SBMV-B (ARK). Sequences from Zambia are in bold letters.

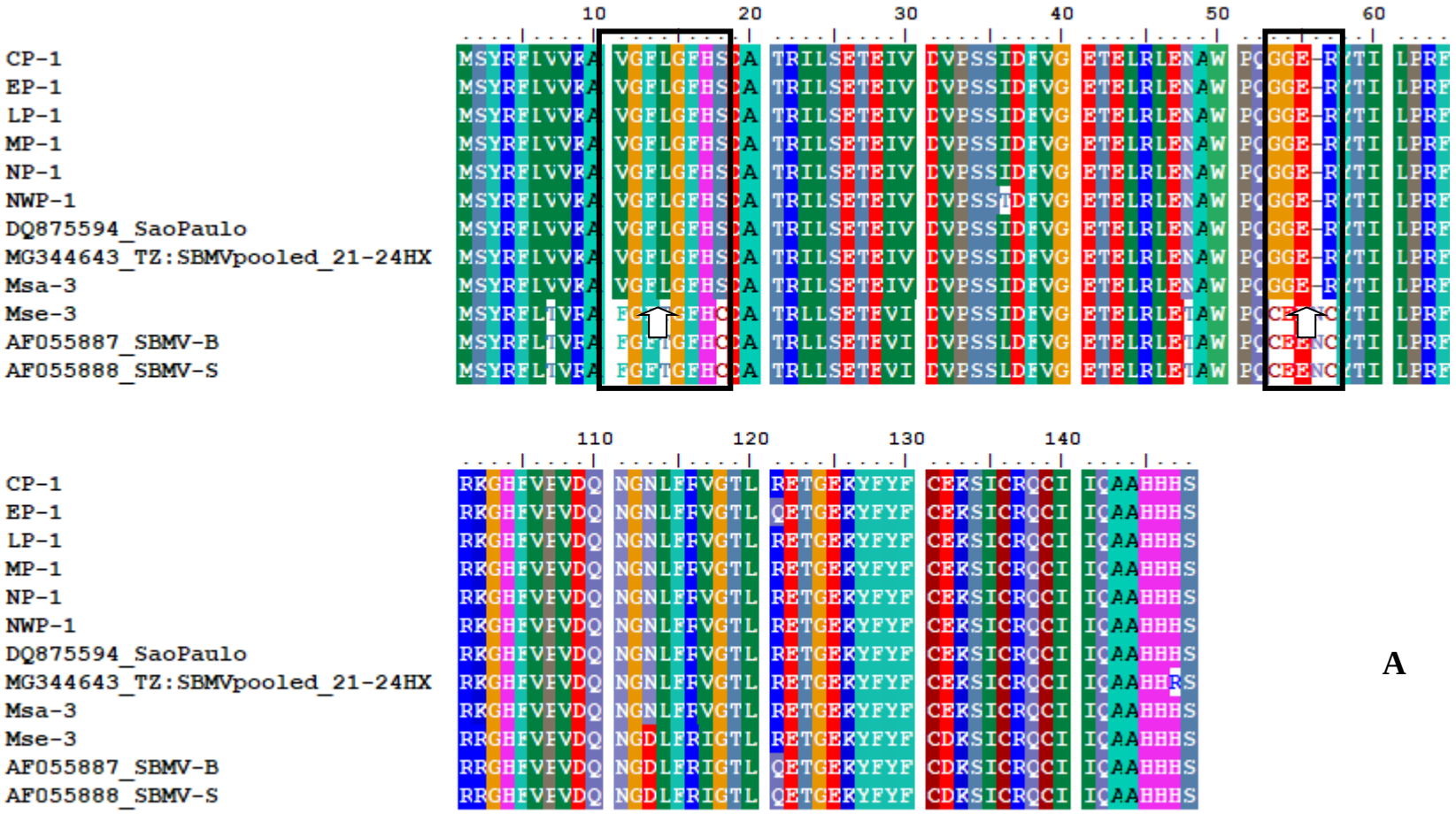


Figure 4.7A. Alignment of the amino acid (aa) sequences of the movement protein (MP) open reading frame 1 (ORF1) genes. Mse-3 is divergent from the isolates from Zambia in the N-terminus of the MP region similar to the SBMV-B(AR) and S homologs. The variable regions are marked on the alignment and shown with an arrow. Differences in sequences are boxed.

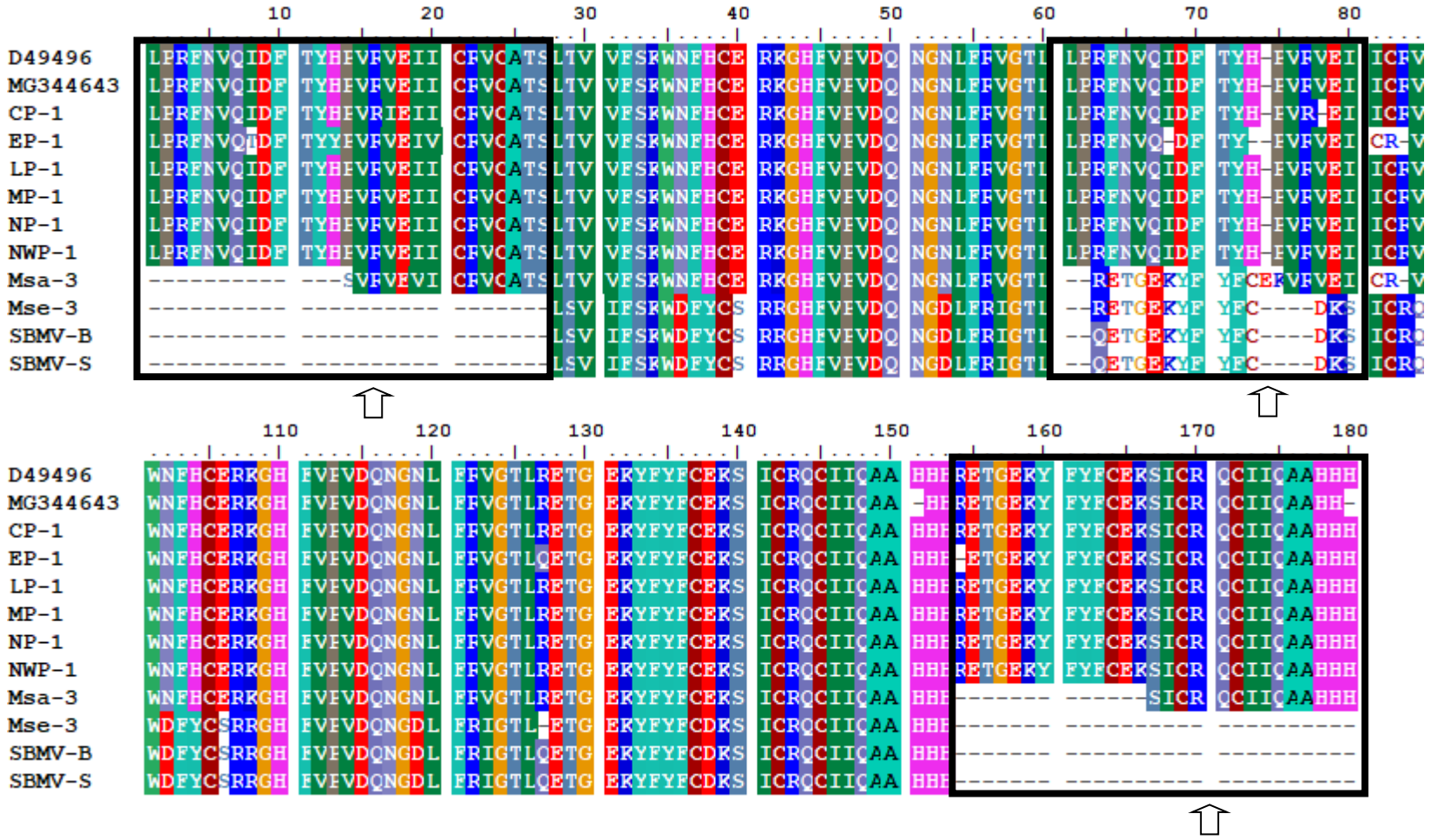


Figure 4.7B. Alignment of the amino acid (aa) sequences of the open reading frame x (ORFx). Mse-3 is divergent from the isolates from Zambia in the N-and C-termini of the Px region similar to the SBMV-B(AR) and S homologs. The N- and C- termini of the ORFx for isolate Mse-3 is shorter (shown by arrow) and are comparable with those of isolates SBMV-B(ARK) and S. The middle portion show deletions (shown by arrow).

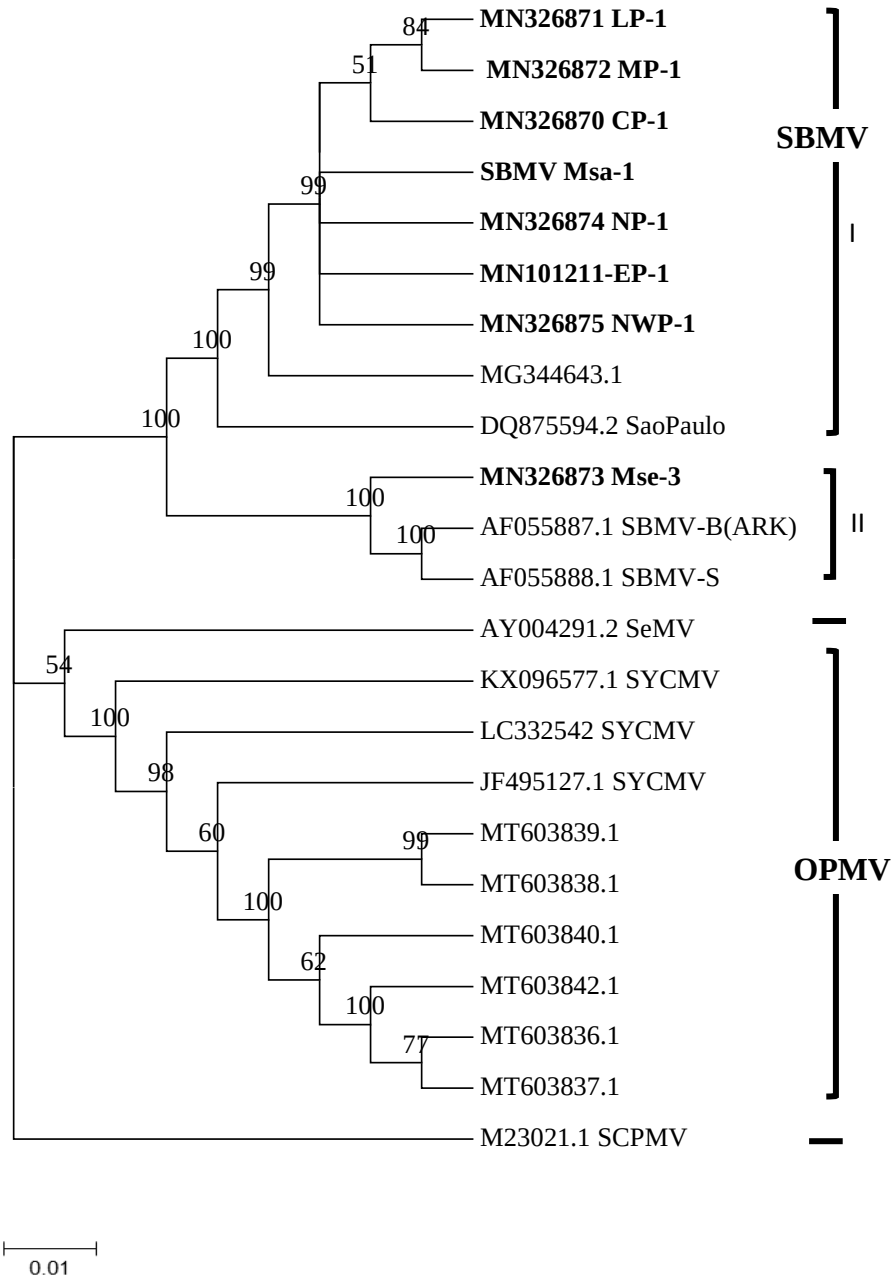


Figure 4.8 Unrooted phylogenetic tree involving 22 sobemoviral genome sequences. The sequences from Zambia (in bold letters) grouped with isolates from Tanzania and Brazil (cluster I) except for Mse-3 (cluster II) that grouped with SBMV-B(ARK) and S from Arkansas, USA. OPMV=Opium poppy mosaic virus; SYCMV=Soyabean yellow mosaic virus; SeMV=Sesbania mosaic virus; SBMV=Southern bean mosaic virus; SCPMV=Southern cowpea mosaic virus.

4.4.1.4 *Cucumber mosaic virus*

Three separate RNA contigs were recovered from the 188,160 HTS generated paired-end reads from composite sample Mse-3 in addition to other contigs that aligned with other viruses. BLASTn search of each of the three RNA sequences in GenBank resulted in hits to RNA1, RNA2 and RNA3 (MN326867-69) of global isolates of CMV thus indicating the viral nature of the CMV RNA segments obtained in this study. BLASTn analysis of the three RNA segments revealed that RNA1 and RNA2 shared (96 to ~97%) nucleotide identities with CMV isolates from Kirinyaga in Kenya (Fig 4.9A and B). However, RNA3 of Mse-3 was more closely (98%) related to strain M48 (D49496) (Fig. 4.9C). Thus, further analysis of the three RNA segments in this study were based on the genome characteristics of their closest homologs.

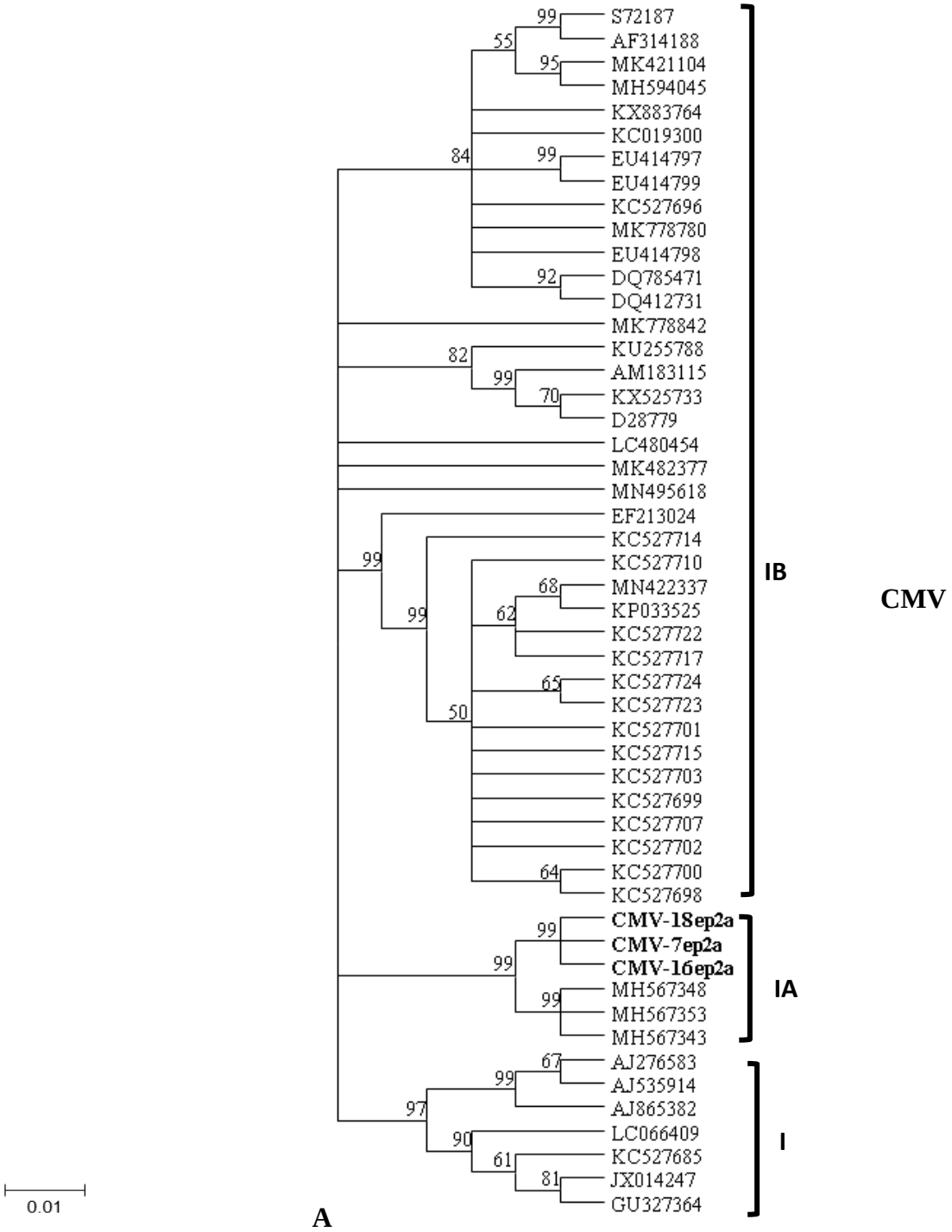
The lengths of the three genomic components of CMV were determined to be 3,343-nt for RNA1, 3,032-nt for RNA2 and 2,179-nt for RNA3. Each RNA segment is incomplete at their 5' and 3' UTRs relative to the sizes of reference sequences. Further analysis revealed that the RNA1 encodes a 2,979 nt replicase complex. The RNA2 encodes a 2,577 nt dicistronic region that consists of 2a (2,285-nt RdRP) and 2b (384 nt long distance MP) genes; the 2b partially overlaps the 2a at the 3´ terminus. The RNA3 encodes a 840 nt movement protein (MP) 3a and a 762 nt coat protein (CP) 3b. Compared with isolates D49496 and Kirinyaga1 (MH567343), the CP gene is 87 nt longer at its 5' terminus. The central intercistronic region is 208 nt long and it is substantially smaller than that of CMV isolates strain M48 (D49496) (296 nt) and Kirinyaga1 (298 nt), respectively.

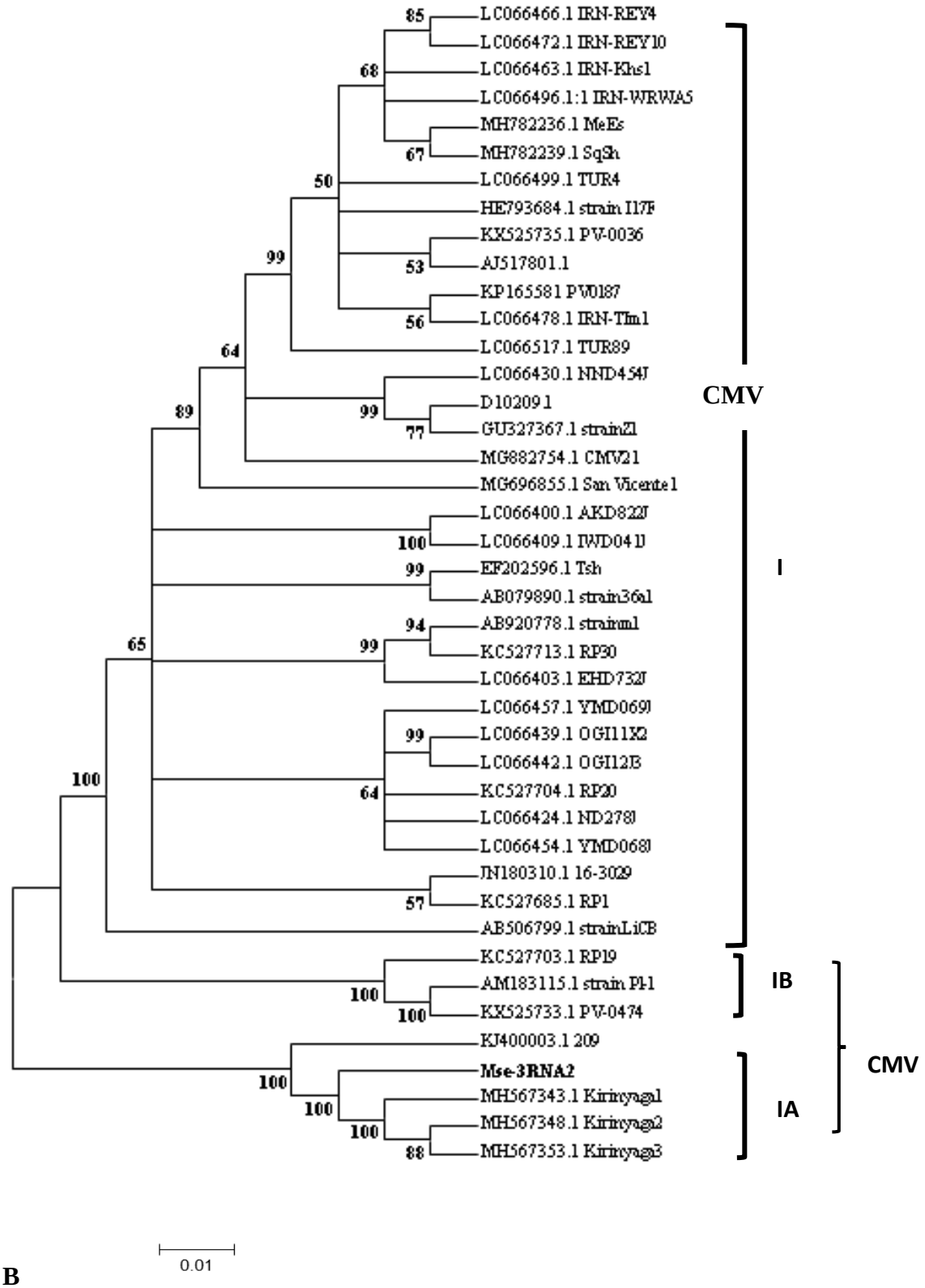
In pairwise comparison, RNA1 shared 96-96.9% nt and 97-98.4% aa identities with the three isolates from Kirinyaga, Kenya whereas its RNA2 segment shared 94 to 97.5% at both nt and aa levels with the same isolates. The RNA3 segment shared 99.3% nt identity with strain

M48 and 95.6% with isolates from Kirinyaga at nt level. Further examination of the 3a and 3b genes revealed that the 3a gene of RNA3 shared 99.8% nt identity to strain M48 and 97.1% nt to the 3a of Kirinyaga isolates. The 3b of isolate RNA3 shared 98.9% nt and 94% aa identities with strain M48 and isolates from Kirinyaga. In phylogenetic analysis, both RNA1 and RNA2 clustered with Kirinyaga isolates (Fig 4.9 A,B) but RNA3 clustered with strain M48 (Fig 4.9C).

4.4.1.5 *Bean common mosaic necrosis virus* and *Bean common mosaic virus*

Three full-length or near complete BCMV genomes (BCMV-EP20, BCMV-EP50 and BCMV-EP95) (MN987557-59) were also assembled from the HTS reads and they ranged in length from 9,937 to 9,958 nt. Each of these sequences also encoded the complete polyprotein sequence of the virus. In pairwise comparison, the three BCMV sequences shared 92 to 95.8% nt and 95 to 97% aa complete polyprotein sequence identities with each other and 95 to 99.3% nt/aa identities with global isolates of the virus. Phylogenetic analyses revealed the clustering of all the three BCMV polyprotein sequences from Zambia into the previously defined R1 clade (Moradi and Mehrvar 2019), along with several recombinant isolates of the virus (Fig. 4.10A). The results showed the presence of genetically diverse sequence variants of BCMV in Zambian common bean fields. Further analysis of the clustering revealed that within the BCMV clade, BCMV-EP-20 tightly clustered with BCMV isolate 1755a (KT175570) a *bc-3* breaking isolate of BCMV and BCMNV in common bean (Feng *et al.,* 2015), BCMV-EP95 with the isolate NL7n (KY057338) and BCMV-EP50 loosely clustered with isolate NY15p (KT175568) (Fig 4.10A). All the three BCMV sequences obtained in this study clustered with PGVI isolates of the virus (Fig 4.10A) but BCMV-EP20 clustered more tightly with the necrotic strain NL-3 which is assigned to PGVIa, a subgroup of PGVI (Fig 4.10A). Three complete or near complete BCMNV genomes (BCMNV-Lsk-1A-F, BCMNV-MP and BCMNV-EP) (MN987554-56) were


99
S72187
55
AF314188
95
MK421104
MH594045
KX883764
KC019300
84
99
EU414797
EU414799
KC527696
MK778780
EU414798
92
DQ785471
DQ412731
MK778842
82
KU255788
99
AM183115
70
KX525733
D28779
LC480454
MK482377
MN495618
EF213024
KC527714
99
KC527710
68
MN422337
99
62
KP033525
KC527722
KC527717
65
KC527724
50
KC527723
KC527701
KC527715
KC527703
KC527699
KC527707
KC527702
64
KC527700
KC527698
IB
CMV
99
CMV-18ep2a
CMV-7ep2a
99
CMV-16ep2a
MH567348
99
MH567353
MH567343
IA
67
AJ276583
99
AJ535914
97
AJ865382
LC066409
90
61
KC527685
81
JX014247
GU327364
I
0.01
A


LC066466.1 IRN-REY4
LC066472.1 IRN-REY10
LC066463.1 IRN-Khs1
LC066496.1:1 IRN-WRWA5
MH782236.1 MeEs
MH782239.1 SqSh
LC066499.1 TUR4
HE793684.1 strain I17F
KX525735.1 PV-0036
AJ517801.1
KP165581 PV0187
LC066478.1 IRN-Tlm1
LC066517.1 TUR89
LC066430.1 NND454J
D10209.1
GU327367.1 strainZl
MG882754.1 CMV21
MG696855.1 San Vicente1
LC066400.1 AKD822J
LC066409.1 IWD041J
EF202596.1 Tsh
AB079890.1 strain36al
AB920778.1 strainml
KC527713.1 RP30
LC066403.1 EHD732J
LC066457.1 YMD069J
LC066439.1 OGI11X2
LC066442.1 OGI12B
KC527704.1 RP20
LC066424.1 ND278J
LC066454.1 YMD068J
JN180310.1 16-3029
KC527685.1 RP1
AB506799.1 strainLiCB
KC527703.1 RP19
AM183115.1 strain Pl-1
KX525733.1 PV-0474
KJ400003.1 209
Mse-3RNA2
MH567343.1 Kirinyaga1
MH567348.1 Kirinyaga2
MH567353.1 Kirinyaga3
CMV
I
IB
IA
CMV
0.01

B

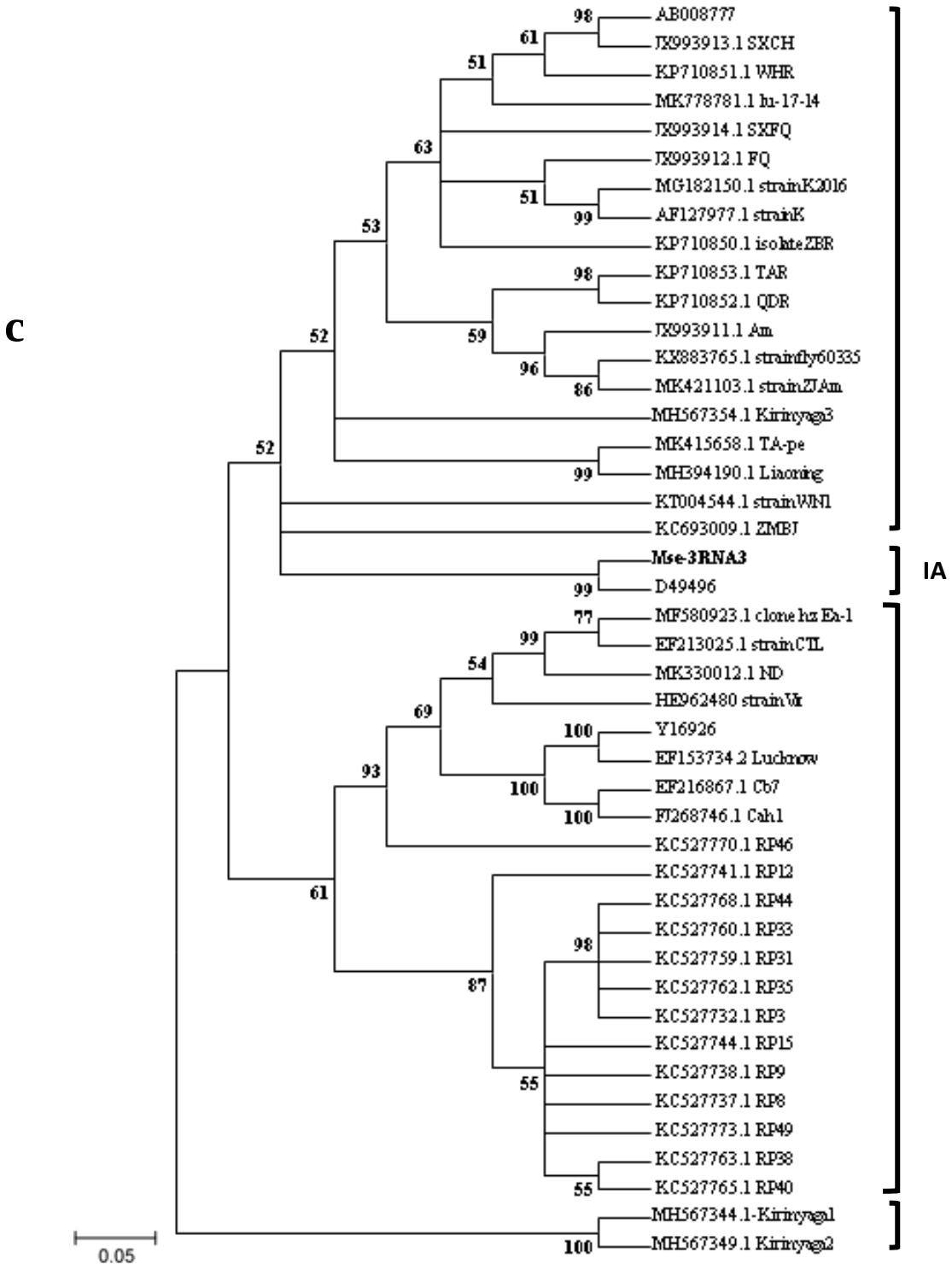
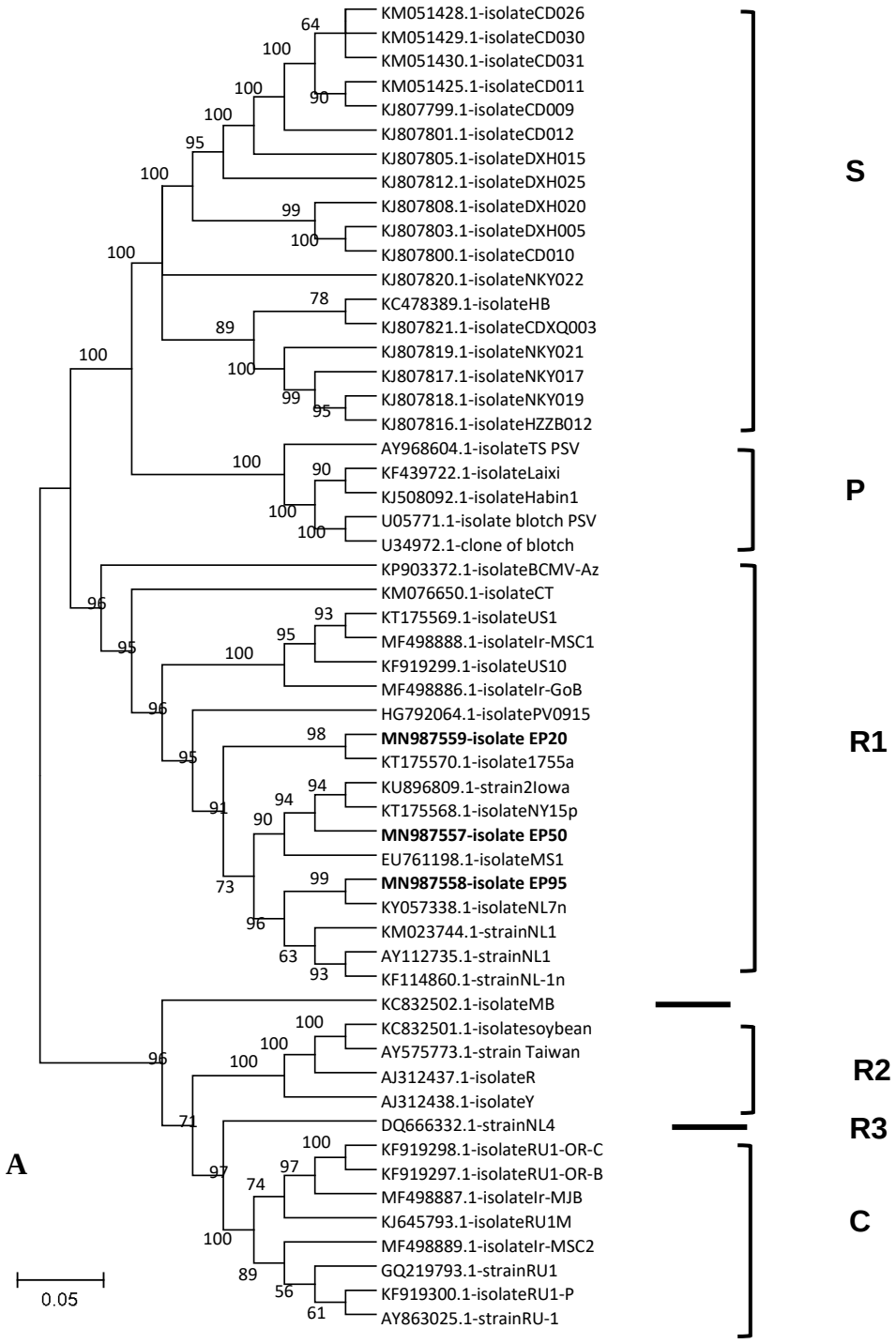
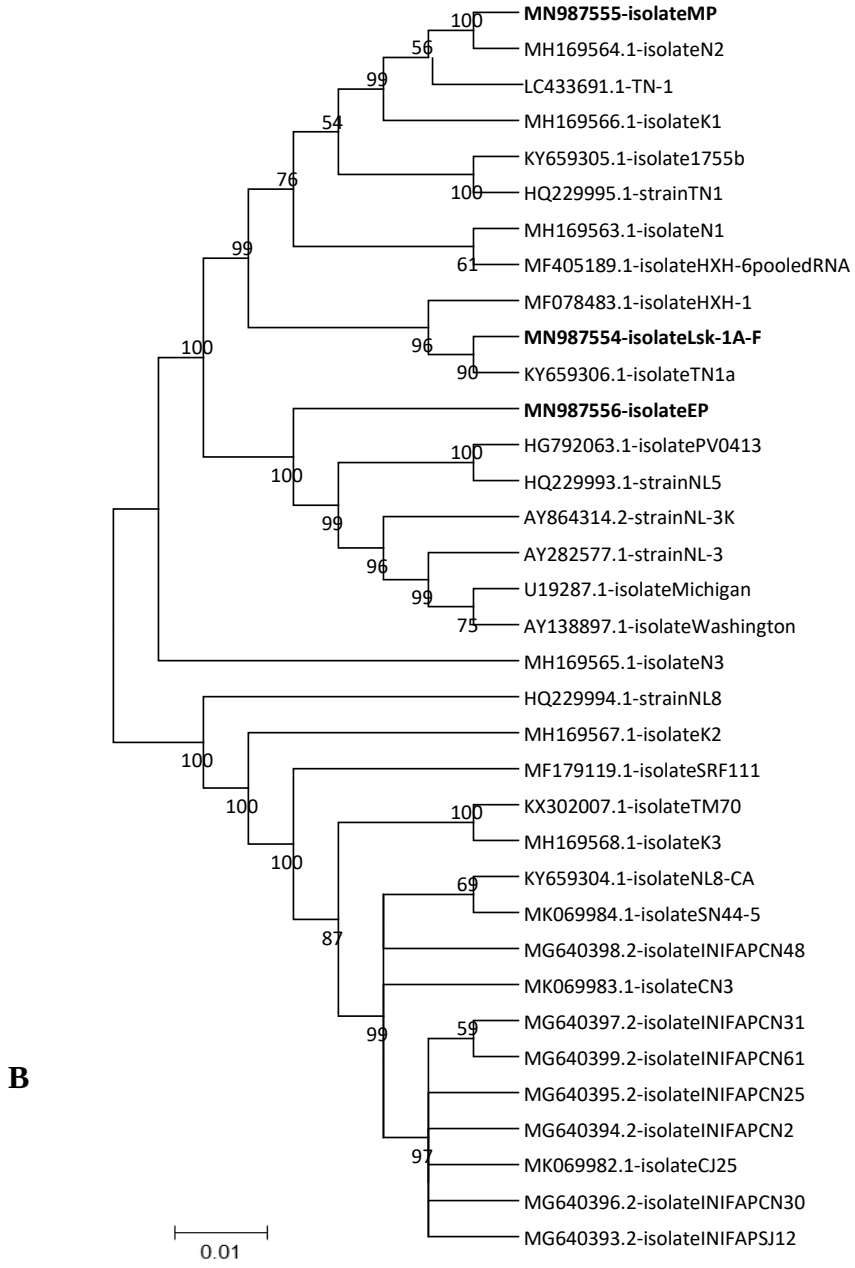


Figure 4.9. Phylogenetic trees of cucumber mosaic virus (CMV). **A.** Phylogenetic analysis of partial sequences of RNA2 segment of CMV obtained from three samples in Zambia along with corresponding virus sequence fragments of global CMV sequences obtained from GenBank. All RNA2 fragments in this study clustered with IA subgroup of CMV isolates both in A (Sanger derived partial sequences) and B (near complete sequences of RNA2). In **C**, near complete sequences of RNA3 clustered with RNA3 sequence of isolate D49496 which is subgroup IB. Isolates from Zambia are in bold letters.

assembled from the HTS reads and they ranged in length from 9,047 to 9,628 nt. Each of these sequences encoded the complete polyprotein sequence of the virus. In pairwise comparison, the three BCMNV sequences shared 98 to 98.9% nt/aa complete polyprotein sequence identities with each other and 98 to 99.6% nt/aa complete polyprotein identities with several BCMNV isolates in GenBank. BCMNV-Lsk-1A-F was most identical to isolate TNA1a (KY659306) with which it shared 99.3% nt and 99.6% aa complete polyprotein identity. BCMNV-MP shared maximum (99.3% nt and 98.8% aa) complete polyprotein sequence identity with isolate N2 (MH169564), and BCMNV-EP was most identical (98.9% nt and 98.9% aa) to strain NL-3 (U19287). All the three BCMNV sequences from Zambia segregated into the previously defined BCMNV phylogroup I (Fig 4.10B) as per Wainaina *et al.* (2019). Two subgroups can be discerned within this clade and whereas BCMNV-Lsk-1A-F and BCMNV-MP both clustered into one subgroup, BCMNV-EP segregated into a different subgroup (Fig 4.10B). The results showed the occurrence of sequence variants of BCMNV across common bean fields in three separate provinces of Zambia.

Since recombination has been implicated in the evolutionary histories of several BCMNV and BCMV isolates, the aligned polyprotein sequence files of the two viruses were analysed with the Neighbor-Net method of the SplitsTree program v4.15.1 (Hudson and Bryant, 2006). Based on the definition given in section 4.3.2, the generated reticulate trees involving BCMV-EP-20, BCMV-EP-50 and BCMV-EP-95 indicate that their evolutionary histories have been shaped by recombination (Fig 4.11). Further, the same alignment files were imported into RDP4 v4.99 (Martin *et al.*, 2015) and scanned for recombination with default settings. The results of both analyses confirmed that BCMV-EP-20, BCMV-EP-50 and BCMV-EP-95 are potential recombinant sequences with putative major and minor parental sequences (Table 4.2). No evidence of recombination was detected for BCMNV-Lsk-1A-F, BCMNV-MP and BCMNV-EP.


KM051428.1-isolateCD026
KM051429.1-isolateCD030
KM051430.1-isolateCD031
KM051425.1-isolateCD011
KJ807799.1-isolateCD009
KJ807801.1-isolateCD012
KJ807805.1-isolateDXH015
KJ807812.1-isolateDXH025
KJ807808.1-isolateDXH020
KJ807803.1-isolateDXH005
KJ807800.1-isolateCD010
KJ807820.1-isolateNKY022
KC478389.1-isolateHB
KJ807821.1-isolateCDXQ003
KJ807819.1-isolateNKY021
KJ807817.1-isolateNKY017
KJ807818.1-isolateNKY019
KJ807816.1-isolateHZZB012
AY968604.1-isolateTS PSV
KF439722.1-isolateLaixi
KJ508092.1-isolateHabin1
U05771.1-isolate blotch PSV
U34972.1-clone of blotch
KP903372.1-isolateBCMV-Az
KM076650.1-isolateCT
KT175569.1-isolateUS1
MF498888.1-isolateIr-MSC1
KF919299.1-isolateUS10
MF498886.1-isolateIr-GoB
HG792064.1-isolatePV0915
MN987559-isolate EP20
KT175570.1-isolate1755a
KU896809.1-strain2Iowa
KT175568.1-isolateNY15p
MN987557-isolate EP50
EU761198.1-isolateMS1
MN987558-isolate EP95
KY057338.1-isolateNL7n
KM023744.1-strainNL1
AY112735.1-strainNL1
KF114860.1-strainNL-1n
KC832502.1-isolateMB
KC832501.1-isolatesoybean
AY575773.1-strain Taiwan
AJ312437.1-isolateR
AJ312438.1-isolateY
DQ666332.1-strainNL4
KF919298.1-isolateRU1-OR-C
KF919297.1-isolateRU1-OR-B
MF498887.1-isolateIr-MJB
KJ645793.1-isolateRU1M
MF498889.1-isolateIr-MSC2
GQ219793.1-strainRU1
KF919300.1-isolateRU1-P
AY863025.1-strainRU-1
S
P
R1
R2
R3
C
A
0.05



B

Figure 4.10. Phylogenetic analysis of A. bean common mosaic necrosis virus (BCMV) and B. bean common mosaic virus (BCMNV) conducted in MEGA7. The evolutionary history was inferred by using the Maximum Likelihood method based on the General Time Reversible (GTR) model. Clustering patterns reported in Moradi and Mehrvar (2019) and Wainaina *et al.* (2019).

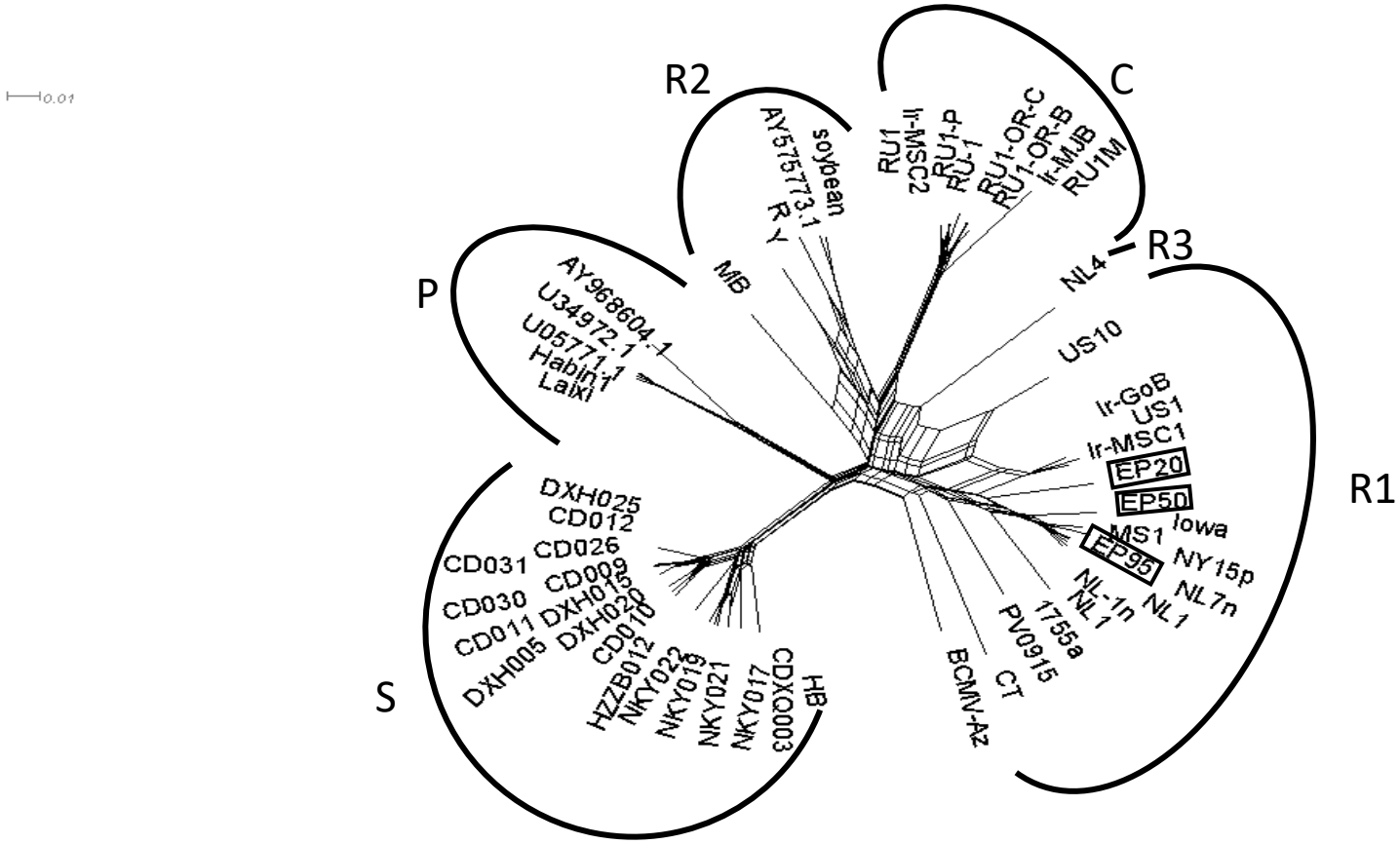


Figure 4.11 Reticulate phylogenetic networks (non-tree-like) generated by scanning a polyprotein alignment file of selected global sequences of bean common mosaic virus and isolate EP-20. The detection was implemented in SplitsTree v4.15.1. Detection of the reticulate networks indicates that the evolutionary histories of BCMVEP_20, BCMV-EP_50 and BCMV-EP-95 may have been shaped by recombination.

Table 4.3. Recombination signals across bean common mosaic necrosis virus (BCMV) performed in RDP4. Recombination was authentic if supported by minimum five detection methods at 0.05 significance level.

Virus	Accession no.	Breakpoint	Putative parental (% Similarity) Major	Minor	Methods	P-Val range
BCMV	MN987557	720-1002	Unknown	EP-95 (97.7)	R G B C S P	3.907×10^{-6} -1.710×10^{-2}
BCMV	MN987558	8385-9010	AY114860	MG640407	R G B M C S T	4.504×10^{-9} – 3.884×10^{-3}
BCMV	MN987559	2542-3011	MH220847	MH220847	R G B M C S T	3.405×10^{-6} – 1.311×10^{-2}
BCMV	MH024840	2504-5787	KY057338 (89.0)	KP903372 (89.3)	R G B M C S	3.534×10^{-41} - 7.342×10^{-6}
BCMV	MH024838	2504-5787	KY057338 (88.9)	KP903372 (88.5)	R G B M C S	3.534×10^{-41} - 7.342×10^{-6}
BCMV	MK069985	2504-5787	KY057338 (92.0)	KP903372 (88.5)	R G B M C S	3.534×10^{-41} - 7.342×10^{-6}
BCMV	MG640400	2505-5788	KY057338 (91.1)	KP903372 (88.9)	R G B M C S	3.534×10^{-41} - 7.342×10^{-6}
BCMV	MG640405	2506-5789	KY057338 (91.1)	KP903372 (88.8)	R G B M C S	3.534×10^{-41} - 7.342×10^{-6}
BCMV	KY057338	2454-4387	MF405191 (98.9)	MH024841 (99.4)	R M C P	4.833×10^{-5} - 2.009×10^{-4}

4.4.1.6 *Peanut mottle virus*

The genome sequences of two PeMoV sequences (CP-com1 and Pemov-20) were determined to be 9,643 nt and 9,693 nt long. Both sequences encode a single polyprotein from nt positions 108 to 9,410 and 123 to 9, 422, respectively. The 107 nt 5´-UTR and 233 nt 3´ UTR of sequence CP-com-1 is presumably missing 15 and 54 nucleotides at both ends, respectively, based on comparison with the PeMoV reference genome isolate Liaoning (MH270528). The putative protease cleavage sites (DKIHQY/S, KHYVVG/G, CLVAYQ/A, ETVRYQ/S, DTVQYQ/S, EPVRYQ/G, DVVATE/G, and DEVRYQ/S) (Adams *et al.,* 2005; Gutiérrez *et al.,* 2016) for the large non-functional polyprotein (9300nt/3100aa) were identified in both CP-com1 and Pemov-20 sequences. Proteolytic cleavage of the polyprotein resulted in the typical 10 potyviral genes that were identified based on comparisons with the corresponding proteins of isolates Liaoning (MH270528) and Pinto (KU708532), respectively.

In pairwise comparison, the complete polyprotein of CP-com1 shared 88 to 89% nt and 93% aa identities with global sequences, respectively (Table 4.3). However, gene to gene comparisons of CP-com1 with global sequences revealed wider identity variations ranging from 81 to 93% nt and 75 to 100% aa levels of identities (Table 4.4; Fig 4.12). The highest percentage nt/aa identities between CP-com1 and other isolates are in the 6K1, CI, NIb and CP cistrons while the lowest were in the P1, 6K2 and NIa-VPg (Fig 4.12, Table 4.4).

The 122 nt 5'-UTR and 233 nt 3' UTR of sequence Pemov-20 is presumably missing 9 and 15 nucleotides at both ends, respectively, based on comparison with the PeMoV reference genome isolate INIFAP SJ8-5 (MG640414). The size of the complete polyprotein of sequence Pemov-20 is 9,297/3099 nt/aa and shares the same length with its PeMoV reference

isolate. In pairwise comparison, the complete polyprotein of Pemov-20 shared 96 to 100% nt and 98 to 99% aa identities with PeMoV global isolates.

In phylogenetic analysis using complete polyprotein nucleotide sequences, sequence CP-Com1 grouped separately from the rest of the PeMoV isolates including PeMoV sequence Pemov-20 (Fig 4.13A) and similar pattern of clustering were detected when the aligned nt sequences of the P1 (Fig 4.13B), CI (Fig 4.13C), NIb (Fig 4.13D), and CP (Fig 4.13E) were used in phylogenetic analyses. The results indicate that CP-com1 is a highly divergent sequence variant of PeMoV.

To further investigate the molecular sequence characteristics of CP-com1, the SNAP was used as described in section 4.3.2. Nine complete polyproteins sequences were included in this analysis and results indicate that synonymous substitutions were more than non-synonymous substitutions (Fig 4.14A). The coding regions contributing to the high synonymous substitutions appear to be NIb and CP (Fig 4.14D, E). However, the P1 and CI were characterized by substantial levels of non-synonymous substitution rates (Fig 4.14B, C). However, the evolutionary pressure across the entire polyprotein of sequence CP-Com1 estimated by the ratio of dN/dS showed an average of 0.0603 implying that the virus was under strong negative or purifying selection.

4.4.1.7 *Cowpea aphid-borne mosaic virus*

In addition to BCMNV, BCMV and PeMoV, two other potyviral sequences were recovered from the HTS data of samples CP-1 and EP-1; they were named as Msekera and EP12, respectively. The length of sequence EP12 was determined to be 9,447-nt (MK906029) whereas that of Msekera was 9,433-nt (MW534369). Both sequences encompass 9159 nt/3053 aa complete polyprotein sequences and their BLASTn hits confirmed their

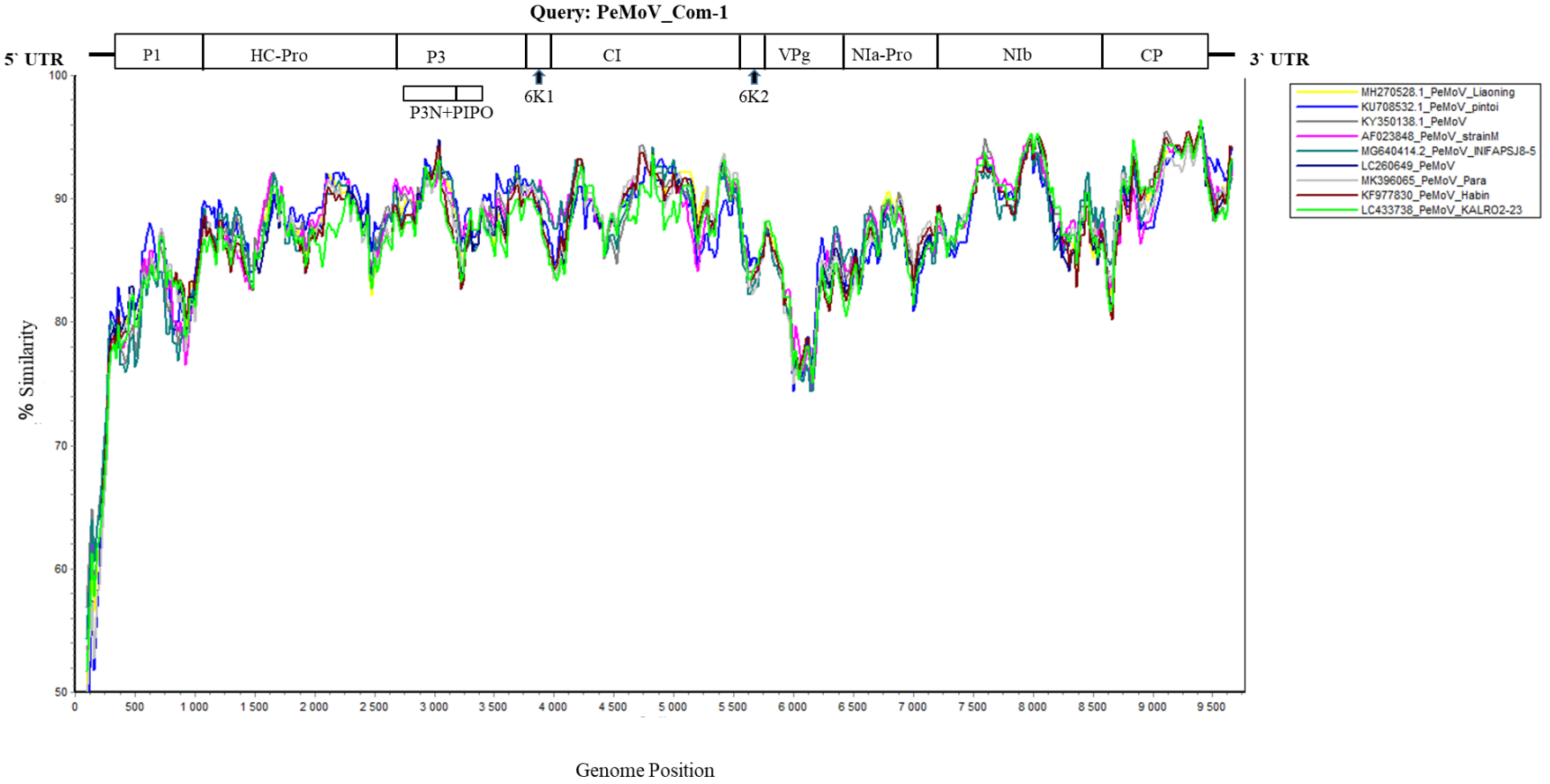
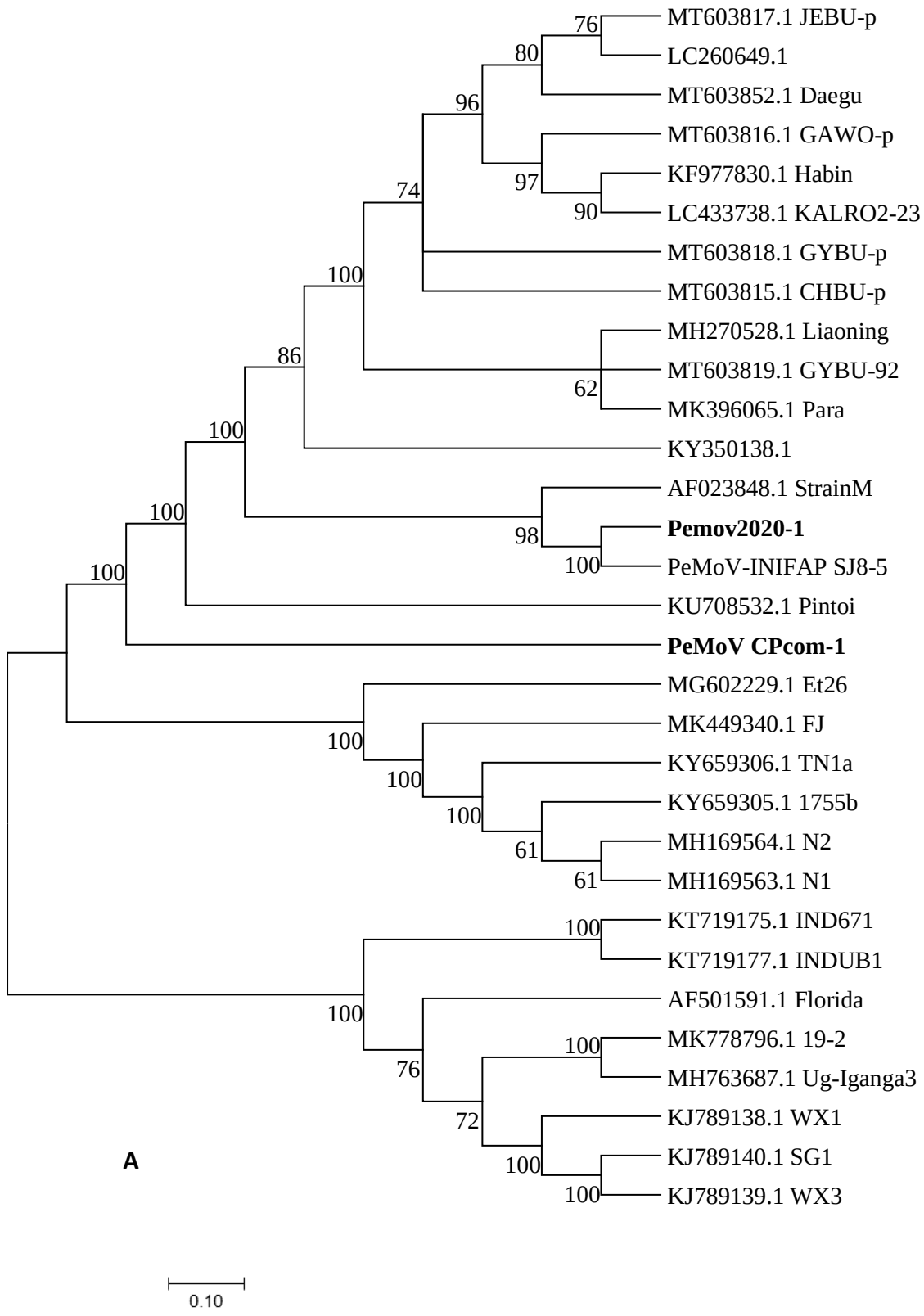


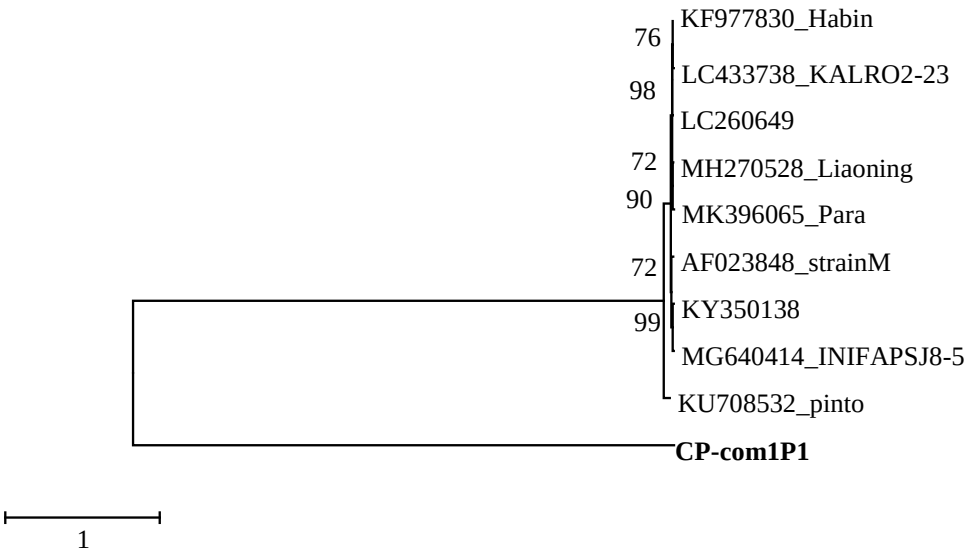
Figure 4.12. Sliding-window SimPlot depicting genome-wide graphical representation of percent similarity between isolate CP-com1 of peanut mottle virus (PeMoV) and nine global isolates of the virus. A diagrammatic representation of genome organisation of PeMoV with locations of the encoded proteins is given. P1=Potyviral protein 1, HC-Pro= Helper component protein, P3=Potyviral protein 3, 6K1= 6 kilodalton protein, CI= Cylindrical inclusion, 6K2=6 kilodalton protein 2, Vpg=Viral protein genome-linked, Nia= Nuclear inclusion-a, Nib=Nuclear inclusion-b and CP= capsid protein CP.

Table 4.4. Pairwise comparison of the nucleotide (nt) and amino acid (aa) identities between polyproteins and the coding proteins of CP-com1 sequence of peanut mottle virus (PeMoV) and global isolates of the virus.

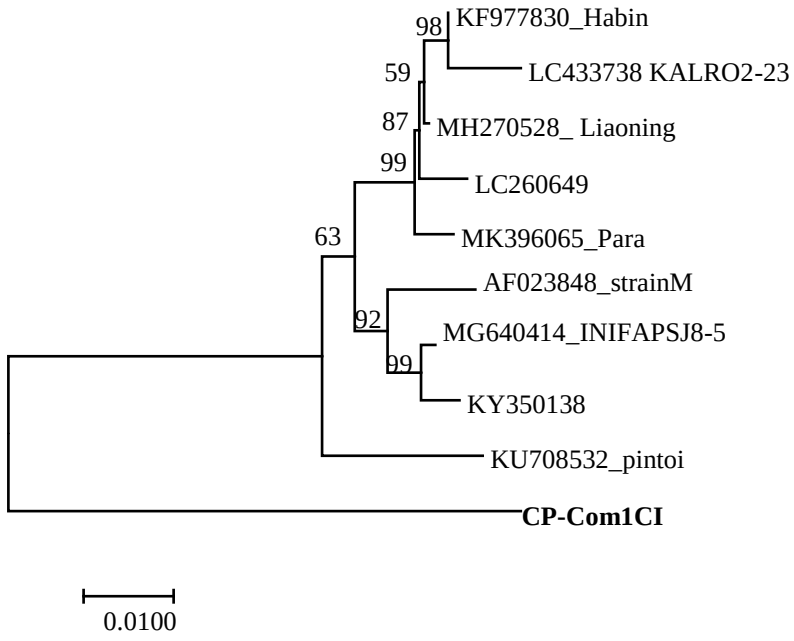
Accession	Polyprotein	P1	HP	P3	6K1	CI	6K2	VPg	NIa	NIb	CP
AF023848	89/93	82/76	89/95	90/91	90/100	91/98	85/90	84/88	88/97	90/96	91/94
KF977830	88/93	82/77	89/95	89/92	90/100	91/98	85/88	84/88	87/97	90/94	92/94
KU708532	89/93	83/78	90/94	90/91	90/98	90/98	86/88	84/90	87/97	90/96	91/94
KY350138	89/93	82/76	89/96	90/92	90/100	91/98	87/88	83/89	88/97	90/96	93/94
LC260649	88/93	82/77	89/94	89/91	89/98	91/97	85/90	84/89	87/97	90/95	92/94
MG640414	88/93	81/75	89/95	89/92	91/100	91/98	84/88	83/89	87/96	90/96	92/95
MH270528	89/93	83/77	89/95	90/92	90/100	91/98	85/90	84/89	87/97	90/95	92/94
MK396065	88/93	82/77	89/95	89/91	90/98	91/98	84/85	84/88	87/97	90/95	92/94

P1=Potyviral protein 1, HC-Pro= Helper component protein, P3=Potyviral protein 3, 6K1= 6 kilodalton protein, CI= Cylindrical inclusion, 6K2=6 kilodalton protein 2, Vpg=Viral protein genome-linked, Nia= Nuclear inclusion-a, Nib=Nuclear inclusion-b and CP= capsid protein CP.

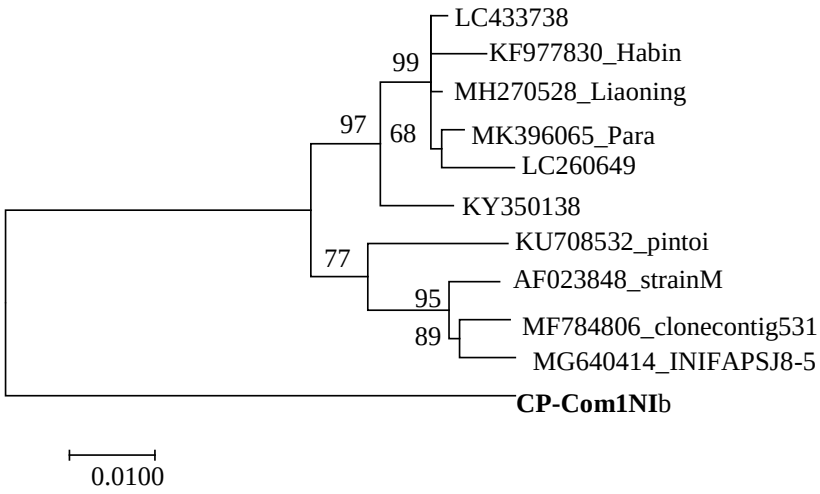

MT603817.1 JEBU-p
LC260649.1
MT603852.1 Daegu
MT603816.1 GAWO-p
KF977830.1 Habin
LC433738.1 KALRO2-23
MT603818.1 GYBU-p
MT603815.1 CHBU-p
MH270528.1 Liaoning
MT603819.1 GYBU-92
MK396065.1 Para
KY350138.1
AF023848.1 StrainM
Pemov2020-1
PeMoV-INIFAP SJ8-5
KU708532.1 Pintoi
PeMoV CPcom-1
MG602229.1 Et26
MK449340.1 FJ
KY659306.1 TN1a
KY659305.1 1755b
MH169564.1 N2
MH169563.1 N1
KT719175.1 IND671
KT719177.1 INDUB1
AF501591.1 Florida
MK778796.1 19-2
MH763687.1 Ug-Iganga3
KJ789138.1 WX1
KJ789140.1 SG1
KJ789139.1 WX3
76
80
96
97
90
74
100
86
62
100
100
98
100
100
100
100
61
61
100
100
76
72
100
100
A
0.10


KF977830_Habin
76
LC433738_KALRO2-23
98
LC260649
72
MH270528_Liaoning
90
MK396065_Para
72
AF023848_strainM
99
KY350138
MG640414_INIFAPSJ8-5
KU708532_pinto
CP-com1P1
1

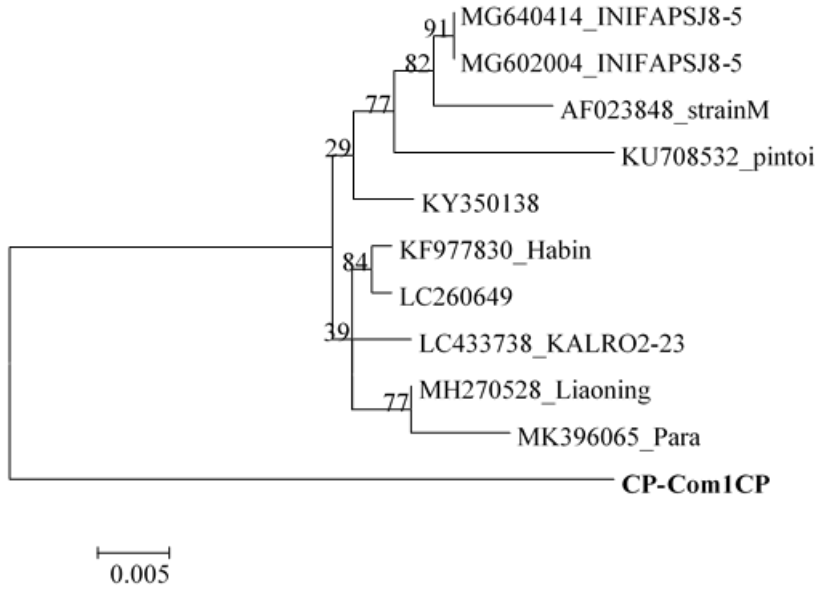
B


98
KF977830_Habin
59
LC433738 KALRO2-23
87
MH270528_ Liaoning
99
LC260649
63
MK396065_Para
AF023848_strainM
92
MG640414_INIFAPSJ8-5
99
KY350138
KU708532_pintoi
CP-Com1CI
0.0100

C



D



E

Figure 4.13. Molecular phylogenetic analysis using maximum likelihood method based on the General Time Reversible model (Nei and Kumar, 2000) that best-fitted the sequence data. The phylogenetic trees depicting relationships between CP-com1 of peanut mottle virus (PeMoV) from Zambia and global sequences as implemented in MEGA7 (Kumar *et al*., 2016) are shown. The trees were generated using the nucleotide sequences of the (A) complete polyprotein, (B) P1 protein, (C) cylindrical inclusion (CI), (D) NIb (the RNA-dependent RNA-polymerase of potyviruses) (CI), (D), and. capsid protein (CP).

Table 4.5. Estimations of substitutions occurring in sequence CPcom-1 of Peanut mottle virus (PeMoV)

Compare	Sequence names	ds	dn	ds/dn	dN/dS
0 1	CP-Com1 KY350138.1	0.5856	0.0345	16.9650	0.0589
0 2	CP-Com1 LC260649	0.5940	0.0370	16.0756	0.0623
0 3	CP-Com1 MG640414	0.6095	0.0350	17.4369	0.0574
0 4	CP-Com1 MH270528	0.5812	0.0354	16.4027	0.0609
0 5	CP-Com1 MK396065	0.5963	0.0374	15.9415	0.0627
0 6	CP-Com1 AF023848	0.5883	0.0352	16.7191	0.0598
0 7	CP-Com1 KF977830	0.6017	0.0370	16.2847	0.0615
0 8	CP-Com1 KU708532	0.5933	0.0352	16.8512	0.0593

Comparison of the averages of the first sequence to others implemented in Synonymous and Nonsynonymous substitution Analysis Program (SNAP) v2.1.1, the Los Alamos National Laboratory HIV-sequence database (http://www.hiv.lanl.gov). First column compares the sequence CP-Com1 with the eight global sequences. **ds=**Synonymous substitution, **dn=** Nonsynonymous substitution. Averages ds = 0.5937, dn = 0.0358. The ratio dN/dS = 0.0603 indicates that there are more nonsynonymous substitutions in CP-com1 compared with the others.

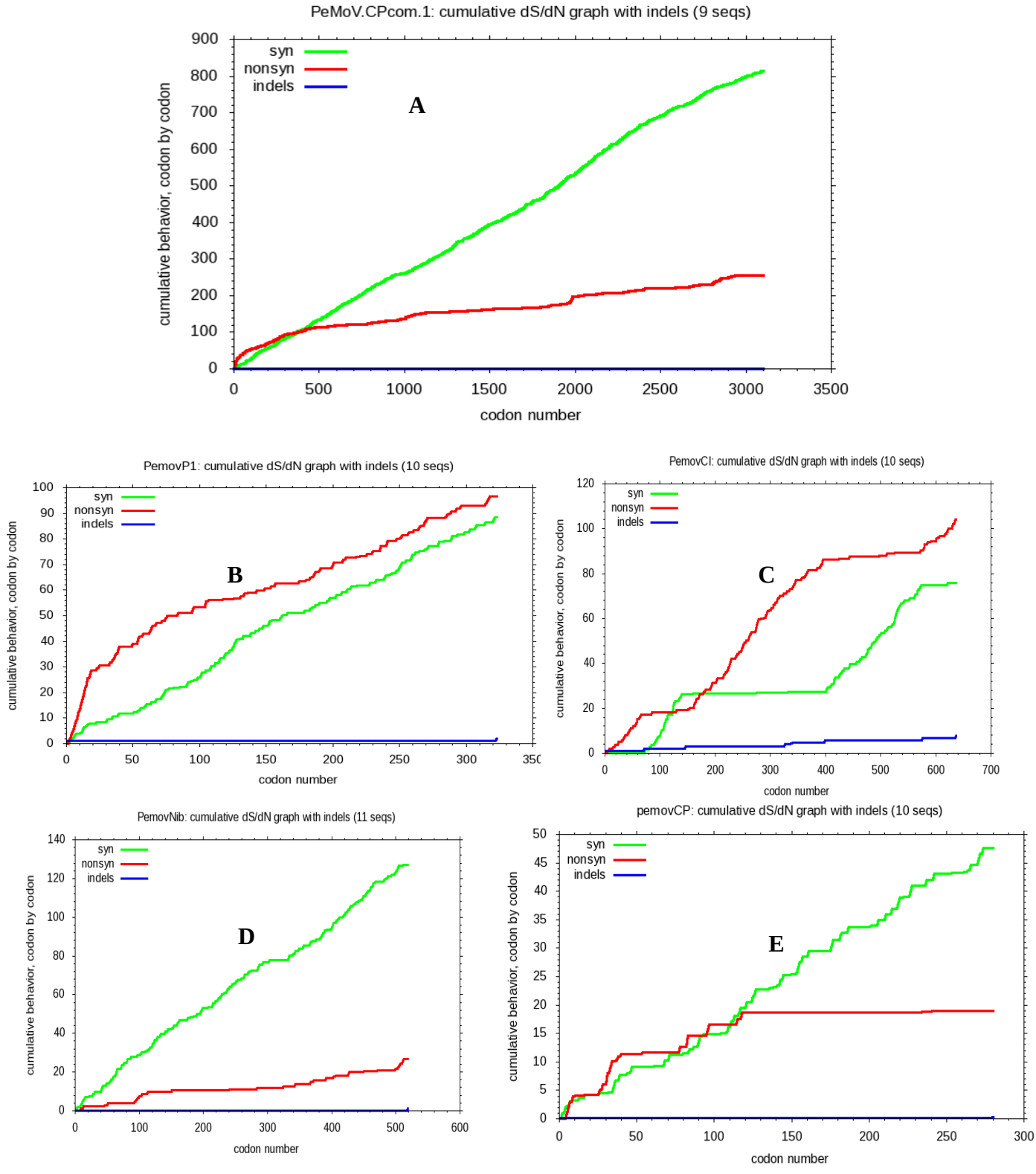


Figure 4.14. Synonymous and non-synonymous substitution in: A=polyprotein sequences of PeMoV global isolates and CP-com1, B=capsid protein (CP), C=cylindrical inclusion (CI), D=P1 and E=Nib (RNA-dependent RNA-polymerase). The analysis was implemented in the Los Alamos National Laboratory HIV-sequence database (http://www.hiv.lanl.gov). Average dS/dN for CP=12.2242, CI=2.9408, NIb=26.738 and P1=4.2640/3.2119.

specificities to cowpea aphid-borne mosaic virus (CABMV).

In pairwise comparison, EP12 and Msekera shared 99.5% complete nt/aa polyprotein sequence identity with each other and 96.6% to 96.8% nt/aa with CABMV isolate Z (AF348210) from Zimbabwe, their closest relative. In phylogenetic analysis the two Msekera and EP12 sequences clustered tightly together within the same phylogroup with isolates CABMV-Z (AF348210) from Zimbawe and QJ (MG995842) from China forming a separate subclade (Fig 4.15), suggesting similar evolutionary histories. The rest of the CABMV were distantly related to both isolates in this study.

4.4.1.8 *Phaseolus vulgaris alphaendonravirus* 1 (PvEV-1) and PvEV-2

The near-complete genome sequences of PvEV-1 (PvEV-1-LP-1) was obtained from composite sample LP-1 from Luapula province with a length of 13,894 nt (MW534367). That of PvEV-2 (PvEV-2-NP-1) was derived from sample NP-1 from Northern province and its length was 14,789 nt (MW534368). The encoded complete polypeptide of PvEV-1-LP-1 is 13,488 nt (4,496 aa) while that of PvEV-2-NP-1 is 14,763 nt (4,921 aa). In pairwise comparison, the complete polypeptide of PvEV-1-LP-1 shared 99% nt/aa identity with the corresponding sequences of isolate JKI ID 31403 (MK948542) from Canada while that of PvEV-2-NP-1 was most identical at the nt/aa levels to the corresponding sequences of a PvEV-2 isolate (AB719398) from Brazil at 98.2%/98.6%. This is the first report of the occurrence of both endornaviruses in Zambia. The combined phylogenetic analyses of complete polyprotein nt sequences of both viruses (2 from this study and 17 from GenBank) showed their segregation into two major virus-specific clades as expected (Fig 4.16), further confirming their phylogeny in respect of global sequences.

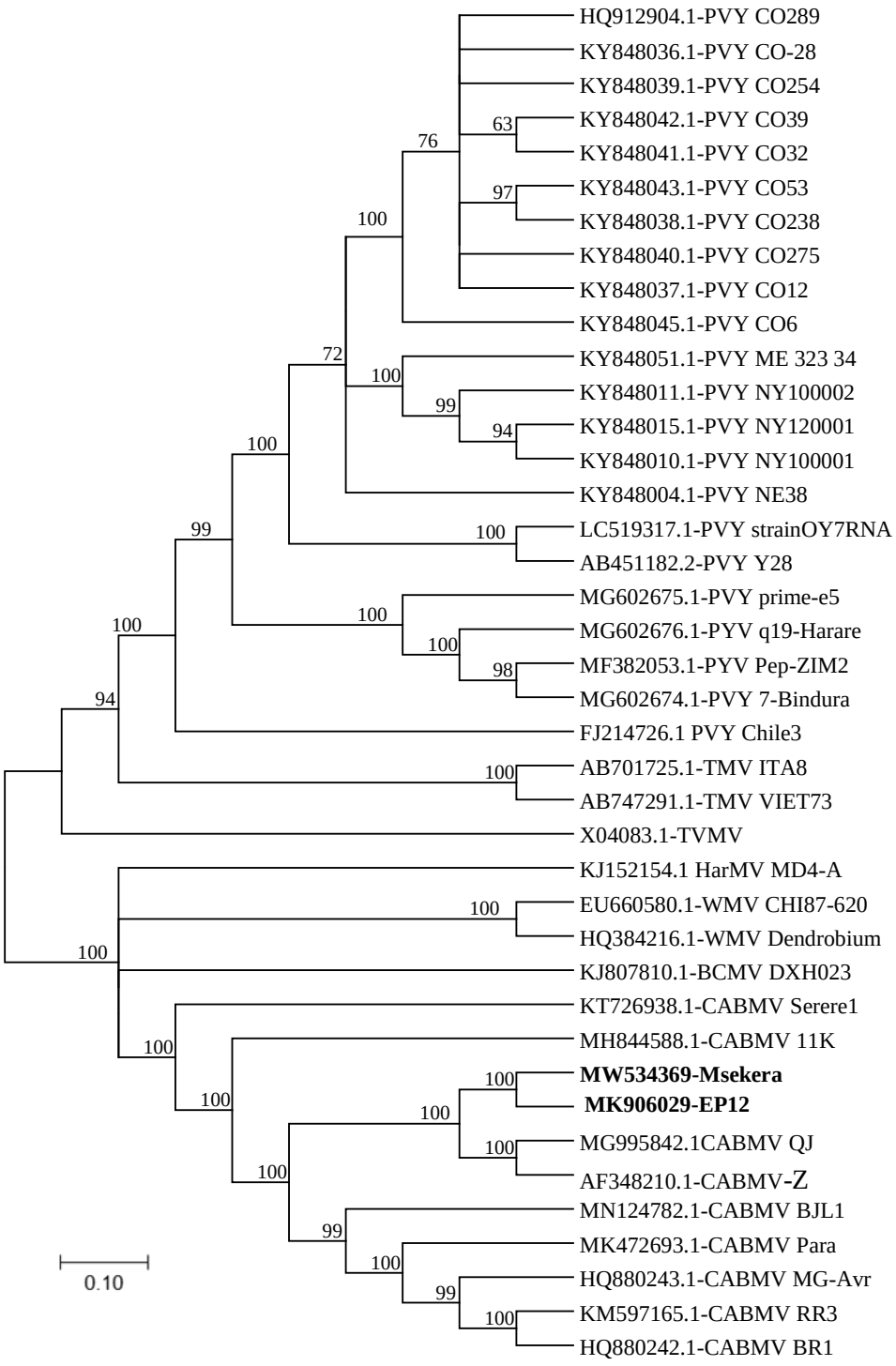


Figure 4.15. Phylogenetic analysis of cowpea aphid-borne mosaic virus (CABMV) showing clustering of Msekera and EP12 sequences with the isolates CABMV-QJ and CABMV-Z. The Phylogenetic analysis was performed using the Maximum Likelihood method and the General Time Reverse (GTR) that best fitted the dataset. The two CABMV isolates from Zambia clustered with a previously characterized cowpea isolate of the virus from Zimbabwe (Mlotshwa *et al.,* 2002) and China (NCBI data). BCMV=Bean common mosaic virus; TMV=Tobacco mosaic virus; PVY=Potato virus Y; WMV=Watermelon mosaic virus.

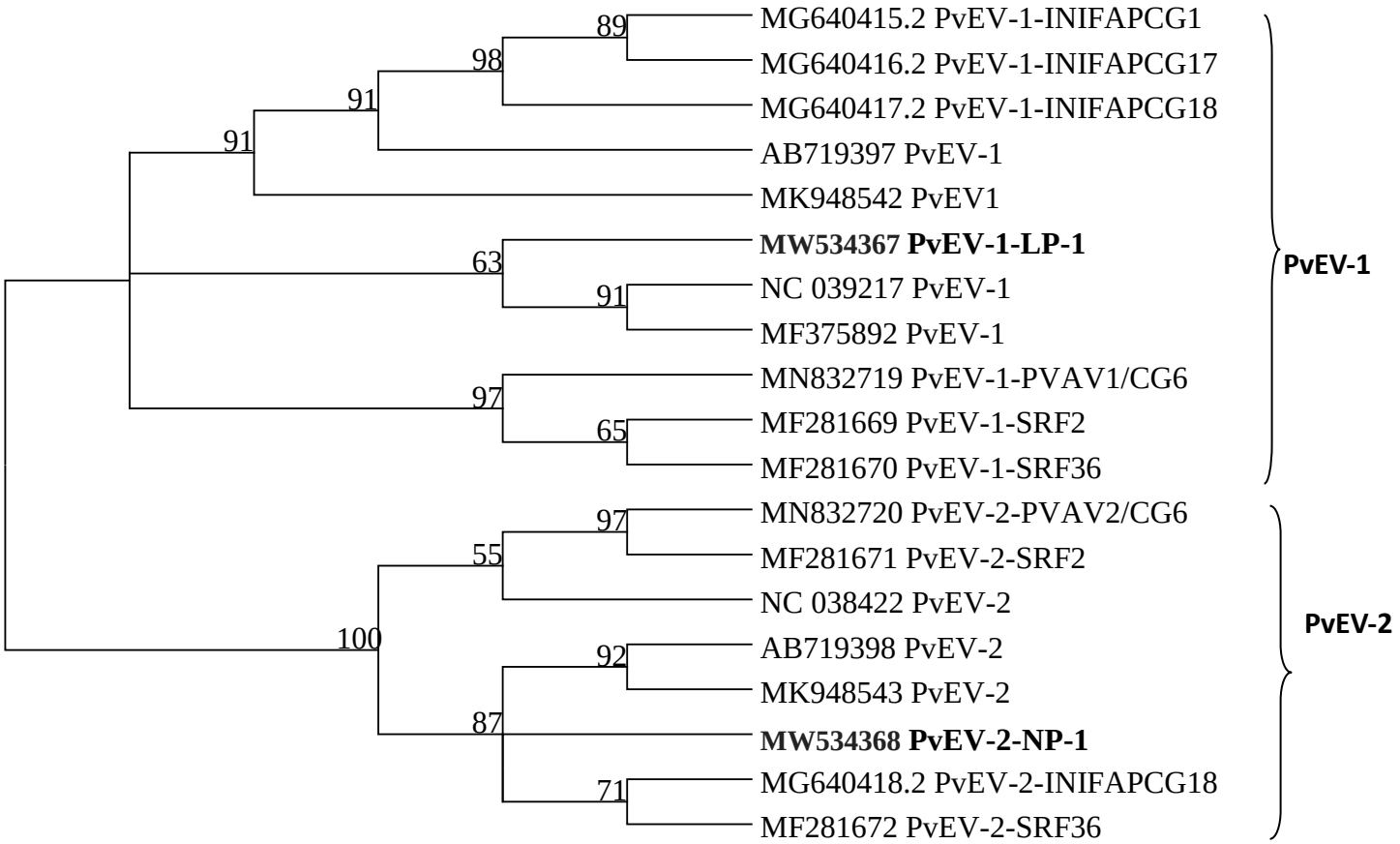


Figure 4.16. Phylogenetic analysis of two sequences determined to be Phaseolus vulgaris endornavirus 1 and 2 (PvEV-1 and 2) conducted in MEGA7. The Maximum Likelihood method was used to infer their evolutionary histories. Sequence PvEV-1ZM clustered with PvEV-1 global isolates (clade I) whereas PvEV-2 clustered with PvEV-2 global isolates (clade II).

4.5 Discussion

Together, the 24 highthroughput derived viral sequences including the CPPV1-Pv that was later recovered from samples that were infected with ETBTV provided insight into the virosphere of the common bean fields in Zambia. In addition to being the first set of sequences for any of the bean-infecting viruses ever reported in Zambia, this study reports for the first time the occurrence of five additional viruses (ETBTV, SBMV, PvEV-1, CPPV1-Pv and PvEV-2) in the country. Until now, ETBTV and its satRNA (ETBTV satRNA) have only been reported in Ethiopia and Kenya infecting members of the family *Solanaceae* (Abraham *et al.,* 2014; Kinoga *et al.,* 2021).

The ETBTV discovered in this study depicted sequence characteristics typical of umbraviruses with the 5' and 3' extremities of the UTRs ending in short sequences of GGG and CCC, respectively, similar to ETBTV isolate 18-2 from Ethiopia. However, the nucleotide sequence identities revealed ETBTV from Zambia is more closely related to the Malawian isolate (Udagawa *et al.,* 2020) than the isolates from Ethiopia and Kenya (Abraham *et al.,* 2014; Kinoga *et al.,* 2021). This suggests molecular diversity among the ETBTV isolates perhaps driven by different evolutionary factors that could be regionally influenced. However, with only five whole genomes available in GenBank to date, it is difficult to determine a well supported evolutionary history and whether there is a regional geographical spread of these viruses from a center of evolution. Based on genome-wide analysis, the low pairwise nucleotide sequence identities between the East African and Southern African ETBTV isolates point to two independently evolving species with a wider natural host range than previously determined (Abraham *et al.,* 2014; Udagawa *et al.,* 2020; Kinoga *et al.,* 2021).

Umbraviruses are satellite viruses that are incapable of coding for the capsid protein and thus need a helper virus for encapsidation and aphid transmission (Ryabov and Taliansky, 2008). Based on published data, helper viruses appear to be unique for respective umbraviruses (Mo *et al.,* 2011; Abraham *et al.,* 2014; Tang *et al.,* 2015). For example, potato leafroll virus (PLRV; genus *Polerovirus*) was associated with ETBTV isolate 18-2 in contrast to cowpea polerovirus1-Pv (CPPV1-Pv; genus *Polerovirus*), the presumed helper virus for ETBTV (CP414-1; MT900845) (part of this study). An unpublished partial sequence of TBTV isolate A2 (AJ704818) recovered from infected tobacco samples in Zimbabwe (GenBank data) was reported in association with tobacco vein distorting virus (TVDV; genus *Polerovirus*) its helper virus. Phylogenetically, TBTV-A2 clustered with ETBTVs characterized to date, similar to the clustering pattern earlier reported by Abraham *et al.* (2014) implying that isolate TBTV-A2 could be a sequence of ETBTV. Thus, the weight of evidence seems to implicate multiple helper viruses of ETBTV than previously thought and that the virus could probably be transmitted by several insect vectors with high probability to infect non *Solanaceous* plants of ETBTV. This poses an epidemiological challenge that needs further interrogation especially that ETBTV can occur in mixed infection with potyviruses (Kinoga *et al.,* 2021).

Southern bean mosaic virus (SBMV) causes varying disease symptoms in common bean (Zaumeyer and Harter, 1943; Tremaine and Hamilton, 1983; Verhoeven *et al.,* 2003; Ghorbani *et al.,* 2010; Mwaipopo *et al.,* 2018). A search in the nucleotide database for complete genomes of SBMV yielded only five genomes, one of which was recently reported from Tanzania, a neighbouring country of Zambia (Mwaipopo *et al.,* 2018). Consequently, the small number of complete or near complete genomes available in the nucleotide database made it difficult to infer the likely geographical origins of the SBMVs.

Most of SBMV sequences (7/8) in this study were genetically closely related to each other (Fig. 4.8). Across the genes, the most variable region noted was the ORF1 (MP) and the recently discovered ORFx (Px). This observation is consistent with reports elsewhere (Ling *et al.,* 2013; Sõmera *et al.,* 2015). Evidently, the MP sequences obtained in this study are largely conserved in most of the isolates that are closely related on the phylogenetic tree (Fig. 4.7a) but this is not so in the Px gene. Although the function of the Px gene is not well studied, mutation changes in turnip rosette virus (TRoV) that disrupted translation of the ORFx also disrupted establishment of infection (Ling *et al*., 2013). Based on this study, it can be hypothesized that changes in this region could have significant biological and epidemiological implications worth considering in future studies. As shown (Figure 4.7a), the SBMV sequence Mse-3 is more identical and phylogenetically related to SBMV-B(ARK) and SBMV-S(ARK) and divergent from other sequences of the virus from Zambia. This implies that SBMVs in Zambia are genetically diverse and that Mse-3 could have evolved separately from the others. Considering that only one divergent sequence of Mse-3 was isolated versus seven that are molecularly similar, it is likely that Mse-3 was a recent introduction. SBMV-S(ARK) is a mutant of the wild-type SBMV-B(ARK) and of the two only the former was capable of systemically infecting common bean cvs. Pinto and Great Northern (Lee and Anderson, 1998). Since the Mse-3 sequence is genetically related to the two isolates from Arkansas (ARK), the biological properties of Mse-3 including its pathogenecity should be investigated in future studies. Results from such a study could help in designing integrated disease management measures.

Another virus that was identified in this study was CMV, first reported in Zambia in the early 1990s (Vetten and Allen, 1991). CMV subgroups IA and II have a wider geographical distribution globally (Nouri *et al.,* 2014) but IB is predominantly of Asian dispersion (Roossinck *et al.,* 1999). Previous reports of CMV in Zambia did not show the occurrence of

any subgroup (Vetten and Allen, 1991) possibly due to limitation in the scope of the study. Wide dispersions of subgroup IB and other subgroups could induce severe pathological reponses in synergistic interactions between CMV and other viruses when in mixed infections (Murphy and Bowen, 2006). Such mixed infections could play significant role in future common bean viral disease outbreaks thereby causing huge yield losses such as reported in USA (Nault *et al.,* 2006; Thompson *et al.,* 2015). The discovery of the subgroup IB of CMV in the current study therefore, is important finding and this information will be useful in addressing diseases of virus etiology in farmers' fields.

Recent metagenomic studies of viruses of common bean in East Africa revealed the widespread occurrences of BCMNV in Kenya and Tanzania (Mutuku *et al.*, 2018; Mwaipopo *et al.,* 2018; Wainaina *et al.,* 2019). The three complete and near complete BCMNV sequences obtained in this study were closely related to three global isolates SRF75 (MF179111), TN1a (KY659306) and NL-3 (U19287). In phylogenetic analyses, the BCMNV sequences from Zambia clustered in the clade I of the three so called I-III clades (Wainaina *et al.,* 2019). Similarly, the three HTS assembled BCMV sequences produced similar clustering patterns forming six phylogroups (S, P, C, R1, R2 and R3 for BCMV) (Fig. 3A and B) consistent with previous reports (Moradi and Mehrvar, 2019). All the BCMV sequences from this study grouped with the recombinant global isolates indicating occurrences of possible recombination within the BCMV species of sequences from this study. This result is consistent with the observation that R1 BCMV isolates are intercontinental in their distribution (Moradi and Mehrvar, 2019). Like any other group of viruses, recombination in potyviruses gives rise to new strains that may induce more severe disease symptoms than those caused by their parental isolates (Feng *et al.*, 2014). For example, isolate NL-3 K that resulted from interspecific recombination between BCMNV NL-3 D and the Russian strain (RU1; accession number GQ219793) produced early and more severe disease symptoms in

common bean (Larsen *et al.,* 2005) and other naturally occurring recombinants are known to overcome resistance (Feng *et al.,* 2014). Recently, occurrence of the RU1 was reported in neighbouring Tanzania (Mwaipopo *et al.,* 2018) suggesting that it may be more widespread across East and Southern Africa. This could present a challenge to common bean growers in terms of disease management and to breeders in terms of finding sources of resistance.

Prior to this study, only eight complete genomes of PeMoV were available in GenBank, all of which were isolated from crops other than common bean. The genomes included PeMoV recently isolated from groundnut (*Arachis hypogaea*) from Kenya (Were *et al.,* 2018). Thus, the near complete PeMoV genomes reported in this study are the first PeMoV sequences isolated from common bean in SSA. The unique clustering of CPcom-1 from global isolates of the virus (Fig 4.13A) suggests that it evolved independently from others. As additional genome sequences of PeMoV isolates are determined from other countries in Africa, a better picture of the evolutionary history of CPcom-1 would emerge. The genome-wide analysis showed that the P1 of CPcom-1 is highly diverged from corresponding P1 regions of other PeMoV isolates, particularly the N-terminus of the peptide, which could explain its distinct phylogeny. Similar divergence in the P1 N-terminus has been reported for plum pox virus (PPV) (Nigam *et al.,* 2019). The P1 of potyviruses is multifunctional (Verchot and Carrington, 1995; Kasschau *et al.,* 2003; Rohozˇkova´ and Navra´til, 2011; Maliogka *et al.,* 2012; Pasin *et al.,* 2014). Considering previous reports which showed that variations in the P1 N-termimus of potyuviruses interfered with host adaptation and vector-mediated transmissibility (Johansen *et al.,* 1996; Nigam *et al.,* 2019), future studies should investigate the biological implications of the P1 N-terminus variations observed for PeMoV-CPcom-1.

CABMV is a frequently detected virus of leguminous plants and can cause significant yield losses of up to 60% (Damiri *et al.,* 2013) in susceptible cowpea genotypes (*Vigna*

unguiculata L. Walp.). In Zambia, CABMV was first detected in cowpea (*Vigna unguiculata* L. Walp.) using virus-specific antisera and five serotypes of the virus were identified (Kannaiyan and Hachiwa, 1993). Whereas documented host range of CABMV includes species in the family *Fabaceae* (Bashir *et al*., 2002), this is the first time CABMV is reported infecting common bean in Zambia. The close phylogenetic relatedness between CABMV isolates from Zambia and isolate CABMV-Z from neighbouring Zimbabwe suggests possible spread of the virus between both countries especially in the border communities perhaps due to unregulated transboundary movement of planting material and aphid vectors for short distance movements. Such observations have been made to explain the incursion of other exotic viruses into Zambia like Ugandan cassava brown streak virus (UCBSV) (Mulenga *et al.,* 2018).

The infection of common bean plants with CABMV adds to the risky probability of exchange of genetic information if BCMNV and CABMV co-infect the same plant. This view is supported by the fact that among potyviruses, intra and interspecies recombinations are common and are associated with the emergency of new strains (Valli *et al.,* 2007; Larsen *et al.,* 2005; Mbanzibwa *et al.,* 2011; Feng *et al.,* 2014; Worrall *et al.,* 2015). For example, an interspecific recombination event between BCMV and soybean mosaic virus (SMV) resulted in a novel strain of watermelon mosaic virus (WMV) (Desbiez and Lecoq, 2004).

The data from this study provides insight into the virosphere of common bean fields in Zambia. The detection of ETBTV in bean plants represents an expansion of the host range of virus to include members of the family *Fabaceae*. The sequence data generated in this study are the first of their kind in Zambia for common bean-infecting viruses and they would serve as useful resource to breeders and plant pathologists in breeding programs and epidemiological studies. The results from this study also provide a snapshot of the genome

characteristics and genetic variability of field isolates of the detected viruses; information generated in this study will be used to design diagnostic tests for detecting viruses infecting common bean both in common bean fields and seed. Scientists may also use this virus sequence information in monitoring evolution of the common bean virome landscape in Zambia and elsewhere.

CHAPTER FIVE

RESPONSE OF COMMON BEAN CULTIVARS TO INFECTION BY BEAN COMMON MOSAIC VIRUS AND BEAN COMMON MOSAIC NECROSIS VIRUS

5.1 Abstract

Occurrence of bean common mosaic necrosis virus (BCMNV) and bean common mosaic virus (BCMV) as different pathotypes is a major constraint to management of these viruses in common bean fields. Five pathogroups (PGs) were reported in Zambia in the early 1990s including PGI, III, VIa and VIb reported in this study. However, previous reports are devoid of information on responses of locally grown common bean cultivars to infection by viruses belonging to the identified PGs. This study aimed to evaluate responses of Zambian common bean cultivars to infection by BCMNV and BCMV. To fill this gap, 67 cultivars (cv) and 13 improved varieties (IV) were planted in rows measuring 1 m x 0.5 m inter-row and 0.1 m intra-row spacing, randomized with three replications for natural infection with BCMNV and BCMV. Results of on-station field screening of the 80 bean genotypes revealed 14 (IV=13 and cv=1) were not infected and subsequently advanced for further evaluation. Independently, 2310 common bean plants were raised from seed collected from farmers' seed bank and individual plants were RT-PCR screened for infections with BCMV and BCMNV. The recovered viruses were pathotyped according to the different pathogenicity profiles exhibited in 16 common bean differential cultivars (DCs). This resulted in the detection of PG I, III, VIa and VIb. Pathogenicity studies involving the four PGs and the 14 genotypes revealed variable responses of the bean genotypes to infection. Further, analysis of resistance (R)-genes in the 14 common bean genotypes by PCR with published random amplified polymorphic DNA (RAPD) primers revealed *bc-1*^{2} R gene as the most prevalent followed by *bc-3* R gene that was detected in two genotypes (Lwangeni and Lunga). Information on the

possible sources of resistance within the Zambian common bean gene pool generated in this study will be useful in marker assisted breeding to develop varieties with multiple resistance to BCMNV and BCMV pathotypes.

5.2 Introduction

Virus symptoms observed in field grown common bean plants are mostly due to mixed virus infections (Mwaipopo *et al.,* 2017). Common bean is susceptible to multiple viruses belonging to several genera, including *Potyvirus*, *Sobemovirus*, *Comovirus*, *Begomovirus*, *Bromovirus*, *Cucumovirus*, *Soymovirus*, *Carlavirus*, *Umbravirus*, *Alphaendornavirus*, *Cytorhabdovirus*, *Crinivirus* and *Alfamovirus* (Vetten and Allen, 1991; Loebenstein and Thottappilly, 2004; Ghorbani *et al.,* 2014; Azizi and Shams-barksh, 2014; Mwaipopo *et al.,* 2018). Specific viruses are considered important if they are widely distributed and can potentially cause deleterious effects on yield (Rubio *et al.,* 2020). In common bean production globally, the two species BCMNV and BCMV are the most important viruses capable of causing considerable yield losses of up to 100% especially in tropical countries (Worrall *et al.,* 2015 and references therein). The management of diseases of viral etiology can be achieved through several approaches, including the use of certified disease free planting materials, control of vectors that transmit viruses, and use of resistant varieties (Worrall *et al.,* 2015). Preferably, development of resistant varieties offers the most durable disease management option (Worrall *et al.,* 2015).

Resistance to BCMNV and BCMV is controlled by the dominant *I* gene and six recessive genes (Drijfhout, 1978; Kelly *et al.,* 1995). Of the six, *bc-u* gene is strain non-specific and the rest are strain specific (Drijfhout, 1978). To fully express, *bc-u* gene is needed except in *I+bc-1*, *I+bc-1*^{2} and *I+bc-3* heterogenic combinations (Drijfhout, 1978; Miklas *et al.,* 2000; Mukeshimana *et al.,* 2005). When the *I*-gene carrying cvs are challenged with BCMV, they

either remain asymptomatic or display temperature dependent symptoms that could include local or systemic necrosis (Feng *et al.,* 2014). The dominant *I*-gene confers immunity to temperature-sensitive BCMV strains provided temperature during the growth cycle of the plant is below 30°C (Naderpour *et al.,* 2009). However, recently whole plant necrosis was reported in I-gene bearing cv infected with BCMV isolate RU1M at <30°C (Feng *et al.,* 2014); hence behaving like BCMNV and further confounding the etiology of common bean viral diseases. Conversely, BCMNV which is endemic in some African countries is temperature insensitive and causes top necrosis (black root) in common bean cvs that are monogenic for the dominant *I* gene. Consequently, there is limited use of the dominant gene in breeding common bean genotypes for resistance in Africa (Morales, 2006; Worrall *et al.,* 2015). Among the recessive genes, *bc-u*/*bc-3* combination confers resistance to most BCMV and BCMNV strains (Drijfhout, 1978; Mukeshimana *et al.,* 2005). The other R genes and their alleles (*bc-1/bc-1*^{2} and *bc-2/bc-2*^{2}) are strain specific (Miklas *et al.,* 2000).

Drijfhout (1978) grouped common bean genotypes into 11 hostgroups (HG-0 to HG-11) based on various genetic combinations of recessive alleles. He then developed a standard procedure for comparing pathogenicity of strains of BCMNV and BCMV using a set of common bean differential cultivars (DCs) that either carry no resistance genes, or may be monogenic or heterogenic for one or more genes. Using the established set of 11 DCs, many strains of BCMNV and BCMV have been characterized worldwide (Worrall *et al.,* 2015 and references therein). BCMNV and BCMV were classified initially into seven pathotypes also called pathogroups (PG) I to VII based on their differential interactions with resistance genes (Table 5.1) possessed by 11 DCs (Drijfhout, 1978; Feng *et al.,* 2014). However, an isolate of BCMV displayed a novel pattern of symptom profile that could not be accommodated within the already described profiles, thus necessitating the formation of a new PG (PG-VIII) (Feng *et al.,* 2015).

The interactions of different PGs of BCMNV and BCMV with the different host groups (HG) of common bean cultivars have been studied extensively (Drijfhout, 1978; Miklas *et al.*, 2000; Mukeshimana *et al.,* 2005; Feng *et al.,* 2014). However, emerging recombinant strains of both viruses continue to challenge known resistance traits of some varieties (Feng *et al.*, 2015). In Zambia, it is common to find virus-like symptoms among plants of several released varieties and landraces of common bean. The resistance gene profiles of the popular varieties in Zambia is not documented and the reference isolates of BCMNV and BCMV PGs that were reported in previous studies (Lana, 1995; Kaitisha, 2003) are unavailable. Thus, the basis for the designation of released common bean varieties in Zambia as resistant, tolerant, or susceptible (Muimui *et al.,* 2016) are unclear. Based on results indicated in Chapter 3, it is evident that southern bean mosaic virus (SBMV) is more widely distributed in Zambia than BCMNV and BCMV and that, depending on the cultivar, SBMV can induce mosaic symptoms resembling those described for BCMNV and BCMV. Thus, symptoms-based assessment of varietal response to viruses (including BCMNV and BCMV) is unreliable. Consequently, this study aimed to: (i) identify pathotypes among BCMNV and BCMV in Zambia, (ii) determine presence of the BCMV and BCMNV resistance gene markers within the gene pool of released common bean varieties in Zambia, and (iii) screen common bean genotypes for resistance to the identified BCMNV and BCMV pathotypes. The information generated from this study could be used to realign breeding priorities and most importantly prioritise marker assisted selection (MAS) in future common bean breeding efforts of the National Agriculture Research Station (NARS).

Table 5.1. Host resistance gene reponses and pathogenicity of bean common mosaic virus (BCMV) and bean common mosaic necrosis virus (BCMNV) isolates. The table was modified from Drijfhout (1978), Spence and Walkey (1995) and Silbernagel *et al.* (2001).

			Pathogroup (PG) x differential cultivar (DC) interactions									
			US1/ NL1	**NL7**	**NL8**	**NL6**	**US5**	**NL2 US2 NY15**	**NY15**	**NL3**	**NL5**	**US6 NL4**
Host group	**Differential Cultivar**	**Strain-specific resistance genes**	**I**	**II**	**III**	**IVa**	**IVb**	**Va**	**Vb**	**VIa**	**VIb**	**VII**
0	Dubbele Witte	0	+	+	+	+	+	+	+	+	+	+
	Sutter Pink	0	+	+	+	+	+	+	+	+	+	+
1	Stringless Green Refugee	*bc-u*	+	+	+	+	+	+	+	+	+	+
2	Redlands Greenleaf C	*bc-u, bc-1*	-	+	-	+	+	+	+	+	+	-
4	Sanilac	*bc-u, bc-2*	-	-	+	-	-	+	+	+	+	-
	Michelite 62	*bc-u, bc-2*	-	-	+	-	-	+	+	+	+	-
	UI-34 Red Mex	*bc-u, bc-2*	-	-	+	-	-	+	+	+	+	-
6	UI-31 Great Northern	*bc-u,bc-1*^{2}, *bc-2*^{2}	-	-	-	-	-	-	-	-	-	+
	Monroe	*bc-u, bc-1*^{2}, *bc-2*^{2}	-	-	-	-	-	-	-	-	-	+
	UI-35 Red Mex	*bc-u, bc-1*^{2}, *bc-2*^{2}	-	-	-	-	-	-	-	-	-	+
8	Black Turtle Soup I	*I*	-	-	+	+	+	+	+	+	+	-
	Widusa	*I*	-	-	+	+	+	+	+	+	+	-
9a	Jubila	*I, bc-1*	-	-	-	+	+	+	+	+	+	-
9b	Topcrop	*I, bc-1*	-	-	-	+	+	+	+	+	+	-
10	Amanda	*I, bc-1*^{2}	-	-	-	+	+	+	+	+	+	-

+ = susceptible; – = resistant

5.3 Materials and Methods

5.3.1 Field screening of common bean cultivars for resistance to viruses

A total of 80 common bean genotypes comprising 67 accessions sourced from the National Plant Genetic Resources Center (NPGRC) of Zambia and 13 improved varieties were included in the field screening experiment. They were planted in an alpha lattice experimental design with rows measuring 1 m long, 0.5 m inter-row and 0.1 m intra-row spacing, randomized incomplete block design with three replications. Known BCMNV and BCMV infected seeds were planted as spreader rows. D-compound fertilizer (10N:20P:10K) was applied at planting at the rate of 150 kg/ha. The plants were rain-fed with supplementary watering during instances of prolonged absence (5 consecutive days) of rain. The experimental fields were weeded regularly, and no insecticides were applied to avoid killing beneficial insects and vectors. Data on incidence and severity of virus and virus-like symptoms was collected commencing seven days post emergence (dpe) and weekly up to 50-60 days after planting (dap) which is the podding stage. Disease incidence was calculated as stated in Section 3.3.1 and symptom severity was scored on a scale of 1-5 as described in section 3.3.7. Aphid densities per plant was scored on a five-point scale: 1 = no aphids, 3 = 1-5 aphids, 5 = 6-10 aphids, 7 = 11-15 aphids, and 9 = more than 15 aphids (modified from Togola *et al.,* 2020). Percentage incidence of aphids was calculated as a proportion of plants having the black bean aphids over the total number of plants assessed per plot.

5.3.2 Screening for viruses using reverse transcription polymerase chain reaction

To profile viruses in experimental plants, leaf samples of symptomatic and nonsymptomatic plants (three plants/cultivar) were collected and used to extract total nucleic acid (TNA) as described in section 3.3.3. The extracted TNA was processed until RT-PCR as described in section 3.3.5. All cultivars that did not show virus and virus-like symptoms or showed mild (severity score=2) during vegetative growth and were negative to BCMNV and BCMV in

RT-PCR were presumed resistant. Seeds from such cvs were collected and advanced for further screening in the screenhouse under artificial infection.

5.3.3 Identification of BCMNV and BCMV inoculum in farmer collected common bean seed

Two hundred and thirty-one seed packs (200 seeds/pack) of beans were collected from four provinces of Zambia; Luapula, Eastern, Northern and Muchinga. Ten seeds per pack were randomly selected and sown in two 20 x 25 cm polythene pots (5 seeds per pot) thus translating into a total of 2,310 seeds. D-compound fertilizer (10N:20P:10K) was applied at the rate of 75 g/pot; an equivalent of 150 kg/ha and the plants were watered every third day. The plants were maintained at a 12 hr photoperiod and day temperature range of 20-35℃. Systemic and contact insecticides and fungicides were applied at rates predetermined by the manufacturer to remedy occurrence of respective pests. Scouting for virus disease symptoms commenced 7-dpe and continued every seven days for 60 days corresponding to the flowering and podding stages. For each accession/variety, two plants with virus symptoms were selected and samples taken from each plant, resulting in 462 symptomatic samples. Additionally, samples with symptoms that were atypical of BCMNV/BCMV infection were collected from 138 plants thus making a total of 600 plants or ~27% of total plants planted. TNA was extracted from the samples as stated above and screened for the presence of the HTS detected viruses (Section 3.3.4). Results from this section were used in subsequent experiments.

5.3.4 Amplification of partial coat protein regions of BCMNV and BCMV and Sanger sequencing

For each sample that was infected with either BCMNV or BCMV alone (section 5.3.3), three separate RT-PCR reactions were performed using newly designed agent-specific primers

targeting the capsid region of both viruses (Table 3.1). Amplified DNA fragments with expected size were gel purified with the Zymoclean® gel DNA recovery kit (Zymoclean, Irvine, CA) and eluates were Sanger sequenced in both orientations. Sequence analyses were performed as in section 3.3.6 and the completed partial sequences were BLASTn queried against the nucleotide database to ascertain virus specificities and obtain homologous sequences.

5.3.5 Mechanical inoculation of common bean samples

Symptomatic leaf samples from plants positive for either BCMNV or BCMV were ground in 0.01M phosphate buffer comprising monopotassium phosphate (KH_2PO_4) and sodium phosphate dibasic dehydrate ($Na_2HPO_4.2H_2O$) pH7.0 to which 0.1% β-mercaptoethanol was added shortly before use. A total of four plants/pot were inoculated with the crude sap of either BCMNV or BCMV by dusting one of the two expanding primary leaves with carborandum powder (400 mesh) and using a sterile swab, the dusted leaves were rub-inoculated with the inoculum homogenate. One control plant was inoculated with buffer alone and excess sap was rinsed off with water. Growth conditions for the inoculated plants were as earlier described. The plants were monitored daily and scored for symptoms beginning seven days post inoculation (dpi), for up to four weeks. This method was used in separate experiments to biologically characterize the PGs of BCMNV and BCMV and determination of responses of common bean cvs to infection by viruses belonging to the different PGs.

5.3.6 Identification of BCMNV and BCMV pathogroups

Leaves from systemically infected plants confirmed by RT-PCR and Sanger sequencing (Section 5.3.4) were collected and used in the infectivity experiment to determine the PGs of BCMNV and BCMV detected in section 5.3.4. Sixteen 16 DCs were obtained from the

United States National Plant Germplasm System (NPGS) and planted in 20 x 25 cm polythene pots at the rate of 10 seeds/pot. After germination, the plants were thinned to five plants/pot. Four of the five plants were inoculated with the BCMNV and BCMV, respectively, as described in section 5.3.5. Infection responses data was collected as earlier indicated (Section 5.3.5). Based on reaction profiles of the 16 DCs and referencing the reaction profiles of the standard pathotype allocation guide for BCMNV and BCMV (Drijfhout, 1978; Silbernagel *et al.*, 2001) (Table 5.1; 5.4), the different isolates of BCMNV and BCMV were allocated to their respective PGs.

5.3.7 Screening of common bean genotypes in controlled environment

Thirteen (13) improved varieties and one accession that showed no symptoms during on-station field evaluations (Section 5.3.1) were advanced to the screenhouse experiment. Ten seeds of each of the 14 test genotypes, three DCs [Dubbele Witte, Amanda and Black Turtle Soup (BTS)] and two F_3 lines introgressed with *bc-3* and *I +bc-3* resistance genes (Courtesy of Dr. Kelvin Kamfwa, Department of Agricultre Sciences, University of Zambia) were planted in 20 x 25 cm polythene pots in a randomized complete block design (RCBD) with three replications. Dubbele Witte (resistance gene *0*), Amanda (resistance gene *I+bc-1*^{2}), BTS (resistance gene- *I*) and F_3 lines (resistance genes *bc-3* and *I +bc*-3) were included as controls. The growth conditions and agronomic practices were as described in section 5.3.3. Inoculation with four PGs identified in section 5.3.6 was achieved following the inoculation method described in section 5.3.5. Reactions of the plants to each of the isolates belonging to four PGs were recorded weekly until 30 dpi. Plants expressing symptoms of systemic mosaic, or necrosis were considered susceptible (S), otherwise they were regarded as resistant (R). Systemic infection with BCMNV and BCMV was confirmed by RT-PCR as described in section 3.3.5.

5.3.8 Screening for genetic markers for BCMNV and BCMV resistance

The common bean genotypes screened in section 5.3.7 were analyzed for the presence or absence of genetic markers (i.e. fragments of DNA showing polymorphism between individuals) flanking resistance genes *I, bc-u, bc-1, bc-1*^{2}, *bc-2, bc-2*^{2} *and bc-3* using published RAPD primers (Table 5.2). Leaf tissue samples were collected from 10-day old common bean plants and DNA extracted from each sample using the CTAB method (Chang *et al.,* 1993) as described in section 3.3.3. Approximately 20 to 50 ng DNA aliquot from each of the samples was used as template in a 12.5 µl PCR reaction volume comprising final concentrations 1x Taq buffer, 0.2 mM dNTPs, 0.2 µM each primer (SW13F/R, SG6F/R, ROC11F/R, and SBD5F/R; Table 5.2), 2 mM $MgCl_2$ and 0.8 U Taq DNA polymerase. Thermal cycling conditions consisted of intital denaturation at 94°C for 3min, 35 cycles of denaturation at 94°C for 30sec, variable annealing temperatures for 30sec, extension at 72°C for 1-2 min and final extension at 72°C for 7 min. Electrophoresis of PCR products was performed in 1% agarose gel stained with ethidium bromide at 10 mg/ml in 0.5x Tris-acetate-EDTA buffer and visualized using a Gel Doc XR System (Bio-Rad Laboratories, Hercules, CA). However, the PCR products obtained with primers SG6F/R were electrophoresed in 1% polyacrylamide gel (Sigma, Steinheim, Germany) to separate the 134 bp lower band (associated with resistance) from the 137 bp upper band (associated with susceptibility).

5.4 Results

5.4.1 Feld screening of common bean accessions/varieties

Diverse virus-like symptoms such as mosaic, leaf necrosis, severe leaf distortion, leaf narrowing, rugosity, curling and vein chlorosis were scored on diseased plants beginning 7 days post inoculation (dpi). Disease prevalence was high (90% or 72 of 80 genotypes) across the cultivars. Disease incidence within genotypes ranged from 0 to 30%, with 66 of 67

Table 5.2. Oligonucleotides used to screen for markers in common bean and detect viruses in the inoculated plants

Primer name	Sequence	Target size (bp)	Reference
SBD5*	GTGCGGAGAGGCCATCCATTGGTG	1300	Miklas *et al*. (2000)
	GTGCGGAGAGTTTCAGTGTTGACA		
ROC11*	CCAATTCTCTTTCACTTGTAACC	350	Johnson *et al*. (1997)
	GCATGTTCCAGCAAACC		
SG6*	GTGCCTAACCGAGTTATCTAGAGT	134/137	Mukeshimana *et al*. (2005)
	GTGCCTAACCCTCCTAAATGACCT		
SW13*	CACAGCGACATTAATTTTCCTTTC	690	Melotto *et al*. (1996)
	CACAGCGACAGGAGGAGCTTATTA		
BCMV-F**	AAGGATGTGAATGCTGGTTCTA	536	This study
BCMV-R	CTCTTGCTCGATCCGATGTT		
BCMNVFcpF2**	GCTGGGGCCGATGAGAG	711	Mwaipopo *et al*. (2018)
BCM-NVcommonR	GTCCCKTGCAGTGTGCCT		

*=Markers used to amplify specific resistance genes in common bean. SBD5 amplified *bc-1*^{2}, ROC11 and SG6 amplified *bc-3* and SW13 amplified *I* gene markers. **= Primers used to confirm presence of Bean common mosaic virus (BCMV), Bean common mosaic necrosis virus (BCMNV) and Cucumber mosaic virus (CMV) in inoculated plants.

accessions recording higher disease incidences than the 13 improved varieties. Symptom severity ranged from 1 to 3.7 across all the 80 genotypes; 1 to 2.6 and 1 to 3.7 for the improved varieties and the 67 accessions, respectively. Seven of the released varieties (Lwangeni, Chambeshi, Lukupa, Kalambo, Kapisha, Sadzu and Lunga) and one NPGR sourced accession ZM3292 did not show virus or virus-like symptoms at all whereas six others showed virus-like symptoms but negative in RT-PCR for any of the viruses tested. Incidence of aphids ranged from 0 to 21.4% but averaged 8.9%. Severity of aphid infestation ranged from 2 to 7.5 with an overall mean of approximately 5.2. Among the improved common bean varieties, the highest number of aphids was recorded on Lunga (Fig. 5.1) with several varieties recording less than 2% incidence whereas among the Genebank collected cultivars, the highest aphid incidence recorded was 19.05%.

RT-PCR screening of three randomly sampled plants per genotype revealed that SBMV was the most prevalent virus, with its occurrence in 60 of the 80 genotypes, followed by BCMNV (n = 44). PvEV-1 (n = 31), CMV (n = 23), PvEV-2 (n = 14), BCMV (n = 5), and PeMoV (n = 5)/CABMV (n = 5) (Table 5.3). None of the 80 genotypes tested positive for ETBTV.

5.4.2 Identification of viruses and assignment of BCMNV and BCMV to pathogroups

Screening of the 600 plants raised from seeds collected from four provinces resulted in the detection of SBMV in 399 plants, BCMNV in 53, PvEV-1 in 40, CMV in 30, PvEV-2 in 18, BCMV in 6, BCMNV+PvEV-1 in 16, BCMNV+CMV in 34 and BCMV+PvEV-2 in 4. All of the 600 plants tested negative for PeMoV, ETBTV and CABMV. The crude extracts prepared from plants that were positive for BCMNV (53) and BCMV (6) were inoculated onto two common bean differential cultivars (DCs); Dubbele Witte (Fig. 5.2) and Sutter Pink (Resistance gene 0), respectively. RT-PCR screening of DCs inoculated with BCMNV (53) and BCMV (6) confirmed single infections in 25 (BCMNV = 23 and BCMV = 2) plants

while the rest were contaminated with either PvEV-1 or PvEV-2. The pathogenicity profiles for the 25 BCMNV and BCMV isolates segregated the 23 isolates of BCMNV into PGVIa=10, PGVIb=5, PGIII=8 and the two BCMV into PGI (Table 5.4).

Partial coat protein (CP) sequences obtained from eight isolates representative of the four identified PGs from Zambia (PGI, PGIII, PGVIa, and PGVIb) shared 97-99.8% nt and 98-100% aa identities with the corresponding global isolates of each pathogroup (PGI = US1/NL1 [MH024841]; PGIII = NL8 [KY659304]; PGVIa = TN1a [KY659306]; PGVIb = TN1 [KY659305]), further supporting the results of the bioassay.

Figure 5.1. Aphid colonization of common bean plants in the open field trial

Table 5.3. Disease incidence and severity of virus-like symptoms and prevalence of viruses detected from testing of common bean accessions/improved varieties from the field trial using reverse transcription polymerase chain reaction (RT-PCR)

Variety/Accessions	Source	Percent mean disease incidence (n = 1,561)	Mean symptom severity	BCMNV	BCMV	SBMV	CMV	PeMoV	CABMV	PvEV-1	PvEV-2	ETBTV
Lunga	NLT	0	1	-	-	-	-	-	-	-	-	-
Kabale	NLT	16.7	2.4	-	-	+(4)	+(1)	-	+(1)	+(1)	+(1)	-
Sadzu	NLT	0	1	-	-	-	-	-	-	-	-	-
Kalambo	NLT	0	1	-	-	-	-	-	-	-	-	-
Kapisha	NLT	0	1	-	-	-	-	-	-	-	-	-
Kabulangeti	NLT	13.3	2.5	-	-	+(3)	+(1)	-	-		+(1)	-
Lukupa	NLT	0	1	-	-	-	-	-	-	-	-	-
Lyambai	NLT	11.7	2.3	-	-	+(1)	+(1)	-	+(1)	+(4)	-	-
Kalungu	NLT	6.7	3	-	-	+(1)	+(2)	-	-	+(1)	-	-
Lungwe bungu	NLT	3.3	2.3	-	-	-	+(2)	-	-	-	-	-
Mbereshi	NLT	16.7	2.6	-	-	+(5)	+(1)	-	-	-	-	-
Chambeshi	NLT	0	1	-	-	-	-	-	-	-	-	-
Lwangeni	NLT	0	1	-	-	-	-	-	-	-	-	-
Accessions (n = 66)	NPGR	0-30	1-3.7	+(44)	+(5)	+(46)	+(15)	+(1)	+(2)	+(25)	+(12)	-
ZM3292 accession	NPGR	0	1	-	-	-	-	-	-	-	-	-

-/+ = negative/positive for the virus, +() = number in parenthesis represents the number of plants that tested positive out of 240 plants tested (3 plants/genotype). NLT=national legumes team, NPGR= National Plant Genetic Resources. BCMNV = bean common mosaic necrosis virus, BCMV = bean common mosaic virus, SBMV = southern bean mosaic virus, CMV = cucumber mosaic virus, PeMoV = peanut mottle virus, CABMV = cowpea aphid-borne mosaic virus, PvEV-1 = Phaseolus vulgaris endornavirus 1, PvEV-2 = Phaseolus vulgaris endornavirus 2, ETBTV = Ethiopian tobacco bushy top virus.

Figure 5.2. Mosaic symptoms induced by bean common mosaic necrosis virus (BCMNV) isolate belonging to pathogroup (PG) III on common bean differential cultivar Dubbele Witte.

5.4.3 Profiling of BCMNV and BCMV resistance genes in the common bean cultivars

Four published gene markers (Table 5.2) were used to profile resistance genes for each of the common bean varieties and accessions evaluated in the screenhouse. The molecular markers that tag *bc-1*^{2} (SBD5: Miklas *et al.,* 2000), *bc-3* (ROC11: Johnson *et al.,* 1997; SG6: Mukeshimana *et al.,* 2005), and *I* (SW13: Haley *et al.,* 1994; Melotto *et al.,* 1996) gene loci (Table 5.2) were deployed in RAPD-PCR. As a guide to interpretation of the results, the

Table 5.4. Reactions of differential cultivars in hostgroups 0 to 10 inoculated with field isolates of bean common mosaic virus (BCMV) and bean common mosaic necrosis virus (BCMNV) in this study.

			BCMV (n = 2)	BCMNV (n = 1)	BCMNV (n = 22)	
Differential cultivar	Hostgroup	Resistance genes	PGI	PGIII	PGVIa	PGVIb
Dubbele Witte	0	0	+	+	+	+
Sutter Pink	0	0	+	+	+	+
Stringless Green Refugee	1	*bc-u*	+	+	+	+
Redlands Greenleaf C	2	*bc-u, bc-1*	-	-	+	+
Sanilac	4	*bc-u, bc-2*	-	+	+	+
Michelite 62		*bc-u, bc-2*	-	+	+	+
UI-34 Red Mex		*bc-u, bc-2*	-	+	+	+
UI-31 Great Northern	6	*bc-u, bc-2, bc-2*^{2}	-	-	-	-
Monroe		*bc-u, bc-I*^{2}*, bc-2*^{2}	-	-	-	-
UI-35 Red Mex		*bc-u, bc-I*^{2}*, bc-2*^{2}	-	-	-	-
Black Turtle Soup I	8	*I*	-	N	N	n
Widusa		*I*	-	N	N	n
Jubila	9a	*I, bc-1*	-	-	N	n
Improved Tendergreen 40031	9b	*I, bc-1*	-	-	N	n
Topcrop		*I, bc-1*	-	-	N	n
Amanda	10	*I, bc-1*^{2}	-	-	N	n

n = systemic necrosis in homogenic (*I* gene) and heterogenic (*I/bc-1*^{2} *and I/ bc-1*) cultivars; + systemic infection (susceptible), - no systemic infection, (resistant). N=local necrosis on the inoculated leaves only

amplification of the size specific DNA band for the marker SBD5 in a sample would indicate presence of the *bc-1*^{2} gene in the genotype and vice versa. Unlike the *bc-1*^{2} and *I* genes, the *bc-3* gene was probed with two markers; SG6 and ROC11. For marker SG6, the presence of the *bc-3* gene was confirmed if the size of amplified DNA band was 134 bp and absent if the size of the band obtained was 137 bp. The presence of *bc-3* gene using ROC11 was indirectly determined such that amplification of the expected 350 bp DNA band implies absence of *bc-3* in the sample while the presence of the gene was determined by non-amplification of the expected band (Hegay *et al.*, 2013). Based on these descriptions, the resistance genes of the 13 common bean varieties and one accession were determined. Of the 14 cultivars tested, nine carry *bc-1*^{2} gene alone, one was heterogenic for *bc-3*+*bc-1*^{2}, two carry *bc-3* alone, and two did not react with any of the markers tested (Table 5.5). SBD5 was amplified in Kabale, Sadzu, Kalambo, Kabulangeti, Lyambai, Lungwebungu, Mbereshi, Chambeshi, Kalungu and ZM3292 indicating that the *bc-1*^{2} marker is widely distributed in these genotypes.

The resistance gene *bc-3* was detected in genotypes Lunga and Lwangeni, with Lunga being a heterogenic (*bc-3* and *bc-1*^{2}) carrier of this gene. None of the 14 genotypes carried the *I*-gene based on non-amplification with the SW13 marker; all the three markers tested were not found in Kapisha and Lukupa (Table 5.5). Data for the 14 genotypes was checked against the positive controls.

5.4.4 Reactions of thirteen common bean varieties and one accession to infections

Fourteen common bean genotypes were challenged via mechanical inoculation with isolates representing the four BCMNV/BCMV pathogroups (PGI = EP-30, PGIII = EP-3, PGVIa = SC-36, and PGVIb = SC-14) identified in this study. As expected, all the 13 varieties and one accession were resistant to EP-30 (PGI). Isolate EP-3 (PGIII) did not induce symptoms in any of the genotypes carrying *bc-1*^{2} gene but necrotic symptoms in differential cultivar Black

Table 5.5. Resistance gene profiles of common bean varieties from Zambia based on screening with published gene markers listed in Table 5.2.

	Markers for different genes in common bean			
Variety	SBD5 (*bc-1*^{2})	ROC11 (*bc-3*)	SG6 (*bc-3*)	SW13 (*I*)
Lunga	+	-	+^{a}	-
Kabale	+	+	+^{b}	-
Sadzu	+	+	+^{b}	-
Kalambo	+	+	+^{b}	-
Kapisha	-	+	+^{b}	-
Kabulangeti	+	+	+^{b}	-
Lukupa	-	+	+^{b}	-
Lyambai	+	+	+^{b}	-
Kalungu	+	+	+^{b}	-
Lungwebungu	+	+	+^{b}	-
Mbereshi	+	+	+^{b}	-
Chambeshi	+	+	+^{b}	-
Lwangeni	-	-	+^{a}	-
ZM3292	+	+	+^{b}	-
F_3-bc-3^{kk}	+	-	+^{a}	-
F_3-I/bc-3^{kk}	+	-	+^{a}	+
Amanda$^+$*	+	-	-	+
DW$^-$*	-	-	-	-

^{kk}=Common bean lines sourced from University of Zambia, courtesy of Dr. Kelvin Kamfwa
$^+$*=Positive control for the *I* gene
$^-$*=Negative control for all genes.
F3-bc-3^{kk} = positive control for bc-3 gene
F3-I/bc-3^{kk} = positive control for *I* gene
+^{a}= Presence of bc-3 gene designated by a band size of 134 bp
+^{b}= Absence of bc-3 gene designated by a band size of 137 bp
Amanda=Differential cultivar (Resistance gene =*I*+*bc-1*^{2})
DW=Dubbele Witte (Resistance gene= *0*)

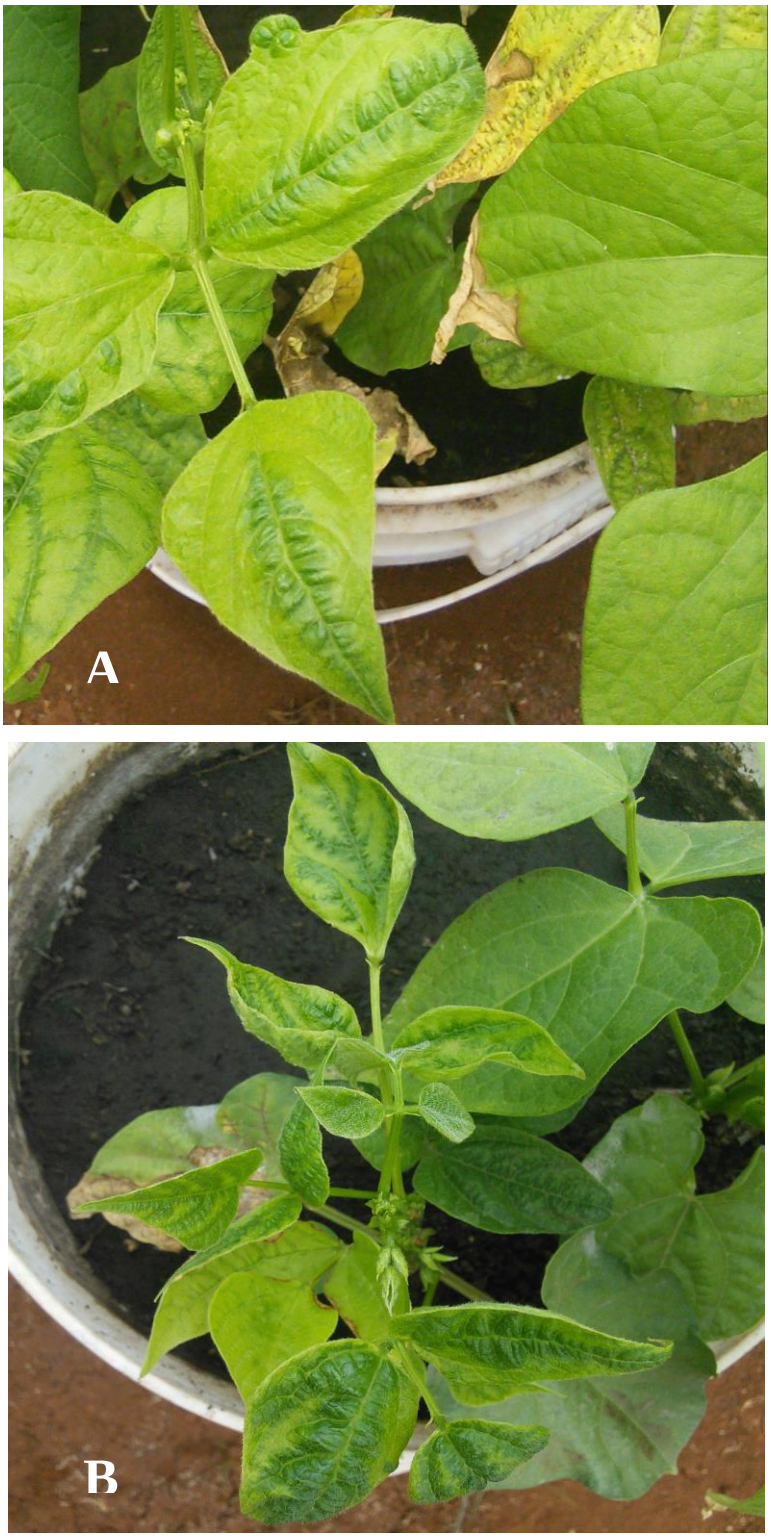


Figure 5.3A. Mosaic symptoms induced by bean common mosaic necrosis virus (BCMNV) isolate belonging to pathogroup (PG) VI on Kabulangeti. **B**. Mosaic symptoms induced by BCMNV isolate PGIII on improved common bean variety.

Table 5.6. Reactions of 14 common bean genotypes to infections by bean common mosaic necrosis and bean common mosaic virus (BCMNV and BCMV) isolates belonging to different pathogroups.

		BCMV	BCMNV		
		Pathogroup			
Variety	Resistance gene	PGI	PGIII	PGVIa	PGVIb
Lunga	*bc-1^2,bc-3*	R	R	R	R
Kabale	*bc-1^2*	R	R	S	S
	bc-1^2	R	R	S	S
Kalambo	*bc-1^2*	R	R	S	S
Kapisha	*un*	S	S	S	S
Kabulangeti	*bc-1^2*	R	R	S	S
Lukupa	*un*	R	R	S	S
Lyambai	*bc-1^2*	R	R	S	S
Kalungu	*bc-1^2*	R	R	S	S
Lungwebungu	*bc-1^2*	R	R	S	S
Mbereshi	*bc-1^2*	R	R	S	S
Chambeshi	*bc-1^2*	R	R	S	S
Lwangeni	*bc-3*	R	R	R	R
ZM3292	*bc-1^2*	R	R	S	S
F_3-bc-3^{kk}	*bc-1^2/bc-3*	R	R	R	R
F_3-I/bc-3^{kk}	*I/bc-1^2/bc-3*	R	R	R	R
BTS$^+$*	*I*	R	+n	+n	+n
DW$^-$*	*0*	S	S	S	S

un = unknown resistance gene
^{kk}=Common bean lines sourced from University of Zambia, courtesy of Dr. Kelvin Kamfwa and used as controls. $^+$*=Positive control for the *I* gene. $^-$*=Negative control for all genes.
F3-bc-3^{kk} = positive control for bc-3 gene. F3-I/bc-3^{kk} = positive control for I gene
BTS=Black turtle soup (Resistance gene =*I*)
DW=Dubbele Witte (Resistance gene= *0*)
R = host resistant to BCMNV or BCMV pathotypes
S= host susceptible to BCMNV or BCMV pathotypes +n = produces necrotic symptoms in the presence of the dominant *I* gene

Turtle Soup (BTS) and mosaic symptoms in differential cultivar Dubbele Witte (DW) (Fig. 5.2a) similar to what was reported for NL8 (PGIII) (Drijfhout, 1978; Spence and Walkey, 1995). Isolates SC-36 (PGVIa) and SC-14 (PGVIb) overcame resistance in the genotypes carrying *bc-1*^{2} gene, inducing systemic mosaic symptoms in the trifoliate leaves of these genotypes (Fig. 5.3a,b); consistent with results obtained in previous studies (Drijfhout, 1978; Feng *et al.*, 2014; Feng *et al.*,2017). The common bean varieties Lwangeni (*bc-3*) and Lunga (*bc-1*^{2}*+bc-3*) were resistant to all four PGs identified in this study (Table 5.6). The varieties Kapisha and Lukupa that lack any of the resistance genes based on marker screenings (Table 5.5) produced different phenotypic responses when challenged with virus isolates from the four PGs. Whereas Kapisha was susceptible to all four PGs, Lukupa was resistant to PGI and PGIII suggesting that it could possess other resistance genes not evaluated in this study.

5.3 Discussion

The objectives of this study were to identify BCMNV and BCMV pathogroups, screen common bean genotypes for resistance to a collection of BCMNV and BCMV pathotypes and determine the molecular markers available within the gene pool of improved varieties and local cultivars. The study revealed the presence of two previously unreported PGs I and III of BCMV and BCMNV, respectively, besides PGs VIa and VIb also reported in previous studies in Zambia (Spence and Walkey, 1995; Lana, 1995). Strains of BCMNV viruses assigned to PGs VIa and VIb (TN1 and TN1a) exhibited pathogenic profiles typical of NL3 and NL5 (Feng *et al.*, 2017) and were reported in previous studies in Kenya (Omunyin *et al*., 1995; Mangeni *et al.*, 2014; Mutuku *et al*., 2018) and Tanzania (Mwaipopo *et al*., 2018). It would appear, therefore that the isolates belonging to PG VI subgroups are ubiquitous in Eastern and Southern African regions. The dispersal of the TN1 virus strains across Easten and Southern Africa could be attributed to many reasons, chief of them being the exchange of planting materials in the region. This is evident in the similarity of names of common bean

cultivars across countries. For instance, a purple coloured medium seed sized common bean cultivar called Kabulangeti in Zambia is called by similar names in Kenya, Malawi and Tanzania (Katungi *et al.,* 2009), thus suggesting common originality.

The responses of common bean genotypes to infection by BCMNV and BCMV has been a subject of extensive study covering different aspects (Drijfhout, 1978; Morales, 1989; Miklas *et al.,* 2000; Mukeshimana *et al.,* 2005; Feng *et al.,* 2014; Feng *et al.,* 2015). Two important factors are critical in these studies: the pathotypes of strains of the virus and the R genes harboured by the hosts. This study showed that majority of the popularly cultivated and released common bean varieties in Zambia carry the *bc-1*^{2} gene and only a few varieties have the *bc-3* or *bc-1*^{2}*/bc-3* gene combinations (Table 5.4). The predominance of the *bc-1*^{2} gene in common bean varieties in Zambia might be due to growing of cultivars of the Andean gene pool (ADP). The ADP cultivars predominantly carry the *bc-1*^{2} gene as illustrated in the study by Mangeni *et al*. (2014). It is also likely that over the years, breeders may have been preferentially selecting for the *bc-1*^{2} gene since it confers resistance to multiple PGs of BCMNV/BCMV (PGI, PGII, PGIII and PGV) (Drijfhout, 1978; Miklas *et al.,* 2000). Unfortunately, results from this study showed that BCMNV/BCMV isolates of PGVI are common in Zambia, which may explain why the commonly grown varieties are susceptible to BCMNV viruses. Hence, the results of this study underscored the need to determine the pathogen genetic diversity and biological properties of viruses of common bean so as to inform future breeding studies. It is known that cultivars or genotypes carrying the *bc-u*/*bc-3* and *I/bc-3* gene combinations provide resistance to all known strains of BCMNV and BCMV (Mukeshimana *et al.,* 2005). The R gene composition of Lunga and Lwangeni varieties (*bc-1*^{2}*/bc-3* and *bc-3*) may explain why the two varieties were asymptomatic in the field screening and subsequently after inoculation with isolates of the different PGs. Nevertheless, since the population compositions of viruses are always in a state of flux, leading to the

evolution of newer strains as documented for a recombinant BCMNV isolate exhibiting atypical pathogenicity profiles (Feng *et al.,* 2015), it is vital to frequently update the knowledge of diversity of viruses circulating in farmers' fields.

The current bean descriptor for Zambia lacks information on the combinations of molecular markers for resistance to BCMNV and BCMV carried by the different common bean varieties that have been released in Zambia. This study profiled the gene combinations in the 13 common bean improved varieties thus providing additional information available to breeders for marker assisted selection (MAS) and that *bc-u/bc-1*^{2} R gene combinations are common and the two *bc-u/bc-3* and *bc-u*/*bc-1*^{2}*/bc-3* R gene combinations probably have limited distribution within the common bean gene pool in Zambia. Generally, the *bc-u*/*bc-3* and *I/bc-3* heterogenic combinations are preferable since such R gene combinations offer the most effective resistance to hitherto known strains of BCMNV and BCMV. Thus, with the existence of the *bc-3* R gene within the bean gene pool in Zambia, breeders have the choice of either pyramiding the dominant necrosis inducing *I* gene and the *bc-3* or *bc-1*^{2}. Either combination would confer broader resistance to BCMNV and BCMV strains detected in this study and many others (Drijfhout, 1978; Miklas *et al.,* 2000). However, in the presence of an epistatic R gene such as *bc-3* the merits of pyramiding the two recessive genes (*bc-1*^{2}*/bc-3*) are difficult to appreciate at the phenotypic level considering that *bc-3* masks *bc-1*^{2} (Kelly *et al.,* 1995). Thus the presence of the *bc-1*^{2}*/bc-3* gene combination found in Lunga does not seem to offer discernable phenotypic merits over others.

Results from this study revealed that Lunga and Lwangeni could be deployed in BCMNV/BCMV hotspots in Zambia since they carry an effective R gene combination compared with other varieties. In chapter three it was shown that Kabulangeti is the most widely cultivated variety in Zambia largely driven by consumer preferences (Sichilima *et al.,*

2016). However, as shown in this study, it is susceptible to BCMNV/BCMV isolates of the most abundant PGVI. There are currently breeding efforts underway at the University of Zambia, Department of Agriculture to pyramid dominant *I* gene and *bc-3* recessive gene into Kabulangeti variety and successful transfer of these genes has been proven in the F_3 lines. This would improve resistance of the widely grown Kabulangeti variety to BCMNV and BCMV and hence increase productivity of common bean in Zambia. Evidently, PGVI isolates of BCMNV/BCMV are common across East and Southern Africa. To combat the deleterious effects of constituent viruses of this PG, legume breeders and virologists could partner in regional breeding research programs through the Southern African Bean Research Netwrok (SABRN) so that superior varieties could be developed for use across the region.

CHAPTER SIX

GENERAL DISCUSSION, CONCLUSION AND RECOMMENDATIONS

6.1 General discussion

A survey of 128 common bean farmers's fields in six provinces of Zambia revealed the widespread presence of virus-like diseases. Results of highthroughput sequencing (HTS), RT-PCR screening of field collected samples and Sanger sequencing resulted in the detection of 10 viruses belong to six genera. The distribution of viruses in the total sample population revealed that collectively, viruses other than BCMNV and BCMV were more prevalent. Such results indicate that combined infections caused by viruses other than BCMNV and BCMV dominate the virome of common bean fields. An intriguing salient finding of this study is that southern bean mosaic virus (SBMV), which hitherto was not reported in Zambia from previous studies, is the most prevalent virus in the country. Considering the similarities in common bean foliar symptoms due to different viruses, it is therefore likely that most of the symptoms attributed previously to BCMNV or BCMV in Zambia may be due to SBMV. This underscores the need to complement symptoms-based field assessments of virus diseases with laboratory diagnosis for the definitive identification of the associated viruses.

Distribution of viruses in the two AEZs reveals the preponderance of common bean-infecting viruses in AEZII compared with AEZIII. The broader diversity of viruses in AEZII is presumably a factor of the influence of ecological systems on factors that promote virus dispersal. Among these factors is the availability of viruliferous arthropod vectors. Admittedly, this study did not include analysis of population dynamics of arthropod vectors and vector activity in the transmission of viruses in common bean fields across the two AEZs. However, considering that there is a dearth of information on barriers to trans-ecological

conveyance of seed, it is plausible to hypothesize that there is limited vector activity in AEZIII leading to asymmetric distribution of the viruses.

Metagenomics analysis of samples collected across six provinces of Zambia revealed a complex virome similar to what was reported in Tanzania (Mwaipopo *et al.,* 2018). The emerging results point to the occurrence of diverse viruses in farmers' fields in East and Southern Africa, which could be contributing to the suboptimal yields experienced among resource poor farmers in both regions. Regardless, BCMNV and BCMV are regarded as the most important yield limiting viruses of common bean in Africa. The importance of the two viruses stems from their ability to cause yield losses of up to 100% (Damayanti *et al.,* 2008; Saqib *et al.,* 2010; Mangeni *et al.,* 2020) and their seed transmission efficiencies of up to 30% (Morales, 1989). Moreover, the two viruses occur as a complex of strains whose evolution is driven by recombination (Karasev and Gray, 2013; Kehoe *et al.,* 2014).

Given the diversity of common bean-infecting viruses detected in this study, the use of HTS as a method of understanding the virosphere of common bean fields as opposed to culture depended methods such as ELISA and PCR was justified. However, the major hindrance to the universal adoption of HTS methods especially in Africa is the complex bioinformatics software and analyses that are involved. Thus, where use of HTS in plant virus detection is constrained by financial resources, primers developed in this study and those in the study by Mwaipopo *et al*. (2018) will be a reliable resource for detecting viruses reported in the two studies.

Data presented in chapters three and four show that the BCMNV and BCMV isolates from Zambia are phylogenetically related to known pathogroups of the two viruses especially those reported in East Africa. For example, isolates BCMV-EP95 clustered with BCMV isolate NL7n (KY057338), BCMV-EP50 with isolate NY15p (KT175568) and BCMV-EP20 with

isolate 1755a (KT175570) all of which are assigned to PGVI. Similarly, the BCMNV isolates from Zambia were phylogenetically related to isolates TN1a, TN1 and 1755b all of which have been shown to display pathogenicity profiles of PGVI members. Thus, the preponderance of PGVI isolates reported in chapter five is consistent with data in the other chapters and further shows the transregional dispersal of BCMNV and BCMV viruses in this PG across Eastern and Southern Africa (Mwaipopo *et al.,* 2018). It would be expected that based on available data on the distribution of the PGs PGVIa and PGVIb in Eastern and Southern Africa (Vetten and Allen, 1991; Spence and Walkey, 1995), the legume breeders in the National Agriculture Research Stations (NARS) of Zambia would have included the PGs in their breeding programmes. Apart from the current effort being conducted at the University of Zambia to introgress resistance genes in Kabulangeti variety, no literature on similar past work was available.

The deployment of resistant genotypes is the most reliable strategy for managing plant virus diseases. The *I*-gene and other resistance genes for BCMNV and BCMV have long been identified and used in MAS breeding strategies. The major challenge in Africa is that BCMNV is endemic especially in East Africa. Therefore, the *I*-gene is not an option for common bean varieties destined for tropical Africa. Consequently, the International Center for Tropical Agriculture (CIAT) introgressed the *bc-3* resistance gene into common bean genotypes suitable for Africa and, where possible, took advantage of the epistatic properties of the *bc-3* gene to mask the phenotypic effects of *I*-gene in genotypes possessing the *I/bc-3* gene combination. The bred cultivars were shared through the Pan African Bean Research Alliance (PABRA) and its networks such as the Southern Africa Bean Research Network (SABRN) that also passed on the materials to the National Agriculture Research programs. It is surprising therefore, that none of the profiled resistance genes in the 14 common bean cultivars studied possessed the *I/bc-3* gene combination. As far as this study shows, the most

prevalent R gene in common bean varieties grown in Zambia is the *bc-1*^{2}. Further, the majority of farmers use landraces (local cultivars) that may not possess the resistance genes against the PGVI isolates. Thus, the yield potentials of commonly grown common bean varieties in Zambia may be difficult to realize, resulting in recurrent shortages of the crop on the Zambian market.

6.2 Conclusion

In conclusion, this study provides molecular evidence of the existence of BCMNV, BCMV, CABMV, ETBTV, CMV, SBMV, PvEV-1, PvEV-2, CPPV1-Pv and PeMoV in Zambia. The discovery of 10 viruses and the generated sequence data will help in designing diagnostic tests for detecting the different viruses. These useful tools are important in monitoring the spread of detected viruses in common bean fields, seedlots and the quick detection of new variants.

It is clear from this study that although the released common bean varieties have adequate resistance to BCMNV and BCMV strains belonging to PGs I, II, III and V, they are susceptible to isolates of PGVI, which are the most prevalent in Zambian common bean fields as shown in chapters three and four. This situation is compounded by the occurrence of SBMV for which there could be no known resistance.

6.3 Recommendations

Given the information generated in this study, the following are the recommendations:

i. Farmers in Eastern Province should be encouraged to grow Lwangeni and Lunga varieties that have resistance to BCMV and BCMNV.
ii. Common bean breeders in Zambia should introduce common bean varieties with known resistance to SBMV and such materials should be promoted across the country.

iii. There is need to understand epidemiological factors limiting the distribution of BCMV and BCMNV across Zambia and use the lessons learnt to manage common bean viral diseases in the farmers' fields.

iv. Given the prevalence of SBMV in common bean fields in Zambia, future studies should focus on assessing its impact on yield reduction of popular common bean cultivars in Zambia and possibly identifying resistance traits in both local cultivars and improved varieties.

v. Yield reduction due to the tripartite infection of ETBTV, SatRNA and helper virus should be investigated in future studies so as to understand its contribution to yield reduction

REFERENCES

Abraham, A. D., Menzel, W., Bekele, B., and Winter, S. (2014). A novel combination of a new umbravirus, a new satellite RNA and potato leafroll virus causes tobacco bushy top disease in Ethiopia. Archives of Virology, 159(12), 3395–3399.

Adams, M. J., Antoniw, J. F., and Beaudoin, F. (2005). Overview and analysis of the polyprotein cleavage sites in the family Potyviridae. Molecular Plant Pathology, 6(4), 471–478.

Akhavan, A., Bahar, M., Askarian, H., Lak, M. R., Nazemi, A., and Zamani, Z. (2013). Bean common bacterial blight: pathogen epiphytic life and effect of irrigation practices. Springer Plus, 2, 41. doi.org/10.1186/2193-1801-2-41.

Alabi, O. J., Kumar, P. L. and Naidu, R. A. (2008). Multiplex PCR for the detection of African cassava mosaic virus and East African cassava mosaic Cameroon virus in cassava. Journal of Virological Methods, 154(1-2), 111–120.

Alicai, T., Ndunguru, J., Sseruwagi, P., Tairo, F., Okao-Okuja, G., Nanvubya, R., Kiiza, L., Kubatko, L., Kehoe, M. A., and Boykin, L. M. (2016). Cassava brown streak virus has a rapidly evolving genome: implications for virus speciation, variability, diagnosis and host resistance. Scientific Reports, 6, 36164. doi.org/10.1038/srep36164.

Ali Rezaian, M., Williams, R. H. V., and Symons, R. H. (1985). Nucleotide sequence of cucumber mosaic virus RNA1. Presence of a sequence complementary to part of the vira; satellite RNA and homologies with other viral RNAs. European Journal of Biochemistry, 150, 331-339.

Aloyce, R. C., Tairo, F., Ndunguru, J., Sseruwagi, P., and Rey, M. E. C. (2013). A single-tube duplex and multiplex PCR for simultaneous detection of four cassava mosaic begomovirus species in cassava plants. Journal of Virology Methods, 189(1), 148-156.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. Journal of Molecular Biology, 215(3), 403-410.

Ampofo, J. K., and Massomo, S. M. (1998). Some Cultural Strategies for Management of Bean Stem Maggots (*Diptera: Agromyzidae*) on Beans in Tanzania. African Crop Science Journal, 6(4), 351-356.

Andret-Link, P., and Fuchs, M. (2005). Transmission specificity of plant viruses by vectors. Journal of Plant Pathology, 87(3), 153–165.

Arli-Sokmen, M., Deligoz, I., and Kutluk-Yilmaz, N. D. (2016). Characterization of Bean common mosaic virus and Bean common mosaic necrosis virus isolates in common bean growing areas in Turkey. European Journal of Plant Pathology, 146(1), 1-16.

Assefa, A., Tana, T., Dechassa, N., Dessalgn, Y., Tesfaye, K., and Wortmann, C. S. (2016). Maize-common bean/lupine intercrop productivity and profitability in maize-based cropping system of Northwestern Ethiopia. Ethiopian Journal of Science and Technology, 9(2), 69-85. doi.org/10.4314/ejst.v9i2.1.

Azizi, A., and Shams-bakhsh, M. (2014). Impact of cucumber mosaic virus infection on the varietal traits of common bean cultivars in Iran. Virus Disease, 25(4), 447–454. doi.org/10.1007/s13337-014-0233-9.

Bashir, M. Ahmad, Z., and Ghafoor, A. (2002). Cowpea aphid-borne mosaic potyvirus: A review, International Journal of Pest Management, 48(2), 155-168. doi.org/10.1080/09670870110118722.

Beebe, S. E., Rao, I. M., Devi, M. J., and Polania, J. (2014). Common beans, biodiversity, and multiple stresses: Challenges of drought resistance in tropical soils. Crop and Pasture Science, 65(7), 667-675. doi.org/10.1071/cp13303.

Beebe, S., Rao, I., Mukankusi, C., and Buruchara, R. (2013). Improving resource use efficiency and reducing risk of common bean production in Africa, Latin America and

the Caribbean. In C. H. Hershey and P. Neate (Ed.). Eco-efficiency, from vision to reality (pp. 117–134): Cali, Colombia: International Center of Tropical Agriculture.

Bellucci, E., Bitocchi, E., Ferrarini, A., Benazzo, A., Biagetti, E., and Klie, S. (2014). Decreased nucleotide and expression diversity and modified coexpression patterns characterize domestication in the common bean. Plant Cell, 26, 1901–12.

Berger, P. H., Wyatt, S. D., Shiel, P. J., Silbernagel, M. J., Druffel, K., and Mink, G. I. (1997). Phylogenetic analysis of the *Potyviridae* with emphasis on legume-infecting potyviruses, Archives of Virology, 142, 1979-1999.

Bitocchi, E., Bellucci, E., Giardini, A., Rau, D., Rodriguez, M., and Biagetti, E. (2013). Molecular analysis of the parallel domestication of the common bean (*Phaseolus vulgaris* L.) in Mesoamerica and the Andes. New Phytology, 197(1), 300–13.

Bitocchi, E., Nanni, L., Bellucci, E., Rossi, M., Giardini, A., and Zeuli, P. S. (2012). Mesoamerican origin of the common bean (*Phaseolus vulgaris* L.) is revealed by sequence data. Proceedings of the National Academy of Science 109, E788–E796.

Blair, M., Gonzales, L. F., Kimani, P. M., and Butare, L. (2010). Genetic diversity, inter-gene pool introgression and nutritional quality of common beans (*Phaseolus vulgaris* L.) from Central Africa. Theoretical and Applied Genetics, 121(2), 237–248.

Boonham, N., Kreuze , J., Winter, S., René van der Vlugt, R., Bergervoet, J., Jenny Tomlinson, J., and Mumford, R. (2014). Methods in Virus Diagnostics: From ELISA to Next Generation Sequencing. Virus Research, 186, 20-31.

Boudreau, A. M. (2013). Diseases in intercropping systems. Annual Review of Phytopathology, 51, 499-519.

Broughton, W. J., Hernández, G., Blair, M., Beebe, S., Gepts, P., and Vanderleyden, J. (2003). Beans (*Phaseolus* spp.) model food legumes. Plant and Soil, 252(1), 55–128. doi.org/10.1023/a:1024146710611

Buruchara, R., Mukankusi, C., and Ampofo, K. (2010). Bean disease and pest identification and management. Handbooks for small-scale seed producers. CIAT Publication no. 371. Kampala, Uganda: CIAT. P. 67.

Celmeli, T., Sari, H., Canci, H., Sari, D., Adak, A., Eker, T., and Toker, C. (2018). The nutritional content of common bean (*Phaseolus vulgaris* L.) landraces in comparison to modern varieties. Agronomy, 8(9), 166-175.

Cervera, M. T., Riechmann, J. L., Martin, M. T., and Garcia, J. A. (1993). 3` terminal sequence of the plum pox virus PS and 06 isolates. Evidence for RNA recombination within the potyvirus group. Journal of General Virology, 74(3), 329-334.

Chacón, S. M. I., Pickersgill, B., and Debouck, D. G. (2005). Domestication patterns in common bean (*Phaseolus vulgaris* L.) and the origin of the Mesoamerican and Andean cultivated races. Theoretical and Applied Genetics, 110, 432–44.

Chang, S., Puryear, J., and Cairney, J. (1993). A simple and efficient method for isolating RNA from pine trees. Plant Molecular Biology Reporter, 11(2), 113–116.

Chen, J., Chen, J., and Adams, M. J. (2001). A universal primer to detect members of Potyviridae and its use to examine taxonomic status of several members of the family. Archives of Virology, 146(2), 757–766.

Chomič, A., Pearson, M. N., Clover, G. R. G., Farreyrol, K., Saul, D., Hampton, G. J., and Armstrong, K. F. (2010). A generic RT-PCR assay for the detection of Luteoviridae. Plant Pathology, 59 (3), 429-442.

Chowdhury, S. R., and Savithri, H. S. (2011). Interaction of sesbania mosaic virus movement protein with VPg and P10: Implication to specificity of genome recognition. PLoS ONE, 6(1), e15609. doi.org/10.1371/journal.pone.0015609.

Chung, B. Y. W., Miller, W. A., Atkins, J. F., and Firth, A. E. (2008). An overlapping essential gene in the Potyviridae. Proceedings of the National Academy of Sciences of the United States of America, 105, 5897–5902.

Cichy, K. A., Porch, T. G., Beaver, J. S., Cregan, P., Fourie, D., Glahn, R. P., Michael, A. G., Kamfwa, K., Katuuramu, D. N., McClean, P., Mndolwa, E., Nchimbi-Msolla, S., Marcial, A., Pastor-Corrales, Miklas, P. N. (2015). A Diversity Panel for Andean Bean Improvement. Crop Science, 55(5), 2149. doi.org/10.2135/cropsci2014.09.0653.

Clark, M. F., AND Adams, A. N. (1977). Characteristics of the microplate method of Enzyme-Linked Immunosorbent Assay for the detection of plant viruses. Journal of General Virology, 34, 475-483.

Damayanti, T., Susilo, D., Nurlaelah, S., Sartiami, D., Okuno, T., and Mise, K. (2008). First report of Bean common mosaic virus in yam bean [Pachyrhizus erosus (L.) Urban] in Indonesia. Journal of General Plant Pathology, 74(6), 438–442.

Damiri, B. V., Abdalla, O. A., and Amer, M. A. (2013). Identification and characterization of Cowpea aphid-borne mosaic virus isolates in Saudi Arabia. Journal of Plant Pathology, 95(1), 79-85.

David, S., and Sperling, L. (1999). Improving technology delivery mechanisms: lessons from bean seed systems research in eastern and central Africa. Agriculture and Human Values, 16(4), 381–388.

Deng, D., Mc Grath, P. F., Robinson, D. J., and Harrison, B. D. (1994). Detection and differentiation of whitefly-transmitted geminiviruses in plants and vector insects by the polymerase chain reaction with degenerate primers. Annals of Applied Biology, 125, 327-336.

Desbiez, C., and Lecoq, (2004). The nucleotide sequence of Watermelon mosaic virus (WMV, Potyvirus) reveals interspecific recombination between two related

potyviruses in the 5' part of the genome Archives of Virology, 149, 1619–1632. doi.org/10.1007/s00705-004-0340-9.

Dickinson, M. (2003). Molecular plant pathology. London, England: Taylor and Francis Group, BIOS Scientific Publishers.

Ding, S. W., Anderson, B. J., Haase, H. R., and Symons, R. H. (1994). New overlapping gene encoded by the cucumber mosaic virus genome. Virology, 198(2), 593–601.

Domingo, E., and Holland, J. J. (1997). RNA virus mutations and fitness for survival. Annual Review of Microbiology, 51(1), 151–8. doi.org/10.1146/annurev.micro.51.1.151.

Drijfhout, E. (1978). Genetic interaction between *Phaseolus vulgaris* and Bean common mosaic virus with implications for strain identification and breeding for resistance. PhD Thesis, Wageningen University.

Drijfhout, E., Silbernagel, M. J., and Burke, D. W. (1978). Differentiation of strains of bean common mosaic virus. Netherlands Journal of Plant Pathology, 84, 13-26.

Dwyer, G. I., Njeru, R., Williamson, S., Fosu-Nyarko, J., Hopkins, R., Jones, R. A., Waterhouse, P. M., and Jones, M. G. (2003). The complete nucleotide sequence of Subterranean clover mottle virus. Archives of Virology, 148(11), 2237–2247.

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research, 32(5), 1792-1797.

Edwardson, J. R., and Christie, R. G. **(**1991). Cucumoviruses. CRC Handbook of viruses infecting legumes. pp 293–319., Boca Raton: CRC Press.

Fang, G. W., Allison, R. F., Zambolim, E. M., Maxwell, D. P., and Gilbertson, R. L. (1995). The complete nucleotide sequence and genome organization of bean common mosaic virus (NL3 strain). Virus Research, 39(1), 13-23.

FAOSTAT, (2018). Online publication. Food and Agriculture Organization of the United Nations. http://www.fao.org/faostat/en/#data/QC. Accessed on 20th January, 2020.

FAO, (2019). FAOSTAT Crops. Online publication. Food and Agriculture Organization of the United Nations. http://www.fao.org/faostat/en/#data/QC

Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U., and Ball, L. A. (2005). Virus taxonomy: classification and nomenclature of viruses. Eighth Report of the International Committee on Taxonomy of Viruses. ICTV 2, 64. doi.org/10.1186/1743-422X-2-64.

Feng, X., Guzm´an, P., Myers, J. R., and Karasev, A. V. (2017). Resistance to Bean common mosaic necrosis virus conferred by the bc-1 gene affects systemic spread of the virus in common bean. Phytopathology, 107(7), 893-900.

Feng, X., Myers, J. R., and Karasev, A. V. (2015). Bean common mosaic virus isolate exhibits a novel pathogenicity profile in common bean, overcoming the *bc-3* resistance allele coding for the mutated eIF4E translation initiation factor. Phytopathology, 105(11), 1487-1495.

Feng, X., Poplawsky, A. R., and Karasev, A. V. (2014). A recombinant of bean common mosaic virus induces temperature-insensitive necrosis in an *I* gene-bearing line of common bean. Phytopathology, 104(11), 1251-1257.

Flores-Este´vez, N., Acosta-Gallegos, J. A., and Silva-Rosales, L. (2003). Bean common mosaic virus and Bean common mosaic necrosis virus in Mexico. Plant Disease, 87(1), 21–25.

Garcia-Arenal, F., Aurora, F., and Malpica, J. M. (2001). Variability and genetic structure of plant virus populations. Annual Review of Phytopathology, 39(1), 157-186.

Garcıa-Arenal, F., Palukaitis, P., Mahy, B. W. J., and Van Regenmortel, M. H. V. (2008). Encyclopedia of Virology (3rd ed. pp. 614–9). Oxford: Academic Press.

Gentry, H. S. (1969). Origin of the common bean, *Phaseolus vulgaris.* Economic Botany, 23 (1), 55-69. doi.org/10.1007/BF02862972.

Ghorbani, S. G. M., Shahraeena, N., and Elahinia, S. A. (2010). Distribution and impact of virus associated diseases of common bean (*Phaseolus vulgaris* L.) in northern Iran. Archives of Phytopathology and Plant Protection, 43(12), 1183-1189. doi.org/10.1080/03235400802366834

Gibbs, M. J., Ziegler, A., Robinson, D. J., Waterhouse, P. M., and Cooper, J. I. (1996). Carrot mottle mimic virus (CMoMV): a second umbravirus associated with carrot motley dwarf disease recognized by nucleic acid hybridization. Molecular Plant Pathology. On-line. http://bspp. org.uk/mppol/1996/1111gibbs.

Givord, L. (1981). Southern bean mosaic virus isolated from cowpea (*Vigna anguiculata*) in Ivory Coast. Plant Disease, 65(9), 755-756

Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinov, V. M. (1989). Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. Nucleic Acids Research, 17(12), 4713–30.

Graham, P. H., and Vance, C. P. (2003). Legumes: Importance and constraints to greater use, Plant Physiology, 131 (3), 872-877. http://dx.doi.org/10.1104/pp.017004.

Gutiérrez, P. A., Jaramillo, H., and Marín, M. (2016). Next generation sequence analysis of the forage peanut (*Arachis pintoi*) virome. Revista Facultad Nacional de Agronomía Medellín, 69(2), 7881-7891. doi: 10.15446/rfna.v69n2.59133.

Ha, C. (2007). Detection and identification of potyviruses and geminiviruses in Vietnam. PhD Thesis. Queensland University of Technology.

Hacker, D. L., and Fowler, B. C. (2000). Complementation of the Host Range Restriction of Southern Cowpea mosaic virus in bean by Southern bean mosaic virus. Virology, 266(1), 140-149.

Hagedorn, D. J., and Inglis, D. A. (1986). Handbook of Bean Diseases. University of Wisconsin Cooperative Extension Publication, Madison. USA.

Haley, S. D., Afanador, L., and Kelly, J. D. (1994). Identification and application of a random amplified polymorphic DNA marker for the *I*gene (potyvirus resistance) in common bean. Phytopathology, 84(2), 157–160.

Hall, T. A. (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series, (41), 95-98.

Hamazakaza, P., Katungi. E., Reyes, B., Maredia, M., Muimui, K., and Ojara, L. (2014). Assessing access and adoption of common bean improved varieties in Zambia. Research Technical Report, Zambia Agriculture Research Institute.

Hanley, K. A., and Weaver, S. C. (2008). Arbovirus evolution. Origin and evolution of viruses (2nd ed.) Academic Press. doi.org/10.1016/b978-0-12-374153-0.00016-3.

Hegay, S., Ortiz, R., Garkava-Gustavsson, L., Persson Hovmalm, H., and Geleta, M. (2013). Marker-aided breeding for resistance to bean common mosaic virus in Kyrgyz bean cultivars. Euphytica. doi.org/10.1007/s10681-013-0928-9.

Helmy, M., Awad, M., and Mosa, K. A. (2016). Limited resources of genome sequencing in developing countries: Challenges and solutions. Applied Translational Genomics, 9, 15-19. doi.org/10.1016/j.atg.2016.03.003.

Harrison, B. D., Swanson, M. M., and Fargette, D. (2002). Begomovirus coat protein: serology, variations and functions. Physiological and Molecular Plant Pathology, 60(5), 257–271.

Hichaambwa, M., Beaver, M., Chapoto, A., and Weber, M. (2009). Patterns of Urban Food Consumption and Expenditure in Zambia. An Overview Report Based on the CSO/MACO/FSRP Food Consumption Survey in Urban Areas of Lusaka, Kitwe, Mansa and Kasama.

Hillocks, R. J., Madata, C. S., Chirwa, R., Minja, E. M., and Msolla, S. (2006). Phaseolus bean improvement in Tanzania. Euphytica, 150(1), 225-231.

Hotz, C., and McClafferty, B. (2007). From harvest to health: challenges for developing biofortified staple foods and determining their impact on micronutrient status. Food and Nutrition Bulletin, 28 (2 Supplmenta), S271–S279.

Hudson, D. H., and Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. Molecular Biology Evolution, 23(2), 54-267.

Hull, R. (1995). Sobemovirus. InVirus taxonomy. Classification and nomenclature of viruses. (F. A. Murphy, C. M. Fauquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Jarvis, G. P. Martelli, M. A. Mayo, and M. D. Summers Ed.). Sixth report of the International Committee on Taxonomy of Viruses. pp. 376–378. Vienna, Austria: Springer-Verlag.

Hull, R. (2002). Plant Virology. (4thEd). New York: Academic Press.

Hull, R. (2014). Plant virology (5th ed.). New York: Academic Press.

Ivanov, K. I., Eskelin, K., Lõhmus, A., and Mäkinen, K. (2014). Molecular and cellular mechanisms underlying potyvirus infection. Journal of General Virology, 95(7), 1415–1429.

Jacquemond, M. (2012). Cucumber mosaic virus. Advances in Virus Research, 84, 439 –504. doi.org/10.1016/B978-0-12-394314-9.00013-0.

Jiang, G. L. (2013). Molecular markers and marker assisted breeding in plants. InTech. https://www.intechopen.com/predownload/40178. doi.org/10.5772/52583.

Johansen, E., Edwards, M. C., Hampton, R. O. (1994). Seed transmission of viruses: current perspectives. Annu Rev Phytopathol. 32, 363-386.

Johansen, I. E., Keller, K. E., Dougherty, W. G., and Hampton, R. O. (1996). Biological and molecular properties of a pathotype P-1 and a pathotype P-4 isolate of pea seed-borne

mosaic virus. Journal of General Virology, 77 (6), 1329–1333. doi.org/10.1099/0022-1317-77-6-1329.

Johary, T., Dizadji, A., and Naderpour, M. (2016). Biological and Molecular characteristics of Bean common mosaic virus isolates circulating in common bean in Iran. Journal of Plant Pathology, 98 (2), 301-310.

Johnson, W. C., Guzman, P., Mandala, D., Mkandawire, A. B. C., Temple, S., Gilbertson, R. L., and Gepts, P. (1997). Molecular tagging of the bc-3 gene for introgression into Andean common bean. Crop Science, 37, 248-254.

Kaitisha, G. C. (2003). Some virus diseases of crop plants in Zambia. In J. d'A. Hughes and B. O. Odu (ed. pp 317-333.). Proceedings in Plant Virology Sub-Saharan Africa. International Institute of Tropical Agriculture, Ibadan, Nigeria.

Kannaiyan, J., and Hachiwa, H. C. (1993). Diseases of food legume crops and the scope for their management in Zambia. FAO Plant Protection Bulletin, 41(2), 73–90.

Karasev, A. V., and Gray, S. M. (2013). Genetic diversity of potato virus Y complex. American Journal of Potato Research, 90, 7-13. doi.org/10.1007/s12230-012-9287-7.

Karkashian, J., Ramos-Reynoso, E. D., Maxwell, D. P., and Ramírez, P. (2011). Begomoviruses associated with bean golden mosaic disease in Nicaragua. Plant Disease, 95(8), 901-906.

Kasschau, K. D., Xie, Z. X., Allen, E., Llave, C., Chapman, E. J., and Krizan, K. A. (2003). P1/HC-Pro, a viral suppressor of RNA silencing, interferes with Arabidopsis development and miRNA function. Developmental Cell, 4(2), 205–217.

Katoh, K., Misawa, K., Kuma, K., and Miyata, T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research, 30(14), 3059–66.

Katungi, E., Farrow, A., Chianu, J., Sperling, L., and Beebe, S. (2009). Common bean in Eastern and Southern Africa: a situation and outlook analysis. International Center for Tropical Agriculture (CIAT), 6, 1–61.

Kehoe, M. A., Coutts, B. A., Buirchell, B. J., and Jones, R. A. C. (2014). Split personality of a potyvirus: to specialize or not to specialize? PLOS ONE 9(8), e105770. doi.org/10.1371/journal.pone.0105770.

Kelly, J. D., Afanador, L., and Haley, S. D. (1995). Pyramiding genes for resistance to bean common mosaic virus. Euphytica, 82, 207-212.

Kim, M.-K., Seo, J.-K., Kwak, H.-R., Kim, J.-S., Kim, K.-H., Cha, B.-J., and Choi, H.-S. (2014). Molecular genetic analysis of Cucumber mosaic virus populations infecting pepper suggests unique patterns of evolution in Korea. Phytopathology, 104(9), 993-1000.

Kinoga, M. N., Kuria, P. K., Miano, D. W., Narla, R. D., and Wasilwa, L. A. (2021). First report of Ethiopian tobacco bushy top virus and its associated satellite RNA in mixed infection with Potato virus Y infecting Solanum betacea in Kenya. New Disease Reports, 1–2. doi.org/10.1002/ndr2.12015.

Koonin, E. V., and Dolja, V. V. (1993). Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. Critical Reviews in Biochemistry and Molecular Biology, 28(5), 375–430.

Kulkarni, H. Y., and Muguga, E. A. (1973). Notes on East African Plant virus diseases. Bean common mosaic virus. East African Agricultural Journal, 39, 72-76.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version7.0 for bigger datasets. Molecular Biology and Evolution, 3(7), 1870-1874.

Kushwaha, N., Singh, A. K., Chattopadhyay, B., and Chakraborty, S. (2010). Recent advances in geminivirus detection and future perspectives. Journal of Plant Protection Sciences, 2 (1), 1-18.

Lacombe, S., Bangratz, M., Vignols, F., and Brugidou, C. (2010). The rice yellow mottle virus P1 protein exhibits dual functions to suppress and activate gene silencing. Plant Journal, 6(3), 371–382.

Lain, S., Riechmann, J., Me´ndez, E., and Garcı´a, J. A. (1988). Nucleotide sequence of the 3`terminal region of plum pox potyvirus RNA. Virus Research, 10(4), 325–341.

Lana, A. F. (1995). Advances in the establishment of the Pan-African bean common mosaic virus disease resistance nurseries in eastern and southern Africa. InFourie, D. and Liebenberg, A.J. (ed.). Proceedings of the 4th SADC Regional Bean Workshop, Potchefstroom, edited by, South Africa, 2–4 October 1995. CIAT African Workshop Series no. 31. CIAT, Cali, Colombia.

Lamptey, P. N. L., and Hamilton, P. I. (1974). A new cowpea strain of southern bean mosaic virus from Ghana. Phytopathology, 64, 1100-1104.

Langmead, B., and Steven, S. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods, 9, 357–359 doi.org/10.1038/nmeth.1923

Larsen, R. C., Miklas, P. N., Druffel, K. L. and Wyatt, S. D. (2005). NL-3 K strain is a stable and naturally occurring interspecific recombinant derived from bean common mosaic necrosis virus and bean common mosaic virus. Phytopathology, 95(9), 1037-1042.

Larsen, R. C., Druffel, K. L., and Wyatt, S. D. (2011). The complete nucleotide sequences of bean common mosaic necrosis virus strains NL-5, NL8 and TN-1. Archives of Virology, 156(4), 729–732.

Lecoq, H., Fabre, F., Joannon, B, Wipf-Scheibel, C., Chandeysson, C., Schoeny, A., Desbiez, C. (2011). Search for factors involved in the rapid shift in watermelon mosaic virus (WMV) populations in south-eastern France. Virus Research, 159, 115–123.

Lee, L., and Anderson E. J. (1998). Nucleotide sequence of a resistance breaking mutant of southern bean mosaic virus. Archives of Virology, 143(11), 2189-2201. doi.org/10.1007/s007050050451.

Li, H. W., Lucy, A. P., Guo, H. S., Li, W. X., Ji, L. H., Wong, S. M., and Ding, S. W. (1999). Strong host resistance targeted against a viral suppressor of the plant gene defence silencing mechanism. EMBO Journal, 18(10), 2683–91.

Ling, R., Pate, A. E., Carr, J. P.; Firth, A. E. (2013). An essential fifth coding ORF in the sobemoviruses. Virology, 446(1-2), 397–408.

Loebenstein, G and Thottappilly G. (2004). Virus and virus-like diseases of major crops in developing countries. Dordrecht: Kluwer Academic Publishers.

Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., Ingersoll, R., Sheppard, H. W., and Ray, S. C. (1999). Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. Journal of Virology, 73(1), 152-160.

Luquet, E., Garner, T. W., Léna, J. P., Bruel, C., Joly, P., Lengagne, T., Grolet, O., Plénet, S. 2012. Genetic erosion in wild populations makes resistance to a pathogen more costly. Evolution, 66(6):1942-52. doi.org/10.1111/j.1558-5646.2011.01570.x.

Maliogka, V. I., Salvador, B., Carbonell, A., Saenz, P., Leon, D. S., and Oliveros, J. C. (2012). Virus variants with differences in the P1 protein coexist in a plum pox virus population and display particular host-dependent pathogenicity features. Molecular Plant Pathology, 13 (8), 877–886. doi.org/10.1111/j.1364-3703.2012.00796.x

Mamidi, S., Rossi, M., Annam, D., Moghaddam, S., Lee, R., and Papa, R. (2011). Investigation of the domestication of common bean (*Phaseolus vulgaris*) using multilocus sequence data. Functional Plant Biology, 38(12), 953–67.

Mandary, M. B., Masomian, M., and Poh, L. C. (2019). Impact of RNA virus evolution on quasispecies formation and virulence. International Journal of Molecular Sciences, 20 (18), 4657. doi.org/10.3390/ijms20184657.

Mangeni, B. C., Abang, M. M., Awale, H., Omuse, C. N., Leitch, R., Arinaitwe,W., Mukoye, B., Kelly, J .D., and Were, H. K. (2014). Distribution and pathogenic characterization of bean common mosaic virus (BCMV) and bean common mosaic necrosis virus (BCMNV) in western Kenya. Journal of Agriculture Food and Applied Science, 2(10), 308–316.

Mangeni, B. C., Were, H. K., Ndong'a, M., and Mukoye, B. (2020). Incidence and severity of bean common mosaic disease and resistance of popular bean cultivars to the disease in western Kenya. Journal of Phytopathology, 1–15. doi.org/10.1111/jph.12928.

Martin, D. P., Murrell, B., Golden, M., Khoosal, A., and Muhire, B. (2015). RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evolution, 1(1), 1-5.

Mbanzibwa, D. R., Tian, Y. P., Tugume, A. K., Patil, B. L., Yadav, J. S., Bagewadi, B., Abarshi, M. M., Alicai, T., Changade, W., Mkumbira, J., Muli, M. B., Mukasa, S. B., Tairo, F., Baguma, Y., Kyamanywa, S., Kullaya,A., Maruthi, M. N., Fauquet, C. M., and Valkonen, J. P. T. (2011). Evolution of cassava brown streak disease-associated viruses. Journal of General Virology, 92(4), 974-987.

McGuire, S., and Sperling, L. (2016). Seed systems smallholder farmers use. Food Security, 8, 179-195.

Melotto, M., Afanador, L., and Kelly, J. D. (1996). Development of a SCAR marker linked to the *I* gene in common bean. Genome, 39(6), 1216–1219.

Miklas, P.N., Larsen, R.C., Riley, R., and Kelly, J. D. (2000). Potential marker assisted selection for *bc-1*^{2} resistance to bean common mosaic potyvirus in common bean. Euphytica, 116, 211–219.

Mlotshwa, S., Verver, J., Sithole-Niang, I., Van Kampen, T., Van Kammen, A., and Wellink, J. (2002). The genomic sequence of cowpea aphid-borne mosaic virus and its similarities with other potyviruses. Archives of Virology, 147(5), 1043-1052. doi.org/10.1007/s00705-002-0800-z.

Mo, X. H., Chen, Z-B., and Chen, J. P. (2011). Molecular identification and phylogenetic analysis of a viral RNA associated with the Chinese tobacco bushy top disease complex. Annals of Applied Biology, 158, 188–193.

Mo, X. H., Qin, X. Y., Wu, J., Yang, C., Wu, J. Y., Duan, Y. Q., Li, T. F., and Chen, H. R. (2003). The complete nucleotide sequence and genome organization of a Chinese isolate of tobacco bushy top virus. Archives of Virology, 148(2), 389–397.

Moradi, Z., and Mehrvar, M. (2019). Genetic variability and molecular evolution of bean common mosaic virus populations in Iran: comparison with the populations in the world. European Journal of Plant Pathology, 154, 673-690.

Morales, F. J. (1989). Bean common mosaic: screening for disease resistance. Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia. 28p.

Morales, F. J., and Kornegay, J. (1996). The use of plant viruses as markers to detect genes for resistance to bean common mosaic and bean common mosaic necrosis viruses. Annual Report of the Bean Improvement Cooperative, 39, 272–273.

Morales, F., Niessen, A., Ramirez, B., Castano, M. (1990). Isolation and partial characterization of a geminivirus causing bean dwarf mosaic. Phytopathology, 80(1), 96-101.

Morales, F. J. (2006). Common beans. In G. Loebenstein and J. P. Carr (ed.). Natural Resistance Mechanisms of Plants to Viruses (pp367-382). Netherland: Springer.

Mucheru-Muna, M., Pypers, P., Mugendi, D., Kung'u, J., Mugwe, J., Merckx, R., and Vanlauwe, B. (2010). A staggered maize-legume intercrop arrangement robustly increases crop yields and economic returns in the highlands of Central Kenya. Field Crops Research, 115(2), 132–139.

Muhire, B. M., Varsani, A, and Martin, D. P. (2014). SDT: A Virus Classification Tool Based on Pairwise Sequence Alignment and Identity Calculation. PLoS ONE, 9(9), e108277. doi.org/10.1371/journal.pone.0108277.

Muimui, K. K. (2015). Zambia bean varieties descriptor. 2nd Edition.

Muimui, K. K., Chirwa, R. M., Zulu, R. M., and Maereka, E. K. (2016). Identifying acceptable improved beans varieties for income and nutrition in Zambia. Poster presentation at the joint Pan-African grain legume and world cowpea conference. 27 February-4 March, Livingstone, Zambia.

Mukeshimana, G., Paneda, A., Rodriguez-Suarez, C., Ferreira, J. J., Giraldez, R., and Kelly, J. D. (2005). Markers linked to the bc-3 gene conditioning resistance to bean common mosaic potyviruses in common bean. Euphytica, 144, 291–299.

Mulenga, R. M., Boykin, L. M., Chikoti, P.C, Sichilima, S., Ng'uni, D., and Alabi, O. J. (2018). Cassava Brown Streak Disease and Ugandan cassava brown streak virus Reported for the First Time in Zambia. Plant Disease, 102(7), 1410-1418. doi.org/10.1094/PDIS-11-17-1707-RE.

Murant, A. F., Rajeshwari, R., Robinson, D. J., and Raschke´, J. H. (1988). A satellite RNA of groundnut rosette virus that is largely responsible for symptoms of groundnut rosette disease. Journal of General Virology, 69(7), 1479–1486.

Murant, A. F., and Kumar, I. K. (1990). Different variants of the satellite RNA of groundnut rosette virus are responsible for the chlorotic and green forms of groundnut rosette disease. Annals of Applied Biology, 117(1), 85–92.

Murant, A.F. (1993). Complexes of transmission-dependent and helper viruses Diagnosis of Plant Virus Diseases. In R.E.F. Matthews (Ed.). Diagnosis of Plant Virus Diseases (pp. 333–357). Boca Raton: CRC Press.

Murphy, J. F., and Bowen, K. L. (2006). Synergistic disease in pepper caused by the mixed infection of Cucumber mosaic virus and Pepper mottle virus. Phytopathology, 96:240-247.

Mutuku, J. M., Wamonje, F. O., Mukeshimana, G., Njuguna, J., Wamalwa, M., Choi, S. K., Tungadi, T., Djikeng, A., Kelly, K., Entfellner, J. B. D., Ghimire, S. R., Mignouna, H. D., Carr J. P., and Harvey, J. J. W. (2018). Metagenomic Analysis of Plant Virus Occurrence in Common Bean (*Phaseolus vulgaris*) in Central Kenya. Frontiers in Microbiology, 9, 2939. doi.org/10.3389/fmicb.2018.02939.

Mwaipopo, B., Nchimbi-Msolla, S., Njau, Deogratius Mark., and Mbanzibwa, D. (2018). Comprehensive surveys of Bean common mosaic virus and Bean common mosaic necrosis virus and molecular evidence of occurrence of other Phaseolus vulgaris viruses in Tanzania. Plant Disease, 102 (11), 2361-2370.

Mwaipopo, B., Nchimbi-Msolla, S., Njau, P, Tairo, F., William, W., Binagwa, P., Kweka, E., Kilango, M., and Mbanzibwa, D. (2017). Viruses infecting common bean (*Phaseolus vulgaris* L.) in Tanzania: A review on molecular characterization, detection and disease management options. Africa Journal of Agricultural Research, 12 (18), 1486-1500.

Myers, J. R., Mink, G. A., and Mabagala, R. (2000). Surveys for bean common mosaic virus in East Africa. Annual Report of Bean Improvement Cooperative and National Dry Bean Council Research Conference, 13–14.

Naderpour, M., Lund, O. S., and Johansen, I. E. (2009). Sequence analysis of expressed cDNA of *Bean common* mosaic virus RU1 isolate. Iran Journal of Virology, 3, 41-43.

Nassary, E. K., Baijukya, F., and Ndakidemi, P. A. (2020). Productivity of intercropping with maize and common bean over five cropping seasons on smallholder farms of Tanzania. European Journal of Agronomy, 113, 125964. doi.org/10.1016/j.eja.2019.125964.

Nault, L. R. (1997). Arthropod transmission of plant viruses: A new synthesis. Annals of the Entomological Society of America, 90(5), 521-541.

Nault, B. A., Shah, D. A., and Taylor, A. G. (2006). Viruses and aphids everywhere in New York snap bean Þelds in 2005, pp. 74 -76. In Proceedings of the 2006 Empire State Fruit and Vegetable EXPO. Cornell Coop. Exten., Syracuse, NY.

Ndomba, O. A. (2002). Control of Aphid-borne viruses of field beans (*Phaseolus vulgaris* L.) using plant extracts. MSc. Thesis, University of Zambia.

Ng, J. C. K., and Falk, B. W. (2006). Virus-vector interactions mediating nonpersistant and semipersistent transmission. Annual Review of Phytopathology, 44, 183–212.

Nigam, D., LaTourrette, K., Souza, P. F. N., and Garcia-Ruiz, H. (2019). Genome-wide variation in potyviruses. Frontiers in Plant Science. 10:1439. doi.org/10.3389/fpls.2019:01439.

Njau, P. J. R., and Lyimo, H. F. J. (2000). Incidence of bean common mosaic virus and bean common mosaic necrosis virus in bean (*Phaseolus vulgaris* L.) and wild legume seed lots in Tanzania. Seed Science Technology, 28(1), 85-92.

Njau, P. J. R., Saria, J. E., Mwakitwange, F. E., and Lyimo, H. F. J. (2006). Detection of mixtures of bean and cowpea viruses by using single and mixed antisera in Enzyme-linked Immunosorbent Assay. Tanzanian Journal of Agriculture Science, 7, 50-55.

Nordenstedt, N., Marcenaro, D., Chilagane, D., Mwaipopo, B., RajamaÈki, M-L., and Nchimbi-Msolla, S. (2017). Pathogenic seedborne viruses are rare but *Phaseolus vulgaris* endornaviruses are common in bean varieties grown in Nicaragua and Tanzania. PLoS ONE, 12(5): e0178242. doi. org/10.1371/journal.pone.0178242.

Nouri, S., Arevalo, R., Falk, B. W., and Groves, R. L. (2014). Genetic structure and molecular variability of cucumber mosaic virus isolates in the United States. PLoS ONE 9(5), e96582. doi.org/10.1371/journal.pone.0096582.

Oana, D., Ziegler, A., Torrance, L., Gasemi, S., and Danci, M. (2009). *Potyviridae* family short review. Journal of Horticulture, Forestry and Biotechnology, 13, 410–421.

OECD (2016). Common bean (*Phaseolus vulgaris*), in safety assessment of transgenic organisms in the environment, Volume 6: OECD Consensus Documents, OECD Publishing, Paris. doi.org/10.1787/9789264253421-7-en.

Olspert, A., Chung, B. W., Atkins, J. F., Carr, J. P., and Firth, A. E. (2015). Transcriptional slippage in the positive-sense RNA virus family. *Potyviridae*. EMBO Reports, 16(8), 995-1004.

Okada, R., Yong, C. K., Valverde, R. A., Sabanadzovic, S., Aoki, N., Hotate, S., Kiyota, E., cMoriyama, H., and Fukuhara, T. (2013). Molecular characterization of two evolutionarily distinct endornaviruses co-infecting common bean (*Phaseolus vulgaris*). Journal of General Virology, 94(1), 220-229.

Omunyin, E. M. (1984). Studies on bean common mosaic viruses on beans in Kenya. MSc. Thesis, University of Nairobi.

Omunyin, M. E., Gathuru, E. M., and Mukunya, D. M. (1995). Pathogenicity groups of bean common mosaic virus isolates in Kenya. Plant Disease, 79(10), 985-989.

Palukaitis, P., Roossinck, M. J., Dietzgen, R. G., and Francki, R. I. B. (1992). Cucumber mosaic virus. Advances in Virus Research, 41, 281–348.

Palukaitis, P., and Garcı´a-Arenal, F. (2003). Cucumoviruses. Advances in Virus Research, 62, 242–323.

Panno, S., Davino, S., Rubio, L., Rangel, E., Davino, M., Garcia-Hernandez, J., and Olmos, A. (2012). Simultaneous detection of the seven main tomato-infecting RNA viruses by two multiplex reverse transcription polymerase chain reactions. Journal Virology. Methods, 186 (1-2), 152-156.

Pasev, G., Kostova, D., and Sofkova, S. (2014). Identification of genes for resistance to Bean common mosaic virus and Bean common mosaic necrosis virus in snap bean (*Phaseolus vulgaris* L.) breeding lines using conventional and molecular methods. Journal of Phytopathology, 162(1), 19–25.

Pasev, G., and Kostova, D. (2015). Occurrence of temperature-independent necrotic strains ofBean common mosaic virus (BCMV) on snap beans (*Phaseolus vulgaris*) in Bulgaria. Acta Horticulturae, 1105(1105), 173-180.

Pasin, F., Simon-Mateo, C., and Garcia, J. A. (2014). The hypervariable aminoterminus of P1 protease modulates potyviral replication and host defense responses. PloS Pathogen. 10 (3), e1003985. doi.org/10.1371/journal.ppat.1003985.

Pecman, A., Kutnjak, D., Gutierrez-Aguirre, I., Adams, I., Fox, A., Boonham, N., and Ravnikar, M. (2017). Next Generation Sequencing for detection and discovery of plant viruses and viroids: Comparison of two approaches. Frontiers in Microbiology, 8, 1998. doi.org/10.3389/fmicb.2017.01998.

Petrović, D., Ignjatov, M., Nikolić, Z., Vujaković, M., Vasić, M., Milošević, M., and Taški-Ajduković, K. (2010). Occurrence and distribution of viruses infecting the bean in Serbia. Archives of Biological Sciences, 62(3), 595–601.

Petry, N., Boy, E., Wirth, J. P., Hurrell, R. F. (2015). Review: The potential of the common bean (*Phaseolus vulgaris*) as a vehicle for iron biofortification. Nutrients, 7(2), 1144-1173.

Pinheiro, V. P., Wilson, J. R., Xu, Y., Zheng, Y., Rebelo, A. R., Fattah-Hosseini, S., Kruse, A., Silva, D. S. R., Xu, Y., Kramer, M., Giovannoni, J., Fei, Z., Gray, S., and Heck, M. (2019). Plant viruses transmitted in two different modes produce differing effects on small RNA-mediated processes in their aphid vector. Phytobiomes Journal, 3, 71-81.

Pita, J. S., Fondong, V. N., Sangaré, A., Kokora, R. N. N. and Fauquet, C. M. (2001). Genomic and biological diversity of the African cassava geminiviruses. Euphytica, 120, 115-125.

Raatz, B., Mukankusi, C., Lobaton, J. D., Male, A., Chisale, V., Amsalu, B., Deidre´ Fourie, Mukamuhirwa, F., Muimui, K., Mutari, B., Nchimbi-Msolla, S., Nkalubo, S., Tumsa, K., Chirwa, R., Maredia, M. K., and He, C. (2019). Analyses of African common bean (Phaseolus vulgaris L.) germplasm using a SNP fingerprinting platform: diversity, quality control and molecular breeding. Genetic Resources and Crop Evolution, 66, 707-722. doi.org/10.1007/s10722-019-00746-0.

Rawal, V., and Navarro, D. K. (2019). The Global Economy of Pulses. Rome, FAO.

Rendón-Anaya, M., Montero-Vargas, J. M., Saburido-Álvarez, S., Vlasova, A., Capella-Gutierrez, S., Ordaz-Ortiz, J. J., Aguilar, M. O., Vianello-Brondani, R. P., Santalla, M., Delaye, L., Gabaldón, T., Gepts, P., Winkler, R., Guigó, R., Delgado-Salinas, A., and Herrera-Estrella, A. (2017). Genomic history of the origin and domestication of

common bean unveils its closest sister species. Genome Biology, 18, 60. doi.org/10.1186/s13059-017-1190-6.

Revathy, K. A., and Bhat, A. I. 2017. Complete genome sequencing of cucumber mosaic virus from black pepper revealed rare deletion in the methyltransferase domain of 1a gene. Virus Disease, 28(3), 309-314. doi.org/10.1007/s13337-017-0386-4.

Revers, F., and García, A. J. (2015). Molecular Biology of Potyviruses. In K. Maramorosch., T. C. Mettenleiter (Ed.). Advances in Virus Research (pp.101-199). Burlington: Academic Press.

Revers, F., Le Gall, O., Candresse, T., Le Romancer, and M., Dunez, J. (1996). Frequent occurrence of recombinant potyvirus isolates. Journal of General Virology, 77(8), 1953-1965.

Riechmann, J., Laı´n, S., and Garcı´a, J. A. (1990). Infectious in vitro transcripts from a plum pox potyvirus cDNA clone. Virology, 177(2), 710–716.

Robertson, N. L., French, R., and Gray, S. M. (1991). Use of group-specific primers and the polymerase chain reaction for the detection and identification of luteoviruses. Journal of General Virology, 72, 1473–7.

Rohozˇkova´, J., and Navra´til, M. (2011). P1 peptidase – a mysterious protein of family Potyviridae. Journal of Biological Sciences, 36(1), 189–200. doi.org/10.1007/s12038-011-9020-6.

Roossinck, M. J. (2001). Cucumber mosaic virus, a model for virus evolution. Molecular Plant Pathology, 2(2), 59–63.

Roossinck, M. J. (2015). Metagenomics of plant and fungal viruses reveals an abundance of persistent lifestyles. Frontiers in Microbiology, 5, 767.

Roossinck, M. J., Bujarski, S. W., Ding, R., Hajimorad, K., Hanada, S., and Tousignant, M. (1999). Family *Bromoviridae*. In M. H. V. van Regenmortel., C. M. Fauquet., D. H. L.

Bishop (Ed.). Virus Taxonomy–Seventh Report of the International Committee on Taxonomy of Viruses (pp. 923–935). San Diego, Califonia: Academic Press

Roossinck, M. J, Zhang, L., and Hellwald, K. H. (1999). Rearrangements in the 5` nontranslated region and phylogenetic analyses of cucumber mosaic virus RNA3 indicate radial evolution of three subgroups. Journal of Virology, 73(8), 6752–8.

Rozanov, M. N., Eugene, V. K., and Alexander, E. G. (1992). Conservation of the putative methyltransferase domain: a hallmark of the 'Sindbis-like' supergroup of positive-strand RNA viruses. Journal of General Virology, 73(8), 2129–2134.

Rubio, L., Galipienso, L., and Ferriol, I. (2020). Detection of Plant Viruses and Disease Management: Relevance of Genetic Diversity and Evolution. Frontiers in Plant Science, 11, 1092. doi.org/10.3389/fpls.2020.01092.

Ryabov, E. V., Oparka, K. J., Santa Cruz, S., Robinson, D. J. and Taliansky, M. E. (1998). Intracellular location of two groundnut rosette umbravirus proteins delivered by PVX and TMV vectors. Virology, 242(2), 303–313.

Ryabov, E. V., Roberts, I. M., Palukaitis, P., and Taliansky, M. (1999a). Host-specific cell-to-cell and long-distance movements of cucumber mosaic virus are facilitated by the movement protein of groundnut rosette virus. Virology, 260(1), 98–108.

Ryabov, E. V., Robinson, D. J. and Taliansky, M. E. (1999b). A plant virus-encoded protein facilitates long-distance movement of heterologous viral RNA. Proceedning of National Academy of Science, 96(4), 1212–1217.

Ryabov, E. V., Robinson, D. J. and Taliansky, M. E. (2001). Umbravirus-encoded proteins both stabilize heterologous viral RNA and mediate its systemic movement in some plant species. Virology, 288(2), 391–400.

Ryabov, E. V., Taliansky, M. E., Robinson, D. J., Waterhouse, P. M., Murant, A. F., de Zoeten, G. A., Falk, B. W., Vetten, H. J., and Gibbs, M. J. (2012). Genus Umbravirus,

In A, King, M.J, Adams, E.C, Carstens, E.J, Lefkowitz (Ed.). Virus taxonomy (pp. 1191–1195). Ninth Report of the International Committee on Taxonomy of Viruses. New York: Elsevier.

Ryabov, E. V., and Taliansky, M. (2008). Umbraviruses (*Tombusviridae*). In B.W, Mahy and M.H.V, Van Regenmortel (Ed.). Encyclopedia of Virology (3rd ed., pp. 209-213): Academic Press Inc: Publishers.

Saqib, M., Nouri, S., Cayford, B., Jones, R. A. C., and Jones, M. G. K. (2010). Genome sequences and phylogenetic placement of two isolates of bean common mosaic virus from *Macroptiliumatropurpureum* in north-west Australia. Australasian Plant Pathology, 39, 184–191.

Sarmiento, C., Gomez, E., Meier, M., Kavanagh, T.A., and Truve, E. (2007). Cocksfoot mottle virus P1 suppresses RNA silencing in Nicotiana benthamiana and Nicotiana tabacum. Virus Research, 123(1), 95–99.

Sastry, K.S. 2013. Methods of Combating Seed-Transmitted Virus Diseases. In: Seed-borne plant virus diseases. Springer, India. doi.org/10.1007/978-81-322-0813-6_8.

Schmutz, J., McClean, P. E. , Mamidi, S., Wu, G. A., Cannon, S. B., Grimwood, J., Jenkins, J., Shu, S., Song, Q., Chavarro, C., Torres-Torres, M., Geffroy, V., Moghaddam, S. M., Gao, D., Abernathy, B., Barry, K., Blair, M., Brick, M. A., Chovatia, M., Gepts, P., Goodstein, D. M., Gonzales, M., Hellsten, U., Hyten, D. L., Jia, G., Kelly, J. D., Kudrna, D., Lee, R., Richard, M. M. S., Miklas, P. N., Osorno, J. M., Rodrigues, J., Thareau, V., Urrea, C. A., Wang, M., Yu, Y., Zhang, M., Wing, R. A., Cregan, P. B., Rokhsar, D. S., Jackson, S. A. (2014). A reference genome for common bean and genome-wide analysis of dual domestications. Nature Genetics, 46(7), 707-713.

Segundo, E., Carmona, M. P., S´aez, E., Velasco, L., Mart´ın, G., Ruiz, L., Janssen, D., and Cuadrado, I. M. (2008). Occurrence and incidence of viruses infecting green beans in south-eastern Spain. European Journal Plant Pathology, 122, 579-591.

Sengooba, T. N., Spence, N. J., Walkey, D. G. A., Allen, D. J., and Lana, F. A. (1997). The occurrence of *Bean common mosaic necrosis virus* in wild and forage legumes in Uganda. Plant Pathology, 46(1), 95–103.

Se´re´me´, D., Lacombe, S., Konate´, M., Pinel-Galzi, A., Traore´, V. S. E., He´brard, E., Traore´, O., Brugidou, C., Fargette, D., Konate, G. (2008). Biological and molecular characterization of a putative new sobemovirus infecting Imperata cylindrica and maize in Africa. Archive of Virology, 153, 1813–1820. doi.org/10.1007/s00705-008-0190-y.

Shahraeen, N., Ghotbi, T., Dezaje, E. A., and Sahandi, A. (2005). A survey of viruses affecting French bean (*Phaseolus vulgaris*) in Iran includes a first report of Southern bean mosaic virus and bean pod mottle virus. Plant Disease, 89(9), 1012.

Shintaku, M. H., Lee, Z., and Palukaitis, P. (1992). A single amino-acid substitution in the coat protein of Cucumber mosaic virus induces chlorosis in tobacco. Plant Cell, 4 (7), 751–7.

Siaw, M. F. E., Shahabuddin, M., Ballard, S., Shaw, J. G., and Rhoads, R. E. (1985). Identification of a protein covalently linked to the 5’ terminus of tobacco vein mottling virus RNA. Virology, 142(1), 134–143.

Sichilima, T., Mapemba, L., and Tembo, G. (2016). Drivers of Dry Common Beans Trade in Lusaka, Zambia: A Trader’s Perspective. Sustainable Agriculture Research, 5 (2), 15-26.

Silbernagel, M. J., Mink, G. I., Zhao, R. L., and Zheng, G. Y. (2001). Phenotypic recombination between bean common mosaic and bean common mosaic necrosis potyviruses *in vivo*. Archives of Virology, 146, 1007–1020.

Singh, S. P., Gepts, P., and Debouck, D. G. (1991). Races of common bean (*Phaseolus vulgaris*, *Fabaceae*). Economic Botany, 45, 379-396.

Singh, S. P., and Schwartz, H. F. (2010). Breeding common bean for resistance to diseases: a review. Crop Science, 50(2), 2199–2223.

Sivakumaran, K., and Hacker, D. L. (1998). The 105-kDa polyprotein of Southern bean mosaic virus is translated by scanning ribosomes. Virology, 246(1), 34–44.

Smith, D. B., McAllister, J., Casino, C., and Simmons, P. (1997). Virus 'quasispecies' making a mountain out of a molehill? Journal of General Virology, 78(7), 1511-1519.

Sõmera, M., Sarmiento, C., and Truve, E. (2015). Overview on Sobemoviruses and a proposal for the creation of the Family Sobemoviridae. Viruses, 7(6), 3076-3115. doi.org/10.3390/v7062761.

Spence, N. J., and Walkey, D. G. A. (1995). Variation for pathogenicity among isolates of bean common mosaic virus in Africa and a reinterpretation of the genetic relationship between cultivars of *Phaseolus vulgaris* and pathotypes of BCMV. Plant Pathology, 44(3), 527–546.

Sseruwagi, P., Sserubombwe, W. S., Legg, J. P., Ndunguru, J., and Thresh, J. M. (2004). Methods of surveying the incidence and severity of cassava mosaic disease and whitefly vector populations on cassava in Africa: a review. Virus Research, 100(1), 129-142.

Strausbaugh, C. A., Miklas, P. N., Singh, S. P., Myers, J. R., and Forster, R. L. (2003). Genetic characterization of differential reactions among host group 3 common bean

cultivars to NL-3 K strain of Bean common mosaic necrosis virus. Phytopathology, 93(6), 683-690.

Suzuki, M., Kuwata, S., Masuta, C., and Takanami, Y. (1995). Point mutations in the coat protein of cucumber mosaic virus affect symptom expression and virion accumulation. Journal of General Virology, 76(7), 1791–9.

Taliansky, M. E., Robinson, D. J., and Murant, A. F. (1996). Complete nucleotide sequence and organization of the RNA genome of groundnut rosette umbravirus. Journal of General Virology, 77(9), 2335–2345.

Taliansky, M. E., Robinson, D. J., and Murant, A. F. (2000). Groudnut rosette disease virus complex: biology and molecular biology. Advances in Virus Research, 55, 357–400.

Taliansky, M. E., and Robinson, D. J. (2003). Molecular biology of umbraviruses: phantom warriors. Journal General Virology, 84(8), 1951–1960. doi.org/10.1099/vir.0.19219-0.

Tamm, T., and Truve, E. (2000). RNA-binding activities of cocksfoot mottle sobemovirus proteins. Virus Research, 66(2), 197–207.

Tamm, T, and Truve, E. (2000). Sobemoviruses. Journal of Virology, 74, 6231–6241.

Tang, J., Lebas, B., Liefting, L., Veerakone, S., Wei, T., and Ward, L. (2015). Opium poppy mosaic virus, a new umbravirus isolated from *Papaver somniferum* in New Zealand. Archives of Virology, 161(1), 197-201. doi.org/10.1007/s00705-015-2651-4.

Thomas, J. E., Massalski, P. R., and Harrison, B. D. (1986). Production of monoclonal anti-bodies to African cassava mosaic virus and differences in their reactivity with other whitefly-transmitted geminiviruses. Journal of General Virology, 67(12), 2739-2748.

Thompson, J. R., Langenhan, J. L., Fuchs, M., and Perry, K. L. (2015). Genotyping of cucumber mosaic virus isolates in western New York State during epidemic years: characterization of an emergent plant virus population. Virus Research, 169–177. doi.org/10.1016/j.virusres.2015.07.028.

Thottappilly, G. (1992). Plant virus diseases of importance to African agriculture. Phytopathology, 134(4), 265-288.

Togola, A., Boukar, O., Servent, A., Chamarthi, S., Tamò, M., and Fatokun, C. (2020). Identification of sources of resistance in cowpea mini core accessions to *Aphis craccivora* Koch (Homoptera: *Aphididae*) and their biochemical characterization. Euphytica: Netherlands journal of plant breeding, 216, 88.

Tremaine, J.H., and Hamilton, R.I. (1983). Southern bean mosaic virus. In CMI/AAB Description of Plant Viruses No. 274; Commonwealth Mycological Institute, Kew, Surrey, and Assocaiation of Applied Biologists: Wellesbourne, Warwick, UK.

Udagawa, H., Koga, K., Shinjo, A., Kitashiba, H., and Takakura, Y. (2020). Reduced susceptibility to a tobacco bushy top virus Malawi isolate by loss of function in host *eIF(iso)4E* genes. Breeding science, 70(3), 313–320. doi.org/10.1270/jsbbs.19135

Valverde, A. R., Khalifa, E. M., Okada, R., Fukuhara, T., Sabanadzovic, S., (2019): ICTV Virus Taxonomy Profile: *Endornaviridae,* Journal of General Virology, 100, 1204–1205.

Valli, A., López-Moya, J. J., and García, J. A. (2007). Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family Potyviridae. Journal of General Virology, 88(3), 1016-1028. doi.org/10.1099/vir.0.82402-0.

Verchot, J., and Carrington, J. C. (1995). Evidence that the potyvirus P1 proteinase functions in trans as an accessory factor for genome amplification. Journal of Virology, 69(6), 3668-3674.

Verhoeven, J. T. H. J., Roenhorst, J. W., Lesemann, D. E., Segundo, E., Velasco, L., Ruiz, L., Janssen, D., and Cuadrado, I. M. (2003). Southern bean mosaic virus the causal agent of a new disease of *Phaseolus vulgaris* beans in Spain. European Journal of Plant Pathology, 109, 935-941.

Vetten, H. J., and Allen, D. J. (1991). Recent progress in the identification of viruses of *Phaseolus vulgaris* in Africa. Annual report of the bean improvement cooperatives and National Dry Bean Council Research Conference. USDA National Agricultural Library pp. 3-4.

Vetten, H. J., Lesemann, D. E., and Maiss E. (1992). Serotype A and B strains of bean common mosaic virus are two distinct potyviruses. In Barnett O.W. (Ed.). Potyvirus Taxonomy. Archives of Virology (Supplementum 5), vol 5. Vienna: Springer. doi.org/10.1007/978-3-7091-6920-9_44.

Wainaina, J. M., Ateka, E., Makori, T., Boykin, L. M., and Kehoe, M. A. (2017). Genomic characterization and phylogenomic relationships of Phaseolus vulgaris endornavirus in common beans of Kenya.https://www.ncbi.nlm.nih.gov/nucleotide.

Wainaina, J. M., Kubatko, L., Harvey, J., Ateka, E., Makori, T., Karanja, D., Boykin, L. M., and Kehoe, M. A. (2019). Evolutionary insights of bean common mosaic necrosis virus and cowpea aphid-borne mosaic virus. PeerJ. doi.org/10.7717/peerj.6297.

Wakarchuk, D. A., and Hamilton, R. I. (1990). Partial nucleotide sequence from enigmatic dsRNAs in *Phaseolus vulgaris*. Plant Molecular Biology, 14(4), 637–639.

Walker, P. J., Siddell, S. G., Lefkowitz, E. J., Mushegian, A. R., Adriaenssens, E. M., Alfenas-Zerbini, P., Davison, A. J., Dempsey, D. M., Dutilh, B. E., García, M. L., Harrach, B., Harrison, R. L., Hendrickson, R. C., Junglen, S., Knowles, N. J., Krupovic, M., Kuhn, J. H., Lambert, A. J., Łobocka, M., Nibert, M. L, Zerbini, F. M. (2021). Changes to virus taxonomy and to the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2021). Archives of Virology. doi.org/10.1007/s00705-021-05156-1.

Wang, Y., Tzfira, T., Gaba, V., Citovsky, V., Palukaitis, P., and Gal-On, A. (2004). Functional analysis of the Cucumber mosaic virus 2b protein: pathogenicity and nuclear localization. Journal of General Virology, 85(10), 3135–47.

Wamonje, F. O., Donnelly, R., Tungadi, T. D., Murphy, A. M., Pate, A. E, Woodcock, C., Caulfield, J., Mutuku, J. M., Bruce, T. J. A., Gilligan, C. A., Pickett, J. A., and Carr, J. P. (2020). Different plant viruses induce changes in feeding behavior of specialist and generalist aphids on common bean that are likely to enhance virus transmission. Frontiers in Plant Science, 10, 1811. doi.org/10.3389/fpls.2019.01811.

Were, H. K., Mukoye, B., Stomeo, F., Were, M. N., Joyce, N., and Torrance, L. (2018). First report of peanut mottle virus (PeMoV) infecting groundnuts in Kenya. Unpublished. https://blast.ncbi.nlm.nih.gov/Blast.cgi#1699888438.

Westwood, J. H., and Stevens, M. (2010). Resistance to aphid vectors of virus disease. Advances in Virus Research, 76, 179–210. doi.org/10.1016/s0065-3527(10)76005-x.

White, A. K. 2020. Tombusvirus-Like viruses (Tombusviridae). Reference Module in Life Sciences. doi.org/10.1016/B978-0-12-809633-8.21260-0.

Wilkus, E. L., Berny Miery Teran, J. C., Mukankusi, C. M., and Gepts, P. (2018). Genetic patterns of common-bean seed acquisition and early-stage adoption among farmer groups in Western Uganda. Frontiers in Plant Science, 9, 586. doi.org/ 10.3389/fpls.2018.00586.

Worrall, A. E., Wamonje, F. O., Mukeshimana, G., Harvey, J. J. W., Carr, J. P., and Mitter, N. (2015). Bean common mosaic virus and bean common mosaic necrosis virus: Relationships, Biology, and Prospects for Control. Advances in Virus Research, 93, 1-46.

Wortmann, C. S. (2006). *Phaseolus Vulgaris* L. (common Bean); Prota: Cereals and pulses/Céréales et légumes secs, http://database.prota.org/PROTAhtml/Phaseolus vulgaris (common bean).

Wortmann, S. C., Kirkley, R. A., Eledu, C. A., and Allen, J. D. (1998). Atlas of common bean (*Phaseolus vulgaris* L.) production in Africa. Centro Internacional de AgriculturaTropica, Cali, Colombia. (CIAT publication; no. 297).

Wu, S. X., Rinehart, C. A., and Kaesberg, P. (1987). Sequence and organization of southern bean mosaic virus genomic RNA. Virology, 161(1), 73–80.

Wyatt, S. D. and Brown, J. K. (1996). Detection of subgroup III geminivirus isolates in leaf extracts by degenerate primers and polymerase chain reaction. Phytopathology, 86, 1288-1293.

Wylie, S.J., Adams, M., Chalam, C., Kreuze, J., López-Moya, J.J., Ohshima, K., Praveen, S., Rabenstein, F., Stenger, D., Wang, A., Zerbini, F.M. (2017). ICTV Virus Taxonomy Profile: *Potyviridae*. Journal of General Virology, 98, 352–354.

Zambia Statistics Agency (ZSA). (2017). Crop focusting survey. Central Statistical Office, Lusaka, Zambia.

Zanardo, L. G., and Carvalho, M. C. (2017). Cowpea mild mottle virus (Carlavirus, Betaflexiviridae): a review. Tropical Plant Pathology, doi.org/10.1007/s40858-017-0168-y.

Zaumeyer, J. W., and Thomas, H. R. (1957). A monographic study of bean diseases and methods for their control. Technical Bulletin no. 868. USA department of Agriculture.

Zaumeyer, W. J., and Harter, L. L. (1943). Inheritance of symptom expression of bean mosaic virus 4. Journal of Agricultural Research, 67, 295–300.

Ziebel, H., Murphy, A. M., Groen, S. C., Tungadi, T., Westwood, J. H., Lewsey, M. G., Moulin, M., Kleczkowski, A., Smith, A. G., Stevens, M., Powell, G., and Carr, J. P.

(2011). Cucumber mosaic virus and its 2b RNA silencing suppressor modify plant aphid interactions in tobacco. Scientific Reports, 1,187.

Zitter, T. A., and Murphy, J. F. (2009). Cucumber mosaic. The Plant Health Instructor. doi.org/10.1094/PHI-I-2009-0518-01.

Zongoma, A. M., Danladi B. Dangora, B. D., Al Rwahnih, M., Bako, S. P., Alegbejo, M., and Alabi, O. J. (2017). First report of grapevine leafroll-associated virus 1 infecting grapevines (Vittis spp.) in Nigeria. Plant Disease, 101(1). doi.org/10.1094/PDIS-07-17-1120-PDN.