

Construction and Characterization of *Ugandan cassava brown streak virus* (UCBSV) Full-Length Infectious Clones

Sarah Nanyiti

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Science



School of Biological Sciences
United Kingdom

October 2016

Word count: ~49,000

Abstract

Cassava is one of the most important staple foods to most of the African populations with over 800 million people depending on cassava as the main source of starch. However, low average yields of cassava are caused by a number of factors including susceptibility to pests and diseases. Among the viral diseases, cassava brown streak disease (CBSD), caused by *Cassava brown streak virus* and (CBSV) and *Ugandan cassava brown streak virus* (UCBSV) family *Potyviridae* and genus *Ipomovirus*, is one of the most devastating. Disease spread has been controlled by planting clean and healthy cuttings. Breeding for resistant varieties has been undertaken, however, there is a need to screen the generated breeding lines for resistance to CSBD, and this has necessitated a need for infectious clones.

The main aim of this PhD work was to construct a full-length infectious clone of UCBSV Kikombe isolate and to study the function of the P1 (first potyviral protein) and HAM1 proteins in the UCBSV genome. Infectious clones of RNA viruses provide a homogenous and reproducible source of viral infection for effective screening of breeding lines.

In order to construct the full-length infectious clone, initially, the complete genome of Kikombe isolate (UCBSV) was amplified in sections by reverse transcription polymerase chain reaction (RT-PCR) and Sanger sequenced. The complete sequence (GenBank accession number KX753356) was found to consist of 9070 nucleotides (excluding the poly-A tail) which translated into a polypeptide of 2902 amino acids.

A full-length *in-vitro* UCBSV infectious clone (GenBank accession KX753357) was generated by cloning the complete genome of the virus in a backbone pYES2.1 vector. An SP6 promoter was introduced at the 5' end of the UCBSV genome to allow for *in-vitro* transcription. The constructed infectious clone was confirmed to infect tobacco and cassava plants as shown by the mosaic like symptoms observed in the infectious clone infected plants. RNA was extracted from the symptomatic plants and used in RT-PCR. PCR and sequencing results confirmed the infection.

The full-length infectious clone was then mutated in the P1 protein to study the function of this protein in UCBSV. Initially, an RNA suppression assay confirmed that P1 protein in UCBSV was an RNA silencing suppressor. A mutation in the P1 encoding sequence that changed LRRA motif to LAAA abolished completely the silencing suppression activity of UCBSV-P1. However, the same mutation in the full-length infectious clone did not affect the infectivity of the infectious clone but the viral titres were observed to be higher in plants inoculated with the mutant infectious clone as compared to the wildtype infectious clone.

The P1 protein was then completely deleted from the UCBSV infectious clone and the P1 mutant was used to infect tobacco. It was observed that the P1 mutant infectious clone was able to move systemically within the plants although RT-PCR showed low levels of the virus in the plants implying that P1 may be possibly be involved in genome amplification.

HAM1 UCBSV infectious clone mutants were used to characterize the role of HAM1 protein in UCBSV genome. It was observed that plants infected with the HAM1 mutant infectious clone developed yellow mosaic symptoms which were different from the mild mosaics observed with the wildtype UCBSV infectious clone implying that HAM1 may be involved in symptom development.

In conclusion, this work has for the first time reported construction of an intron-less, *in-vitro* infectious clone of the full length cDNA of the UCBSV. Attempts to construct *in-vivo* infectious clones failed and that was probably caused by positions selected for insertions of introns and suggest the sequences of UCBSV when placed under strong constitutive promoters are toxic to *E. coli* cells. Further studies are recommended to investigate appropriate positions for insertions of the introns for construction of *in-vivo* infectious clones

Dedication and Acknowledgements

At this moment in time, I start my acknowledgments by wholeheartedly thanking my family, especially my late mother Mrs Wassanyi Rose (RIP) and my elder sister Nabasirye Lillian who supported me emotionally when I took a decision to leave a one year old baby son Ssali Ethan in Uganda to come to the UK and pursue my PhD studies. It was a very trying moment imagining the time I was going to be away from my son but I am very grateful that today I have come to the end of the road and I will forever be grateful for the support I have received throughout the years. To Ethan, I must say that even if I left you behind, I loved you so much as my son and I promise that I will try my best to be the mother I am supposed to be to you and I hope that we shall be able to bond as mother and son.

I heartedly thank my supervisors Prof Gary Foster and Dr Andy Bailey for the excellent job of guiding me in my studies. I also thank my industrial supervisor Dr Titus Alicai for the support and encouragement he has offered especially in encouraging me to work hard so that I was able to finish my PhD in time. Prof Sue Seal is thanked for her tireless efforts in ensuring that I received the materials that I needed to do my PhD project, starting from collecting and sending the original infected cassava leaves that I used for my work and also bringing over cassava plantlets to University of Bristol. I also thank Dr Celia Duff-Farrier and Dr Deusdedith Mbanzibwa for the assistance offered to me as I worked on my PhD project. I have learnt a lot from the mentorship I have received while working on my PhD project.

I wish to thank my sponsors the National Agricultural Research Organisation (NARO) through the EAAPP project that has funded my PhD studies. I also acknowledge the partial studentship funding that I received from the University of Bristol. I thank Dr George Lukwago the EAAPP co-ordinator in Uganda for the scholarship opportunity.

Many thanks, go to all the members of the Molecular Plant Pathology Laboratory (Lab 321) at the University of Bristol for providing a conducive environment for my work. I thank Suhad Al Salihi and Dr Kathryn Ford for the support they provided that made my laboratory stay very conducive. Dr Fabrizio Alberti is also thanked for providing constructive ideas when requested for assistance. For all the other members, Farhana, Nisa, Murnita, Katie, Luis, Anna, Alice, Lloyd, Bew, Dr Hamidun Bunawan, Dr Dan Maxwell, Dr Solaf Ali and Dr Rose Murray, thank you all for the team work.

To my Family, I once again thank late Mum (RIP) for looking after my son while I was away in the UK and my sister Lillian for being a mother to my son. I know Ethan will always call Lillian his mummy and I am happy that Lillian filled in for me. I thank Dr Henry Mukasa for checking on our son Ethan while I was away. Thanks to Mr John Bosco Otema for the role played in Ethan's life all through the years.

I also acknowledge my late father Mr James Wassanyi (RIP) for his efforts in educating me.

To my sisters Lillian Nabasirye, Juliet Nankanja and Ruth Nabuuso, thank you so much for being there for Ethan, all you people have provided a lot of support that enabled me to comfortably stay in UK. My nieces and Nephews Nicole Blessing, Nathan Nicolas, Eugene, Eureka, Estella, Elliana, Hellen, Duncan, Mark, Marvin, Maria - I have really missed you a lot. For my brother Ronald Kasirye, it is time for us to be together again. All my cousins, aunties and uncles are thanked for all the support.

I thank my friends, Aisha Birungi, Patricia Ndugga, Joselyn Atuhairwe, Agnes Namuli, Lillian Nangozi, Sam Opio, Chris Kizza, Alice Mulondo and Dr Kateregga John. My friend Rukia Karama and her family is thanked for having provided a home for me when I first came to the UK and also for all the support I have received throughout the years.

Finally, I thank the members of the Holy Trinity Church in Hotwells Bristol that provided a family for me whilst in UK. I especially thank Valerie Pearson and her husband, Christina, Sophia, Chloe, and Allan to mention but a few. Thank you all.

This work is dedicated to my son Ssali Ethan and my late mother Mrs Wassanyi Rose (RIP).

"To God be the glory for the great things he has done."

Author's Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

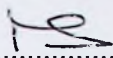
SIGNED:  DATE: 14/10/2016

Table of Contents

Abstract.....	i
Dedication and Acknowledgements	ii
Author's Declaration	iv
Table of Contents.....	v
List of Tables	xii
List of Figures.....	xiv
Chapter 1: Introduction.....	1
1.1 Plant pathogens and effects on food security.....	1
1.2 Viruses	2
1.3 Properties of plant viral genomes	3
1.4 Cassava	4
1.4.1 Cassava production.....	5
1.4.2 Cassava propagation	6
1.5 Challenges to cassava production	7
1.5.1 CBSD symptoms	9
1.6 Diagnosis of cassava brown streak disease.....	11
1.6.1 Distribution of cassava brown streak disease	12
1.6.2 Characterization of cassava brown streak viruses	13
1.7 Other viral diseases of cassava	15
1.7.1 Cassava mosaic geminiviruses (CMG).....	15
1.8 Genera of the <i>Potyviridae</i> family.....	16
1.9 The genus <i>Ipomovirus</i>	17
1.9.1 <i>Sweet potato mild mottle virus</i> (SPMMV)	18
1.9.2 <i>Cucumber vein yellowing virus</i> (CVYV)	18
1.9.3 <i>Squash vein yellowing virus</i> (SqVYV).....	19
1.9.4 <i>Cassava brown streak virus</i> (CBSV) and <i>Ugandan cassava brown streak virus</i> (UCBSV).....	19
1.10 Functions of the different proteins in the typical potyviral genome.....	22
1.10.1 P1	22
1.10.2 HC-Pro.....	22
1.10.3 P3	23
1.10.4 6K1 and 6K2.....	24
1.10.5 CI	24

1.10.6	VPg	24
1.10.7	Nla	25
1.10.8	Nlb	25
1.10.9	HAM1	26
1.10.10	Coat Protein (CP).....	27
1.10.11	5' and 3' untranslated regions (UTR)	27
1.11	Plant host pathogen interaction in virus infections	28
1.12	Resistance to viral pathogens	28
1.12.1	Pathogen derived resistance (PDR)	28
1.12.2	Host derived resistance	29
1.13	RNA silencing.....	30
1.13.1	RNA silencing mechanisms.....	30
1.13.2	Viral suppressors of RNA silencing	30
1.14	Construction of infectious clones.....	32
1.15	Applications of infectious clones of viruses	34
1.15.1	Potyvirus infectious clones	34
1.16	Aims and objectives of the study	35
CHAPTER 2:	Materials and Methods	37
2.1	Plant material and virus isolates used in this study.....	37
2.1.1	Plant material	37
2.1.1.1	<i>Nicotiana species</i>	37
2.1.1.2	Cassava plants (<i>Manihot esculenta</i>).....	38
2.1.2	Virus isolates	38
2.1.2.1	Kikombe UCBSV isolate source	38
2.2	Virus preparations.....	39
2.2.1	Virus inoculum preparation and inoculation	39
2.2.1.1	Herbaceous plants	39
2.2.1.2	Cassava infections.....	39
2.3	Microbiological methods	40
2.3.1	Strains	40
2.3.1.1	<i>Escherichia coli</i> strains	40
2.3.1.2	<i>Agrobacterium tumefaciens</i> strains.....	40
2.3.1.3	<i>Saccharomyces cerevisiae</i> strains	40
2.4	Media	41
2.4.1	Bacterial media	41

2.4.1.1	<i>E. coli</i> and <i>Agrobacterium</i> media	41
2.4.2	<i>Saccharomyces cerevisiae</i> media.....	41
2.4.3	Media supplements.....	41
2.4.3.1	Antibiotics.....	41
2.4.4	Culture maintenance and storage.....	42
2.4.4.1	<i>E. coli</i> culture maintenance.....	42
2.4.4.2	<i>S. cerevisiae</i> maintenance	43
2.5	Plasmid Extraction.....	43
2.5.1	<i>E. coli</i> plasmid extraction	43
2.5.2	Yeast plasmid extraction	43
2.6	Molecular Techniques.....	44
2.6.1	RNA extraction.....	44
2.6.2	First-strand cDNA synthesis.....	44
2.6.3	Restriction digestion of DNA	45
2.6.4	PCR Primers	46
2.6.5	Polymerase Chain Reaction.....	46
2.6.5.1	Conventional PCR	46
2.6.5.2	Colony PCR	46
2.6.5.3	High fidelity PCR	47
2.6.6	Agarose Gel Electrophoresis	48
2.6.6.1	Loading and running gels.....	48
2.6.6.2	Visualisation of Nucleic acids	49
2.6.7	Cloning of PCR products for sequencing	50
2.6.8	Ligation of DNA fragments into vectors.....	51
2.7	Preparation of chemically competent <i>E. coli</i> cells.....	51
2.7.1	Transformation of chemically competent <i>E. coli</i>	52
2.8	Preparation of electrocompetent <i>E. coli</i> cells	52
2.8.1	Transformation of electrocompetent <i>E. coli</i>	53
2.9	Construction of plasmids by homologous recombination in yeast	54
2.10	<i>In-vitro</i> Transcription of the infectious clone	55
2.10.1	Inoculation of plants with full length viral transcript.....	56
2.11	Genome Assembly with Sequencher	57
2.11.1	Annotation of the UCBSV Kikombe genome	57
2.12	Serological methods.....	58
2.12.1	Triple Antibody Sandwich ELISA	58

2.12.2	ELISA buffers.....	58
2.12.3	ELISA protocol.....	58
CHAPTER 3 Initial characterization of <i>Ugandan cassava brown streak virus</i> (UCBSV)		
“Kikombe” Isolate		60
3.1	Introduction.....	60
3.1.1	Classification of <i>Ugandan cassava brown streak virus</i> (UCBSV)	60
3.1.1.1	UCBSV genome organization.....	61
3.1.2	Use of plant indicators in virus infections	62
3.1.3	Symptom and host range of CBSV and UCBSV	62
3.1.4	Transmission of <i>Ugandan cassava brown streak virus</i> (UCBSV).....	63
3.1.5	Aims of this chapter.....	64
3.2	Results.....	66
3.2.1	The UCBSV isolate “KIKOMBE”	66
3.2.2	Host indicator plants infection to assess symptoms	67
3.2.3	Designing primers for sequencing the complete genome of UCBSV	70
3.2.3.1	Amplification of UCBSV in sections (1-8)	73
3.2.3.2	Amplification of the 3’end of UCBSV genome	73
3.2.3.3	Amplification of 5’ end of UCBSV genome	74
3.2.3.4	Amplification of section 4 (position 3476-4918).....	74
3.2.3.5	RT-PCR to amplify the different sections of UCBSV.....	75
3.2.3.6	Complete genomic sequence assembly.....	76
3.2.3.7	Description of assembled UCBSV Kikombe genome	78
3.2.3.8	Genome confirmation and translation.....	78
3.2.3.9	Annotation of the assembled UCBSV Kikombe genome.....	79
3.2.4	Comparison between the individual proteins in UCBSV Kikombe (GenBank accession KX753356) against the reference UCBSV (NC-014791.1)	79
3.2.5	Cleavage site analysis in the UCBSV genome	80
3.2.6	Bioinformatic analysis.....	82
3.2.6.1	Complete genome comparison between UCBSV Kikombe (KX753356) and other isolates	82
3.2.6.2	Comparison of P1 and HAM1 protein between isolates.....	84
3.3	Discussion	87
Chapter 4: Construction of an <i>in-vitro</i> and <i>in-vivo</i> infectious clone of UCBSV Kikombe isolate.....		
4.1	Introduction.....	89
4.1.1	Applications of infectious clones	90

4.1.1.1	Viral protein functional studies.....	90
4.1.1.2	Study viral cell –to- cell movement.....	91
4.1.2	Methods of generating infectious clones.....	91
4.1.2.1	<i>E. coli</i> cloning system.....	91
4.1.2.2	<i>In-vitro</i> ligation of PCR fragments.....	92
4.1.2.3	Insertion of Introns.....	92
4.1.3	<i>In-vitro</i> and <i>in-vivo</i> infectious clones.....	93
4.1.4	Aims of this Chapter.....	94
4.2	Results.....	95
4.2.1	Designing specific primers based on UCBSV genome to construct <i>in-vitro</i> infectious clone.....	95
4.2.2	PCR to amplify different sections of UCBSV.....	96
4.2.3	Yeast homologous recombination.....	97
4.2.4	Construction of pYES2.1 with 5' end UCBSV genome.....	98
4.2.5	Construction of pYES2.1 with 3' end UCBSV genome.....	99
4.2.6	Construction of UCBSV infectious clone plasmid.....	101
4.2.7	Checking the integrity of the constructed UCBSV infectious clone.....	104
4.2.7.1	PCR to check for the individual sections (Section 1-8).....	104
4.2.7.2	Restriction analysis to confirm integrity.....	105
4.2.7.3	Sequencing of the UCBSV infectious clone genome.....	106
4.2.8	Comparison of the genome of the assembled infectious clone and wildtype UCBSV genome.....	108
4.2.9	<i>In-vitro</i> transcription and infection experiments.....	109
4.2.9.1	<i>In-vitro</i> transcription to generate transcripts.....	110
4.2.9.2	Inoculation of plants with infectious clone.....	111
4.2.9.3	Confirmation of UCBSV infection.....	112
4.2.9.4	Sequencing the PCR products.....	114
4.2.10	Testing the infectious clone in cassava plants.....	115
4.2.10.1	Sequencing the PCR products to confirm UCBSV infection in cassava plants.	118
4.2.11	Construction of <i>in-vivo</i> UCBSV Infectious Clone.....	119
4.2.11.1	Construction of UCBSV genome in PCAMBIA 0380 backbone.....	119
4.2.11.2	Introducing restriction sites at the 5' end of the UCBSV genome and 3' end of the UCBSV genome.....	121
4.2.11.3	Introduction of introns into the UCBSV genome.....	123
4.3	Discussion.....	126

CHAPTER 5: CHARACTERIZATION OF P1 and HAM1 PROTEINS IN UCBSV KIKOMBE ISOLATE.....	129
5.1 Introduction.....	129
5.1.1 Functions of UCBSV proteins.....	129
5.1.1.1 RNA silencing in viral infections	130
5.1.1.2 P1 as an RNA silencing suppressor	131
5.1.1.3 HAM 1 protein.....	131
5.1.2 Aims of the Chapter.....	133
5.2 Results.....	135
5.2.1 Identification of LRRA motif in P1 sequences of different UCBSV isolates .	135
5.2.2 Primer design for site directed mutagenesis	136
5.2.3 Construction of mutated P1 UCBSV infectious clone by yeast recombination	137
5.2.4 Experiments of mutated P1 infectious clone infections in <i>N. benthamiana</i>	139
5.2.4.1 RT-PCR to confirm UCBSV infection in the plants.....	141
5.2.5 Experiments with the UCBSV P1 complete knockout infectious clone	142
5.2.5.1 Construction of P1 complete knockout infectious clone	142
5.2.5.2 Infection of P1 Knockout infectious clone plasmid in <i>N. benthamiana</i> , <i>N. tabacum</i> and <i>N. clevelandii</i>	144
5.2.6 Construction of P1 agroinfiltration vectors for RNA silencing suppression studies	146
5.2.6.1 Agroinfiltration assays to assess RNA silencing suppression	148
5.2.7 HAM1 UCBSV knockout infectious clone experiments.....	151
5.2.7.1 Verification of the integrity of the HAM1 constructs.....	151
5.2.7.2 <i>In-vitro</i> transcription with the HAM1 knockout constructs.....	152
5.2.7.3 Infection of <i>N. benthamiana</i> , <i>N. clevelandii</i> and <i>N. tabacum</i> with UCBSV HAM1 knockout constructs.....	153
5.2.7.4 Infecting the HAM1 deletions in transgenic HAM1 <i>N. tabacum</i> plants to study complementation.....	155
5.2.7.5 RT-PCR to detect UCBSV infection in Transgenic HAM1 plants.....	157
5.3 Discussion.....	159
CHAPTER 6: GENERAL DISCUSSION.....	163
6.1 GENERAL DISCUSSION	163
6.1.1 Genome sequencing and characterization of UCBSV Kikombe isolate	164
6.1.2 UCBSV Kikombe isolate infectious clone construction	166
6.1.3 Characterization of P1 and HAM1 proteins in the UCBSV genome	170

6.2	Recommendations.....	174
6.3	Conclusions.....	175
REFERENCES		176
Appendix 1: Complete genome of UCBSV wildtype Kikombe isolate (9096 Nucleotides) GenBank Accession Number KX753356.....		197
Appendix 2: Complete genome of UCBSV Infectious clone (9096 nucleotides) GenBank Accession Number KX753357		198
Appendix 3: Alignment of the Nucleotide sequence between the UCBSV wildtype and UCBSV infectious clone genomes (99% homology)		199
Appendix 4: Alignment of amino acid sequence (ORF) between UCBSV Infectious clone and UCBSV wildtype		202
Appendix 5: Translation of UCBSV Genome with cleavage sites highlighted		203
Appendix 6: Nucleotide alignments of P1 sequences from different UCBSV and CBSV isolates		204
Appendix 7: Amino acid alignments between P1 protein sequences of different UCBSV and CBSV isolates.....		206
Appendix 8: Nucleotide alignments of HAM1 sequences from different UCBSV and CBSV isolates		207
Appendix 9: Amino acid alignments between HAM1 sequences between UCBSV and CBSV isolates		209

List of Tables

Chapter 2: Materials and Methods

Table 2.1: Antibiotics used in this study	42
---	----

Chapter 3: Initial characterization of *Ugandan cassava brown streak virus* (UCBSV) “Kikombe” isolate

Table 3.1: The identities and origins of selected UCBSV isolates in the NCBI database that were used in designing degenerate primers for sequencing the genome of the uncharacterized UCBSV Kikombe isolate collected from Mikocheni area, Tanzania.	70
Table 3.2: List of primers used to amplify the sections of the genome of UCBSV Kikombe isolate.....	71
Table 3.3: Sequence identity of each protein coding region between UCBSV Kikombe isolate (KX753356) with UCBSV (NC_014791.1).....	80
Table 3.4: Amino acid sequences at the nine predicted cleavage sites in the UCBSV Kikombe isolate polyprotein	81

Chapter 4: Construction of an *in-vitro* and *in-vivo* infectious clone of UCBSV Kikombe isolate

Table 4.1: Specific primers used to construct the UCBSV infectious clone. The primers are based on the UCBSV Kikombe genome (KX533356). The sequences in italics represent the pYES2.1 vector backbone overlapping sequence to allow for homologous recombination in yeast	96
Table 4.2: Single nucleotide polymorphisms (SNP's): The positions of single nucleotide polymorphisms observed in the genomes of the infectious clone (UCBSV I.C) (GenBank accession KX753357) and the wildtype (UCBSV WT) (GenBank accession KX753356) genome in comparison with selected UCBSV isolates in NCBI database listed by accession numbers.	108
Table 4.3: Amino acid changes at the positions with single nucleotide polymorphisms (SNPs) in the wildtype UCBSV (GenBank accession KX753356) genome and constructed UCBSV infectious clone (GenBank accession KX753357).	109

Chapter 5: Characterization of P1 and HAM1 proteins in UCBSV Kikombe isolate

Table 5.1: Oligonucleotide primers used for site directed mutagenesis of the LRRA motif to LAAA in the UCBSV P1 protein. The underlined bases represent the region targeted in the mutation analysis and the italicised amino acids represent the mutations.	136
Table 5.2: PCR sections generated in constructing UCBSV infectious clone with a mutation change in P1 protein. Sequences for primers are in Table 4.1 and Table 5.1.	137
Table 5.3: UCSBV P1 knockout primers	142

List of Figures

Chapter 1: Introduction

- Figure 1.1:** Cassava plant showing the tuberous roots, stem and palmate compound leaves (Source <https://rauman.wikispaces.com/cassava>)..... 5
- Figure 1.2:** Symptoms of cassava brown streak disease: a and b) Brown streaks on young stem, c) Different forms of chlorosis on leaves (note progressive symptoms changes as leaf ages), d) constrictions on roots and e) root necrosis. (Adapted from Patil et al. 2015). 11
- Figure 1.3:** Distribution of CBSV and UCBSV in Uganda and Tanzania as detected by RT-PCR. Adapted from Mbanzibwa et al. (2011a). 13
- Figure 1.4:** Comparison of genome organization of UCBSV with other virus members in the family Potyviridae. Potyvirus; Potato virus Y (PVY), Rymovirus ; Ryegrass mosaic virus (RGMV), Tritimovirus; Wheat streak mosaic virus (WSMV); Ipomovirus; Squash vein yellowing virus (SqVYV), Cucumber vein yellowing virus (CVYV) and Ugandan cassava brown streak virus (UCBSV). Adapted from Mbanzibwa et al. (2009a). 21

Chapter 2: Materials and Methods

- Figure 2.1:** 5 µl of HyperLadder 1 were run on the first lane of each agarose gel to determine DNA product sizes and to estimate concentration..... 49
- Figure 2.2:** pJET Plasmid Map. The vector allows for cloning of inserts less than 10 kb into the MCS. 51

Chapter 3: Initial characterization of *Ugandan cassava brown streak virus* (UCBSV) “Kikombe” isolate

- Figure 3.1:** UCBSV infection confirmation in the original CBSD infected cassava plants. Lane 1 and 8 contain 5 µl of HyperLadder (Bioline). Lane 2-6; 440bp RT- PCR product obtained from CBSD infected cassava leaves. Lane 7 is negative control (healthy plant) with no amplification. 66
- Figure 3.2:** NCBI BLAST of the sequence data for the 440 bp RT-PCR fragment. The sequence was identified as the coat protein (CP) of UCBSV in the database..... 67
- Figure 3.3:** Asymptomatic *N. tabacum* (A and B) and *N. benthamiana* (C and D) plants infected with wildtype UCBSV Kikombe isolate..... 69
- Figure 3.4:** Detection of wildtype UCBSV infection in *N. benthamiana* plants: Lane 1 and 8 contain 5 µl of HyperLadder 1. Lane 2 and 3 contain 1000bp PCR product amplified from *N. benthamiana* plants infected with wildtype UCBSV Kikombe isolate. Lane 6 positive control 9 (template from the original UCBSV infected

- cassava plants), Lane 7 contains the negative (healthy plant) with no amplification. 69
- Figure 3.5:** Schematic representation of overlapping genome sequencing strategy for the UCBSV genome: 8 overlapping PCR sections encompassing the UCBSV genome from the 5' UTR to the poly (A) tail were amplified. 72
- Figure 3.6:** Purified PCR amplifications of the different regions in the UCBSV Kikombe genome. **A:** Lane 1 and 6 contains 5 µl of HyperLadder 1, Lane 2 a 2.7 kb PCR fragment amplified with CBSV2F1 and CBSV2R (section 1), Lane 3 a 1069bp PCR fragment amplified by 5'1KBFW and 5'1KBREV (section 2), Lane 4 a 867bp PCR fragment amplified by 5'867bpFwd and 5'867bpRev (section 3), Lane 5 a 1125bp PCR fragment amplified by 5'1125bp F and 5'1125bp R (section 4). **B:** Lane 1 and 6 contains 5 µl of HyperLadder 1, Lane 2 a 1000bp section amplified by Mid 1KB F and Mid 1KB R (section 5). Lane 3 a 1.5Kb PCR fragment amplified with 3'Mid 1.5KB F and 3'Mid 1.5KB R (section 6), Lane 4 a 1400bp PCR fragment amplified 3'1.4KB F and 3'1.4KB R (section 7), Lane 5 a 780bp PCR fragment amplified 922bpFS and poly(A) (section 8). 75
- Figure 3.7:** Alignment of different clones to assemble the genome of the UCBSV genome in Sequencher Software. 4-5 different clones were sequenced for each region of the genome. The accession number of the consensus sequence is GenBank KX753356 77
- Figure 3.8:** Bootstrapped Neighbour-Joining tree (1000 replicates) constructed from the alignment of 30 complete genomes of UCBSV and CBSV in the NCBI database. The new sequenced UCBSV Kikombe isolate (KX753356) is highlighted in red. The two CBSV isolates available at the University of Bristol are highlighted in blue 83
- Figure 3.9:** Phylogenetic tree of the P1 protein sequences of UCBSV Kikombe with other UCBSV and CBSV isolates in the NCBI database. A number of representative potyviruses were included in the analysis. The numbers at the forks are the bootstrap values which indicate the number of times out of 100 times that this grouping occurred after bootstrapping the data. The scale bar shows the number of substitutions per base. 85
- Figure 3.10:** Phylogenetic analysis by Neighbour-joining methods of UCBSV Kikombe HAM1, HAM1 from other UCBSV and CBSV isolates in the NCBI database and the HAM1 in other representative organisms. The unicellular fungus (yeast) *S. cerevisiae*, the plants rice (*Oryza sativa*) and *Arabidopsis thaliana* and animals mouse (*Mus musculus*), domestic cow (*Bos taurus*) and human (*Homo sapiens*) were included in the analysis. 86

Chapter 4: Construction of an *in-vitro* and *in-vivo* infectious clone of UCBSV “Kikombe” isolate

Figure 4.1: PCR amplification of different sections of UCBSV Kikombe sections for yeast recombination; **A:** Lane 1 5µl of HyperLadder 1, Lane 2 section 1 (2700bp). **B:** Lane 1 contains 5µl HyperLadder 1, Lane 2 section 2 (1069bp) Lane 3 section 4 (1125bp), Lane 4 section 3 (867bp), Lane 5 section 6 (1500bp), Lane 6 section 7 (1400bp) **C:** Lane 1 5µl of HyperLadder 1, Lane 2 section 5 (1041bp). 97

Figure 4.2: **A:** Constructed plasmid of pYES2.1 with 5’ end UCBSV genome. The blue arrow represents 5’ end (sections 1, 2, 3, and 4) of the UCBSV genome and the pink arrow represents pYES2.1 backbone vector. The Xho1 restriction sites present in the constructed plasmid are also shown on the map **B:** Xho1 restriction, Lane 1 HyperLadder 1, Lane 2 and 3 Xho1 restriction of two different E. coli plasmid extractions on the constructed pYES2.1 containing section 1,2,3 and 4 (5’ end half of the UCBSV genome)..... 99

Figure 4.3: **A:** Constructed plasmid of pYES2.1 with 3’ end UCBSV genome. The blue arrow represents the 3’ end (sections 5, 6, 7 and 8 of the UCBSV genome). The pink arrow represents the pYES2.1 backbone vector. The Spe1 restriction sites present in the plasmid are also shown on the map. **B:** Lane 1 contains 10 µl of Spe1 restriction reaction on the constructed pYES2.1 containing section 5,6,7 and 8 (3’ end half of the UCBSV genome). Lane 2 Spe1 restriction on the backbone pYES 2.1 vector backbone. 100

Figure 4.4: Constructed plasmid of pYES2.1 containing the complete genome of UCBSV (Infectious clone plasmid): The blue arrow represents the complete UCBSV genome from base position number 1 to the end of the poly (A) tail (9096 nucleotides). The pink arrow represents the pYES2.1 backbone vector. The recombinant plasmid is the infectious clone plasmid of UCBSV cloned inside pYES2.1. The Xho1 restriction sites present in the plasmid are shown on the map. The Kpn1 and Not1 restriction sites are also shown. 102

Figure 4.5: Summary of schematic flow of steps taken to construct UCBSV *in-vitro* infectious clone. F1 to F8 refer to the eight sections that were homologously recombined in yeast. The complete UCBSV genome was initially divided into two halves which were separately recombined in pYES2.1 vector (pYES2.1 5’ end UCBSV and pYES2.1 3’ end UCBSV). Once stability was confirmed, the separate halves were recombined to generate pYES2.1 UCBSV genome. 103

Figure 4.6: Checking the presence of all the recombined sections in the UCBSV infectious clone: Lane 1 and 18 contain 5 µl of HyperLadder 1. Lane 2 is section 1 (2700bp) amplified from infectious clone plasmid and Lane 3 is positive control for section 1. Lane 4 is section 2 (1069bp) amplified from infectious clone plasmid and Lane 5 is positive control for section 2. Lane 6 is section 3 (867bp) amplified from infectious clone plasmid and Lane 7 is positive control for section 3, Lane 8 is section 4 (1442bp) amplified from infectious clone plasmid and Lane 9 is positive control for section 4, Lane 10 is section 5 (1044bp) amplified from infectious clone plasmid and Lane 11 is positive control for section 5, Lane 12 is section 6 (1500bp) amplified from infectious clone plasmid and Lane 13 is positive control for section 6, Lane 14 is section 7 (1400bp) amplified from infectious clone

- plasmid and Lane 15 is positive control for section 7, Lane 16 section 8 (922bp) and lane 17 is positive control for section 8. 104
- Figure 4.7:** Restriction analysis to check the integrity of the different clones of constructed UCBSV infectious clone. **A:** KpnI/Not I restriction, Lane 1 and 6 contain 5 µl of HyperLadder 1. Lane 2 contains 5 µl of undigested infectious clone plasmid. Lanes 3, 4 and 5 contain three different clones showing the 10 kb UCBSV genome and the 6 kb pYES2.1 backbone vector released from the KpnI/NotI restriction digestion on the infectious clone plasmid. **B.** XhoI restriction. Lane 1 and 6 contain 5 µl of HyperLadder 1, Lanes 2, 3 and 4 contain three different clones of XhoI restriction on different clones of UCBSV infectious clone showing the expected sizes. Lane 5 contains pYES2.1 backbone plasmid. 105
- Figure 4.8:** Schematic view of the UCBSV infectious clone genome assembly in Sequencher. Each region of the pYES2.1 containing the UCBSV genome was sequenced 4-5 times and the overall consensus sequence assembled. The assembled infectious clone genome has been submitted to NCBI under GenBank accession KX753357. 107
- Figure 4.9:** *In-vitro* transcription reactions inoculated on plants. Lane 1 contains 5 µl of Hyper Ladder Lane 2 and 3 contain 10 µl of *in-vitro* reactions containing the RNA transcripts. Lane 4 is the positive control *in-vitro* reaction for the Promega Kit. Lanes 5 is the *KpnI* digested backbone template used in the *in-vitro* transcription reactions. 111
- Figure 4.10:** Symptoms on systemic leaves of the infectious clone infected plants: **A:** *N. benthamiana* inoculated with *in-vitro* RNA transcripts (Infectious clone virus) showing mild mosaic symptoms, **B:** *N. benthamiana* infected with *in-vitro* RNA transcripts (Infectious clone virus) with mild mosaic symptoms, **C:** *N. benthamiana* plant inoculated with wildtype UCBSV inoculum (Positive control) **D:** *N. benthamiana* plant rubbed with potassium phosphate buffer (Negative control)..... 112
- Figure 4.11:** RT-PCR on the infectious clone infected plants: Lane 1 and 10 contain 5 µl of HyperLadder 1, Lane 2: RT-PCR product using cDNA from infectious clone infected *N. benthamiana* plant inoculated with *in-vitro* infectious clone, Lane 3 and 4: RT-PCR product using cDNA from *N. clevelandii* plant inoculated with *in-vitro* infectious clone, Lane 5 and 6: wild type UCBSV infected *N. benthamiana* plant, Lane 7: Positive control. Lane 8 and 9: Negative control plants..... 114
- Figure 4.12:** Multiple alignment of the sequences from UCBSV Infectious clone infected *N. benthamiana* plants aligned/compared against the genomes of the assembled wild type UCBSV genome and Infectious clone genome. Section 3 (867bp) of UCBSV genome was targeted. Kikombe wild is the genome of the wildtype UCBSV isolate, Infectious c is the genome of the UCBSV infectious clone. I.C Beth pla is the sequence recovered from RT-PCR and sequencing from 2 different *N. benthamiana* plants infected with the UCBSV infectious clone..... 115
- Figure 4.13:** Confirming UCBSV infection in cassava plants. **A:** Amplification of section 5 (Position 4758-5798) approximately 1 kb fragment of UCBSV from cassava plants. Lane 1 contains 5 µl of HyperLadder 1, Lane 2-5 contain 1000bp RT-PCR amplification products from the four different cassava plants infected with

UCBSV infectious clone. **B:** Amplification of RuBisCo from four cassava plants amplifying a 619bp PCR product. 117

Figure 4.14: Systemic leaves of cassava plants infected with UCBSV infectious clones showing characteristic mosaic symptoms. **A:** Healthy negative control plant. **B-G:** Leaves from different cassava plants inoculated with UCBSV infectious clone. 117

Figure 4.15: Multiple alignment of sequences from UCBSV infectious clone infected cassava plants. The 1000bp RT-PCR product amplified from two cassava plants infected with UCBSV infectious clone: The sequences aligned in the expected regions of the UCBSV wildtype and infectious clone genomes. Wildtype UCB is the genome of the wildtype UCBSV Kikombe isolate, UCBSV infect is the genome of the UCBSV infectious clone. Cassava I.C is the sequence recovered from RT-PCR of cassava plants infected with the UCBSV infectious clone. 118

Figure 4.16: UCBSV *in-vivo* clone map: The UCBSV genome is represented by a blue arrow. The pink arrow represents the yeast adapted pCAMBIA0380 backbone vector. The Kan resistance, 2 μ m Ori and *ura3* genes is a separate red arrows on the map. 120

Figure 4.17: *HindIII/EcoRI* restriction of extracted plasmid. Lane 1 and 14 contain 5 μ l of HyperLadder1, Lanes 2, 3, 4, 5, 7 and 8: different restrictions from different clones showing the expected 700bp. Lanes 6, 9 and 10 released an unexpected 400bp fragment. Lane 12 and 13 contain undigested circular pCAMBIA0380YA backbone plasmid. 121

Figure 4.18: *XmaI* and *KpnI* restriction of the pYES2.1 containing UCBSV plasmid with created *XmaI* restriction site at the 5' end of genome. Lane 1 and 5 contain 5 μ l of HyperLadder1. Lane 2, 3 and 4 contain different clones of *XmaI* and *KpnI* restricted plasmids. 122

Figure 4.19: Intron insertion in the different genomic regions of UCBSV: Introns were inserted in the P3, CI and NIb regions of UCBSV genome. 124

Figure 4.20: Amplification of intron containing PCR products from UCBSV infectious clone: **A:** Lane 1 and 7 contain 5 μ l of HyperLadder1: Lane 2, 3, 4, 5 and 6 contain intron containing UCBSV PCR products. 124

Figure 4.21: Undigested plasmids recovered from the *E. coli* cells after transferring the intron containing UCBSV sections into yeast adapted pCAMBIA0380. The plasmids recovered were of incorrect sizes (less than 10 kb) whilst the expected size of the recombinant plasmids was over 10 kb. 125

Chapter 5: Characterization of P1 and HAM1 proteins in UCBSV “Kikombe” isolate

- Figure 5.1:** The LRRA motif in the aligned amino acid partial P1 sequence of UCBSV Kikombe and other published sequences of CBSV and UCBSV isolates. The conserved LRRA motif that was targeted for mutation in this study is underlined. Accession numbers for the isolates are; UCBSV Kikombe (KX753356), Strain Ugandan (FJ185044), Ug 23 (FN434109), MLB3 (FJ039520), Ma 42 (FN433932), Ke 125(FN433930), Ke 54 (FN 433931), Mo 83 (FN434436), Strain Tanzanian (GQ329864). 135
- Figure 5.2:** Amplification of PCR fragments from UCBSV infectious clone to generate new modified sections for introducing LRRA to LAAA mutation. Lane 1 and 6 contain 5 µl of Hyperladder 1. Lane 2 contains 5 µl of 459bp Section 1a (mutated P1). Lane 3 contains 5 µl of 2254bp Section 1b (mutated P1). Lane 4 contains 5 µl of 2497bp Section 2, 3and 4. Lane 5 contains 5 µl of 4312bp Section 5, 6, 7 and 8. 138
- Figure 5.3:** ELISA analysis of viral titres between the plants infected with P1 mutated UCBSV infectious clone and wild type infectious clone showing the levels of UCBSV viral protein titres with time. 140
- Figure 5.4:** RT-PCR amplification of P1 protein from plants infected with P1 mutated UCBSV I.C and wildtype UCBSV I.C. Lane 1 and 7 contain 5 µl of HyperLadder 1, Lane 2 and 3 1000bp RT-PCR product from *N. benthamiana* plants infected with wildtype UCBSV infectious clone. Lane 4 and 5 1000bp RT-PCR product from *N. benthamiana* plants infected with mutated UCBSV infectious clone. Lane 6: Healthy mock inoculated *N. benthamiana* plant. 141
- Figure 5.5:** Alignment of the sequences from RT-PCR of plants infected with wildtype P1 UCBSV infectious against those infected with the LAAA mutated P1 UCBSV infectious clone. The mutated UCBSV infectious clone still has the mutation. . 142
- Figure 5.6:** Schematic representation of genome of UCBSV P1 knockout virus: The complete P1 coding sequence was deleted from the genome..... 143
- Figure 5.7:** PCR confirmation of deletion of P1 protein in the UCBSV infectious clone: Lane 1 and 10 contains 5 µl Hyperladder 1: Lane 2, 3,4,5,6 and 7: 1700bp PCR products generated from different screened plasmids recovered from UCBSV P1 knockout plasmid extractions. Lane 8: 2700 bp PCR product obtained from the P1 containing UCBSV infectious clone. Lane 9 negative control with no amplification. 144
- Figure 5.8:** Symptoms in different indicator plants infected with UCBSV P1 knockout I.C: A: *N. benthamiana* plant infected with P1 UCBSV Knockout infectious clone B: *N. benthamiana* plant infected with full-length UCBSV infectious clone, C: *N. clevelandii* plant infected with P1 UCBSV knockout infectious clone D: *N. clevelandii* plant infected with full-length UCBSV infectious clone. E: *N. tabacum* plant infected with P1 knockout infectious clone.F: *N. tabacum* plant infected with full-length UCBSV infectious clone. G: Healthy mock inoculated *N. benthamiana* plant, H, Healthy *N. clevelandii* plant mock inoculated with buffer. 145
- Figure 5.9:** Amplification of PCR products from P1 UCBSV knockout I.C infected plants: Lane 1and 12 contain HyperLadder 1, Lane 2, 3, 4 and 5 contain PCR products

from *N. benthamiana* plants infected with P1 knockout plasmid. Lane 6 and 7 contain PCR products from *N. clevelandii* infected plants infected with P1 knockout plasmid. Lane 8 and 9 contain PCR products from *N. tabacum* plants infected with P1 knockout plasmid. Lane 10 contains the PCR from *N. benthamiana* plant infected with wildtype UCBSV infectious clone. Lane 11 is the positive control plasmid..... 146

Figure 5.10: (A) Constructed recombinant pCAMBIA2300 with P1 protein for agroinfiltration assays. (B) Schematic representation of the different constructs used in the agroinfiltration assays: A: pCAMBIA2300 cloned with the wildtype P1 UCBSV, B: pCAMBIA2300 cloned with LAAA mutated P1 UCBSV, C: pCAMBIA2300 cloned with positive control PVY HC-Pro, D: pCAMBIA2300 cloned with GFP for co-infiltration. 147

Figure 5.11: Restriction analysis of recombinant colonies digested with *Bam*H1 and *Kpn*1; Lane 1 contains 5 µl of HyperLadder1. Lane 2 and 4 contain undigested recombinant pCAMBIA 2300 cloned with LAAA mutated P1. Lane 3 and 5 contain *Bam*H1/*Kpn*1 restrict digested recombinant pCAMBIA 2300 cloned with LAAA mutated P1, a 1000bp product was released on restriction digestion of recombinant pCAMBIA2300 with *Bam*H1 and *Kpn*1 restriction enzymes. 148

Figure 5.12: Silencing suppression of wildtype UCBSV P1 and LAAA mutated P1: WT: Wildtype *N. benthamiana* plants. LAAA mutated P1 UCBSV; No GFP green fluorescence observed with the mutated P1 infiltrated leaves, P1 UCBSV; GFP green fluorescence observed in the wildtype UCBSV P1 protein infiltrated leaves, HC-Pro; Positive control, GFP; Negative control, P1 Nampula; No green fluorescence from P1 of CBSV Nampula isolate (internal negative control). 150

Figure 5.13: Schematic presentation of the genomes of the different UCBSV HAM1 knock out constructs. A: Genome of full-length UCBSV infectious clone, B: Genome of HAM1 full deletion UCBSV infectious clone. C: Genome of HAM1 partial deletion UCBSV infectious clone..... 151

Figure 5.14: PCR to confirm the integrity of the UCBSV HAM1 knockout constructs: Lane 1: 5 µl of HyperLadder 1. Lane 2: 700bp PCR product from the HAM1 full deletion construct confirms actual deletion of HAM1. Lane 3: 1500bp PCR product from HAM1 containing UCBSV infectious clone. A reduction in the product size from 1500bp in the HAM1 containing UCBSV infectious clone to 700bp in the HAM1 full-deletion UCBSV infectious clone confirms deletion of HAM1. 152

Figure 5.15: Plasmids recovered from *E. coli* cells for the different infectious clone constructs: Lane 1: 5 µl of HyperLadder 1, Lanes 2, 3, 4 and 5 UCBSV HAM1 full deletion plasmids. Lanes 6, 7, 8 and 9 UCBSV HAM1 partial deletion plasmid. Lanes 10, 11, 12 and 13 full-length UCBSV infectious clone plasmid.153

Figure 5.16: Symptoms on plants infected with UCBSV HAM1 knockout infectious clones: A: HAM1 partial deletion *N. benthamiana* infected plant. B: HAM1 full deletion *N. benthamiana* infected plant. C: UCBSV full-length infectious clone infected *N. benthamiana* plant. D: Healthy mock inoculated *N. benthamiana*..... 154

Figure 5.17: RT-PCR on HAM1 knockout infected plants. Lane 1: Contains 5 µl of HyperLadder 1. Lane 2 and 3 1000bp RT-PCR product from UCBSV full length infected *N. benthamiana* plants. Lane 4 is *N. benthamiana* plant infected with

HAM1 full deletion UCBSV infectious clone. Lane 5 <i>N. benthamiana</i> plant infected with HAM1 partial deletion UCBSV infectious clone. Lane 6 is the positive control plasmid.....	155
Figure 5.18: Comparison of amino acid sequence between UCBSV Kikombe HAM1 and CBSV Nampula HAM1. The overall amino acid identity between the two isolates was 45%.....	156
Figure 5.19: Transgenic CBSV HAM1 lines of <i>N. tabacum</i> infected with different constructs: A and B: HAM1 full deletion UCBSV infectious clone infected plants, C and D: HAM1 partial deletion UCBSV infectious clone infected plants, E and F: Full length UCBSV infectious clone infected plants. G and H: Mock infected plants.....	157
Figure 5.20: RT-PCR to detect UCBSV infection in transgenic HAM1 lines: Lane 1 and 10: HyperLadder 1. Lane 2: 867bp PCR product from transgenic line CBSV HAM1 <i>N. tabacum</i> infected with UCBSV HAM1 full deletion infectious clone. Lane 3: Transgenic line CBSV HAM1 infected with UCBSV HAM1 full deletion infectious clone. Lane 4 and 5: HAM1 transgenic lines infected with UCBSV HAM1 partial deletion infectious clone. Lane 6 and 7: HAM1 transgenic lines infected with full length UCBSV infectious clone. Lane 8: <i>N. tabacum</i> non transgenic plant infected with full-length UCBSV infectious clone. Lane 10: Positive control plasmid.	158

Chapter 1: Introduction

1.1 Plant pathogens and effects on food security

With the increasing world population estimated to reach 9.7 billion by the year 2050, there is an urgent need for measures to be put in place for the provision of food for the growing population (Melrose *et al.*, 2015). Most of the developing countries especially in Africa are at risk if food production does not increase by 2050. However, food production has been challenged by a number of factors including climate change, pests and plant pathogens that affect most of the cultivated crops that feed most of the global population. The diseases affecting crops include viral, bacterial and fungal diseases (Alfano & Collmer, 1996). Current measures to control some of the effects of plant pathogens on crop production have included generating disease resistant crops through breeding and understanding the pathogenicity of the disease pathogens in order to devise control strategies.

Cassava brown streak disease (CBSD) caused by the *Cassava brown streak virus* (CBSV) and *Ugandan cassava brown streak virus* (UCBSV), the latter of which is the pathogen studied in this thesis, causes high yield losses in cassava production in Africa (Patil *et al.*, 2015). In order to control CBSD, a number of breeding programs have aimed to generate disease resistant varieties of cassava (Yadav *et al.*, 2011; Patil *et al.*, 2011; Wagaba *et al.*, 2016). The wild relatives of cassava *Manihot galziovii* have been used as sources of resistance for generating resistant varieties (Allem, 2002). However, there is a need to screen the generated resistant varieties for disease resistance. Infectious clones of RNA viruses provide a homogenous and reproducible source of viral infection for effective screening of breeding lines. In this PhD work, infectious clones of *Ugandan cassava brown streak virus* (UCBSV) were generated and

will ultimately be used for screening breeding lines and understanding the pathogenesis of this virus and thus devise appropriate control measures for CBSD.

1.2 Viruses

Identification of viruses as causal agents for disease dates back to the work of Martin Beijerinck in 1898 who recognized that viruses were contagious elements as they could easily be transmitted from non-diseased plants to healthy plants (Beijerinck, 1898). Electron microscopic analysis of some viruses shows them to appear as flexuous particles. Koch's postulates, in which the theory of causation is studied showing the steps in which a given organism can be considered as causal, has been applied in understanding causation theory for viruses. Koch's postulates involves the isolation of an organism, growth of the organism in pure culture, introducing this organism into healthy plants and then the isolated organism is expected to cause similar disease. Using Koch's postulates, it was realized that the noxious organisms thought to cause disease were called "*contagium vivum* or *contagium vivum fixum*"; and they were liquid or soluble and not corpuscular as they could diffuse through agar (Beijerinck, 1898). The "*contagium vivum*" was thus regarded as a "virus". Viruses reside inside the cells of the host and use the host biosynthetic apparatus for their replication and because of this, viruses may thus be considered as "not living" thus the term "virus life cycle" could as well be referred to as "virus infection cycle" (Hull, 2014). Because of the fact that the viruses utilize the host factors for their successful transmission, it is important that the virus does not compromise and destroy the host too quickly as this acts to their disadvantage. It is for this reason that a number of mechanisms have evolved in the transcription and translation processes of the virus in relation to the host to ensure that the host is maintained in a living state as long as possible for the continual survival of the virus.

1.3 Properties of plant viral genomes

There is a variation in the genetic composition of different viruses with some viruses encoding single-stranded or double-stranded, DNA or RNA which can be linear or circular. Single-stranded DNA viruses infecting plants include the viruses in the family *Geminviridae* like the *African cassava mosaic virus* (ACMV) and *Tomato yellow leaf curl virus* (TYLCV) (Scholthof *et al.*, 2011). Double-stranded DNA viruses include *Cauliflower mosaic virus* (CaMV) (Pallas & Garcia, 2011). Most of the plant-infecting viruses are single-stranded RNA viruses and the nucleic acid can either be (+) sense ssRNA like the *Tobacco mosaic virus* (TMV) or (-) sense ssRNA like the *Tospoviruses* (Baltimore, 1971). For the (+) sense RNA viruses, the genome is effectively one large mRNA or segment allowing translation of some proteins before any replication occurs. However for all the viruses with the different types of nucleic acids, translation normally goes through the mRNA so the viruses encode a number of different enzymes involved in the transcription to mRNA which include, RNA dependent RNA polymerases (RdRp) or DNA dependent RNA polymerases depending on genome structure among other enzymes. The genomes of plant viruses may exist as a single piece of nucleic acid in which case they are referred to as monopartite genomes or the genome may exist as two or more different pieces as is the case for bipartite genomes. For the bipartite genomes, it may be important to determine if there are two distinct pieces of the genome. This distinction can be done by hybridization experiments estimating sequence complexity. However, to confirm that the genome of a virus exists in two pieces of nucleic acids, this can be done by cloning the full-length of each piece of nucleic acid as separate sections and demonstrating that each section of the genome is required for infection to occur (Hamilton *et al.*, 1983).

1.4 Cassava

Cassava (*Manihot esculenta* Crantz Family *Euphorbiaceae*) (Figure 1.1) is an essential staple food to approximately 800 million people around the world (Thresh, 2006). Cassava has different names in different communities, some of the known names for cassava include manioc, yucca, or tapioca (Bredeson *et al.*, 2016). In most of the developing countries, especially in Africa, cassava is the third largest source of carbohydrate in the diet after rice and maize. It is mainly grown as a subsistence crop by farmers, however in some countries, cassava production is also for industrial purposes which includes production of starch or biofuel among other applications (Thresh, 2006).

Cassava is presumed to have originated from South America (Brazil) and was introduced East Africa via West Africa (Hillocks, 1997). The reason for the assumption that cassava originated from Brazil is because most of the *Manihot* species have their centre of biodiversity in this region (Olsen & Scaal, 1999). Spread of cassava from the centre of origin was mainly by migration of Portuguese navigators who took cassava from Brazil to west coast of Africa (Jones, 1959) and later to East Africa through Madagascar (Patil *et al.*, 2015). Cassava is known to have originally been a wild shrub *Manihot esculenta flabellifolia* (Olsen & Scaal, 1999; Allem, 2002) and there were a number of modifications that occurred in the domestication of the crop which included changes in the storage roots becoming tubers. It is also known that the wild relatives of cassava have many flowers while the domesticated cassava has few flowers (Nassar, 2002).



Figure 1.1: Cassava plant showing the tuberous roots, stem and palmate compound leaves
(Source <https://rauman.wikispaces.com/cassava>)

1.4.1 Cassava production

Different countries produce cassava for different purposes but the estimated world annual yield is about 270 million tonnes (FAOSTAT, 2014). Africa is known to be largest centre of production contributing 54.5% (145 million tonnes). In Brazil, cassava production is mainly for industrial purposes while cassava is mainly produced for export in Thailand, but in Africa, cassava is mainly grown as a food crop for consumption by most farmers (Fauquet & Fargette, 1990). In most African countries, cassava is grown mainly by traditional small scale farmers but there are some industrial plantations in Liberia, Nigeria, and Togo. Nigeria is known to be world's largest producer of cassava while Thailand is the largest exporter of dried cassava with a total of 77% of world export in 2005.

There is a variation in the amount of time it takes for the cassava plants to achieve their maximum yield, although for most of the grown cultivars harvesting can be done from 6 months after planting, although there are some varieties that may take 18 months before they can be harvested. With the current improvements in plant breeding of cassava varieties, most of the improved varieties are able to attain their maximum yield between 12-15 months after planting (El-Sharkawy, 2003). With the increasing world population, there is a need to increase production of affordable food crops to cater for the growing population, so it is for this reason that total production of cassava in Africa increased from 55 million tonnes in 1984 to over 145 million tonnes in 2014. African farmers usually grow cassava in monoculture or together with maize (*Zea mays*) and bananas (*Musa acuminata*), peanuts (*Arachis hypogea*), rice (*Oryza sativa*), beans (*Phaseolus vulgaris*) or other crops.

Furthermore, cassava has the characteristic of being a hardy crop in addition to being a drought tolerant crop. With these favourable characteristics, cassava has been considered as a food security crop for the developing world. Cassava is able to grow and yield in conditions where other crops fail to grow as it tolerates drought, is able to grow in soils with low nutrient capacity, but also responds well to higher rainfall conditions or to irrigation and fertilizers (El-Sharkawy, 2003).

1.4.2 Cassava propagation

Cassava is propagated vegetatively using stems (stakes). In this case, there is need to obtain enough planting material for each planting season. In order to achieve this, several cuttings or plantings are obtained from a single mother plant at the age of 8-18 months (Keating *et al.*, 1988). This mode of propagation requires obtaining bulky planting materials and in some cases, the cuttings deteriorate before they are planted especially if they are not adequately dried

immediately after harvest. For this reason, there is need to find alternative modes of propagation for cassava which include being able to propagate cassava by seeds which allows for outcrossing and genetic variation as they would give genetically variable progeny, however, no commercially viable propagation system via true seeds is currently available (Leihner, 2002).

Furthermore, because farmers are able to use cuttings from one planting season for the next planting season, this may sometimes lead to accumulation of viruses in the fields as the plants in the previous planting season may have been infected with viruses and this is carried over to the next planting season (Legg *et al.*, 2015).

Approaches to obtain clean cassava plants and reduce viral diseases include use of tissue culture based methods such as thermotherapy, cryotherapy and chemotherapy (Kanju *et al.*, 2002). Also farmers try to select 'the best'. In addition, tissue culture allows for massive propagation of planting material proved to be virus free. It is recommended that most of the planting materials are virus indexed before they are planted (Ndunguru & Kapinga, 2007), however, virus indexing is expensive and not readily available to subsistence farmers. It is important that virus free planting material is used because planting infected cuttings has an effect on the overall yield obtained at harvesting which further affects the production output of cassava.

1.5 Challenges to cassava production

The low average yields in cassava production are caused by a number of factors among which are susceptibility of the commonly grown varieties to major diseases and pests and variability in climate patterns. Among the major diseases, viral diseases are the most important in tropical Africa although there are cases of problems with bacterial diseases like cassava bacterial blight

(CBB) and fungal diseases. Some of the major pests affecting cassava plantations include cassava green mites, elegant grasshopper, cassava mealybug and root mealybug among others (Hillocks & Jennings, 2003). Of the viral diseases affecting cassava production, cassava brown streak disease (CBSD) and cassava mosaic disease (CMD) are the major diseases. CBSD is estimated to cause up to 70% yield loss and thus threatening the livelihoods of farmers in Africa (Hillocks & Jennings, 2003). Some of the impacts of CBSD on the affected plants include loss in root weight but the most important economic effect of the disease is root necrosis which makes the storage roots unsaleable (Hillock *et al.*, 2001). CBSD is caused by either *Cassava brown streak virus* (CBSV) or *Ugandan Cassava brown streak virus* (UCBSV). The viruses are transmitted by the white fly vector, *Bemisia tabaci* (*Gennadicus*) in a non-persistent manner. The vector *Bemisia tabaci* (*Gennadicus*) was confirmed in experiments in which the CBSV virus could be successfully transmitted by the whitefly between cassava plants (Maruthi *et al.*, 2005). In order to control CBSD, there is a need to understand how the disease is manifested and with this, be able to develop measures against the disease. A number of methods have been put in place to control the spread and damage caused by CBSD.

Some of the available options to control virus spread include the use of chemicals to control virus vectors (insects). Because the cassava crop is vegetatively propagated, the virus can also be spread by infected propagating cuttings and by tools used during propagation (Maruthi *et al.*, 2005). Vertical transmission can be reduced by removing possible reservoirs for the viruses (for instance by roguing infected plants) and planting disease free cuttings. Alternatively transgenic lines with resistance to CBSD have been generated using RNAi constructs (Patil *et al.*, 2015), but these are not yet available commercially and being genetically modified, they may not be acceptable in all countries.

1.5.1 CBSD symptoms

CBSD is often unrecognized in the field because of the inconspicuous foliar symptoms compared to those of cassava mosaic disease (CMD) and cassava green mite (CGM). Symptoms of CBSD on the leaf can take different forms depending on the cassava variety, age of the plant and the conditions under which it is growing (Hillocks & Jennings, 2003). The CBSD gets the name from the term “brown streak” referring to the brown lesions that are observed on the cortex after stripping off the bark of the stem in infected plants (Hillocks & Jennings, 2003). Plants infected with CBSV/UCBSV sometimes show symptoms of the disease on the leaves, stems and root tubers (Figure 1.2). However, in the naming of the CBSD, Storey, (1936a) only considered the brown streaks of the stem, which according to Nichols, (1950) are not the most characteristic symptoms of CBSD (Hillocks *et al.*, 2001).

Root necrosis (Figure 1.2e) and radial constrictions of roots (misshapen) (Figure 1.2d) are two root symptoms of economic significance associated with CBSD as they can reduce the overall yield by over 70% (Hillocks *et al.*, 2001). Formation of root lesions (necrosis) makes the root unconsumable and it may cause a complete loss of crop (Storey, 1936a) and this is especially true for the susceptible varieties like “Albert” where starch tissues are fully affected. However, in some CBSD resistant cassava cultivars like “Kaleso”, the infected plants are symptomless (Maruthi *et al.*, 2014). Root necrosis symptoms are only observed after the crop is harvested but it greatly affects the marketability of cassava. In some cases, plants without aerial leaf symptoms may present with root necrosis symptoms (Hillocks & Jennings, 2003). Root symptoms may also appear as pits in the surface back of the tuber (Figure 1.2d) although the necrotic symptoms are only observed after the root is cut. Plants without foliar symptoms for CBSD may still show root necrosis at harvesting, which is otherwise common for plants with foliar symptoms (Hillocks *et al.*, 1996). Storey, (1936a) described the colour of root lesions as

brown while Nichols, (1950) suggested that dark brown (sepia) was most common, other colours such as yellow streaks and green necrosis were also associated with root necrosis.

According to Nichols, (1950), there are two types of leaf symptoms (Figure 1.2c) which are both characterized by some forms of chlorosis that occur on mature lower leaves. The first type is the yellowing that occurs first along the secondary veins and later spreads to tertiary veins. The second type is associated with mottling in cases of mild infection but the lamina may largely turn chlorotic and eventually the entire leaf is blotched for severely infected cassava plants (Nichols, 1950). The second type of leaf symptoms is not associated with the veins. Lack of distortion of leaves in both types of CBSD leaf symptoms distinguishes the CBSD from CMD foliar symptoms. However, leaf symptoms are not always correlated with yield reduction or tuber necrosis (Hillocks & Jennings, 2003).

Stem lesions have been observed on CBSD affected plants (Figure 1.2 a and b), according to (Hillocks & Thresh, 2000), the purple or brown lesions appear on the exterior surface of young green stems. This is followed by formation of sepia necrotic lesions in the leaf scars (Nichols, 1950). The other notable stem symptom is the death of node and internodes which results in the so called “die back” (Nichols, 1950).

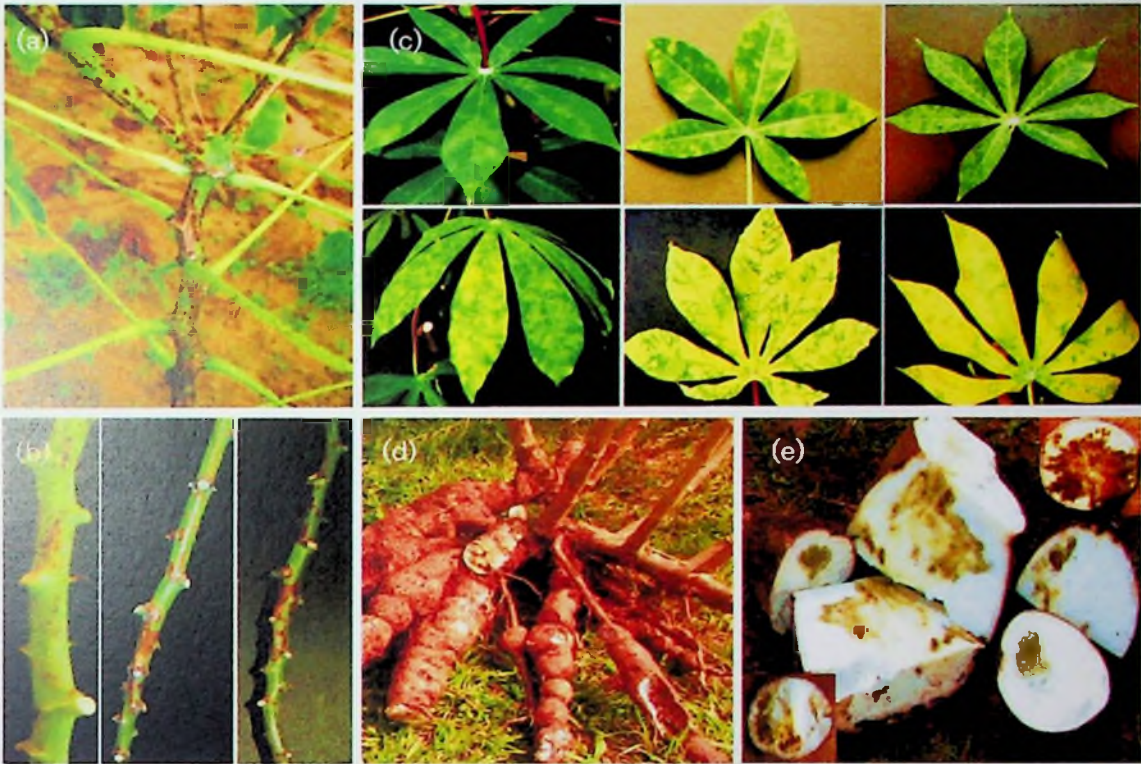


Figure 1.2: Symptoms of cassava brown streak disease: a and b) Brown streaks on young stem, c) Different forms of chlorosis on leaves (note progressive symptoms changes as leaf ages), d) constrictions on roots and e) root necrosis. (Adapted from Patil *et al.*, 2015).

1.6 Diagnosis of cassava brown streak disease

Cassava brown streak disease (CBSD) diagnosis using symptoms is complicated by the fact that some infected plants may not present with any phenotypic disease symptoms on the leaves and furthermore, there is a variation in disease symptom presentation depending on the cultivar of cassava that is infected with CBSD (Hillocks *et al.*, 1996; Hillocks & Jennings, 2003). In some cases, there is slow speed of presentation of symptoms and later recovery of foliage. In addition, a number of other factors affect the symptom outcome which include the environmental conditions in which the cassava plant is growing, temperature, age and soil fertility among others (Nichols, 1950; Lister, 1959). For instance, it is observed that infected cassava plants may lose most of their leaves during the dry season and the new growth leaves

may show no sign of the disease but still present with infected tubers. With the challenges in diagnosing CBSD using symptoms, there is a need to develop more reliable and reproducible tools that can allow for easy and efficient diagnosis of the disease and hence enable sustainable measures to be put in place for the control of the spread of CBSD (Hillocks & Jennings, 2003). In order to control the spread of a given virus, there is need to have information about the virus, the different variants or strains that exist and their distribution, if reliable resistance breeding methods are to be adequately developed.

1.6.1 Distribution of cassava brown streak disease

Despite the fact that CBSD was first reported from the coastal region of Tanzania in the 1930's (Storey, 1936b), the disease was not initially considered as a serious threat, compared to other diseases affecting cassava like cassava mosaic disease (CMD). This is because CBSD was originally restricted to low land areas of East Africa (Nichols, 1950; Hillocks & Jennings, 2003; Hillocks *et al.*, 1999), however, as of 2004, CBSD has been reported to be spreading rapidly in East and Central Africa, thus affecting the food security of a number of cassava farmers (Alicai *et al.*, 2007; Legg *et al.*, 2011). Cassava brown streak disease is believed to have originated from East Africa and is now spreading westwards towards Central and West Africa (Legg *et al.*, 2011).

Interest in CBSD dates back in the late 1980's when the symptoms were observed in Tanzania and Malawi. Extensive surveys on distribution of cassava virus diseases conducted in Tanzania (Legg & Raya, 1998) and Zanzibar (Thresh and Mbwana, 1998) showed that CBSD was found to occur along the coast, mainly at altitudes <1000m where the incidence was less than 36%. With the recent advent in a number of molecular diagnostic tools for CBSD (Mbanzibwa *et al.*, 2011), the two causal viruses *Cassava brown streak virus* (CBSV) and *Ugandan cassava*

brown streak virus (UCBSV) have been found to be rapidly spreading in East Africa and with some regions having mixed infections of the two viruses (Figure 1.3). There are reports of the disease in Burundi, Rwanda, DR Congo and Malawi (Bigirimana *et al.*, 2011; Legg *et al.*, 2011; Mahungu *et al.*, 2003; Mbewe *et al.*, 2015).

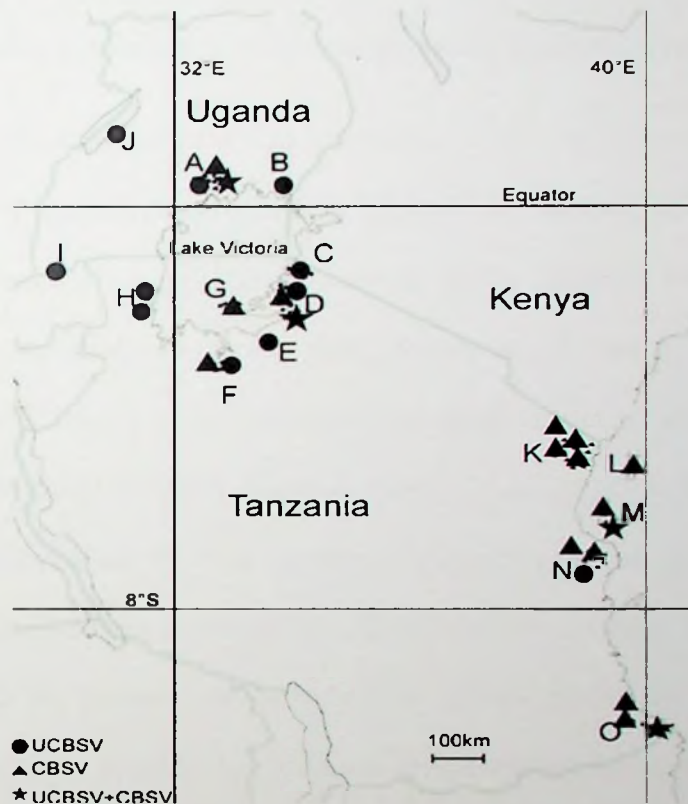


Figure 1.3: Distribution of CBSV and UCBSV in Uganda and Tanzania as detected by RT-PCR. ●: Represents regions with UCBSV infections, ▲: Represents regions with CBSV infections whilst ★: Represents regions with mixed infections of UCBSV and CBSV. Adapted from Mbanzibwa *et al.*, (2011a).

1.6.2 Characterization of cassava brown streak viruses

Initial work on the identification and characterization of the virus responsible for CBSD involved virus purification from infected plants and using electron microscopy to try and identify the virus. Basing on the observed *Carlavirus* of length (650 ± 690 nm) particles, CBSD was thought to be caused by a *Carlavirus* (Brunt *et al.*, 1996). Lennon *et al.*, (1986) revealed

the presence of pinwheel inclusion bodies in the cytoplasm of CBSD infected leaves, a characteristic which had been reported for potyviruses. It was then deemed important to determine if CBSD was caused by a *Carlavirus* or a *Potyvirus* or a virus complex of both. However, most of the experiments to clarify the actual identity of the CBSD causal virus using herbaceous hosts were unsuccessful. In the work of Monger *et al.*, (2001), the coat protein sequence of the cassava brown streak virus was sequenced and it was proved to be a *Potyvirus*.

Recently, a number of serological and molecular tools have been developed for the identification and characterization of CBSV/UCBSV in infected plants. The serological methods include Enzyme Linked Immunosorbent Assay (ELISA) which depends on the development of monoclonal antibodies that identify a number of different strains of CBSV/UCBSV. The ELISA Kits for cassava brown streak viruses are commercially available and have been reliably used in CBSD diagnosis (Osogo *et al.*, 2014). With the recent advances in sequencing (Radford *et al.*, 2012), a number of different genomes for UCBSV and CBSV with over 30 complete genome sequences of different isolates available in the NCBI database, conserved regions in the genomes of the isolates have been used in the designing of primers for the simultaneous virus specific detection and differentiation between UCBSV and CBSV isolates (Mbanzibwa *et al.*, 2011a). These primers were designed by using the genome sequence information available for a number of UCBSV and CBSV isolates. In the alignments of these viruses, it was observed that there is a variation in the length of the 3' UTR of CBSV and UCBSV with the UCBSV isolates having a longer 3'UTR. Diagnostic primers were thus designed in the conserved regions of the 3' UTR forward primer CBSDDF2 (5'-GCTMGAAATGCGGRTAYACAA -3') and reverse primer CBSDDR (5'-GGATATGGAGAAAGRKCTCC-3). These primers give different size products for CBSV and UCBSV isolates with the CBSV isolates giving a 344nt product while the UCBSV isolates give a 438-440nt product. The primers have been reliably used for the accurate and

reproducible diagnosis of UCBSV and CBSV infections in cassava and other herbaceous hosts. Other reported diagnostic methods include reverse transcription Loop mediated isothermal amplification (RT-LAMP) (Tomlinson *et al.*, 2013) and real time RT-PCR (Adams *et al.*, 2013).

1.7 Other viral diseases of cassava

1.7.1 Cassava mosaic geminiviruses (CMG)

Apart from cassava brown streak disease (CBSD), cassava is infected by a number of other viral diseases mainly caused by cassava mosaic geminiviruses (CMG's) (*Geminiviridae*, *Begomovirus*) (Calvert & Thresh, 2002). Of the cassava mosaic geminiviruses, cassava mosaic disease (CMD) caused by *Cassava mosaic virus* (CMV) is considered the most important due to the high yield losses it inflicts on the infected crops. Despite having been originally reported in 1894 in Tanzania, CMD did not receive much attention until there was an epidemic in the late 1980's in Tanzania, Rwanda and Kenya (Legg & Fauquet 2004). CMV is spread by the whitefly vector *Bemisia tabaci* (Genn) and is also spread through infected propagated cuttings. The genome of CMV consists of two components of single-stranded DNA (Harrison *et al.*, 1995). Most of the initial measures to control cassava mosaic disease focused on developing resistant germplasm by a number of breeding programs in Africa. The breeding programs mainly relied on inter-crossing of cultivated cassava with the wild relative tree species of *Manihot glaziovii* to obtain disease resistant cultivars (Jennings, 1994).

Two forms of viruses isolated from CMD infected plants were named according to studies in which Koch's postulates were analysed for the viruses, initially, the isolated virus could not be back-inoculated into cassava plants and for this reason, it was referred to as cassava latent virus (CLV). In later experiments, the virus could be back transmitted from *Nicotiana benthamiana*

to cassava plants and this virus was referred to as *African cassava mosaic virus* (ACMV) (Bock & Woods, 1983).

Cassava mosaic geminiviruses (CMG's) have been serologically characterized and grouped according to the geographical regions in which the viruses are known to be found, two forms that occur in Africa have been named *African cassava mosaic virus* (ACMV) and *East African cassava mosaic virus* (EACMV) while the form occurring in India is referred to as *Indian cassava mosaic virus* (ICMV).

Recombination events have been reported between different geminiviruses leading to more severe forms of the viruses. For example, recombination between the *African cassava mosaic virus* (ACMV) and *East African cassava mosaic virus* (EACMV) DNA fragments gave rise to *East African cassava mosaic Cameroon virus* (EACMCV) and *South African cassava mosaic viruses* (SACMV) (Fondong *et al.*, 2000).

1.8 Genera of the *Potyviridae* family

The *Potyviridae* family to which CBSV and UCBSV belong, contains the largest number of positive single-stranded RNA viruses infecting plants (Fauquet *et al.*, 2005; Domingo & Holland, 1997) with almost 180 members of plant infecting viruses belonging to this group. Electron microscopic analysis of the viruses in the *Potyviridae* family show the viruses to appear as flexuous filamentous particles which are about 12 to 15 nm in diameter and 500 to 600 nm in length. One characteristic feature of potyviruses is that they encode a protein that aggregates in the cytoplasm of infected cells to form pin wheel inclusion bodies (Shukla & Ward, 1988). These cylindrical inclusion bodies (CI) are formed by a virus-encoded protein and are considered as the most important criterion for assigning viruses to the *Potyviridae* family. The genome of most potyviruses is about 10,000 nucleotides in length encoding a single

large polyprotein that is autoproteolytically cleaved into individual proteins that serve different functions in the virus infection cycle. The potyviral genome is polyadenylated at the 3' end and has a 5' terminal genome linked protein (VPg). Most of the potyviruses are transmitted by vectors in a non-persistent, non-circulative manner.

The *Potyviridae* family is divided into the genus *Bymovirus* which have bipartite genomes and the genera *Ipomovirus*, *Macluravirus*, *Potyvirus*, *Rymovirus*, *Brambyvirus*, *Poacevirus* and *Tritimovirus*, with monopartite genomes. Of the eight genera, *Potyvirus* contains by far the most numerous and most important plant viruses. Most viruses that cause severe diseases of crops belong to this genus (Agrios, 2005). Examples of important plant viruses that cause severe diseases are *Potato virus Y* (PVY), *Watermelon mosaic virus* (WMV) and *Turnip mosaic virus* (TuMV). The type member of genus *Potyvirus* is *Potato virus Y* (PVY) with a genome size of 9.7 kb (Robaglia *et al.*, 1989). Most of the *Potyvirus* transmission is performed by aphid vectors although some potyviruses are transmitted by mites and whiteflies.

1.9 The genus *Ipomovirus*

Ipomoviruses (genus *Ipomovirus*) are whitefly transmitted viruses in the *Potyviridae* family. There are currently five official members including *Sweet potato mild mottle virus* (SPMMV), *Cucumber vein yellowing virus* (CVYV), *Squash vein yellowing virus* (SqVYV), *Cassava brown streak virus* (CBSV) and *Ugandan cassava brown streak virus* (UCBSV) (Dombrovsky *et al.*, 2014). The ipomoviral genome sizes range from about 8995 to 11,000 nucleotides and they encode a polyprotein ranging from 2902 to 3011 amino acids (Dombrovsky *et al.*, 2014).

1.9.1 *Sweet potato mild mottle virus (SPMMV)*

Sweet potato mild mottle virus (SPMMV) is the type member of the genus *Ipomovirus*. SPMMV is the third most prevalent virus infecting sweet potato (*Ipomea batatas*). Studies have shown East Africa as the origin of SPMMV (Loebenstein and Thottapilly 2013; Tairo *et al.*, 2005). SPMMV is known to infect hosts from different families with *Convolvulaceae* being the main host family although *Hewittia sublobata* and *Lepistemon owariensis* are also known hosts (Tugume *et al.*, 2010). The SPMMV is transmitted by the whitefly vector *B. tabaci* and it has the largest genome in the genus *Ipomovirus* with a genome size of 10,818 nucleotides excluding the 3' terminal poly (A) tail (Colinet *et al.*, 1998).

1.9.2 *Cucumber vein yellowing virus (CVYV)*

Cucumber vein yellowing virus (CVYV) is a whitefly borne virus transmitted by *Bemisia tabaci* vector in a non-persistent manner (Lecoq *et al.*, 2000). The virus mainly infects host plants in the family *Cucurbitaceae* which include squash (*Cucurbita pepo*), watermelon (*Citrullus lanatus*), *Citrullus colocynthis* and *Lagenaria siceraria* (Harpaz & Cohen, 1965; Mansour & Al-Musa, 1993). The virus is thought to have originated from Israel (Harpaz & Cohen, 1965) and Jordan (Al-Musa *et al.*, 1985). There is variation in the presentation of symptoms between the different CVYV isolates with the Jordanian isolates presenting more severe stunting symptoms in infected cucumber (Lecoq *et al.*, 2000) and watermelon plants (Galipienso *et al.*, 2012) as compared to the isolates from Israel. The genome of CVYV has been found to consist of a positive sense single-stranded RNA consisting of 9751 nucleotides which encodes a 3148 amino acid polyprotein (Galipienso *et al.*, 2012; Janssen *et al.*, 2005). Interestingly, CVYV is a relatively unusual member of the family *Potyviridae* as the genome lacks the multifunctional HC-Pro protein known to be present in most of the sequenced isolates of potyviruses and instead has a duplicated P1 protein (Valli *et al.*, 2008; Janssen *et al.*, 2005) with the resulting

proteins named as P1a and P1b. In the CVYV, it has been shown the P1b protein serves the function of the RNA silencing that is needed by the virus in order to overcome the host defence system and thus cause infection (Valli *et al.*, 2008).

1.9.3 *Squash vein yellowing virus (SqVYV)*

Squash vein yellowing virus (SqVYV) is an *Ipomovirus* that is transmitted by the whitefly vector *Bemisia tabaci*. The virus was originally discovered in Florida and has been found to mainly infect plants in the *Cucurbitaceae* family, including *Citrullus lanatus* (watermelon), *Cucumis melo* (muskmelon), *Cucurbita pepo* (squash) and *Cucurbita maxima* (pumpkin) (Webb *et al.*, 2003; Adkins *et al.*, 2008). SqVYV is also known to infect two wild cucurbits hosts *Momordica charantia* (balsam apple) and *Melothria pendula* (creeping cucumber) which may serve as reservoirs hosts for SqVYV, especially between planting seasons (Adkins *et al.*, 2008). The symptoms of SqVYV in infected plants include transient vein yellowing, vein yellowing or systemic wilt and necrosis. The SqVYV genome encodes 9836 nucleotides which translate into a polyprotein consisting of 3172 amino acids and just like the CVYV, the SqVYV genome also lacks the HC-Pro protein but has a duplicated P1 protein (Li *et al.*, 2008).

1.9.4 *Cassava brown streak virus (CBSV) and Ugandan cassava brown streak virus (UCBSV)*

Cassava brown streak disease (CBSD) is caused by two *Ipomovirus* species which have been described as *Cassava brown streak virus (CBSV)* and *Ugandan cassava brown streak virus (UCBSV)* (Mbanzibwa *et al.*, 2011b). Earlier reports referred to *Ugandan cassava brown streak virus (UCBSV)* as *Cassava brown streak Uganda virus (CBSUV)* (Patil *et al.*, 2011; Mohammed *et al.*, 2012; Mbanzibwa *et al.*, 2009a). Other names were also proposed (Winter *et al.*, 2010). However, the name *Ugandan cassava brown streak virus (UCBSV)* has been

finally ratified by the International Committee for Taxonomy of Viruses (ICTV). The original identification of the two different viruses causing CBSD was by molecular characterization and comparison of different sequences of cassava brown streak viruses in which there was a clustering of isolates in two different clades (Mbanzibwa *et al.*, 2009a; Winter *et al.*, 2010; Monger *et al.*, 2010). However, recent sequencing of more complete genomes of UCBSV and CBSV has revealed further speciation within the isolates showing six different clades (Ndunguru *et al.*, 2015). The presence of different forms of the virus pose serious challenges to breeding programs against CBSD which need to generate improved germplasm with resistance to all the CBSD viruses.

Both CBSV and UCBSV genomes are positive sense single-stranded RNA viruses transmitted by whiteflies in a non-persistent manner (Maruthi *et al.*, 2005). However, there is variation in the genome sizes of UCBSV and CBSV with the complete genome of *Ugandan cassava brown streak virus* (UCBSV) consisting of 9070 nucleotides excluding the poly(A) tail and is predicted to produce a polyprotein of 2902 amino acids (Mbanzibwa *et al.*, 2009a) while the genome of CBSV consists of 8997 nucleotides encoding 2912 amino acids (Winter *et al.*, 2010).

Both viruses encode a polyprotein that is autocatalytically cleaved into ten mature proteins by the viral proteinases. Another protein PIPO (Pretty Interesting Potyviral ORF) is translated in the +2 reading frame within the P3 region. The PIPO protein is expressed by polymerase slippage (Chung *et al.*, 2008). A unique property of these two viruses, as compared to other fully described members of genus *Ipomovirus* (and family *Potyviridae*) is that they have a HAM1 protein and no HC-Pro. Unlike the CVYV and SqVYV, the P1 of CBSV and UCBSV is not duplicated despite of lack of HC-Pro. The Maf/HAM1 proteins are nucleoside triphosphate (NTP) pyrophosphatases that act by reducing mutagenesis through intercepting noncanonical NTP's thus preventing their incorporation into DNA or RNA (Mbanzibwa *et al.*,

2009a). The HAM1 gene is incorporated in the 3'-proximal part of the viral genome between the viral replicase (Nlb) and coat protein (CP) encoding sequences (Winter *et al.*, 2010; Mbanzibwa *et al.*, 2009a). A HAM1-like sequence at the corresponding genomic position has also been identified in *Euphorbia Ringspot virus* (EuRSV) that is placed in the genus *Potyvirus* based on the phylogeny of CP encoding sequences. In addition, the UCBSV/CBSV genomes contain a single P1 proteinase and are lacking the helper component proteinase (HC-Pro) at the N-proximal part of the polyprotein (Mbanzibwa *et al.*, 2009a). Other ipomoviruses described contain HC-Pro for example *Sweet potato mild mottle virus* (SPMMV), the type species of *Ipomovirus*, or are lacking HC-Pro but contain two P1 proteinases like in *Cucumber vein yellowing virus* (CVYV) and *Squash vein yellowing virus* (SqVYV). (Figures 1.4) shows the comparison of the genomes of CBSV with some members in the *Potyviridae* family.

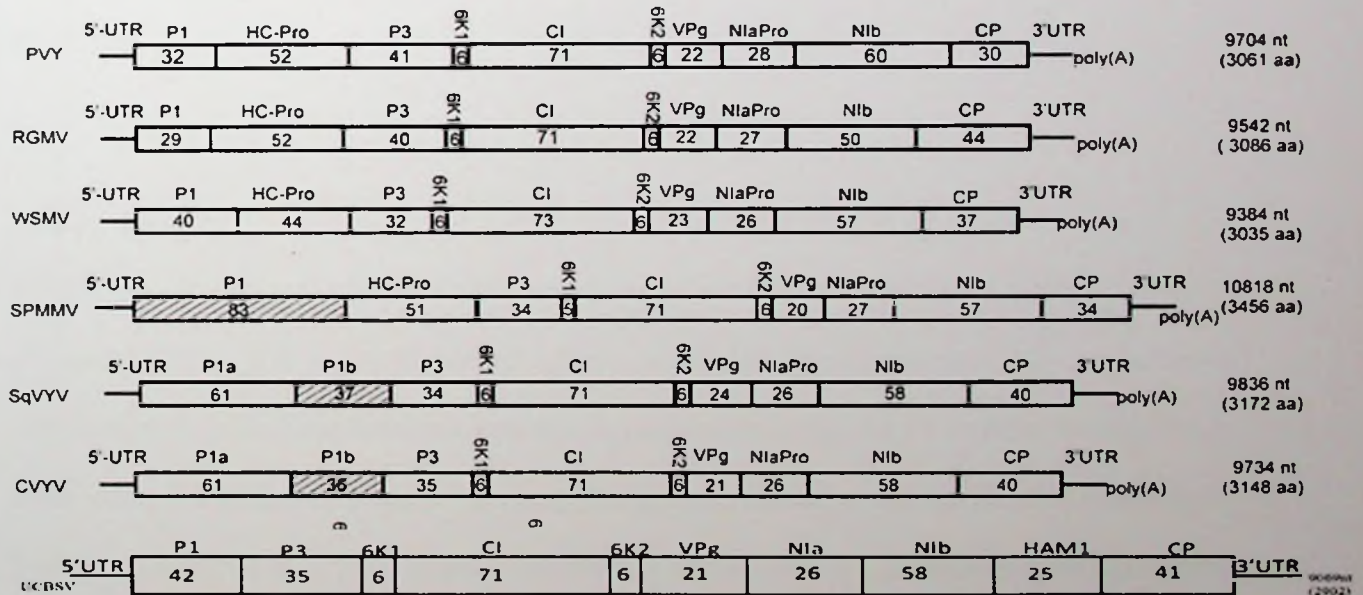


Figure 1.4: Comparison of genome organization of UCBSV with other virus members in the family *Potyviridae*. *Potyvirus*; *Potato virus Y* (PVY), *Rymovirus*; *Ryegrass mosaic virus* (RGMV), *Tritimovirus*; *Wheat streak mosaic virus* (WSMV); *Ipomovirus*; *Squash vein yellowing virus* (SqVYV), *Cucumber vein yellowing virus* (CVYV) and *Ugandan cassava brown streak virus* (UCBSV). Adapted from Mbanzibwa *et al.*, (2009a).

1.10 Functions of the different proteins in the typical potyviral genome

1.10.1 P1

P1 is also known as the first potyviral protein and has been reported to be a serine protease (classified as MEROPS clan SA, family S30). The P1 protein acts as a protease cleaving the P1 protein at its C-terminal (Verchot *et al.*, 1991). There is variation in the sequence length of P1 (ranging from 40-80 kDa) and number of P1 proteins between different viruses in the *Potyviridae* family, for example, the genome of *Cucumber vein yellowing virus* (CYVY), contains a duplicated P1 protein designated P1a and P1b (Valli *et al.*, 2008). P1 protein has been shown to be involved in suppression of RNA silencing activities in CBSV and UCBSV (Giner *et al.*, 2010; Mbanzibwa *et al.*, 2009a). In *Sweet potato feathery mottle virus* (SPFMV; *Potyvirus*), the P1 protein has an extra embedded P1-N-PISPO (Pretty interesting sweet potato open reading frame) protein that is expressed by polymerase slippage and is translated in the -1 opening reading frame and has been reported to be involved in RNA suppression (Mingot *et al.*, 2016).

1.10.2 HC-Pro

The HC-Pro protein is the helper component protease protein in the potyviral genome with size of about 50 kDa. It is a multifunctional protein which has been reported to play a number of different roles in the virus infection process. In conjunction with the P1 protein, the HC-Pro is involved in N terminal cleavage of the potyviral polypeptide cleaving the HC-Pro from the P3 protein (Verchot *et al.*, 1991). The C- terminal of the HC-Pro is a proteinase (MEROPS clan CA, family C6) while the N-terminal has been reported to play a role in aphid transmission of the virus (Maia *et al.*, 1996; Verchot *et al.*, 1991; Torres-Barcel *et al.*, 2008; Klein *et al.*, 1994). Furthermore, the HC-Pro acts as a pathogenicity enhancer that activates replication of heterologous viruses by reducing or eliminating the processing of dsRNA by Dicer (Pruss *et*

al., 1997; Vance *et al.*, 1995; Dunoyer *et al.*, 2004; Mallory *et al.*, 2002; Valli *et al.*, 2014). Cell-to-cell movement which allows the virus to systemically spread through the host has been reported to be facilitated by HC-Pro by transferring newly synthesized genomes from sites of replication to the intracellular transport system (Carrington *et al.*, 1996; Klein *et al.*, 1994). Because of the RNA binding properties of HC-Pro, the protein has also been implicated to be involved in viral replication processes within the host by allowing interaction between the viral genomic RNA and viral protein (Merits *et al.*, 1998; Urcuqui-Inchima *et al.*, 2000).

1.10.3 P3

The third potyviral protein, P3, is about 30 kDa. One interesting feature in the P3 protein of potyviruses is that it contains an overlapping open reading frame PIPO (Pretty interesting potyviral ORF). The PIPO is embedded within the P3 cistron of the potyviral polyprotein and is translated in the +2 reading frame. A signature motif G₁₋₂A₆₋₇ facilitates ribosomal frame shift or transcriptional slippage for the expression of PIPO (Chung *et al.*, 2008). The PIPO is known to be expressed as an N-terminal fusion protein with the P3 protein (P3N-PIPO). Existence of overlapping reading frames in potyviral genomes is justified by the fact that most viruses are under selective pressure to encode all genes required for translation, replication, transmission, and movement in the minimum sequence space. Because of the RNA binding properties of the P3 protein, it has been implicated to be involved in genome amplification and replication as it allows for interaction between viral genomic RNA and viral proteins (Merits *et al.*, 1998). The P3 protein is however one of the least characterized proteins in the potyviral genome, and PIPO is even less characterized.

1.10.4 6K1 and 6K2

The 6K1 and 6K2 proteins are also known as 6 kDa proteins due to their sizes. The proteins have been reported to be endoplasmic reticulum association factors, this function is related to the fact that most viral genomes are known to replicate in close association with membranes, the membrane association is important in increasing the surface area for RNA synthesis. The proteins therefore serve as a membrane anchor facilitating the viral replication process (Schaad *et al.*, 1997).

1.10.5 CI

The protein is also known as the cylindrical inclusion protein in the potyviral genome. The CI is the largest protein among all the potyviral encoded proteins with the size ranging at about 70 kDa. Just like the HC-Pro protein, the CI is a multifunctional protein that has been reported to be involved in many different processes in the viral infection cycle. One characteristic feature of the CI is that it encodes a protein that aggregates in the cytoplasm of infected cells forming pin-wheel inclusion bodies a feature that is used in assigning viruses to the *Potyviridae* family (Riechmann *et al.*, 1992). CI has been reported to have *in-vitro* RNA dependent ATPase and helicase activities by unwinding double-stranded polyribonucleotides (Revers *et al.*, 1999). Furthermore, due to the RNA binding properties of the CI protein, the protein has been studied to be involved in genome replication as it is present in replication complexes of viruses (Deng *et al.*, 2015; Fernández *et al.*, 1995).

1.10.6 VPg

The protein is also known as viral genome linked protein with a size of approximately 24 kDa. The VPg is covalently attached to the RNA genome at the 5' termini and has thus acting like

the 5' cap on mRNA has been suggested to be involved in RNA synthesis initiation and replication (Schaad *et al.*, 1997; Satheshkumar *et al.*, 2005; Ivanov *et al.*, 2014; Thivierge *et al.*, 2008). The VPg interacts with the host eukaryotic initiation factor 4E (eIF4E) to allow for virion disassembly and initiate potyviral genome translation (Wang, 2015; Thivierge *et al.*, 2008). Translation and replication are two key steps needed for effective potyviral infection and spread. Although potyviruses lack a dedicated movement protein, VPg is involved in cell-to-cell movement of the virus within the host (Schaad *et al.*, 1997). The interaction between Vpg and RNA dependant RNA polymerase (RdRp) has been reported to aid in potyviral replication (Daròs *et al.*, 1999).

1.10.7 NIa

Also known as Nuclear Inclusion a protein, the NIa is found to aggregate and form crystalline inclusion in the nucleus of the host cell and is of a size of approximately 26 kDa. NIa is the major proteolytic enzyme in the potyviral genome and it catalyzes the cleavage of six of the polyprotein cleavage sites (Adams *et al.*, 2005). The C-terminal half of the NIa protein is a serine type (Trypsin like proteinase MEROPS clan PA(C), family C4) and the N-terminal is Vpg linked (Dougherty & Carrington, 1988; Mathur *et al.*, 2012). Most of the biochemical cleavage sites for the NIa protease have been enzymatically characterized for some potyviruses and for others, these cleavage sites are normally inferred from sequence comparison with the characterized viruses (Adams *et al.*, 2005).

1.10.8 NIb

Also known as Nuclear Inclusion b protein, just like NIa aggregates and forms crystalline inclusion bodies in the host cell and is of a size of approximately 58 kDa. The NIb is the RNA

polymerase protein in the potyviral genome. The RNA dependent RNA polymerase function of the NIb protein was inferred by comparison of the NIb sequence with the sequences of other known RNA dependent RNA polymerases (Domier *et al.*, 1987). The NIb contains a conserved GDD motif known to be involved in the polymerase activities of the protein (Hong & Hunt, 1996).

1.10.9 HAM1

The HAM1 protein is novel protein that has been found in the genomes of UCBSV and CBSV isolates but is absent in the genomes of other reported ipomoviruses SPMNV, SqVYV and CVYV (Mbanzibwa *et al.*, 2009a). The HAM1 is found incorporated between the NIb and coat protein (CP) regions in the genomes of all UCBSV and CBSV sequenced isolates. The NIa enzyme cleavage sites have also been found to be adjacent at the ends of the N terminal and C terminal of the HAM1 coding sequence suggesting that the HAM1 protein is cleaved from the polyprotein and may play a role in the virus infection cycle. HAM1 has also been found in the genomes of *Euphorbia Ringspot virus*. Recently, a cassava research group at the International Center for Tropical Agriculture (CIAT), reported presence of Maf/Ham1 domain in the genomic sequence of *the Cassava torrado-like virus* (CsTLV; *Torradovirus*; *Secoviridae*) (Carvajal-Yepes, unpublished). BLASTP searches using the HAM1 sequence query shows the HAM1 protein to be related to the Maf/HAM1 family of proteins that have been reported to act in the reduction of mutations by reducing the incorporation of non-canonical nucleotides in the DNA or RNA (Mbanzibwa *et al.*, 2009a). It is not clear how a virus with limited sequence space can evolve to acquire an extra protein in its genome and more studies are underway to understand and elucidate the function of HAM1 protein in the CBSV/UCBSV genome. However, it is interesting to note that all viruses that have been found to encode this protein infect cassava or other plants in the family *Euphorbiaceae* as their natural hosts.

1.10.10 Coat Protein (CP)

The potyviral coat protein has the size of approximately 40 kDa. The coat protein has been reported to play the roles of a movement protein allowing for systemic movement of the virus from the initially infected cells through the plasmodesmata and thus spread throughout the plant (Revers *et al.*, 1999; Doija *et al.*, 1994). The coat protein has also been reported to aid in aphid transmission of the viruses in the genus *Potyvirus* which is needed for virus spread between hosts. The N terminus of the coat protein of the potyvirus contains a conserved DAG motif which has been implicated in successful transmission of potyviruses by aphids (Atreya *et al.*, 1990).

1.10.11 5' and 3' untranslated regions (UTR)

The UTRs encompass the non-coding regions of the genome. There is a lot of variation in the length of different potyviral 5' and 3' UTR. Despite the fact that the UTR's are non-coding, they have been reported to contain regions or sequences that control the expression of genes in the coding regions for instance, in *Tobacco etch virus* (TEV), the long 5' non-translated region has been found to substitute the 5' cap in the interaction with the 3' poly (A) and by this it acts as a translational enhancer by increasing translation competency (Gallie *et al.*, 1995). Furthermore, in the non-polyadenylated *Tobacco mosaic virus* (TMV), the 3'UTR has been found to replace the poly (A) tail in enhancing translational efficiency by promoting translation initiation (Gallie & Kobayashi 1994). In potexviruses, the 5' non-translated region was found to play a role in the synthesis of RNA and viral plasmodesmal movement while the 3' UTR regulates plus and minus strand RNA synthesis (Verchot-Lubicz *et al.*, 2007).

1.11 Plant host pathogen interaction in virus infections

Viruses are obligate parasites that depend on the host machinery for their successful infection and eventual spread inside the host. A successful virus infection depends on an interplay between the virus and host factors (Wang, 2015). It is thus important to understand the network of virus host interaction mechanisms involved in the infection process and this may guide in developing appropriate antiviral strategies. There are mainly two lines of plant defence against virus infection. The first line of defence is the natural defence which depends on barriers like the rigid cell wall and lignification of plant cells and the second line of defence is the active defence which depends on recognition of pathogens. This defence is associated with hypersensitive response (HR) which destroys surrounding cells at the site of infection by programmed cell death (PCD) leading to small localized necrotic lesions (Goldbach *et al.*, 2003).

1.12 Resistance to viral pathogens

1.12.1 Pathogen derived resistance (PDR)

Pathogen derived resistance is normally based on exploiting a section of the viral genome and using it as a bait to control virus spread by generating transgenic lines expressing the protein encoding the viral protein. In deciding the choice of viral proteins to target in transgenic plant virus resistance, it is important to use mild, symptomless, or attenuated strains of the virus in generating transgenic plants (Wang *et al.*, 2013; Zhang *et al.*, 2011). Another important factor to consider is that some viral proteins have been reported to be involved in synergistic interactions of viral infection and therefore these proteins should be avoided. The coat protein (CP) has mainly been used for this purpose to achieve CP-mediated resistance or coat protein mediated protection (CPMP) (Bendahmane *et al.*, 2002; Powell *et al.*, 1990; Dun *et al.*, 1988;

Prins *et al.*, 2008). CP mediated resistance has mainly been achieved in RNA viruses including TMV, PVX and CMV.

For cassava brown streak disease, resistant lines have been generated by exploiting RNAi and hairpin constructs. Transgenic lines expressing RNAi constructs targeting different sequences in the coat protein were generated and the lines were found to be resistant to disease on challenging them with CBSVs (Yadav *et al.*, 2011; Odipio *et al.*, 2013; Patil *et al.*, 2011; Ogwok *et al.*, 2012).

1.12.2 Host derived resistance

In some cases, host factors may be targeted for the development of inheritable recessive genetic resistance by deletion of target host factor needed in viral infection. Host derived resistance can be achieved through advanced biotechnological means which include mutation, silencing or down regulation of the host factor genes (Wang and Krishnaswamy, 2012). A number of host factors have been reported to interact with viral proteins in yeast two hybrid protein-protein interaction assays. Host proteins reported to aid in the viral infection process include eukaryotic translation initiation factors (eIF4E) or eIF(iso)4E which have been reported to interact with the viral genome linked protein (VPg) to aid in the initiation of viral translation processes (Schaad *et al.*, 1997; Estevan *et al.*, 2014). In some studies, the host genes like (eIF4E) or eIF(iso)4E are targeted and edited by the current advanced genome editing using clustered regularly interspaced short palindromic repeats (CRISPR) and the associated Cas 9 proteins techniques as a measure to control viral disease (Wang *et al.*, 2013).

1.13 RNA silencing

1.13.1 RNA silencing mechanisms

As a defensive strategy to combat plant viral infections, plants have evolved to acquire defence responses which involve a number of proteins among which are the Dicer like protein (DCL) and RNA induced silencing complex (RISC) (Lecellier & Voinnet, 2004). RNA silencing can also be referred to as RNA interference in animals or posttranscriptional gene silencing (PTGS) in plants (Voinnet, 2005). RNA silencing is a surveillance mechanism used by plants for defence against viruses. RNA silencing is induced in plants by double-stranded RNA (dsRNA), which are normally produced from the replicative forms of the virus during the viral replication process (Mallory *et al.*, 2002; Dunoyer *et al.*, 2004). The dsRNA then acts as a template for degradation by the dsRNA specific Dicer like endonuclease (DCL) which digests the dsRNA into fragments of 21-24 nucleotides designated as small interfering RNA (si-RNA) (Garcia-Ruiz *et al.*, 2010; Llave, 2010). When the si-RNA are from viral genomes, they are referred to as viral small interfering RNAs (vsiRNA). The si-RNA then assemble into Argonuate (AGO) proteins forming the RNA induced silencing complex (RISC) which specifically targets and inactivates invading nucleic acids such as RNA viruses with sequence homology with the original viral RNA (Baulcombe, 2004). A high accumulation of viral RNAs and vsiRNAs are correlated in virus diseased, susceptible plants that cannot recover from virus infection (Gammelgard *et al.*, 2007).

1.13.2 Viral suppressors of RNA silencing

As a counter-defensive strategy, plant viruses have evolved proteins in their genomes that protect them from the RNA silencing defence mechanism in plants (Anandalakshmi *et al.*, 1998; Brigneti *et al.*, 1998; Kasschau & Carrington 1998; Jiang *et al.*, 2012). This explains the reason why viruses are able to multiply and spread in plants despite the presence of the RNA

silencing mechanism. The encoded RNA silencing suppressors (RSS) act by affecting a number of steps involved in the RNA silencing pathway for instance, the RSS may sequester or bind the siRNA molecules or they may affect the enzyme activities of the RNA endonuclease activities of the Dicer-like enzymes (DCL) thereby inhibiting the synthesis of siRNA (Nakahara & Masuta, 2014). RNA silencing suppressors act by affecting anti-viral defence and interfering with other processes that utilize RNA silencing like pathogenesis and physiological cellular activities (Valli *et al.*, 2008).

A lot of work has been done on understanding the different RNA silencing suppressor proteins in different plant viruses and reports have confirmed the HC-Pro of potyviruses as a major RSS. In *Cucumber mosaic virus* (CMV), the 2b protein has been confirmed as the RSS (Zhang *et al.*, 2006; Rashid *et al.*, 2008) while p19 protein is the main RSS in tombusviruses (Lakatos *et al.*, 2004). A number of other proteins have also been reported to act as RSS for example the P1 protein in CBSV genome has been confirmed as an RSS (Mbanzibwa *et al.*, 2009a).

HC-Pro is the most characterized poytviral suppressor of RNA silencing, however, in some viruses, it is either weak or not suppressing thus the P1 protein has evolved as a suppressor for instance in *Sweet potato mild mottle virus* (SPMMV), the type member of the genus *Ipomovirus*, the P1 protein has been shown to interact with Argonaute (AGO1) thereby blocking the formation of RNA-induced silencing complex (RISC) that specifically targets and cleave viral RNA (Giner *et al.*, 2010).

Winter *et al.*, (2010) suggests that, there is a difference in the sequences found in P1 genes as well as symptoms shown in CBSD infection, this may suggest P1 protein acting as viral suppressor of RNA silencing (Mbanzibwa *et al.*, 2009a) and could be involved in infection, symptom development and pathogenicity. In cucumber, the necrosis caused by *Cucumber*

mosaic virus (CMV) was found to be a result of protein-protein interaction between the 2b protein which is the RNA-silencing suppressor and a host catalase (Inaba *et al.*, 2011).

1.14 Construction of infectious clones

Infectious clones of full-length cDNA are very important tools for studying the molecular biology of plant RNA viruses (López-Moya & García 2000; Nagyová & Šubr 2007). Because most of the current biotechnological tools are biased towards manipulating DNA sequences, this makes it complicated to engineer RNA genomes of RNA viruses (Verchot *et al.*, 2007). By making infectious clones of RNA viruses, it is possible to generate DNA copies of RNA viruses. The cloned DNA can then be used in an *in-vitro* transcription reactions to generate transcripts which are able to cause infection similar to the wildtype RNA virus.

The construction of infectious clones of RNA viruses involves generating a cDNA copy of the complete genome of an RNA virus cloned inside a suitable backbone vector. The RNA is reverse transcribed into cDNA using a primer that binds to the 3' end of the genome to generate single-stranded DNA. The single-stranded DNA is then converted to double-stranded cDNA by using primers that encompass the whole genome to the 5' end. A suitable promoter is inserted upstream of the cDNA to allow for transcription of the cDNA. There are a number of different promoters that have been used in infectious clone construction among which the bacteriophage SP6, T7, or *Cauliflower mosaic virus* CaMV 35S are the most commonly used. Depending on the promoter used, an infectious clone may either be transcribed *in-vitro* to generate RNA transcripts as is the case for infectious clones generated with SP6 or T7 promoters. Alternatively the infectious clone may be introduced into the host plant by agroinfiltration or biolistics (Fakhfakh *et al.*, 1996) for the *in-vivo* transcription and translation from infectious clones containing the CaMV 35S promoter (Boyer & Haenni, 1994). For *in-*

in vitro transcribed cDNA infectious clones, the plasmid with the cloned complete genome of the RNA virus can act as a template in an *in-vitro* transcription reaction to generate RNA transcripts that can cause an infection similar to the wildtype RNA virus. Inoculation of susceptible host plants with DNA of many plant virus genomes including numerous RNA viruses, can result in the establishment of viral infections (Boyer & Haenni, 1994). It is known that for successful infection to occur, the complete genome of the virus must be present although there have been reports of the possibility of generating infectious transcripts from incomplete genomes (Davis *et al.*, 1989). The construction of full-length cDNA clones from which infectious transcripts can be synthesized *in-vitro* is sometimes long and tedious, this may be due to toxicity of some of the viral proteins in the *Escherichia. coli* host through which the plasmids are regenerated. For potyviruses, the toxicity has been reported to be mainly due to the P3 and CI proteins present in the potyviral genomes (Desbiez *et al.*, 2012; Boyer & Haenni, 1994; Jakab *et al.*, 1997).

Moreover, cDNA synthesis, cloning strategies and the design of sequences bordering the viral insert can have a strong influence on the infectivity of the resulting transcript due to the presence of non-viral nucleotides at the 5' end of viral transcripts which may affect infectivity (Domier *et al.*, 1989). A serious limiting factor for some viral genomes appears to be the synthesis of full-length first cDNA strand probably because the polymerization step is hampered by secondary structures on the viral template RNA. It is noteworthy that even for a given virus, some research groups but not others have succeeded in obtaining infectious clones of potyviruses (Boyer & Haenni, 1994; Nagyová & Šubr 2007; Riechmann *et al.*, 1990).

1.15 Applications of infectious clones of viruses

The infectious clones offer valuable tools for viral functional genomics in which the roles of the different proteins encoded in the viral genome can be studied, are helpful in studying the expression and replication of RNA viruses using (Boyer & Haenni, 1994). With infectious clones, it is possible to undertake *in-vitro* mutagenesis (deletions, insertions, substitutions) and complementation, as well as understanding the mechanisms of defect interfering RNA and satellite RNA genesis (Nagyová & Šubr, 2007). Furthermore, with infectious clones, it is possible to study the molecular biology of viral proteins and the functions that these proteins serve in the viruses in order to develop new resistance strategies (Revers & García, 2015). In addition, infectious clones can be tagged to reporter genes like the green fluorescent protein (GFP) and the tagged virus used to study virus –host interactions (e.g intracellular transport) (Nagyová & Šubr 2007; Bendahmane *et al.*, 2002). The infectious clones are a pool of viral genes that can be used in the development of antiviral strategies. In some cases, the infectious clones may act as viral vectors through which trans genes can be expressed in plants (Boyer & Haenni, 1994).

1.15.1 Potyvirus infectious clones

Potyraviruses represent over 25% of known plant viruses. The very first potyviral infectious clone was obtained in 1989 (Domier *et al.*, 1989). There are now over 40 different infectious clones belonging to 15 different virus species although such tools are missing for some of the major agronomically important viruses. Among the infectious clones available for potyraviruses include those *Tobacco vein mottling virus* (TVMV) (Domier *et al.*, 1989), *Zucchini yellow mosaic virus* (ZYMV) (Gal-On *et al.*, 1991), *Potato virus Y* (PVY) (Jakab *et al.*, 1997), *Tobacco etch virus* (TEV) (Riechmann *et al.*, 1992). *Watermelon mosaic virus* (WMV) (Desbiez *et al.*, 2012), *Papaya ringspot virus* (PRSV) (Desbiez *et al.*, 2012), *Potato virus A*

(PVA) (Kekarainen *et al.*, 2002), *Bean yellow mosaic virus* (BYMV) (Nakahara *et al.*, 2015), *Clover yellow vein virus* (CYVV) (Nakahara *et al.*, 2015) and *Tobacco vein banding mosaic virus* (TVBMV) (Gao *et al.*, 2012). Efforts to generate infectious clones for most viruses are challenged by the difficulties associated with the process of making infectious clones. Some of the problems include the fact that the length of the viral RNA (about 10 kb, making one step cloning difficult), the highly deleterious effects of adding any extra nucleotides in the 5' or 3' extremity of the genome during the cloning process (Boyer & Haenni, 1994) and the toxicity of some viral proteins in *Escherichia coli*, resulting in spontaneous deletions of some clones in bacterial cells. The most toxic proteins appear to be P3 and CI encoded in the central region of the genome (Jakab *et al.*, 1997). In order to overcome these difficulties, several techniques have been used, including cloning the genome in two halves that have to be religated *in vitro* before inoculation (Jakab *et al.*, 1997), adding introns in the cDNA to prevent expression in *E. coli* (Johansen, 1996) or removing cryptic prokaryotic promoter elements in the P3 coding region by site directed mutagenesis (Chikh *et al.*, 2011). Homologous recombination in yeast rather than *in vitro* restriction and ligation has been shown as a rapid way to obtain infectious clone of potyviruses (Desbiez *et al.*, 2012).

1.16 Aims and objectives of the study

Infectious clones are valuable tools for understanding the pathogenicity of viruses and thus devise control methods. For cassava brown streak disease, some of the control methods have relied on generating transgenic resistance using RNAi and hairpin constructs. However, screening the generated breeding lines has been challenged by the poor transmissibility of the virus by whitefly vectors thus necessitating the need for infectious clones which provide a homogenous source of viral infection for screening breeding lines in addition to providing valuable tools for studying UCBSV viral pathogenicity.

Considering the importance of infectious clones in understanding UCBSV and CBSV pathogenicity, the main aim of this PhD project was to make infectious clones of UCBSV Kikombe isolate which were used for *in-vitro* transcription using SP6 promoter. The specific aims of the study included.

- Sequencing the complete genome of UCBSV Kikombe isolate and understanding the infection status of the wildtype virus in indicator plants.
- Construction of *in-vitro* and *in-vivo* infectious clones of UCBSV Kikombe isolate.
- Using the infectious clones to study the functions of P1 and HAM1 proteins in UCBSV genome.

Chapter 2: Materials and Methods

General materials and methods common to this study are explained in this chapter, whilst the details for the specific studies are explained in the respective chapters (3-5).

All plastic, glassware, solutions, and media were purchased sterile or sterilized before use by autoclaving at 120°C for 15mins.

All chemicals were purchased from Sigma-Aldrich or Fisher and were analytical grade. Primers were purchased from Integrated DNA Technologies (IDT) and commercial kits were purchased from Omega Bio-Tek, Thermo-Scientific, Promega and Zymo Research, as stated. New England Biolabs (NEB) and Thermo Scientific supplied all enzymes and buffers whilst Melford supplied antibiotics. DNA ladder markers were purchased from Bioline. Sterile deionized water (SDW) was used throughout.

2.1 Plant material and virus isolates used in this study

2.1.1 Plant material

2.1.1.1 *Nicotiana species*

Nicotiana benthamiana, *Nicotiana tabacum* (SR1), *Nicotiana glutinosa* and *Nicotiana clevelandii* plants were grown from in-house raised seed and planted in 10 cm pots containing multipurpose compost (William Sinclair). The plants were grown in the greenhouse with conditions of 25°C with a 16 h light and 8 h dark pattern. The plants were inoculated with virus inoculum at 3-4 weeks old and were kept in growth cabinets at 28°C with a 16 h light and 8 h dark pattern.

N. benthamiana 16C line used in the RNA silencing suppression studies was a gift from Prof David Baulcombe of the University of Cambridge.

2.1.1.2 Cassava plants (*Manihot esculenta*)

The cassava plants were generated by tissue culture at the Natural Resources Institute (NRI) in London and they were virus indexed to confirm that were free of any UCBSV or CBSV infection. The plants were transported to the University of Bristol. The cultivar/variety of cassava used was “Columbian” which is known to be susceptible to cassava brown streak disease (CBSD). These plants were grown in Sinclair multipurpose potting compost in 16 cm pots and kept in the growth cabinets at the University of Bristol at 28°C with a 16 h light and 8 h dark pattern.

2.1.2 Virus isolates

2.1.2.1 Kikombe UCBSV isolate source

The UCBSV Kikombe isolate source was cassava leaves from symptomatic CBSD infected cassava plants showing characteristic foliar CBSD symptoms collected from Tanzania (East Africa), transported to the Natural Resources Institute (NRI) in London (UK), and kept at -80°C.

2.2 Virus preparations

2.2.1 Virus inoculum preparation and inoculation

2.2.1.1 Herbaceous plants

Infected leaves showing CBSD symptoms were ground in liquid nitrogen using a pestle and mortar. Freshly prepared ice cold 0.01M potassium phosphate ($K_2HPO_4 + KH_2PO_4$) prepared from a 0.1M stock of K_2HPO_4 and 0.1M stock of KH_2PO_4 pH 7.0 buffer containing 0.2% (w/v) sodium sulphite and 0.01M β mercaptoethanol (Ogwok *et al.*, 2010; Fauquet, 2010) was added during the grinding at a ratio of 1:6 (w/v) tissue to buffer. The inoculum was kept on ice during the inoculation process. Test indicator plants (*N. benthamiana*, *N. tabacum*, *N. glutinosa* and *N. clevelandii*) were dusted with carborundum (particle size 180 grit) powder to act as an abrasive and approximately 500 μ l of extract was rubbed gently on the leaf surface using gloved hands. Inoculation was done on the most expanded leaf from the top of the plant. After inoculation, the plants were left for 10 min then gently sprayed with water to remove carborundum. The inoculated plants were maintained in the growth room at 28°C. After 4 weeks, systematically infected leaves were collected and RNA extracted for cDNA synthesis and subsequent virus detection.

2.2.1.2 Cassava infections

Cassava plants were infected with the infectious clone to test the infectivity of the clone in these plants. The inoculum for infection was prepared by setting up *in-vitro* transcription reactions using the generated UCBSV infectious clone plasmid as template. The *in-vitro* transcription reactions were set up as explained in (section 2.10). To infect the cassava plants, 50 μ l of the *in-vitro* transcription reaction was added to 200 μ l of freshly prepared ice cold 0.01M Potassium phosphate ($K_2HPO_4 + KH_2PO_4$) pH 7.0 buffer containing 0.2% (w/v) sodium sulphite and 0.01M β mercaptoethanol. The resulting mix was then used to infect the youngest

expanded leaf of the cassava plant by mechanical inoculation through rubbing the infectious transcript.

2.3 Microbiological methods

2.3.1 Strains

2.3.1.1 *Escherichia coli* strains

Escherichia coli strain DH5- α (- F - λ - endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ 80dlacZ Δ M15 Δ (lacZYA-argF) U169, hsdR17 (rK - mK +) was used for standard plasmid propagation and maintenance.

The cells SURE (e14-(McrA-) Δ (mcrCB-hsdSMR-mrr)171 endA1 supE44 thi-1 gyrA96 relA1 lac recB recJ sbcC umuC::Tn5 (Kanr) uvrC [F' proAB lacIq Z Δ M15 Tn10 (Tetr) Amy Camr) (Stratagene, USA), C43 (DE3) (Lucigen, USA), ABLE[®] K (*E. coli* C lac(LacZ ω -) [Kanr McrA- McrCB- McrF- MrrHsdR(rK - mK -)] [F' proAB lacIq Z Δ M15 Tn10 (Tetr) (Stratagene, USA), XL10-Gold (Tetr Δ (mcrA)183 Δ (mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F' proAB lacIq Z Δ M15 Tn10 (Tetr) Amy Camr (Stratagene, USA), DH5- α , and α -select (Bioline, UK) cells were used to optimize for good plasmid yield.

2.3.1.2 *Agrobacterium tumefaciens* strains

Agrobacterium tumefaciens strain LBA-1126 was used for the preparation electrocompetent *A. tumefaciens* cells for use in agro-infiltration.

2.3.1.3 *Saccharomyces cerevisiae* strains

All homologous recombination in *Saccharomyces cerevisiae* (yeast) was carried out using yeast strain Y10000 (MAT α , his3 Δ 1; leu2 Δ 0; lys2 Δ 0; ura3 Δ) (Sherman, 2002).

2.4 Media

2.4.1 Bacterial media

2.4.1.1 *E. coli* and *Agrobacterium* media

Luria-Bertani Broth (LB: 10g/l NaCl, 10g/l tryptone, 5g/l yeast extract, pH 7) and **Luria-Bertani Agar** (LBA: LB supplemented with 15g/l agar) were used for general culture of *E. coli* and *Agrobacterium*.

2.4.2 *Saccharomyces cerevisiae* media

YPD (10g/l yeast extract, 20g/l peptone, 20g/l D-glucose, pH 6.5) and **YPDA** (YPD supplemented with 15g/l agar) were used for growth and maintenance of *S. cerevisiae*.

Yeast Synthetic Drop out media (YSDM: 1.7g/l yeast nitrogen base, 5g/l ammonium sulphate, 0.77g/l yeast drop out mix, 20g/l agar) was used for selection of *S. cerevisiae* transformants. Yeast strain Y10000 is a uracil mutant and the pYES2.1 vector used in this study contains the *ura3* gene conferring uracil production. Transformants were therefore selected on YSDM uracil deficient media.

2.4.3 Media supplements

2.4.3.1 Antibiotics

All antibiotics were supplied by Melford. Stock solutions were prepared and stored as stated (Table 2.1). Antibiotics prepared with sterile distilled water were sterilized by filtration through a 0.22 μ m filter (Millipore); those prepared with Methanol were not filter sterilised.

Table 2.1: Antibiotics used in this study

Antibiotic	Diluent	Stock concentration (mg/ml)	Working concentration (µg/ml)	Storage (°C)
Ampicillin	Water	100	100	-20
Kanamycin	Water	100	100	-20
Carbenicillin	Water	50	50	-20
Rifampicin	Methanol	20	20	-20

2.4.4 Culture maintenance and storage

2.4.4.1 *E. coli* culture maintenance

E. coli cells were kept frozen at -80°C. Cells were gently spread on plates without antibiotics and then used to make chemically competent or electrocompetent cells.

Glycerol stocks of transformed *E. coli* cells were made by adding 900 µl overnight cultured cells in LB media to 900 µl of 30% glycerol. The vials were snap frozen in liquid nitrogen and kept at -80°C. Plates were used for temporary maintenance of the transformed bacteria.

E. coli was routinely grown on LBA plates with antibiotics added as required and incubated at 37°C, overnight. Liquid cultures were established in 5 ml LB, containing appropriate antibiotics for selection, by inoculating with a single colony from an agar plate using flame-sterilized inoculating loop and incubated in a shaking incubator overnight at 37°C. 50ml cultures were used when required for large-scale plasmid extractions.

2.4.4.2 *S. cerevisiae* maintenance

Saccharomyces cerevisiae was routinely grown on YPDA and incubated at 28°C for 2-3 days. Liquid cultures were established in YPD by inoculating with a single colony from an agar plate using flame sterilized inoculating loop and incubated by shaking overnight at 28°C.

2.5 Plasmid Extraction

2.5.1 *E. coli* plasmid extraction

Small-scale extraction of plasmid DNA (approximately 1 µg/ml) was performed using 1.5 ml of an overnight bacterial culture with a Miniprep Plasmid Extraction Kit (Thermo Scientific). Large-scale plasmid extractions (approximately 1 mg/ml) were performed using 50 ml of overnight liquid culture using the Midiprep Plasmid Extraction Kit (Thermo Scientific). Both kits were used according to the manufacturers' instructions. The extracted plasmid DNA was stored at -20°C.

2.5.2 Yeast plasmid extraction

Plasmids were extracted from 1.5 ml overnight culture of *S. cerevisiae* grown in YPD and supplemented with the appropriate antibiotics (Ampicillin for pYES2.1 and Kanamycin for pCAMBIA 0380), using Zymoprep Yeast Plasmid Miniprep II Kit (Zymo Research). Incubation of pelleted yeast cells with the Zymolase enzyme allows the breakdown of the fungal cell walls. This was followed by a lysis step, then clearing the lysates via centrifugation. The extracted yeast plasmid was then used to transform electrocompetent *E. coli* cells by electroporation.

2.6 Molecular Techniques

2.6.1 RNA extraction

RNA was extracted from cassava leaves and herbaceous plants using the CTAB method as described previously (Alicai *et al.*, 2007). Briefly, about 200 mg of leaf tissue was ground in liquid nitrogen and the powder was added to a 2 ml tube. 800 μ l of prewarmed (65°C) CTAB buffer (100 mM Tris-HCl, pH 8.0, 1.4 M NaCl, 20 mM EDTA, 2% CTAB (w/v), 2% PVP40 (w/v), 0.2% β -mercaptoethanol (v/v) was added and mixed vigorously. The mixture was incubated at 65°C for 15-30 min with periodic vortexing at 10 min intervals and then centrifuged at 12000 *g* for 10 min.

About 750 μ l of the aqueous phase was collected and an equal volume of phenol:chloroform:isoamylalcohol (25:24:1) pH 8.0 was added and centrifuged. The uppermost phase was collected into a 1.5 ml tube and the RNA was precipitated using equal volume (750 μ l) of cold isopropanol at -20°C. The RNA was resuspended in 30 μ l of DEPC treated water. DNase treatment was done by setting up a 10 μ l reaction containing 1 μ g (RNA), 1 μ l of 10X MgCl₂ buffer and 1 unit of DNase enzyme. The treated RNA was then used for first strand cDNA synthesis.

2.6.2 First-strand cDNA synthesis

First-strand cDNA was synthesized using Molony Murine Leukemia Virus Reverse Transcriptase (M-MuLV RT) (Thermo Scientific, Lithuania) according to the manufacturer's instructions and oligo (dT)₁₈ was used as the initial primer to synthesise the first strand of cDNA. The synthesized cDNA was then used as template in a PCR reaction. First-strand cDNA synthesis reactions were performed using First Strand cDNA Synthesis Kit (Thermo Scientific, Lithuania); the reaction was set up as follows;

Template RNA (Total RNA)	0.1- 5 μ g
Oligo (dT) ₁₈ Primer (10 μ M)	1 μ l
Nuclease Free Water	11 μ l total volume

The samples were then incubated at 65°C for 5 minutes, centrifuged and then chilled on ice.

To the reaction above, the following were added in order;

5X Reaction Buffer	4 μ l
RiboLock RNase inhibitor	1 μ l (20U/ μ l)
10mM dNTP Mix	2 μ l
M-MuLV Reverse Transcriptase	2 μ l (20U/ μ l)
Nuclease Free Water	20 μ l total volume

The reaction was incubated at 37°C for 60 minutes then terminated by incubation at 70°C for 15 min.

2.6.3 Restriction digestion of DNA

Restriction digestion of DNA was carried out on plasmid DNA before proceeding with yeast-based homologous recombination and in order to linearize the backbone vector. Restriction digestion was also done routinely to check correct identity of plasmid vectors after plasmid DNA purification from *E. coli*. Restriction enzymes were purchased from Thermo Scientific or from New England Biolabs and reactions were carried out according to the manufacturer's protocols. Unless otherwise specified, restriction digestion reactions were carried out mixing the following components; 1 μ l 10X buffer (specific for the enzyme used, according to the manufacturer's instructions), 1 μ g plasmid DNA, 1 μ l restriction enzyme, and sterile deionized water up to a final volume of 10 μ l. The reaction was incubated at 37°C for 30 minutes for Fast Digest™ (Thermos Scientific, Lithuania) enzyme restrictions, or for 2 hours when New England Biolabs enzymes were used.

2.6.4 PCR Primers

PCR and sequencing primers were initially designed based on the UCBSV genomes in the NCBI database for use in sequencing and assembling the UCBSV Kikombe genome. Once the UCBSV genome was assembled, the subsequent primers were designed based on the assembled UCBSV Kikombe genome (GenBank accession KX753356). All designed primers were synthesized at Integrated DNA Technologies (IDT). Details are in the appropriate chapters.

2.6.5 Polymerase Chain Reaction

2.6.5.1 Conventional PCR

For analytical PCR, low fidelity PCR was performed in a 20 μ l final reaction volume which contained 7 μ l of sterile distilled water, 10 μ l of 2X PCR ready mix containing Taq DNA polymerase (Dream Taq DNA polymerase, Thermo Scientific) and 1 μ l each of the forward (10 μ M) and reverse (10 μ M) primers. The appropriate positive and negative controls were included for each of the PCR.

PCR was carried out in a PTC-200 Engine thermocycler (Bio-Rad) with the program as an initial denaturation at 95°C for 2 min, followed by 36 cycles of denaturation at 95°C for 30 sec, annealing at $T_{\text{annealing}}$ (depending on the primers) for 30 seconds and extension at 72°C for 1 minute per kb (depending on the length of the expected PCR product) and a final extension at 72°C for 10 min.

2.6.5.2 Colony PCR

Colony PCR was used to test *E. coli* transformant colonies growing on the appropriate selective plates for presence of the desired recombinant plasmid after transformations with plasmid

DNA. The same 2X DreamTaq Green Mix prepared as described for conventional PCR was used also in this assay. The PCR reactions were set up using the following reagents;

10 μ l 2X DreamTaq Green Mix, 1 μ l of 10 μ M Forward primer, 1 μ l of 10 μ M Reverse primer, and sterile deionised water up to a final volume of 20 μ l. Half a colony was picked from the plate and placed into the PCR tube. The same cycles as for conventional PCR were followed, although the initial denaturation step of 95°C was extended to 10 minutes. If the PCR reaction gave amplification of an amplicon of the expected size, then the other half of the colony remaining on the plate was used to set up an overnight culture in the appropriate selective liquid medium with the appropriate conditions in order to perform purification of the plasmid DNA.

2.6.5.3 High fidelity PCR

High fidelity PCR was used in the amplification of PCR products for sequencing. Phusion High Fidelity DNA Polymerase (Thermo Scientific) was used to avoid amplification errors as Phusion has a proof reading mechanism. A 50 μ l final reaction was set up as follows;

5X Phusion buffer (HF)	10 μ l
10 mM dNTPs	1 μ l
Forward primer (10 μ M)	2 μ l
Reverse Primer (10 μ M)	2 μ l
Template cDNA	5 μ l
Phusion DNA polymerase (2 U/ μ l)	0.5 μ l
Water	29.5 μ l

The following program was used to carry out PCR reactions. An initial denaturation at 98 °C for 30 sec, followed by 34 cycles of the specific cycle depending on the expected size of the PCR product (denaturation 98 °C for 10 sec, annealing at $T_{\text{annealing}}$ for 30 sec, extension at 72 °C for 30 sec per kb depending on the length of the expected PCR product) and final extension at 72°C for 10 min.

2.6.6 Agarose Gel Electrophoresis

Agarose gel electrophoresis was used to visualize PCR products, plasmid extractions, nucleic acid extractions, restriction digestions, nucleic acid synthesis reactions and to separate DNA fragments prior to purification.

A 1% (w/v) agarose (Type 1 Molecular Biology grade, Bioline) gel in 1X TAE buffer (40 mM Tris, 20 mM acetic acid and 1 mM EDTA pH (8.0) was used to separate the DNA fragments. The agarose was microwaved to dissolve and cooled to approximately 50°C to 60°C. The gel was prestained in ethidium bromide (2 µl from a 5 mg/ml stock solution per 50 ml of gel) or Midori Green Nucleic acid staining solution (2 µl from a 20,000X stock solution per 50 ml of gel). The solution was poured into UV-transparent gel tray (Bio-Rad) enclosed in a gel cast (Bio-Rad) and a comb was inserted in order to form wells. Once set (polymerized), the gel tray was released from the cast and the comb removed, then placed in an electrophoresis tank (Bio-Rad) filled with 1XTAE buffer.

2.6.6.1 Loading and running gels

A 5X loading dye (Bioline) was used with all nucleic acid samples except PCR products generated using Dream Taq ® DNA polymerase (Thermo Scientific) as the buffer supplied with this polymerase contains the loading dye. The gel was run at 100V for 45 min. Molecular weight ladder (HyperLadder 1) shown in (Figure 2.1) was run beside the DNA samples in order to determine the sizes and concentration of the expected DNA fragments.

HyperLadder I

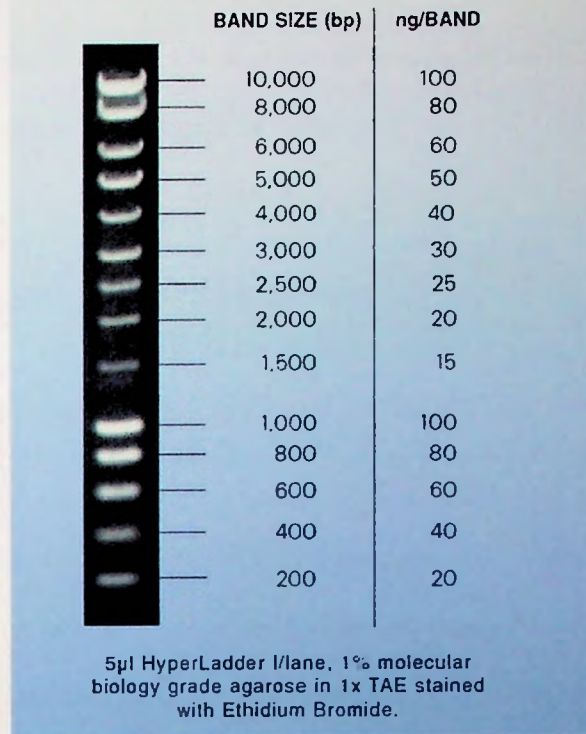


Figure 2.1: 5 µl of HyperLadder 1 were run on the first lane of each agarose gel to determine DNA product sizes and to estimate concentration.

2.6.6.2 Visualisation of Nucleic acids

The Ethidium Bromide or Midori Green, present in the agarose gel, intercalates into DNA and can be visualized using 2000™ imaging system (Bio-Rad Laboratories, Segrate, Milan Italy), which encompasses a UVP transillumination system and camera to allow recording of the image. DNA bands were excised from the agarose gels and purified using a GeneJET® Gel Extraction Kit (Thermo Scientific) as according to the manufacture's provided protocol. The purified DNA was then used for cloning, yeast recombination, and sequencing

2.6.7 Cloning of PCR products for sequencing

Purified DNA fragments were cloned into the pJET cloning vector for sequencing. The cloning was done using the CloneJET PCR cloning Kit (Thermo Scientific). Ligations were set up according to the manufacturer's protocols as follows;

2X Reaction Buffer	10 μ l
PCR product	1 μ l (0.15 pmol ends)
pJET1.2 (50ng/ μ l)	1 μ l (0.05 pmol ends)
T4 DNA Ligase	1 μ l
SDW	to 20 μ l

The ligation reaction was incubated at room temperature for 15 min and 5 μ l of the reaction was transformed into competent *E. coli* cells. The transformed cells were plated on LB + ampicillin (50 μ g/ml) plates and incubated overnight at 37° C. Colonies were screened for the presence of the desired insert by colony PCR using the pJET forward and pJET reverse sequencing primers. Positive colonies were grown in liquid LB ampicillin (50 μ g/ml) and shaken overnight at 37°C, and plasmids was extracted from the overnight culture using the Thermo Scientific Plasmid Miniprep Kit following the manufacturer's instructions. The plasmids were then sent for sequencing with either the pJET forward or pJET reverse sequencing primers.

The pJET vector shown in (Figure 2.2) allows the insertion of a blunt ended PCR products generated using the Phusion DNA polymerase at the multiple cloning sites (MCS) and has a lethal gene for non-transformed cells (thus there is no need for blue white screening that requires use of X-gal).

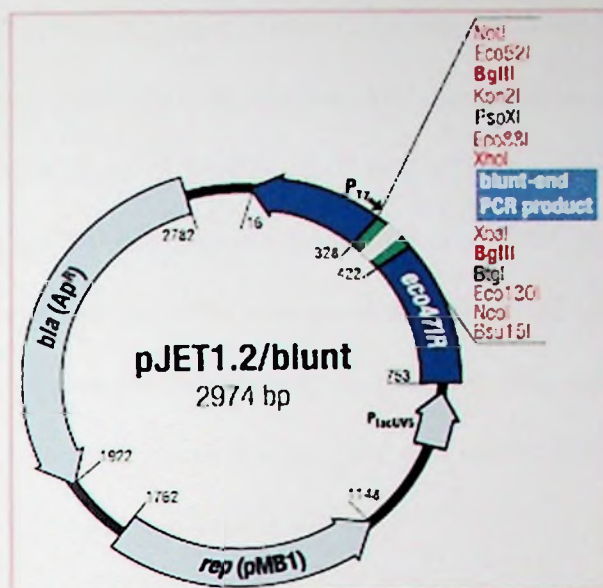


Figure 2.2: pJET Plasmid Map. The vector allows for cloning of inserts less than 10 kb into the MCS.

2.6.8 Ligation of DNA fragments into vectors

Ligations were conducted using T4 DNA ligation Kit (NEB Biolabs). Ligations were set up as follows;

10X T4 DNA Ligase Buffer	2 μ l
Vector DNA	50 ng (0.020 pmol)
Insert DNA	37.5ng (0.060 pmol)
T4 DNA Ligase (5U/ μ l)	1 μ l
SDW	to 20 μ l

The ligation reaction was incubated at room temperature for at least 20 min or incubated at 16°C overnight. 5 μ l was used to transform competent *E. coli* cells.

2.7 Preparation of chemically competent *E. coli* cells

Chemically competent *E. coli* cells were prepared using the CaCl_2 method described in (Dagert & Ehrlich, 1979). Briefly the *E. coli* cells were grown overnight with shaking at 250 g at 37°C. 100 ml of LB medium was inoculated with 1 ml of overnight saturated culture and shaken at 37°C until O.D₆₀₀ was 0.4 (usually 2-3 hours). The cells were placed in an ice bath for 10 min

and transferred to 50 ml prechilled falcon tubes. Cells were pelleted by centrifugation at 2700 g for 10 min at 4°C, the supernatant was removed and the pellet resuspended with 1.6 ml ice cold MgCl_2 - CaCl_2 solution (0.08 M MgCl_2 +0.02 M CaCl_2) by swirling on ice gently. Cells were incubated on ice for 30 min and then centrifuged at 2700 g for 10 min at 4°C. The supernatant was removed and the pellet resuspended with 1.6 ml ice cold 100 mM CaCl_2 by swirling on ice gently and then incubated on ice for 30 min. The cells were combined to one tube and 0.5 ml of ice-cold 80% glycerol was added and mixed. Cells were aliquoted in 100 μl , snap frozen in liquid nitrogen and kept at -80°C.

2.7.1 Transformation of chemically competent *E. coli*

An aliquot of competent cells was removed from -80°C and thawed on ice. 50 ng of DNA was added to 100 μl of chemically competent *E. coli* cells and incubated on ice for 30 min. The cells were heat shocked in a 42°C water bath for 45 sec and then placed back on ice for 2 min. 1 ml of pre-warmed (37°C) SOC was added and the cells were then incubated at 37°C with 230 g shaking for one to allow for the expression of the antibiotic resistance gene. An appropriate volume (typically 100 μl per plate) was spread onto LBA plates, supplemented with the appropriate antibiotic. Plates were allowed to dry, then inverted and incubated overnight at 37°C.

2.8 Preparation of electrocompetent *E. coli* cells

Electrocompetent *E. coli* cells were prepared according to the procedure described in (Sharma & Schimke, 1996) which involves several washes to remove salts on cells and treatment with 10% glycerol. Briefly an *E. coli* culture was prepared by inoculating 5 ml of liquid LB medium with a single colony and shaking overnight at 37°C. 0.5 ml aliquots of this starter culture were used to inoculate 10 flasks each containing 50 ml of liquid LB media. The ten flasks were

incubated for 4 h at 37°C with shaking. 0.5 ml aliquots of this starter culture were used to inoculate 10 flasks each containing 50 ml of liquid LB media. The flasks were incubated for 4 hours at 37°C with shaking. The cultures were transferred to 50 ml tubes and centrifuged at 4000 g for 10 min at 4°C. All the following steps were performed on ice with ice cold solutions. The pellets were resuspended in 2 ml of cold sterile deionised water (SDW) then topped up with chilled SDW to 20 ml. Cell suspensions were combined pairwise, resulting in five tubes each containing 40 ml each. These tubes were centrifuged at 5000 g for 10 min at 4°C. The resuspension procedure was repeated and the cell suspensions were combined in two tubes each containing 50 ml of the cell suspension. Centrifugation was repeated and the two pellets were each resuspended in 25 ml of ice cold sterile glycerol (10%). These two were combined into one 50 ml falcon tube and centrifugation was repeated. The pellet was resuspended in 1.5 ml of 10% cold sterile glycerol. 50 µl aliquots were transferred to 1.5 ml microcentrifuge tubes and snap frozen in liquid nitrogen and then kept at -80°C.

2.8.1 Transformation of electrocompetent *E. coli*

A 50 µl aliquot of *E. coli* was removed from -80°C and thawed on ice. 2 µl low salt concentration DNA was added (high salt concentrations can cause the sample to arc, killing the *E. coli*). The cells and DNA were transferred to an ice cold sterile 0.2 cm electroporation cuvette (Bio-Rad). Excess moisture was removed from the outside of the cuvette, to prevent arcing, which was placed in a Electroporator (Bio-Rad). A 2.5kV pulse was applied according to the manufacturers' instructions. 200 µl of LB was added to the cuvette following electroporation and the suspension transferred to a 1.5 ml eppendorf tube. This was incubated for 30 min at 37°C, 225 rpm. 100 of the transformation mixture was plated onto LBA plates supplemented with the appropriate antibiotic and left to dry. The plates were then inverted and

incubated at 37°C for 16 h. The transformed colonies were then used for plasmid extraction using Mini Prep Plasmid extraction Kit (Thermo Scientific)

2.9 Construction of plasmids by homologous recombination in yeast

Yeast homologous recombination exploits the yeast *Saccharomyces cerevisiae* to join together DNA fragments with overlapping homologous DNA sequences. The technique involves yeast gap repair cloning which joins overlapping DNA sequences. Backbone pYES2.1 plasmid was extracted from *E. coli* and linearized through restriction digestion with *EcoR*I and *Sac*I restriction enzymes. The method described by (Gietz and Woods 2002) was used to transform the yeast cells.

A starter culture of *S. cerevisiae* was grown from a single colony from a YPDA plate (10 g yeast extract, 20 g peptone, 0.048 g adenine hemisulphate, 20 g glucose and 20 g agar in 1 L sterile distilled water) in 10 ml YPD (10 g yeast extract, 20 g peptone, 0.048 g adenine hemisulphate and 20 g glucose in 1 L sterile distilled water) and grown overnight with shaking at 200 rpm at 28°C. This culture was added to 40 ml YPD in a 250 ml flask topped with a foam bung and incubated at 28°C for 5 hours at 220 rpm. The culture was decanted into a 50 ml falcon tube and centrifuged for 5 minutes at 4000 g at 4°C. The supernatant was discarded and the pellet resuspended in 25 ml sterile distilled water and the centrifugation was repeated. After discarding the supernatant, the cells were resuspended in 1 ml 0.1M LiOAc and transferred to a 1.5 ml Eppendorf tube, pelleted at 13,000 g for 15 sec and the supernatant was discarded. The cells were resuspended in 400 µl of 0.1M LiOAc, vortexed and transferred in 50 µl aliquots as required to new Eppendorfs. The cells were pelleted for 15 seconds, the supernatant was removed and then the following solutions were added in order into each Eppendorf: 240 µl 50% PEG 3350, 36 µl 1 M LiOAc, 50 µl salmon sperm DNA (2 mg/ml, boiled for 5 min at 95°

and then chilled on ice until use), 34 μ l of mixed DNA fragments in appropriate concentrations containing the different sections for each plasmid being created. The cells were vortexed for 1 min to resuspend the pellet and incubated at 30°C for 30 min and then 42°C for 30 min. The cells were pelleted at 6000 g for 15 seconds and the supernatant was removed. The cells were resuspend in 500 μ l SDW by pipetting up and down gently and 100 μ l was spread onto SM plates (1.7 g yeast nitrogen base, 5 g ammonium sulphate, 20 g glucose, 0.77 g yeast dropout mix (-uracil), 20 g agar) and incubated upside down at 28°C for 3 days.

Colony PCR was used to test the transformant yeast cells for the presence of the insert, positive colonies were used to inoculate 10 ml of YSDM and incubated overnight at 28°C with shaking. Yeast cells were harvested and plasmid extracted using the Zymoprep II yeast Plasmid Miniprep Kit according to the manufacturer's instructions. In order to have a high number of copies, the plasmid extracted from yeast was transformed into electrocompetent *E. coli* cells and then purified again from the *E. coli*.

2.10 *In-vitro* Transcription of the infectious clone

Plasmids were linearized by *Kpn*I digestion prior to transcription. The linearization of the plasmid allows the production of "run-off" transcripts which all stop at same place after the poly (A) from the template plasmid DNA. All enzyme and associated buffers were removed using the GeneJET PCR Purification Kit (Thermo Scientific). *In-vitro* transcripts were generated using a Riboprobe SP6 *in-vitro* Transcription system (Promega) according to the manufacturer's guidelines. The transcription was done in conjunction with a Ribo m⁷G Analog (Promega). The cap structure is incorporated into RNA synthesized *in-vitro* to mimic the capped structure of mRNA.

Briefly, the *in-vitro* transcription reaction was set up as follows;

Transcription Optimized 5X buffer	10 µl
DTT, 100mM	5 µl
Recombinant RNasin Ribonuclease inhibitor	10 µl
Nucleotides (2.5 µl each of 10mM rATP, r CTP, rUTP plus 2.5 µl of 1mM r GTP)	10 µl
Ribo m7G Cap analog, (5 mM)	5 µl
Linearized Template DNA (1 mg/ml in water or T.E buffer)	16 µl
SP6 RNA polymerase	40U
Nuclease free water to a final volume of	50 µl

The reactions were incubated at 37°C for 1 hr. To increase the RNA yield, an additional 40 units of RNA polymerase was added and incubated for an additional 1 hr.

The *in-vitro* transcription reactions were then used as inoculum to infect *N. benthamiana* and *N. clevelandii*, *N. tabaccum* and cassava plants.

2.10.1 Inoculation of plants with full length viral transcript

The *in-vitro* transcription reactions (containing RNA transcripts) were used as a source of inoculum to infect plants. Freshly prepared ice-cold 0.01M potassium phosphate (K_2HPO_4 + KH_2PO_4) pH 7.0 buffer containing 0.2% sodium sulphite and 0.01M β -mercaptoethanol was used as the infection buffer. The 50 µl *in-vitro* transcription reaction was mixed with 200 µl of infection buffer and the inoculum was kept on ice during the inoculation process. Test indicator plants were dusted with carborundum powder to act as an abrasive and the *in-vitro* RNA transcripts in infection buffer were rubbed gently on the leaf surface using gloved hands. After inoculation, the plants were gently sprayed with water to remove carborundum. The inoculated plants were maintained in the growth cabinet at 28°C. After 2 weeks, systemically infected leaves showing mosaic like symptoms were collected and RNA extracted for cDNA synthesis and subsequent virus detection by RT-PCR and sequencing.

2.11 Genome Assembly with Sequencher

DNA sequencing was done at GATC Biotech (London). The assembly of ABI sequence files was performed using Sequencher Version 5.0 Software. To increase the accuracy of the assembled viral genome, an average of four different clones were considered for each region of the genome. The sequences were trimmed to remove ambiguous data, data that had low confidence scores and data contaminated with vector sequences. The default trimming criterion in Sequencher were used in trimming the sequences and good quality sequence data was used for the genome assembly. Contigs were assembled using the "Assemble to reference" parameter in Sequencher using the UCBSV NC_014791.1 isolate in the GenBank as the reference sequence in the assembly. The consensus contig was exported and the sequence saved as the complete assembled genome of the UCBSV Kikombe isolate. The contig was subjected to a Blastx and the Blast hits on the query sequence were polyprotein of UCBSV isolates in the NCBI database in addition to the predicted potyviral polyprotein domains.

2.11.1 Annotation of the UCBSV Kikombe genome

The open reading frame of the assembled UCBSV genome was predicted by using the online ExPASy translate tool and the genome annotation done by aligning the translated sequences of the the UCBSV isolates in the NCBI alongside the UCBSV Kikombe genome. Predicted cleavage sites in the alignment were identified and the genome annotated by comparing the predicted cleavage sites of other UCBSV isolates in the NCBI database.

2.12 Serological methods

2.12.1 Triple Antibody Sandwich ELISA

ELISA was done to confirm the presence of UCBSV infection in the infectious clone infected plants and to analyse viral titres in the plants. High protein affinity microtitre plates were used. UCBSV specific IgG and Alkaline Phosphate conjugated antibodies purchased from DSMZ were used in the ELISA analysis.

2.12.2 ELISA buffers

Coating buffer (1.5g Na_2CO_3 , 2.93g NaHCO_3 , 0.2g NaN_3 dissolved in 1000ml deionised water, pH 9.6). PBS-Tween (8.0 g NaCl, 0.2 g KH_2PO_4 , 0.15g Na_2HPO_4 , 0.2g KCl, 0.2g NaN_3 + 0.5 ml Tween-20 per liter). Sample extraction buffer (0.05 M Tris containing 0.06 M sodium sulphite pH 9.8). Conjugate buffer (PBS-Tween + 2% PVP + 0.2% egg albumin). Substrate buffer (100 mM Tris base, 100 mM NaCl, 5 mM MgCl_2). All the buffers were sterilised by autoclaving at 121°C and stored at 4°C.

2.12.3 ELISA protocol

Purified IgG (1:250) was added to the coating buffer and 100 μl added to each well of a microtitre plate. The plates were covered and incubated at 37°C for 2-4 h. The plates were washed three times with PBS-Tween. 100 μl of 2% skimmed milk (blocking reagent) in PBS-Tween was added to each well and the plates incubated for 30 min at 37°C. The blocking reagent was removed and the plates tapped to dry.

Samples were extracted (1:20 (w/v) in sample extraction buffer. 200 μl aliquots of the test sample were added in duplicate to the wells and incubated at 4°C overnight. The plates were washed three times with PBS-Tween and 100 μl of monoclonal antibody (1:50) diluted in

conjugate buffer was added to each well. The plates were incubated at 37°C for 1 h and the primary antibody was removed and the plates washed three times. 100 µl of alkaline phosphate conjugated secondary antibody (1:1000) was added in conjugate buffer and the plates incubated at 37°C for 1 h. The plates were washed three times and 100 µl of freshly prepared substrate (10 mg p-nitrophenyl phosphate) dissolved in 10 ml of substrate buffer.

Chapter 3: Initial characterization of *Ugandan cassava brown streak virus* (UCBSV) “Kikombe” Isolate

3.1 Introduction

3.1.1 Classification of *Ugandan cassava brown streak virus* (UCBSV)

Ugandan cassava brown streak virus (UCBSV) is one of the viruses that has been reported to be the causal agent for cassava brown streak disease (CBSD), the other virus being *Cassava brown streak virus* (CBSV). Cassava brown streak disease (CBSD) is one of the main diseases affecting cassava production leading to yield losses of up to 70% (Manyong *et al.*, 2010). Initial work showed that CBSD was caused by one virus (CBSV), however, further studies and characterizations have identified another closely related virus to CBSV which has reduced sequence similarity to other CBSV strains and differs in the symptomatology presented in a range of different infected host plants (Yadav *et al.*, 2011; Abarshi *et al.*, 2012). The second virus was originally referred to as *Cassava brown streak Uganda virus* (CBSUV) (Patil *et al.*, 2015), however, the name *Ugandan cassava brown streak virus* (UCBSV) has been confirmed by the International Committee for Taxonomy of Viruses (ICTV). Although it has been named *Ugandan cassava brown streak virus* (UCBSV) suggesting its occurrence in Uganda, the virus is not only restricted to Uganda and has been found in different agro-ecological zones in East and central Africa and the origin remains undetermined (Ndunguru *et al.*, 2015). Both CBSV and UCBSV have been placed in the family *Potyviridae*, genus *Ipomovirus* (Mbanzibwa *et al.*, 2009b; Winter *et al.*, 2010; Ndunguru *et al.*, 2015). The classification and placement of the CBSVs in the *Ipomovirus* genus was based on phylogenetic relationships and sequence homology with *Sweet potato mild mottle virus* (SPMMV), *Cucumber vein yellowing virus* (CVYV) and *Squash vein yellowing virus* (SqVYV) (Monger *et al.*, 2001). *Sweet potato mild mottle virus*

(SPMMV) is the type member in the *Ipomovirus* genus. Most of the viruses infecting plants belong to the *Potyviridae* family. Electron microscopic analysis of extracts from cassava brown streak disease infected leaves shows the viruses to appear as filamentous and flexuous particles with a clear modal length of 650nm with a helical symmetry. Most of the CBSVs characterization was initially based on the coat protein (CP) sequences (Monger *et al.*, 2001), however the recent advances in the sequencing and assembling of a number of complete genomes for CBSV and UCBSV isolates has allowed a clear classification and identification of the viruses (Monger *et al.*, 2010; Mbanzibwa *et al.*, 2009b; Ndunguru *et al.*, 2015).

3.1.1.1 UCBSV genome organization

The UCBSV genome is a monopartite positive stranded RNA. For most of the sequenced and published UCBSV isolates in the NCBI database, the complete UCBSV genome (genus *Ipomovirus*, family *Potyviridae*) has been found to consist of 9070 nucleotides which are translated to produce a polyprotein of 2902 amino acids that is autoproteolytically cleaved by virus encoded proteases to yield ten mature proteins that function in different roles in the virus infection cycle. The 5' UTR consists of 134 nucleotides while the 3' UTR is 226 nucleotides. All the genomes have conserved amino acid sequences representing the positions at which proteolytic cleavage occurs in the mature protein. The cleavage sites have been mainly studied for the main protease enzyme NIa in the viral genome (Adams *et al.*, 2005). Furthermore, another protein referred to as PIPO (Pretty interesting potyviral open reading frame) is found in the +2 reading frame in the P3 protein of the UCBSV genome (Chung *et al.*, 2008).

3.1.2 Use of plant indicators in virus infections

Herbaceous hosts have provided valuable systems for studies on different viruses mainly because the different plants may sometimes respond differently in the symptomatology when infected with similar or related viruses. In this case, the host plants can thus be used in the diagnosis of viral infection especially if there is a difference in symptoms between different isolates (Mohammed *et al.*, 2012). In some cases, a given virus strain may be able to infect one host but not the other. Furthermore, experiments with plant host indicators allows for the selection of susceptible laboratory hosts in which constructs for generating transgenic lines with viral disease resistance can be applied in the “proof of concept” laboratory studies before the constructs can be transformed into ultimate field hosts. This has mainly been applied in generating transgenic *N. benthamiana* lines for resistance to cassava brown streak disease which mainly targets coat protein (CP) of the virus (Taylor *et al.*, 2004). In this way, the model laboratory host plants are initially tested for their response to the transgenes before embarking on the laborious task of generating transgenic cassava plants.

3.1.3 Symptom and host range of CBSV and UCBSV

A study by Mbanzibwa *et al.*, (2009b) in which a range of indicator plants including *N. rustica*, *N. tabacum* and *N. benthamiana* were challenged with different CBSV/UCBSV isolates showed that there is a variation in the symptoms presented when the two viruses are infected in different solanaceous hosts. Differences are mainly observed in the *N. benthamiana* host in which the CBSV virus isolates are found to cause severe necrosis and death of the plant within weeks after infection. On the other hand, the UCBSV isolates cause mild leaf mosaic symptoms in *N. benthamiana* plants (Mohammed *et al.*, 2012; Winter *et al.*, 2010). The symptoms induced on *N. rustica* and *N. tabacum* include necrotic local

lesions, stunted growth and mottling of leaves but were not different between isolates of CBSV and UCBSV. CBSVs are reported to infect the following hosts: *Nicotiana tabacum*, *N. rustica*, *N. glutinosa*, *N. debneyi*, *N. benthamiana*, *Datura stramonium*, *Petunia hybrida*, *Chenopodium quinoa*, and *C. amaranticolor*. *N. benthamiana* shows a differential reaction to the virus species (Mohammed *et al.*, 2012). *Petunia hybrida* has been reported to be a particularly susceptible host (Lister, 1959). Local necrotic lesions in *N. tabacum* SR1 plants are a response to infection with CBSV isolates which eventually results in plant deterioration and death whereas UCBSV isolates cause chlorotic mottle and leaf curl symptoms (Winter *et al.*, 2010; Monger *et al.*, 2001).

3.1.4 Transmission of *Ugandan cassava brown streak virus* (UCBSV)

Most of the studies on understanding the pathogenicity of CBSVs have been challenged by the low transmission potential of the virus under laboratory conditions and the lack of infectious clones of the virus (Ogwok *et al.*, 2010; Patil *et al.*, 2015). For CBSD, Storey, (1936b) showed that the virus was graft transmissible and that symptoms could be produced on plants free of mosaic symptoms. For plants grafted with *Cassava brown streak virus* (CBSV) alongside *African cassava mosaic virus* (ACMV), both disease symptoms could be distinguished on the leaf.

Studies by Maruthi *et al.*, (2005) have shown transmission of CBSV by two whitefly species *Bemisia afer* (Priesner & Hosny) and *B. tabaci* (Gennadius). Adults of whiteflies were given access to CBSV by containing them in clip cages on CBSV-infected cassava plants. Whiteflies were then transferred, together or separately, to CBSD-susceptible cassava plants of variety Albert in a controlled environment. Transmission of CBSV was found to be sporadic and occurred in three of seven experiments when inoculated by *B. afer* and *B.*

tabaci or *B. tabaci* alone, but not by *B. afer* alone. The challenge with vector transmission is that it is considered to be slow under laboratory conditions (Maruthi *et al.*, 2005).

One of the most commonly used technique for characterising plant viruses is mechanical transmission, achieved mainly by sap inoculation. Efficient transmission of CBSV or UCBSV to the common laboratory host *N. benthamiana* has been reported to be essential for both basic and applied studies of the virus. However, there are a number of factors affecting mechanical transmission efficiency of CBSV and UCBSV to *N. benthamiana* and *N. debneyi* which include the effects of buffer composition, antioxidants, inoculum concentration, plant age and temperature (Ogwok *et al.*, 2010). It has also been reported that higher temperatures (30°C) favoured rapid symptom initiation compared to low temperatures (21°C). Lister, (1959) also found that CBSD viruses were sap transmissible between cassava plants and from cassava to range of herbaceous indicator plants.

3.1.5 Aims of this chapter

The UCBSV Kikombe isolate used in this study had not been characterised at sequence level so hence initial work involved sequencing the complete genome in order to use the genome sequence in constructing an infectious cDNA clone of the virus. A viral genome overlapping strategy was used to amplify the full-length nucleotide sequence of the UCBSV Kikombe isolate under this study. Furthermore, this chapter shows the study of different host ranges of the UCBSV Kikombe isolate. Before constructing an infectious clone of a new uncharacterized UCBSV virus, it was important to determine that the wildtype virus was able to be mechanically transmitted to herbaceous plants that were to be used in the experiments to test the constructed infectious clone. This chapter thus shows the preliminary experiments that were undertaken to study the best host indicator plants which could be infected with the

UCBSV Kikombe wildtype isolate used in this study which later formed a basis for the choices of plants that were ultimately chosen to test the constructed UCBSV infectious clone.

3.2 Results

3.2.1 The UCBSV isolate “KIKOMBE”

Cassava leaves from CBSD infected cassava plants showing characteristic foliar CBSD symptoms collected from Tanzania (East Africa) were used as the original source of isolate. The leaves were infected with a *Ugandan cassava brown streak virus* (UCBSV) Kikombe isolate and this was confirmed by RT-PCR using CBSDDF2 (5' GCT MGA AAT GCY GGR TAY ACA A 3') and CBSDDR (5' CGA TAT GGA GAA AGR KCT CC 3') degenerate diagnostic primers which give a characteristic 440bp fragment with UCBSV isolates and a 340bp PCR product with CBSV isolates (Mbanzibwa *et al.*, 2011a). The obtained 440bp PCR product (Figure 3.1) was cloned into pJET cloning vector and sequenced. The sequence that was obtained was used in BLAST search (blastn) in the NCBI and was closely related to the coat protein (CP) encoding sequence and the 3' UTR of the UCBSV (Figure 3.2). After confirming the identity of the virus source, the infected cassava leaves were thus used as the original UCBSV Kikombe virus source for all the subsequent work which included inoculation of different host indicator plants, sequencing the complete genome of the isolate and constructing infectious clone of the isolate.

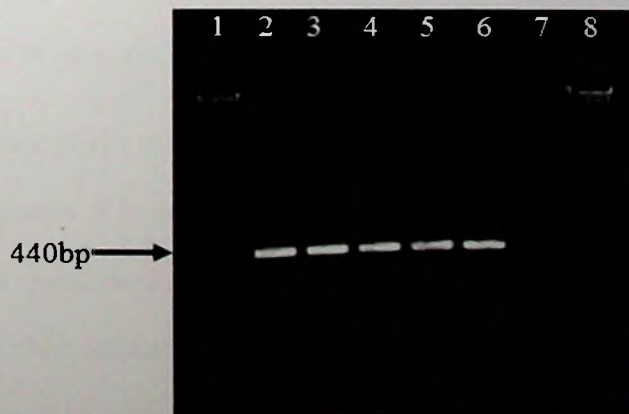


Figure 3.1: UCBSV infection confirmation in the original CBSD infected cassava leaves. Lane 1 and 8 contain 5 µl of HyperLadder (Bioline). Lane 2-6; 440bp RT-PCR product obtained from CBSD infected cassava leaves. Lane 7 is negative control (healthy plant) with no amplification.

Ugandan cassava brown streak virus partial CP gene for polyprotein, Coat Protein region, clone 11-clone5KeNTST
Sequence ID: [embL1560277.1](#) Length: 1670 Number of Matches: 1

Range 1: 1235 to 1670 [GenBank](#) [Graphics](#)

Score	Expect	Identities	Gaps	Strand
730 bits(395)	0.0	424/438(97%)	2/438(0%)	Plus/Plus
Query 463	GAAATGCCGGGTACACAAGTATTGAAAATAAATGGCTGGGTATAGATTTTCTTGCTGAAG			522
Sojct 1235	GAAATGCTGGATACACAAGTATTGAAAATAAATGGCTAGGTATAGATTTTCTTGCTGAAG			1294
Query 523	CTCAATTGTCCCAAAGTCAACTTGATATCAAGCATCAGATTTTGGCAGCGAATGTTGGAA			582
Sojct 1295	CTCAATTGTCCCAAAGTCAACTTGATATCAAGCATCAGATTTTGGCAGCGAATGTTGGAA			1354
Query 583	GGACAAGGACAAGGCTGTTTGCAATTAGCAGCTCCTGGAGATGATAACAATGTAGATAAAG			642
Sojct 1355	GGACAAGGACAAGGCTGTTTGCAATTAGCAGCTCCTGGAGATGATAACAATGTAGATAAAG			1414
Query 643	AAAGGCACACGACACATGATGTTAGTGCAACAGACACAGTTACAGTGGTGCCGCAATTG			702
Sojct 1415	AAAGGCACACGACACATGATGTTAGTGCAACAGGACAGTTACAGTGGTGCCGCAATTG			1474
Query 703	AATAAATAAATATTAGTGTCTTAATTGAATTTGAGTGAGATTATGGCTTAGTTAGAAA			762
Sojct 1475	AATAAATAAATATTAGTGTCTTAATTGAATTTGAGTGAGATTATGGCTTAGTTAGAAA			1534
Query 763	ACCAAGTTATATTTCAAATTTAAATTCAGTATTTAAATATTATATTTTATTTTCGAGTT			822
Sojct 1535	ACCGAGTTATATTTCAAATTTAAATTCAGTATTTAAATGTTATATTTTATTTTCGAGTT			1594
Query 823	TGGTGGAGTTTTGGATAGCCAATTATATTTTGGTTGGATAAGCCAAAGATGTATTTTGGAG			882
Sojct 1595	TGGTGGAGTTTTGGATAGCCAATTATATTTTGGTTGAGTAAGCCAAATGTATTTTGGAG			1654
Query 883	CCCTTTTCTCCATATCC 928			
Sojct 1655	CCCTTT--CTCCATATCC 1670			

Figure 3.2: NCBI BLAST of the sequence data for the 440 bp RT-PCR fragment. The sequence was identified as the coat protein (CP) of UCBSV in the database.

3.2.2 Host indicator plants infection to assess symptoms

Different host indicator plants were mechanically inoculated with UCBSV virus inoculum following the protocol in section (2.2.1.1).

It was observed that none of the tested different host indicator plants developed any obvious symptoms after four weeks of monitoring (Figure 3.3). However, earlier reports had shown that UCBSV isolates presented mild mosaic and rugosity in *N. benthamiana* plants (Mbanzibwa *et al.*, 2009a; Mohammed *et al.*, 2012). RNA was extracted from systemic leaves of all the infected host plants and cDNA synthesized for UCBSV virus detection. Diagnostic primers reported previously (Mbanzibwa *et al.*, 2011a) were used to detect infections in the inoculated plants. Despite the fact that none of the tested host indicator plants showed any symptoms, the UCBSV infection was confirmed to be present in all tested

plants by RT-PCR (Figure 3.4). A 1000bp PCR product was amplified from the infected plants using the degenerate forward primer 5'1KB F and 5'1KB R (Table 3.2). The PCR product was cloned and sequenced and the results confirmed the presence of UCBSV infection. The reason for the failure of symptom development could not be explained as the initial virus source had been confirmed to contain UCBSV infection. Thus, this experiment was repeated with optimizations in a number of factors which included variations in the temperature, buffer composition among others (Ogwok *et al.*, 2010).

After the repetitions, of the indicator plants tested by mechanical inoculation infection with the wild type UCBSV Kikombe isolate, it was observed that mild mosaic CBSD symptoms were obtained in *N. benthamiana* plants. The UCBSV infection could also be detected in the systemic leaves of symptomatic *N. benthamiana* plants by RT-PCR. The mild symptoms observed with the wild type UCBSV Kikombe isolate were consistent with the symptoms reported for UCBSV isolates which include mild systemic mosaics but no local lesions (Winter *et al.*, 2010). In the *N. clevelandii*, *N. tabacum* *N. rustica* and *N. glutinosa* plants there were still no observable UCBSV symptoms with the wild type UCBSV Kikombe isolate although RT-PCR and sequencing confirmed the presence of UCBSV infection in these plants. The *N. benthamiana* and *N. clevelandii* plants were thus chosen to be used in the subsequent experiments to test the infectivity of the constructed *in-vitro* UCBSV infectious clone in indicator plants.

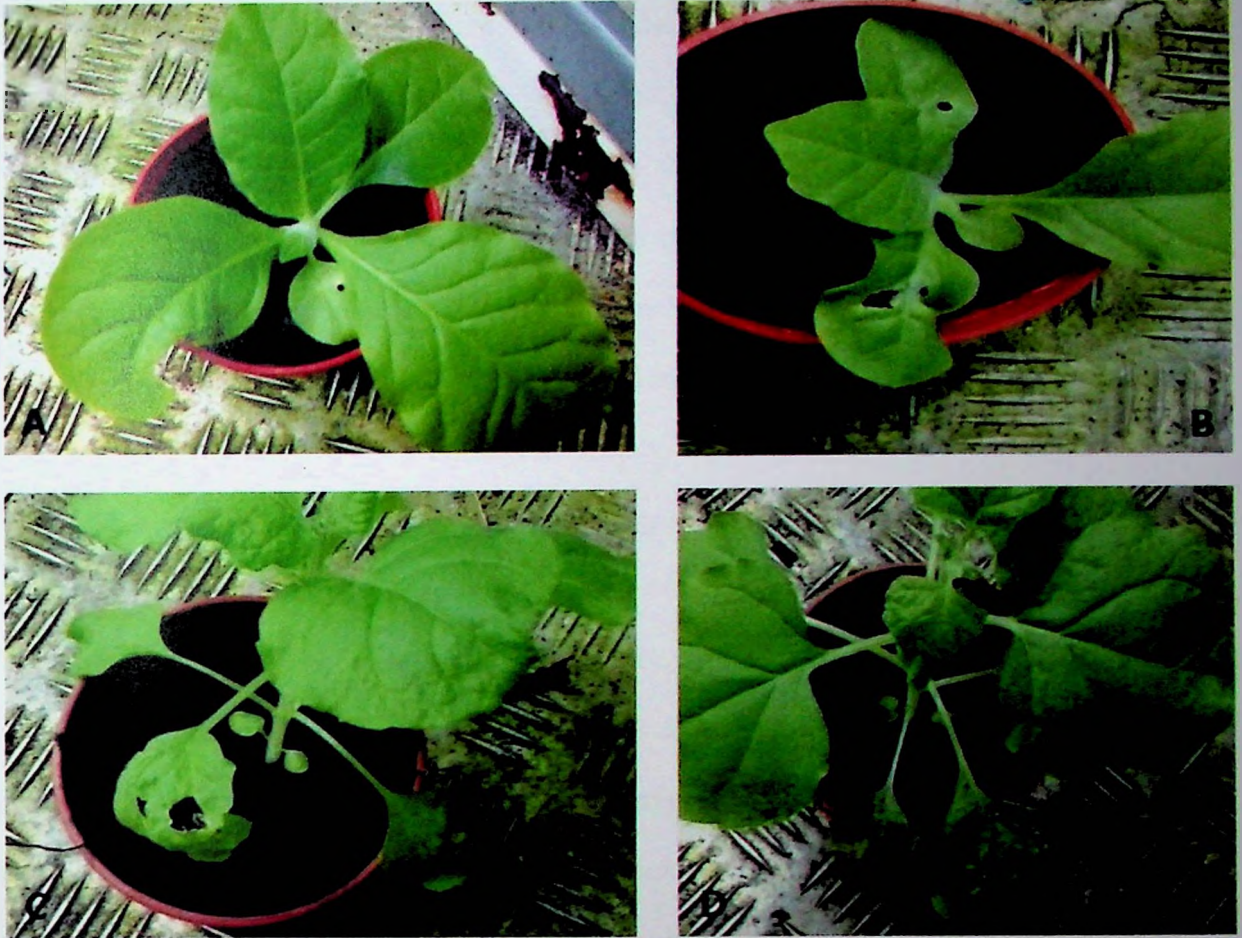


Figure 3.3: Asymptomatic *N. tabacum* (A and B) and *N. benthamiana* (C and D) plants infected with wildtype UCBSV Kikombe isolate.



Figure 3.4: Detection of wildtype UCBSV infection in *N. benthamiana* plants: Lane 1 and 8 contain 5 μ l of HyperLadder 1. Lane 2, 3 and 4 contain 1000bp PCR product amplified from *N. benthamiana* plants infected with wildtype UCBSV Kikombe isolate. Lane 5 is empty. Lane 6 positive control 9 (template from the original UCBSV infected cassava plants), Lane 7 contains the negative (healthy plant) with no amplification.

3.2.3 Designing primers for sequencing the complete genome of UCBSV

The UCBSV Kikombe isolate used in this study had not been sequenced so it was necessary to sequence the complete genome in order to use the genome sequence in constructing a cDNA infectious clone of the virus. Degenerate primers were designed to amplify the different overlapping sections of UCBSV Kikombe isolate as shown in (Figure 3.5) aiming for products of approximately 1-1.5 kb as amplification of larger products was unreliable. These were designed based on conserved regions of nucleotide sequences of UCBSV isolates in the NCBI database. The accession numbers and the sources of the UCBSV sequenced isolates used in the sequence alignments to design degenerate primers are shown in (Table 3.1). These genomes were aligned in Clone Manager Software and conserved regions in the genomes were used to design degenerate primers manually. The template used in the PCR reaction to amplify the different overlapping sections for sequencing the genome was oligo d(T)₁₈ primed cDNA synthesized from RNA extracted from cassava leaves confirmed to be infected with UCBSV Kikombe isolate.

Table 3.1: The identities and origins of selected UCBSV isolates in the NCBI database that were used in designing degenerate primers for sequencing the genome of the uncharacterized UCBSV Kikombe isolate collected from Mikocheni area, Tanzania.

Virus	Isolate Name	Country of origin	Accession number
UCBSV	MLB3	Muleba, Tanzania	FJ039520
UCBSV	Strain Ugandan	Namulonge, Uganda	FJ185044
UCBSV	Ma-43	Salima, Malawi	FN433933
UCBSV	Ke-125	Kilifi, Kenya	FN433930
UCBSV	Ke-54	Kilifi, Kenya	FN433931
UCBSV	Ma-42	Chitendze, Malawi	FN433932
UCBSV	Ug-23	Namulonge, Uganda	FN434109

Table 3.2: List of primers used to amplify the sections of the genome of UCBSV Kikombe isolate

Primer	Sequence	Position in NCBI UCBSV genomes	Expected product size
CBSV2F	5' AAA TAA ACA TGA CAT AAG AAT AC3'	3 to 25	
CBSV2R1	5' CAT ATG TCA TGA TTG TAA TAG 3'	2665 to 2685	2682bp
GAP 5' F	5' GAG TGC GAT GAC ATG TTT GA 3'	960 to 979	
GAP5'R	5' ATC AAC ATC TCT ATT CCC GGT TTG 3'	1806 to 1829	869bp
5' 1KB FWD	5' ATT GAR CCA ACC ACR AGY GGA CCY ATG 3'	2382 to 2408	
5' 1KB REV	5' ACT GAC CCT ATA TCY ACA TCA TCA AC 3'	3426 to 3451	1069bp
5'867bp F	5' GAT TTT GGY GAA CGC ATA AGT CCA AA 3'	3195 to 3220	
5' 867bp R	5' CTC CAT TTA AYT GGT TCA AGA GYA ACT C 3'	4035 to 4062	867bp
5'1125bpF	5' GCT GAG TGT GTB AAA GCT GCA AAG CC 3'	3750 to 3775	
5'1125bp R	5' ITG TTR RCT CTG GYT TWG CRA CYT GAA CTT C3'	4845 to 4875	1125bp
5'MID1KBF	5' GTA GCT CAT AAA ATA AGT CCA AA 3'	3822 to 3844	
5'MID 1KB R	5' CCA AGR YGY GCY CTR CCA AGT CTA ACC3'	4895 to 4921	1099bp
MID FF 1KB	5' GCS CCA ATG GGA TTY CCA GAA GAA GAG GG 3'	4758 to 4586	
MID RR 1KB	5' CAC TGA YCC AAC TCG AAC WGG CTC 3'	5775 to 5799	1041bp
3'MID 1.5KB	5' GGA CAT TYA ATC CTG ATC TTG TTG ART GG 3'	5461 to 5489	
3'MID 1.5KB	5' CCT TCY ACA CGC TTT CTA TCC 3'	6950 to 6970	1500bp
3'1.4KB F	5' GGG CTR CCG RRA AAG GWT GTG 3'	6916 to 6936	
3'1.4KB R	5' CCT GTA TAT TAA CTC CAT ATG CTT TAG C 3'	8304 to 8331	1400bp
922bpFS	GGGCAATTGAGGTTGCCAAAGC	8290 to 8311	
3' Final	TTTTTTTTTTTTTTTTTTTTTGGAC	Poly(A) tail	780bp
Kik440F	5'TCG AAA TGC CGG GTA CAC AAG 3'	8597	
Kik440R	5' GGA TAT GGA GAA AAG GGC TCC 3'	9034	440bp
CBSDDF2	5'GCT MGA AAT GCY GGR TAY ACA A 3'	8595	
CBSDDR	5' CGA TAT GGA GAA AGR KCT CC 3'	9035	440bp
3'667bp F	5' GCC TCC AAA RGT DTC ACA RTC AGC 3'	8138 to 8161	
3'667bp R	5' GCA CTG ACA TCR TGT GTY GTG TGC CC 3'	8782 to 8806	667bp

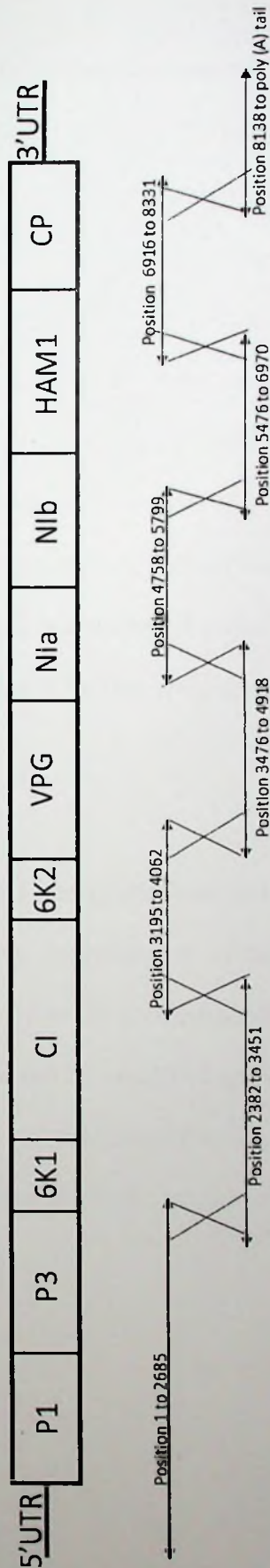


Figure 3.5: Schematic representation of overlapping genome sequencing strategy for the UCBSV genome: 8 overlapping PCR sections encompassing the UCBSV genome from the 5' UTR to the poly (A) tail were amplified.

3.2.3.1 Amplification of UCBSV in sections (1-8)

To sequence the UCBSV Kikombe genome, initially, the genome was divided into overlapping sections designated section (1-8) (Figure 3.5). PCR fragments of sizes ranging from 800bp to 2700bp were amplified from cassava cDNA. The sizes of the PCR fragments were considered optimal for PCR amplification from cDNA and subsequent cloning and sequencing. However, most of the PCR fragments were about 1000bp (1 kb) with the exception of section 1 which was about 2.7 kb (Figure 3.5). The expected sizes of the PCR products were section 1, 2.7 kb (Position 1 to 2685); section 2, 1069bp (Position 2382 to 3451); section 3, 867bp (Position 3195 to 4062); section 4, 1125bp (Position 3750 to 4875); section 5, 1099bp (Position 3822 to 4921); section 6, 1500bp (Position 5461 to 6970); section 7, 1400bp (Position 6916 to 8331) and section 8, 922bp (Position 8138 to poly A tail).

3.2.3.2 Amplification of the 3'end of UCBSV genome

Despite optimization on the quality of cDNA template used in the PCR reactions, some sections of the genome could not be PCR amplified from the synthesized cDNA. The 3' end could not be amplified using the Oligo dT primers and successful PCR amplification was only obtained when the Oligo dT primer was stabilized by the addition of a *Kpn*I restriction enzyme site to the primer sequence.

3.2.3.3 Amplification of 5' end of UCBSV genome

The 5' end of the UCBSV genome was amplified with primers CBSV2 F and CBSV2R (Table 3.2) after optimising the Phusion PCR conditions to setting up a 50 µl reaction containing 5X Phusion buffer (GC), 1 µl 10 mM dNTPs, 2 µl of the CBSV2 F and 2 µl of CBSV2R, 5 µl of cDNA template and 29.5 µl of water. The annealing temperature was optimized to 50°C and the extension time optimized to 2 min. With the optimizations of the genome amplification strategy for the 5' end of the genome, the 5' Random amplification of cDNA ends (RACE) procedure was therefore not performed.

3.2.3.4 Amplification of section 4 (position 3476-4918)

Amplification of section 4 (position 3476 to 4918) from cassava cDNA using the degenerate primers forward primer 5'GTAGCTCATAAAATAAGTRTAGG3' and reverse primer 5'CCAAGRYGYGCYCTRCCAAGTCTAACC3') was also problematic. The reason for the failure could not be easily explained as the 5' end of the genome had successfully been earlier amplified confirming that full-length first strand cDNA had been generated. Optimizations for successful amplification of this region included changing the primers used in the first-strand cDNA synthesis from using oligo d(T)₁₈ to using random hexamers. However, this did not enable amplification of this region. Finally, after obtaining the sequence results for the sections adjacent to section 4, this allowed for designing of specific primers to the UCBSV Kikombe genome which allowed amplification of this region.

3.2.3.5 RT-PCR to amplify the different sections of UCBSV

All the PCR fragments were finally amplified from the cassava cDNA (Figure 3.6) using Phusion Taq polymerase (section 2.6.5.3). The expected product were excised from the gel and purified and cloned into pJET vector. Plasmids were recovered from the transformed *E. coli* cells and checked by restriction with *Bgl*III enzyme to confirm that they contained the correct insert size before they were sent for sequencing. At-least 4-5 individual clones were sequenced for each fragment (Figure 3.7) and the consensus sequence representing the complete genome obtained (Appendix 1, GenBank accession KX753356).

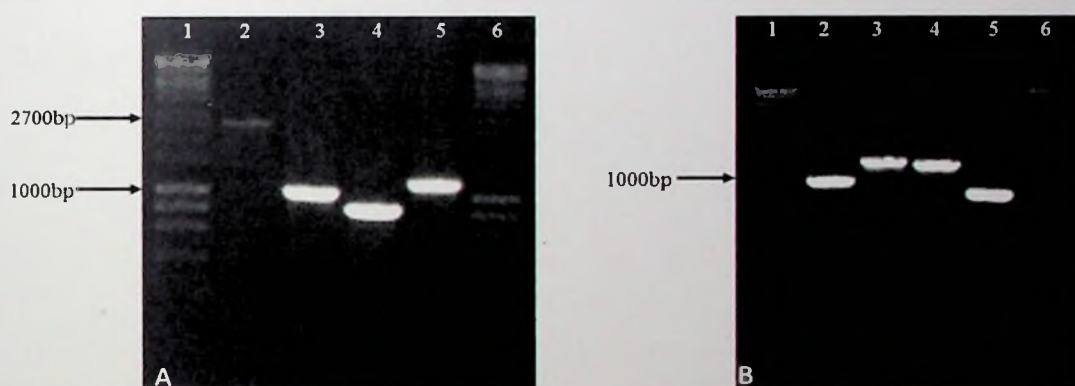


Figure 3.6: Purified PCR amplifications of the different regions in the UCBSV Kikombe genome. **A:** Lane 1 and 6 contains 5 μ l of HyperLadder 1, Lane 2 a 2.7 kb PCR fragment amplified with CBSV2F1 and CBSV2R (section 1), Lane 3 a 1069bp PCR fragment amplified by 5'1KBFWD and 5'1KBREV (section 2), Lane 4 a 867bp PCR fragment amplified by 5'867bpFwd and 5'867bpRev (section 3), Lane a 1125bp PCR fragment amplified by 5'1125bp F and 5'1125bp R (section 4). **B:** Lane 1 and 6 contains 5 μ l of HyperLadder 1, Lane 2 a 1000bp section amplified by Mid 1KB F and Mid 1KB R (section 5). Lane 3 a 1.5Kb PCR fragment amplified with 3'Mid 1.5KB F and 3'Mid 1.5KB R (section 6), Lane 4 a 1400bp PCR fragment amplified 3'1.4KB F and 3'1.4KB R (section 7), Lane 5 a 780bp PCR fragment amplified 922bpFS and poly(A) (section 8).

3.2.3.6 Complete genomic sequence assembly

The complete genomic sequence of the UCBSV Kikombe isolate was assembled using Sequencher Software Version 5.0. Firstly, the individual sequence data reads for each section of the UCBSV Kikombe isolate were imported into Sequencher Version 5.0 Software and the sequences trimmed to remove ambiguous data, data that had low confidence scores and data contaminated with vector sequences. Then contigs were assembled using the UCBSV NC_014791.1 isolate as the reference sequence. The consensus contig (Figure 3.7) was exported and the sequence saved as the complete assembled genome of the UCBSV Kikombe isolate (Appendix 1). The consensus contig was subjected to a blastx using the tools at the NCBI and the blast hits on the query predicted to be a polyprotein of UCBSV isolates in the NCBI database.



Figure 3.7: Alignment of different clones to assemble the genome of the UCBSV genome in Sequencher Software. 4-5 different clones were sequenced for each region of the genome. The accession number of the consensus sequence is GenBank KX753356

3.2.3.7 Description of assembled UCBSV Kikombe genome

The complete genome of UCBSV Kikombe isolate was found to consist of 9096 nucleotides including 26 nucleotides of the poly (A) tail (Appendix 1, GenBank accession KX753356). The 5'UTR consisted of 134 nucleotides as reported for most UCBSV isolates in the NCBI database. The 3'UTR consisted of 226 nucleotides and this was also consistent with reported UCBSV isolates (Mbanzibwa *et al.*, 2009a; Winter *et al.*, 2010). The genome consisted of polyprotein that was predicted to autocatalytically cleave into ten mature proteins present in other earlier sequenced UCBSV and CBSV isolates in the NCBI database namely; First Potyviral protein (P1), Third Potyviral protein (P3), First 6-kDa protein (6K1), cylindrical inclusion (CI), Second 6-kDa protein (6K2), Viral Genome Linked protein (VPg), Nuclear Inclusion a (NIa), Nuclear Inclusion b (Nib), HAM1 protein and the Coat protein (CP). Consistent with other CBSV genomes, the genome of UCBSV Kikombe isolate lacked the HC-Pro protein but contained the HAM1 gene incorporated between the Nib and coat protein (Mbanzibwa *et al.*, 2009a). The nucleotide sequence identity between the assembled UCBSV isolate across the genomic sequences of other UCBSV isolates in the NCBI database was between 86-92% while the protein identity was between 64-81%.

3.2.3.8 Genome confirmation and translation

The sequenced and assembled complete genome of the *Ugandan cassava brown streak virus* (UCBSV) Kikombe isolate was translated using the online expasy translate tool (<http://web.expasy.org/translate>) to obtain an open reading frame (ORF). The translated sequence was checked for the presence of the predicted cleavage sites that have been studied in potyviruses. Once all these predicted conserved sequences were identified in the genome, the genome was annotated (Appendix 5).

3.2.3.9 Annotation of the assembled UCBSV Kikombe genome

Genome annotation was carried out by aligning the translated sequences of the the UCBSV isolates in the NCBI alongside the UCBSV Kikombe genome. Predicted cleavage sites in the alignment were identified and the genome annotated by comparing the predicted cleavage sites of other UCBSV isolates in the NCBI database. The ORF consisted of 2902 amino acids (Appendix 5) and all predicted cleavage sites were present (Table 3.4).

3.2.4 Comparison between the individual proteins in UCBSV Kikombe (GenBank accession KX753356) against the reference UCBSV (NC-014791.1)

To confirm the genome identity of the UCBSV Kikombe isolate, a comparison was done between the encoded proteins against the sequence of the UCBSV (NC_014791) in the NCBI database. (Table 3.3) shows the percentage identity when individual nucleotide sequences and protein sequences were aligned and compared using Clone Manager Software (Version 5.0). The high percentages in polyprotein identities 94% obtained confirmed the origin of the UCBSV Kikombe isolate. A comparison of the protein identity between isolates is important as it also shows an indication of the selective pressure that is occurring in the genomes of these isolates which is an indication of positive or negative selection occurring in the genome.

Table 3.3: Sequence identity of each protein coding region between UCBSV Kikombe isolate (KX753356) with UCBSV (NC_014791.1).

Coding region		% identity between coding regions	
UCBSV Kikombe isolate and NC_014791.1			
Position in UCBSV genome	Encoded protein	% Nucleotide identity	%Amino acid identity
1-134	5'UTR	89	-
135-1220	P1	89	90
1221-2102	P3	93	95
2103-2258	6K1	93	100
2259-4142	CI	93	96
4143-4298	6K2	94	98
4299-4853	VPg	91	93
4854-5555	NIa	93	96
5556-7061	NIb	93	94
7062-7739	HAM1	89	89
7740-8843	CP	93	95
8844-9070	3'UTR	91	-
1-9070	Genome	92	-
135-8843	Polyprotein	92	94

3.2.5 Cleavage site analysis in the UCBSV genome

For most of the plant viruses in the *Potyviridae* family, the viruses express their protein as a large polypeptide that is cleaved into 10 mature proteins by virus encoded proteases with the NIa protein cleaving at most of the sites and the P1 protein cleaving at the P1/P3 site. A series of conserved amino acids at defined positions at the cleavage site are specified for accurate cleavage of the polyprotein. The predicted cleavage sites have been experimentally determined in a number of biochemical proteolysis assays for a few species and the sites are generally inferred by sequence comparison for a number of related species (Adams *et al.*, 2005). For this reason, the UCBSV assembled genome was checked for the presence of predicted cleavage sites for the NIa protease (Table 3.4). This was done by aligning the sequence of the assembled

UCBSV amino acid polyprotein with the polyproteins of annotated UCBSV genomes in the NCBI database. In the alignment, the conserved amino acids could be identified in other UCBSV genomes and inferred for the UCBSV Kikombe genome. All the predicted cleavage sites were as expected especially for positions that are amino acid specific for accurate cleavage for instance, there was a conserved glutamine (Q) at the -1 position for all the NIa cleavage sites (Table 3.4). The conserved Y/S cleavage site for the P1 protease was also present in the genome (Appendix 5).

Table 3.4: Amino acid sequences at the nine predicted cleavage sites in the UCBSV Kikombe isolate polyprotein

Junction	Amino acid position (from N terminus to C terminus) in UCBSV Kikombe genome						
	-6	-5	-4	-3	-2	-1	+1
P1/P3	D	L	I	D	L	Y	S
P3/6K1	G	F	V	E	V	Q	G
6K1/CI	D	F	I	E	K	Q	G
CI/6K2	E	Y	L	E	S	Q	C
6K2/VPg	E	L	V	E	K	Q	A
VPg/NIa	D	E	I	E	V	Q	V
NIa/NIb	N	T	I	T	V	Q	A
NIb/HAM1	C	Y	V	D	I	Q	T
HAM1/CP	L	T	I	D	V	Q	V
NIa consensus	X	X	V/I/L	E/D/T	X	Q	X

3.2.6 Bioinformatic analysis

3.2.6.1 Complete genome comparison between UCBSV Kikombe (KX753356) and other isolates

Phylogenetic analysis using Neighbour-Joining tree (MEGA6.0 Software) was performed to compare the complete genome of the sequenced UCBSV Kikombe isolate with the complete genomes of earlier sequenced UCBSV and CBSV isolates in the NCBI database (Figure 3.8). The phylogenetic tree gives an evolutionary relationship between the viruses and allows to understand the molecular diversity of isolates. A total of 27 complete genomes were retrieved from the NCBI database. An additional two genomes that have been sequenced in the laboratory at the University of Bristol but not published for CBSV “Tanza” and “Nampula” isolates were included in the phylogenetic analysis. The complete genomes were aligned using the ClustalW alignment tool in MEGA 6.0 Software and an un-rooted phylogenetic tree was generated which separated the viruses into two clades representing UCBSV and CBSV. The tree also indicates that the CBSV isolates were earlier ancestors than UCBSV isolates. The UCBSV Kikombe genome clustered with the UCBSV clade in the phylogenetic analysis confirming that the sequenced isolate was a UCBSV isolate.

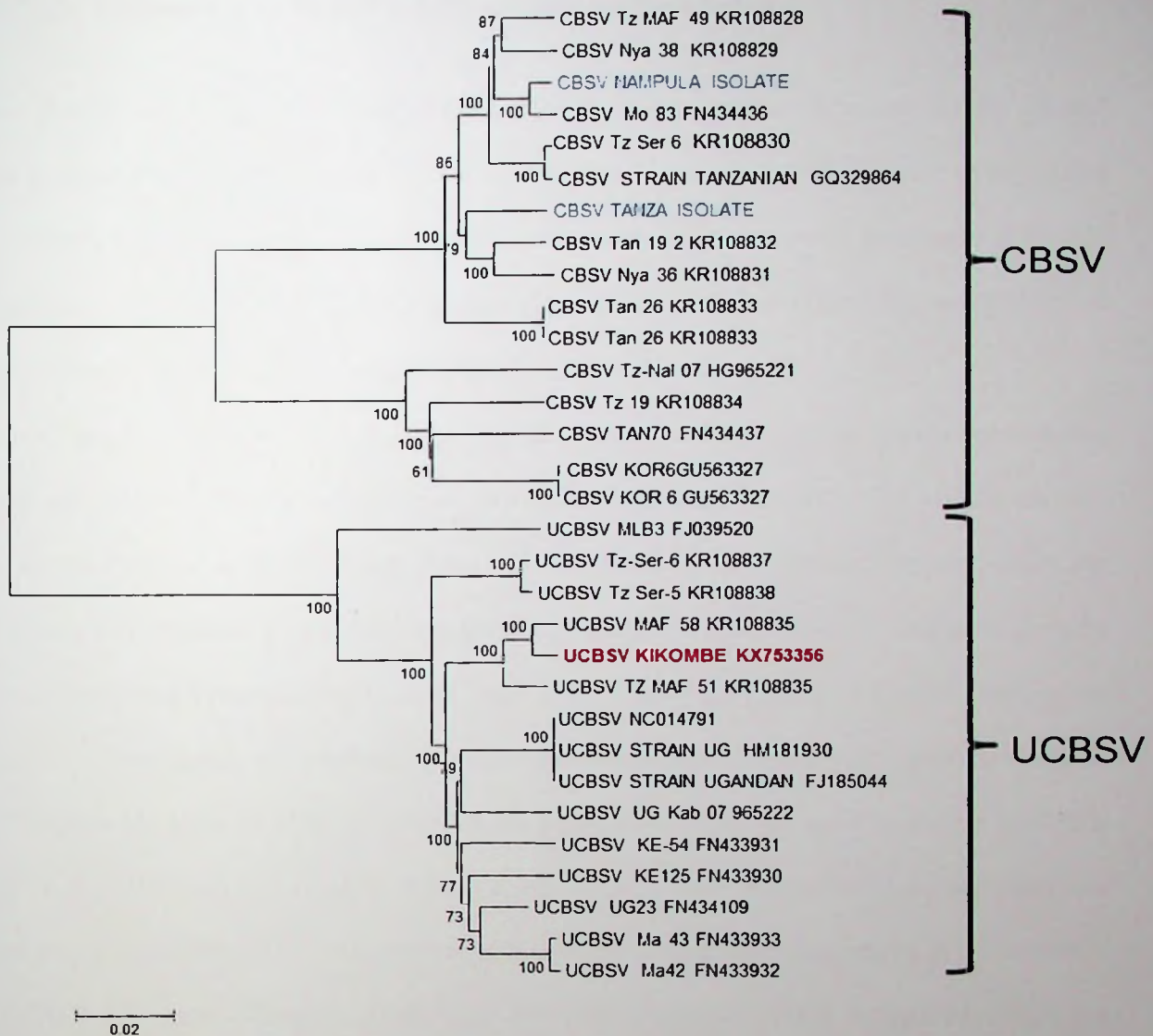


Figure 3.8: Bootstrapped Neighbour-Joining tree (1000 replicates) constructed from the alignment of 30 complete genomes of UCSV and CBSV in the NCBI database. The new sequenced UCSV Kikombe isolate (KX753356) is highlighted in red. The two CBSV isolates available at the University of Bristol are highlighted in blue

3.2.6.2 Comparison of P1 and HAM1 protein between isolates

Further phylogenetic analysis was performed for the individual protein sequences for the P1 and HAM1 proteins between different CBSV and UCBSV isolates in the NCBI database as compared to the UCBSV Kikombe (KX753356) P1 and HAM1 proteins. The reason for the choice of HAM1 protein in this study was because HAM1 is a novel protein that has been identified in CBSVs thus it was considered important to try and elucidate the role of this protein.

For the P1 proteins, the generated phylogenetic tree showed different subgroupings between the UCBSV and CBSV proteins indicative of two different clades. However, the results showed differing distributions of the bootstrap values as compared to the bootstrap support values for phylogenetic tree obtained with complete genomes although the individual proteins still clustered in two different clades representing UCBSV and CBSV (Figure 3.9) meaning that the comparison based on individual genes were reflective of the comparison based on the entire genomes (ORF). The P1 protein has been considered as one of the genes responsible for rapid evolution of CBSV (Alicai *et al.*, 2016) and according to Winter *et al.*, (2010), the variation in sequences between isolates may explain the variation in symptoms. It was observed that the sequences of P1 of CBSV and UCBSV clustered differently from other potyviral sequences which is indicative that the cassava brown streak viruses are distinct strains from other potyviruses. The P1 sequences of CBSV and UCBSV isolates were more related to P1b of *Cucumber vein yellowing virus* (CVYV) than P1 of other potyviruses. This was consistent with previous results which showed that CBSV and UCBSV isolates are more related to the viruses in the genus *Ipomovirus* than to viruses in other genera of the family *Potyviridae* and thus their placement into the genus *Ipomovirus*.

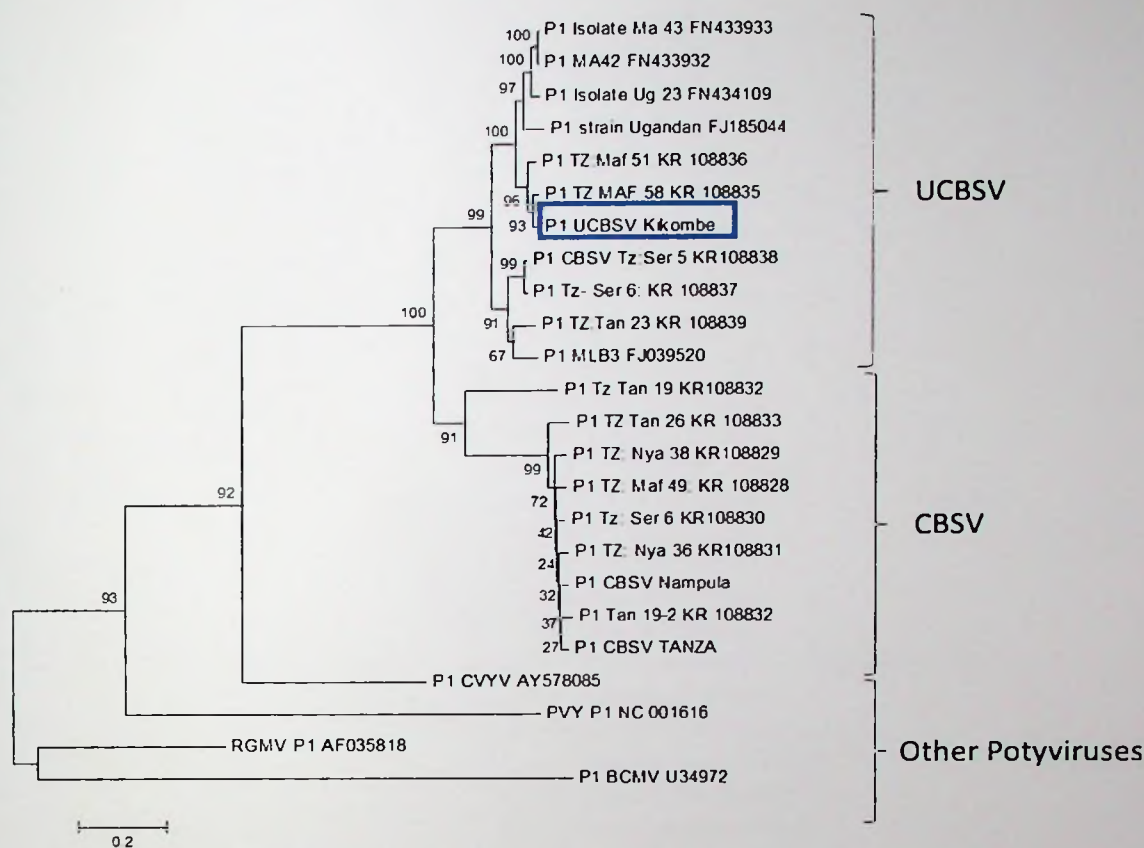


Figure 3.9: Phylogenetic tree of the P1 protein sequences of UCBSV Kikombe with other UCBSV and CBSV isolates in the NCBI database. A number of representative potyviruses were included in the analysis. The numbers at the forks are the bootstrap values which indicate the number of times out of 100 times that this grouping occurred after bootstrapping the data. The scale bar shows the number of substitutions per base.

Phylogenetic analysis was also performed to compare the HAM1 protein in the UCBSV Kikombe isolate with the HAM1 proteins of other UCBSV and CBSV isolates whose sequences have been submitted in the NCBI database. Other HAM1 containing representative organisms (Figure 3.10), and it was observed that the UCBSV and CBSV HAM1 proteins formed distinct clades different from the HAM1 of other organisms.

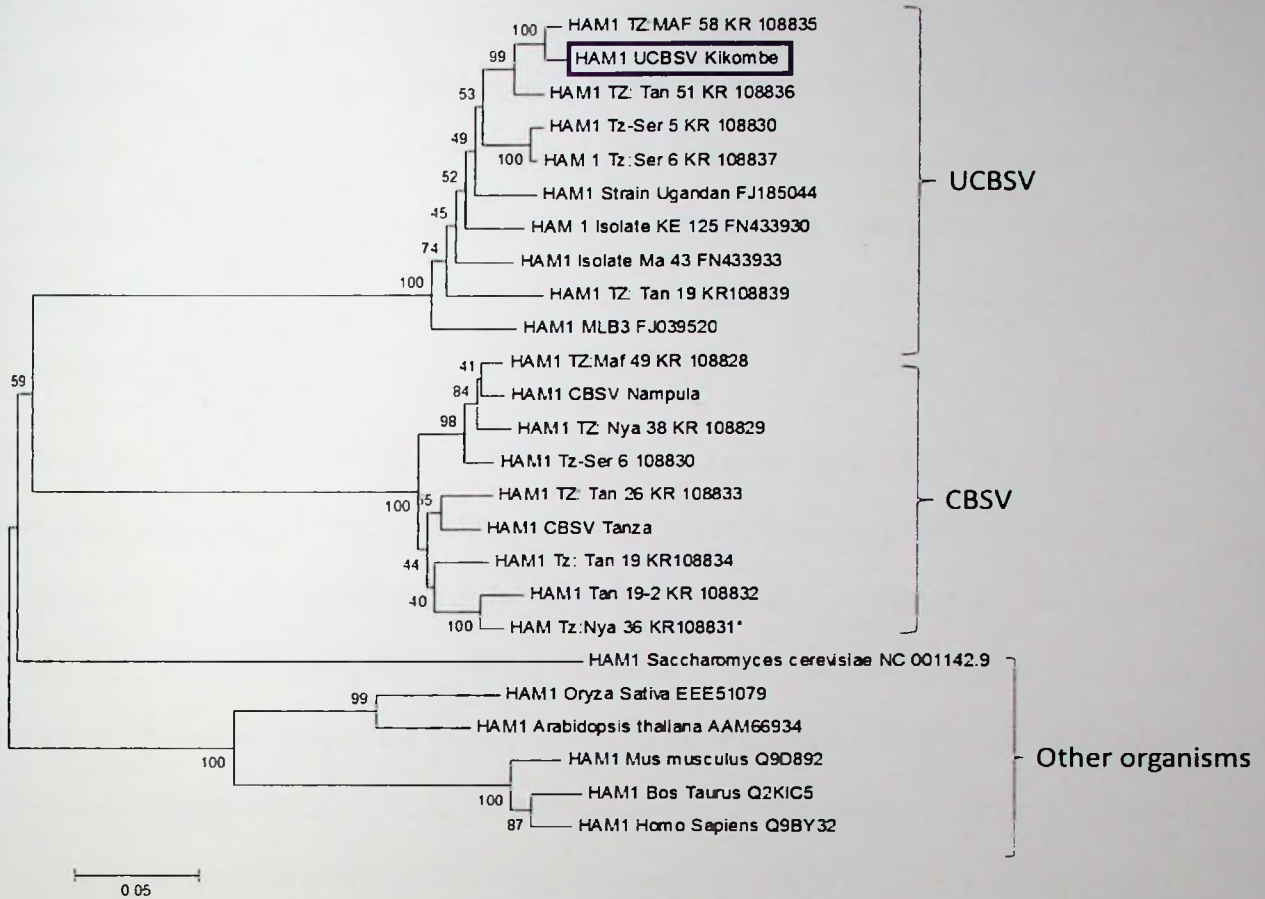


Figure 3.10: Phylogenetic analysis by Neighbour-joining methods of UCBSV Kikombe HAM1, HAM1 from other UCBSV and CBSV isolates in the NCBI database and the HAM1 in other representative organisms. The unicellular fungus (yeast) *S. cerevisiae*, the plants rice (*Oryza sativa*) and *Arabidopsis thaliana* and animals mouse (*Mus musculus*), domestic cow (*Bos taurus*) and human (*Homo sapiens*) were included in the analysis.

3.3 Discussion

In this chapter, the complete genome of UCBSV Kikombe isolate was sequenced and assembled using an overlapping DNA strategy where the genome was divided into sections and different individual sections sequenced and assembled. Furthermore, this chapter demonstrated the successful infection of the wildtype UCBSV Kikombe isolate in a number of model host plants.

Although earlier studies have shown UCBSV isolates to cause mild symptoms in herbaceous host plants (Ogwok *et al.*, 2010; Mohammed *et al.*, 2012), it was very difficult to get similar observations with the UCBSV Kikombe isolate. A comparison was done by infecting host plants with UCBSV and CBSV isolates (Tanza and Nampula isolates available in the laboratory). However, despite the fact that CBSV isolates have been reported to cause severe symptoms (Mohammed *et al.*, 2012) in host plants, the CBSV isolates Tanza and Nampula could not cause the reported symptoms. The mild mosaics observed in UCBSV infections could also not be observed in *N. benthamiana* plants however, a number of different optimisations were performed including changing the growth temperatures from 25°C to 28°C and mild mosaics were later obtained for *N. benthamiana* plants. For *N. clevelandii*, *N. glutinosa* and *N. tabacum* plants, there were no observable mosaics with UCBSV Kikombe infection. All the plants were checked for the presence of UCBSV infection by RT-PCR and sequencing and they were confirmed to be infected confirming that the wildtype UCBSV is infectious in these different hosts.

The assembled UCBSV Kikombe complete genome contained all the genes coding for the proteins that have been reported to be present in the UCBSV genomes in the NCBI database (Mbanzibwa *et al.*, 2009a; Monger *et al.*, 2010). A number of bioinformatic analyses were performed on the sequenced and assembled UCBSV genome to confirm that it was accurate,

could translate into a functional polyprotein with all the predicted cleavage sites confirmed to be present. Just like the genomes of other Ipomoviruses, *Squash vein yellowing virus* (SqVYV) and *Cucumber vein yellowing virus* (CVYV), the UCBSV Kikombe genome lacked the HC-pro protein and the P1 protein was immediately followed by the P3 protein with the predicted Y/S cleavage site (Adams *et al.*, 2005). The HAM1 protein that has been reported in the genomes of other UCBSV and CBSV isolates was also present in the UCBSV Kikombe genome (Mbanzibwa *et al.*, 2009a). Phylogenetic analysis of the P1 and HAM1 individual proteins separated the UCBSV and CBSV isolates into two distinct clades with the UCBSV Kikombe protein sequences clustering with UCBSV isolates confirming the genome origin of the isolate. It was important that the genome is accurate as an infectious clone was to be generated using pJET plasmids cloned with the different sections of the genome as a template. Any errors, frame-shift mutations, insertions and deletions in the genome would ultimately affect the infectivity of the constructed infectious clone. The UCBSV Kikombe genome reported in this chapter was submitted NCBI database (GenBank accession number KX753356), increasing the number of complete genomes of UCBSV isolates. Complete genomes allow for easy characterisation and diagnosis of viruses in addition to understanding the evolutionary relationships of viruses (Ndunguru *et al.*, 2015).

Now that the genome was confirmed to be correct and accurate, the next chapter presents the construction of an infectious clone of UCBSV isolate using homologous recombination in yeast. The UCBSV Kikombe genome will be cloned inside a pYES2.1 vector to generate an infectious clone of this virus.

Chapter 4: Construction of an *in-vitro* and *in-vivo* infectious clone of UCBSV Kikombe isolate

4.1 Introduction

Infectious full-length clones of plant RNA viruses are valuable tools for studying a number of aspects of the molecular biology of viruses (Verchot-Lubicz *et al.*, 2007; Revers & García, 2015). Understanding a virus and all the different strains, and variants is very important in breeding programs. This is because the generated resistant plant material is tested for the response to the different variants of the virus (Garcia-Arenal *et al.*, 2001; Domingo & Holland, 1997). Infectious clones provide a uniform/homogenous source of the virus as opposed to the mixed populations in the wildtype viruses. Furthermore, infectious clones allow for genetic studies to be undertaken on RNA viruses (Desbiez *et al.*, 2012; López-Moya & García, 2000; Garcia-Arenal *et al.*, 2001). With infectious clones there is the conversion of the RNA viruses into cDNA which is maintained within a plasmid backbone vector for ease of propagation and manipulation (Desbiez *et al.*, 2012; Boyer & Haenni, 1994; Revers & García, 2015). In the cDNA form of an infectious clone, a number of genetic manipulations can be undertaken which include mutations, site directed mutagenesis and deletions among others which allow studying the functions of the different proteins present in the viral genome. To generate an infectious clone of an RNA virus, the RNA is initially reverse transcribed into cDNA using a primer that binds to the 3' end of the genome to generate single- stranded DNA, for viruses that contain a poly (A) tail sequence at the 3' end, an oligo d (T) primer can be used in the synthesis of the first strand cDNA. The synthesized single-stranded DNA is then converted to double-stranded cDNA by using primers that encompass the whole genome to the 5' end. A suitable promoter can be inserted upstream of the cDNA to allow for transcription of the cDNA. There are a number of different promoters that have been used in infectious clone construction among

which the bacteriophage SP6 (Butler & Chamberlin, 1982), T7, or *Cauliflower mosaic virus* (CaMV) 35S are the most commonly used promoters. An infectious clone may either be transcribed *in-vitro* to generate RNA transcripts as is the case for infectious clones generated with SP6 or T7 promoters (Meshi *et al.*, 1986., Vos *et al.*, 1988) or the infectious clone may be introduced into the host plant by particle bombardment (Gal *et al.*, 1997) or agroinfiltration for the *in-vivo* transcription and translation for infectious clones containing the CaMV 35S promoter (Verchot-Lubicz *et al.*, 2007). For *in-vitro* transcribed cDNA infectious clones, the plasmid with the cloned complete genome of the RNA virus can act as a template in an *in-vitro* transcription reaction to generate RNA transcripts that can cause an infection similar to the wildtype RNA virus. Inoculation of susceptible host plants with cloned material of many plant virus genomes including numerous RNA viruses, can result in the establishment of viral infections (Revers & García, 2015).

4.1.1 Applications of infectious clones

4.1.1.1 Viral protein functional studies

Infectious clones provide information on the study of genetic expression and replication of RNA viruses through enabling their manipulation through mutagenesis, deletions, and insertions and by complementation experiments. Mutations can be introduced into infectious clones to allow for functional analysis of the different proteins in the viral genomes, for instance site directed mutagenesis in the conserved motif of the FRNK sequence of the HC-Pro gene in the infectious clone of *Zucchini yellow mosaic virus* (ZYMV) allowed for the effects of the mutations on the symptom expression of the infectious clone to be studied (Gal-On, 2000). Chimeric forms of infectious clones of aggressive and wildtype isolates can allow the study and identification of pathogenic determinants in viruses (Duff-Farrier *et al.*, 2015; Fernandez-

Delmond *et al.*, 2004). Because infectious clones provide a homogenous source of viral infection, they can be used for testing disease resistant germplasm in breeding programs. Infectious clones of CBSVs are very vital in cassava breeding programs as they provide a uniform homogenous source of the virus for testing breeding lines against resistance to CBSD.

4.1.1.2 Study viral cell –to- cell movement

Infectious clones can be fused to fluorescent proteins like the green fluorescent protein GFP of *Aequorea victoria* and the movement of the virus within the host studied by following the movement of the fluorescent reporter protein by microscopy. For instance, the GFP gene has been inserted between the Nib and wildtype CP and mutated CP encoding regions of the *Tobacco mosaic virus* (TMV) infectious clone and the virus movement studied by comparing the levels of fluorescence between the wildtype and mutated CP infectious clones (Bendahmane *et al.*, 2002).

4.1.2 Methods of generating infectious clones

4.1.2.1 *E. coli* cloning system

E. coli has been used to generate infectious clones of stable potyviruses (Boyer & Haenni, 1994). However, for some potyviruses, it has been found that some of the potyviral proteins are toxic in *E. coli* and thus it is not possible to propagate the infectious clone plasmids through the *E. coli* unless introns are inserted to intercept the undesired translation of toxic viral proteins (Johansen 1996; Fakhfakh *et al.*, 1996). The CI gene which encodes the viral helicase has been reported as the main toxic protein challenging propagation of potyviral infectious clones through *E. coli* (Lain *et al.*, 1990).

4.1.2.2 *In-vitro* ligation of PCR fragments

Because of the problems associated with toxicity of some of the viral proteins in the *E. coli* propagation system which is one of the main systems used in the generation of infectious clones, alternatives that bypass the need for using these systems are needed. These include performing RT-PCR to generate different fragments of the RNA genome which are later ligated *in-vitro* using appropriate restriction enzyme sites (Jakab *et al.*, 1997). Alternatively, full length first-strand cDNA can be synthesized using optimized PCR conditions that can amplify long DNA fragments (Cheng *et al.*, 1994). One of the challenges associated with infectious clones generated in this way is that they cannot be utilized for further genetic manipulation for functional studies.

4.1.2.3 Insertion of Introns

Another way of generating stable infectious clones is by insertion of intragenic sequences, or introns in the viral genome. Introns occur naturally in the genomes of most organisms and they are spliced out from the mRNA during post-transcriptional modifications. Once the introns are spliced out, the exons are joined together and the protein is expressed without the introns. Interrupting the viral open reading frame by introduction of introns into the proteins normally known to be toxic in *E. coli* has been applied in the generation of infectious clones previously (Johansen, 1996). Expression of toxic proteins often occurs from cryptic promoters present in the viral genome. The introns interfere with the open reading frame of the viral polyprotein by disrupting the reading frame and creating stop codons thus preventing the expression of toxic proteins in *E. coli*. In this way, the infectious clone can be successfully propagated. Once the infectious clone is recovered from the *E. coli* cells, it can be transformed into *Agrobacterium tumefaciens* and can be introduced into plants by use of an *Agrobacterium* vector via agro-

infiltration (Johansen, 1996). Once inside the plant, the infectious clone is transcribed and the introns are spliced out during post-transcriptional modification of the mRNA. Because *in-vivo* clones need splicing, they can only work *in planta* and not *in-vitro*. In this way, the viral polyprotein is expressed after intron-splicing in a way that is similar to the wildtype infection.

4.1.3 *In-vitro* and *in-vivo* infectious clones

In general, the construction of cDNA infectious clones consists of reverse transcribing the viral RNA into a single-stranded DNA using a primer hybridizing specifically to the 3' end of the viral genome. After elimination of the viral RNA, the single-stranded DNA is converted into its double-stranded form by initiating DNA synthesis with a second primer encompassing the sequence corresponding to the 5' end of the of the viral RNA (Boyer & Haenni, 1994). Typically, the sequence for an RNA polymerase promoter like the SP6 or T7 is fused to the viral sequence at the 5' end and is included in the second primer. In this form, infectious transcripts can be generated *in-vitro*. Alternatively, placing the sequence under the control of a host promoter such as CaMV 35S, allows for *in-vivo* transcription. A serious limiting factor for some viral genomes appears to be the synthesis of the full-length first cDNA strand probably because the polymerization step is hampered by secondary structures on the viral template RNA. It is noteworthy that even for a given virus, some research groups but not others have succeeded in obtaining full-length first cDNA strand (Meshi *et al.*, 1986; Dawson *et al.*, 1986). The advantages of an *in-vivo* infectious clone as compared to an *in-vitro* infectious clone include the fact that with an *in-vivo* clone, infectivity is less dependent on RNA degradation since it presumably occurs only within cells where the RNAs are synthesized and the replication process can overcome detrimental effects resulting from degradation. *In-vivo* infectious clones can also accommodate inclusion of introns for the stabilization of the clones.

In addition, *in-vivo* clones do not need the costly reagents such as Cap analogues and RNA polymerases (Boyer & Haenni, 1994).

4.1.4 Aims of this Chapter

The aim of this chapter is to generate infectious clones of UCBSV Kikombe isolate. In the previous chapter, the complete sequence of the UCBSV isolate was sequenced and assembled. The sequence was needed to allow for the construction of the full-length cDNA clone of the UCBSV isolate. It was very important that the assembled genome was checked for accuracy as any errors in the genome could have affected the infectivity of the constructed infectious clone. The constructed UCBSV *in-vitro* infectious clone was tested in tobacco and cassava. The strategies employed in constructing the UCBSV *in-vitro* infectious clone are described in this chapter and the challenges encountered in constructing the *in-vivo* UCBSV infectious clone are also discussed.

4.2 Results

4.2.1 Designing specific primers based on UCBSV genome to construct *in-vitro* infectious clone

Once the sequenced and assembled UCBSV genome data had been confirmed with no insertions, deletions, or frame-shifts in the sequence (see Chapter 3), specific primers were designed to amplify the different sections of UCBSV Kikombe isolate to use in constructing an infectious clone of the virus by homologous recombination in yeast (Table 4.1). The primers were designed to reverse the orientation of the full-length infectious clones avoiding any “leaky” read through that may occur via the pGAL promoter present in the pYE2.1 vector backbone. Sequence inversion thus eliminated any likelihood of expression and toxicity that could otherwise prove lethal (Johnston *et al.*, 1994). In addition, a *Kpn*I restriction site was incorporated at the 3' end of the full-length infectious clone to allow for plasmid linearization and *in-vitro* RNA production. The *Kpn*I restriction site is absent in the UCBSV Kikombe genome and thus the enzyme can be appropriately used to linearize the infectious clone plasmid without digesting into the UCBSV genome. In order to allow for *in-vitro* transcription, a bacteriophage SP6 promoter sequence was added at the 5' end primer upstream of the 5'UTR.

Table 4.1: Specific primers used to construct the UCBSV infectious clone. The primers are based on the UCBSV Kikombe genome (KX753356). The sequences in italics represent the pYES2.1 vector backbone overlapping sequence to allow for homologous recombination in yeast

Primer name	Primer sequence	Position in genome UCBSV genome (KX753356)
Sec1A p1-Vec F	<i>GAGCGGCCGCCAGTGTGATGGATATCTGCAAAAAATAAACATGACATAAGAATAC</i>	1 to 25 with pYES2.1 overhan
Vec SP62	<i>GAGCGGCCGCCAGTGTGATGGATATCTGCA</i> <i>AATTAGATCTATTTAGGTGACACTATG AAAAAATAAACATGACATAAGAATAC</i>	Primer for SP6 promoter (bold)
Sec1B C1 Kikombe R2	<i>CATATGTCATGATTGTAATTG</i>	2665 to 2685
5'1KB FWD S	<i>ATTGAGCCAACTACGAGTGGACCTATG</i>	2382 to 2408
5'1KB REV S	<i>ACTGACCCTATATCTACATCATCAAC</i>	3426 to 3451
5'867bp fwd S	<i>GATTTTGGTGAACGCATAAGTCCAAA</i>	3195 to 3220
5'867bp Rev S	<i>GAGTTACTCTTGAATCAATTAATGGAG</i>	4035 to 4062
5'1442F	<i>AAAGAATACATTAGAATCGA</i>	3651 to 3670
5'1442 R Vec	<i>CTCACTATAGGGAATATTAGCTTGGTACCAATTGGTACCACGCG</i> <i>GGATTAGTTGGCTCTGGTTTTGCAACCTGAACCTC</i>	4845 to 4879 with pYES2.1 overhang
5'1442 R S	<i>GGATTAGTTGGCTCTGGTTTTGCAACCTGAACCTC</i>	4845 to 4879
MIDFF1KB Vec	<i>GAGCGGCCGCCAGTGTGATGGATATCTGCGCTCCAATGGGATTCCAGAAGAAGA</i>	4758 to 4783 with pYES2.1 overhang
MIDRR1KB S	<i>CACTAACCCAACCTGAACTGGCTC</i>	5775 to 5798
3'MID 1.5KB S	<i>GGACGTTTAATCCTGATCTTGTTGAGTGG</i>	5461 to 5489
3'MID' 1.5KB S	<i>TCCAGCAATGGATACTCACAA</i>	6930 to 6952
3'1.4KB F S	<i>GGGCTGCCGAAAAAGGTTGTG</i>	6916 to 6936
3'1.4KB R	<i>CCTGTATATTAACCTCATATGCITTAGC</i>	8304 to 8331
922bp F S	<i>GGGCAATTGAGGTTGCCAAAGC</i>	8290 to 8311
3' Final	<i>CTCACTATAGGGAATATTAGCTTGGTACCAATTGGTACCACGCG</i> <i>TTTTTTTTTTTTTTTTTTTGGAC</i>	To polyA tail with pYES2.1 overhang
5'1442 R S	<i>GGATTAGTTGGCTCTGGTTTTGCAACCTGAACCTC</i>	4845 to 4879
MIDFF1KB S	<i>GCTCCAATGGGATTTCAGAAGAAGA</i>	4758 to 4783
3' Final S	<i>CTCACTATAGGGAATATTAGCTTGGTACCG TTTTTTTTTTTTTTTTTTTTGGAC</i>	To PolyA tail with pYES2.1 overhang
3'Final C	<i>CTCACTATAGGGAATATTAGCTTGGTACCGAATTGGTACCACGCG</i> <i>TTTTTTTTTTTTTTTTTTTGGAC</i>	To polyA tail with pYES2.1 overhang and <i>KpnI</i> Restriction site (bold)

4.2.2 PCR to amplify different sections of UCBSV

The specific primers listed in (Table 4.1) were used to amplify the different overlapping sections of the UCSBV Kikombe genome. PCR primers were designed in such a way that sequences for different sections to be recombined were overlapping by at least 30 nucleotides. Vector sequences for the pYES2.1 backbone were also introduced on viral sequence primers to allow for homologous recombination between viral sequence and the pYES2.1 vector. The viral templates used in the PCR were the originally cloned pJET vectors containing the

different sections of the UCBSV genome that had been used in the genome sequencing project. The expected sizes of the PCR products (Figure 4.1) were, section 1 (2685bp), section 2 (1069bp), section 3 (867bp), section 4 (1125bp), section 5 (1041bp), section 6 (1500bp) section 7 (1400bp) and section 8 (922bp). All PCR reactions were performed using Phusion Taq polymerase enzyme to avoid amplification sequence errors.

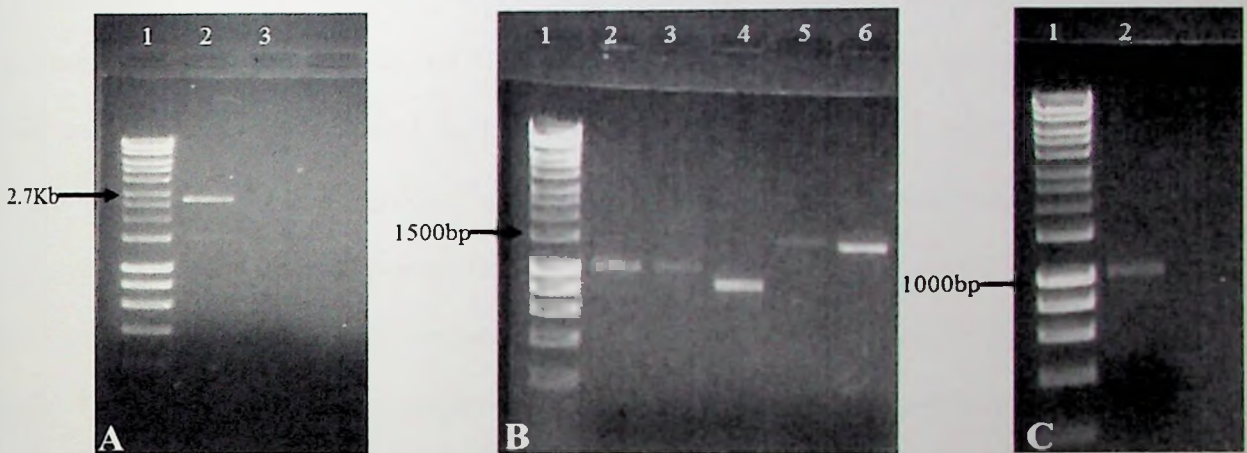


Figure 4.1: PCR amplification of different sections of UCBSV Kikombe sections for yeast recombination; A: Lane 1 5 μ l of HyperLadder 1, Lane 2 section 1 (2700bp). B: Lane 1 contains 5 μ l HyperLadder 1, Lane 2 section 2 (1069bp) Lane 3 section 4 (1125bp), Lane 4 section 3 (867bp), Lane 5 section 6 (1500bp), Lane 6 section 7 (1400bp) C: Lane 1 5 μ l of HyperLadder 1, Lane 2 section 5 (1041bp).

4.2.3 Yeast homologous recombination

Homologous recombination in yeast was used to assemble the different overlapping UCBSV genome sections. The pYES2.1 backbone vector was linearised by digestion with *Eco*R1 and *Sac*I restriction enzymes. In order to check for the stability of the UCBSV infectious clone in *E. coli* cells, the UCBSV genome was initially divided into two halves of approximately 5 kb designated the 5' end and 3' end plasmids and constructed separately. Once stability was confirmed, the two halves were recombined to generate the full-length UCBSV infectious clone. The protocol used for the yeast recombination is as explained in (section 2.9).

4.2.4 Construction of pYES2.1 with 5' end UCBSV genome

The 5' end of the genome consisted of four sections designated section 1 to 4 (comprising base positions 1-4879 of the UCBSV genome (Figure 4.2A) cloned in backbone pYES2.1 with a size of 10.7 kb. PCR was performed to amplify the different UCBSV sections (Figure 4.1) using the pJET plasmids that contained the cloned UCBSV sections that had been constructed in the genome sequencing project as template. These different sections were recombined in pYES2.1 plasmid by homologous recombination in yeast. As a positive control, yeast cells were transformed with undigested pYES2.1 backbone vector. Negative controls consisted of yeast transformation with *EcoR1/Sac1* digested pYES2.1 vector and water. Yeast cells transformed with undigested pYES2.1 generated about 520 colonies confirming the viability of the yeast cells for transformation. The yeast cells transformed with *EcoR1/Sac1* digested pYES2.1 generated about 17 colonies and there were no colonies on water transformed control plates. The experimental plates transformed with different UCBSV sections (section 1-4) generated about 340 colonies confirming successful recombination. Ten yeast colonies were individually selected and grown in SM media. Yeast plasmid was extracted and then rescued into *E. coli* cells. Plasmid was then extracted from two *E. coli* colonies and analysed by restriction analysis and sequencing for their integrity. For the 5' end of the genome, the pYES2.1 vector containing sections 1-4 was digested with *Xho1* restriction enzyme and the expected product sizes (6617bp, 2394bp, 1131bp, and 589bp) as predicted by the *in-silico* restriction analysis with Clone Manager Software Suite 7 software were obtained (Figure 4.2B). The constructed plasmid was sequenced in the region containing the UCBSV genome using specific primers (Table 4.1) to further confirm its integrity.

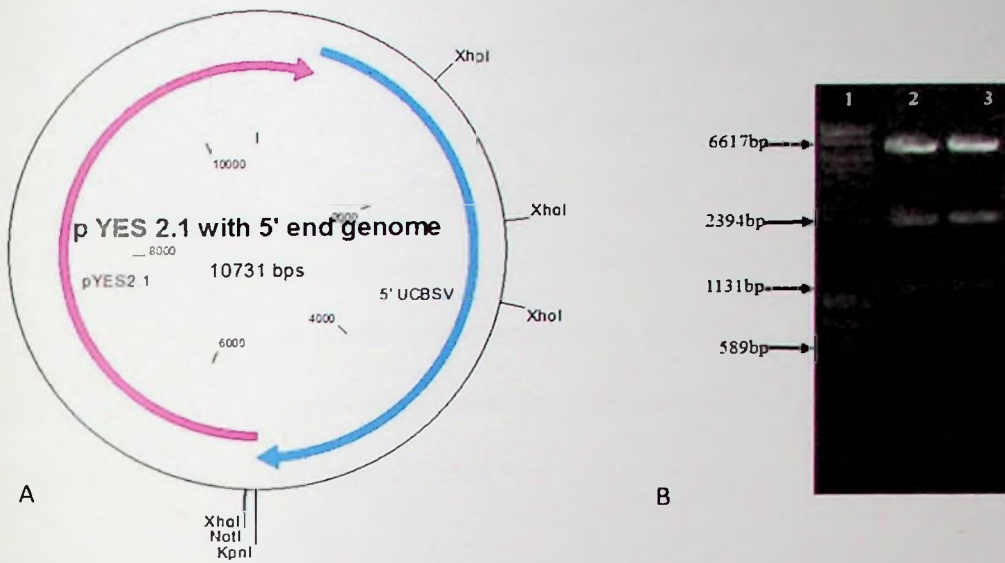


Figure 4.2: A: Constructed plasmid of pYES2.1 with 5' end UCBSV genome. The blue arrow represents 5' end (sections 1, 2, 3, and 4) of the UCBSV genome and the pink arrow represents pYES2.1 backbone vector. The *XhoI* restriction sites present in the constructed plasmid are also shown on the map B: *XhoI* restriction, Lane 1 HyperLadder 1, Lane 2 and 3 *XhoI* restriction of two different *E. coli* plasmid extractions on the constructed pYES2.1 containing section 1,2,3 and 4 (5' end half of the UCBSV genome).

4.2.5 Construction of pYES2.1 with 3' end UCBSV genome

The 3' end consisted of four sections designated section 5 to 8 comprising base positions 4758 to the polyA tail (26 nucleotides) of the UCBSV genome cloned in backbone pYES2.1 with a size of 10 kb (Figure 4.3A). These different sections were also recombined in pYES2.1 vector by homologous recombination in yeast. As a positive control, yeast cells were transformed with undigested pYES2.1 backbone vector. Negative controls consisted of yeast transformation with *EcoR1/Sac1* digested pYES2.1 vector and water. Yeast cells transformed with undigested pYES2.1 generated about 500 colonies confirming the viability of the yeast cells for transformation. The yeast cells transformed with *EcoR1/Sac1* digested pYES2.1 generated about ten colonies and there were no colonies on water transformed control plates. The experimental plates transformed with different UCBSV sections (section 5-8) generated about

200 colonies confirming successful recombination. Ten yeast colonies were individually selected and grown in SM media. Yeast plasmid was extracted and then rescued into *E. coli* cells. Plasmid was then extracted from *E. coli* cells and analysed by restriction analysis and sequencing for their integrity. The different PCR sections (section 5 to 8) were recombined in pYES2.1 backbone vector by recombination in yeast and successful recombination was confirmed by *Spe*I enzyme restriction and the expected sizes of 5326bp, 5530bp, and 1213bp were obtained (Figure 4.3B). The constructed plasmid was sequenced in the region containing the UCBSV genome using specific primers (Table 4.1).

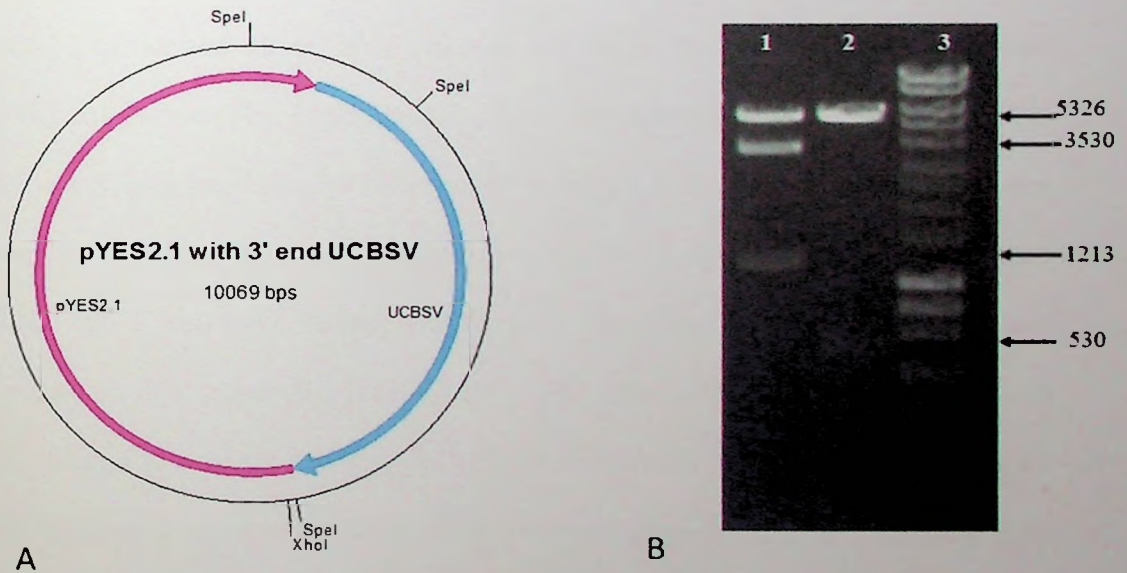


Figure 4.3: **A:** Constructed plasmid of pYES2.1 with 3' end UCBSV genome. The blue arrow represents the 3' end (sections 5, 6, 7 and 8 of the UCBSV genome). The pink arrow represents the pYES2.1 backbone vector. The *Spe*I restriction sites present in the plasmid are also shown on the map. **B:** Lane 1 contains 10 µl of *Spe*I restriction reaction on the constructed pYES2.1 containing section 5,6,7 and 8 (3' end half of the UCBSV genome). Lane 2 *Spe*I restriction on the backbone pYES 2.1 vector backbone.

4.2.6 Construction of UCBSV infectious clone plasmid

Once the sequences in the 5' end and 3' end of the genome had been confirmed to be stable with no sequence insertions or deletions as determined by restriction analysis and Sanger sequencing, the 5' end and 3' end independent half genome sections were recombined to obtain the full-length complete genome cloned inside a pYES2.1 vector (Figure 4.4). The different 5 kb PCR fragments encompassing the complete genome were recombined in pYES2.1 backbone vector and the plasmid extracted from yeast and rescued into *E. coli* cells. As a positive control, yeast cells were transformed with undigested pYES2.1 backbone vector. Negative controls consisted of yeast transformation with *EcoR1/Sac1* digested pYES2.1 vector and water. Yeast cells transformed with undigested pYES2.1 generated about 430 colonies confirming the viability of the yeast cells for transformation. The yeast cells transformed with *EcoR1/Sac1* digested pYES2.1 generated about nine colonies and there were no colonies on water transformed control plates. The experimental plates transformed with different UCBSV sections (section 1-8) generated about 320 colonies confirming successful recombination. Ten yeast colonies were individually selected and grown in SM media. Yeast plasmid was extracted and then rescued into *E. coli* cells. Plasmid was then extracted from *E. coli* cells and analysed by restriction analysis and sequencing for their integrity. Further analyses were performed on the recovered plasmid to confirm the integrity of the infectious clone plasmid. The summary of the steps undertaken in constructing the UCBSV infectious clone are shown in the schematic diagram in (Figure 4.5).

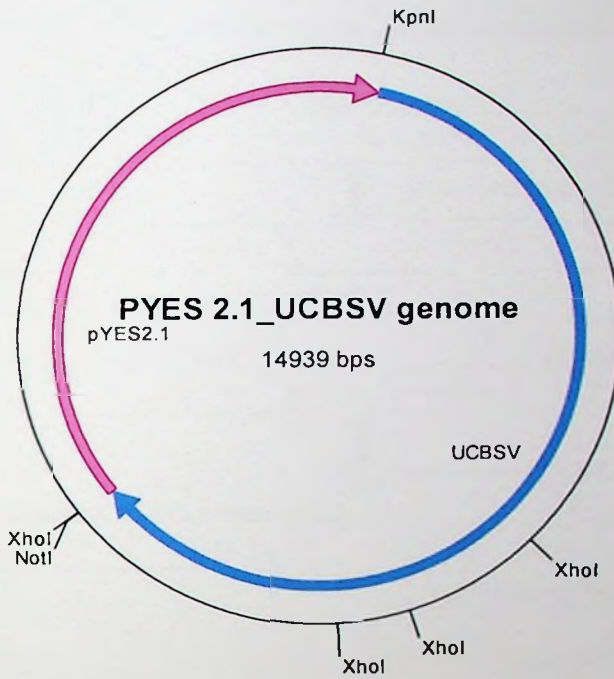


Figure 4.4: Constructed plasmid of pYES2.1 containing the complete genome of UCBSV (Infectious clone plasmid): The blue arrow represents the complete UCBSV genome from base position number 1 to the end of the poly (A) tail (9096 nucleotides). The pink arrow represents the pYES2.1 backbone vector. The recombinant plasmid is the infectious clone plasmid of UCBSV cloned inside pYES2.1. The *XhoI* restriction sites present in the plasmid are shown on the map. The *KpnI* and *NotI* restriction sites are also shown.

UCBSV Kikombe infectious clone construction

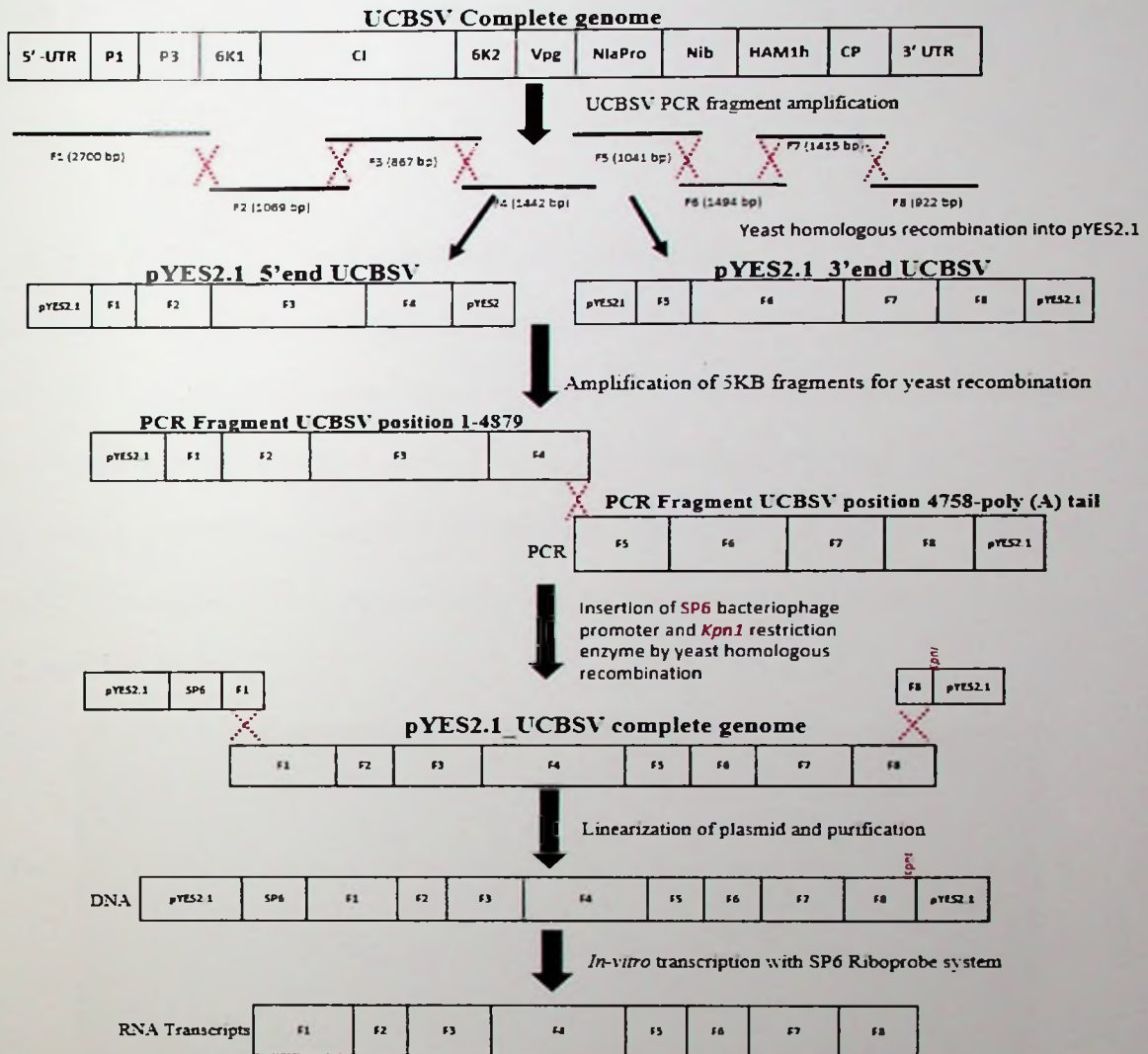


Figure 4.5: Summary of schematic flow of steps taken to construct UCBSV *in-vitro* infectious clone. F1 to F8 refer to the eight sections that were homologously recombined in yeast. The complete UCBSV genome was initially divided into two halves which were separately recombined in pYES2.1 vector (pYES2.1 5' end UCBSV and pYES2.1 3' end UCBSV). Once stability was confirmed, the separate halves were recombined to generate pYES2.1 UCBSV genome.

4.2.7 Checking the integrity of the constructed UCBSV infectious clone

Checks on the integrity of the clone were performed by PCR, restriction analysis and sequencing. It was important that all the sequences in the infectious clone were in-frame without any insertions or deletion as these might affect the infectivity of the constructed infectious clone.

4.2.7.1 PCR to check for the individual sections (Section 1-8)

PCR amplification was performed using the constructed infectious clone plasmid as template in the reactions. From the PCR results, it was shown that all the eight UCBSV sections were present in the recombined pYES2.1 vector (Figure 4.6). The expected sizes were section 1 (2700bp), section 2 (1069bp), section 3 (867bp), section 4 (1442bp), section 5 (1000bp), section 6 (1500bp), section 7 (1400bp) and section 8 (922bp) were all able to be amplified from the infectious clone plasmid.



Figure 4.6: Checking the presence of all the recombined sections in the UCBSV infectious clone: Lane 1 and 18 contain 5 μ l of HyperLadder 1. Lane 2 is section 1 (2700bp) amplified from infectious clone plasmid and Lane 3 is positive control for section 1. Lane 4 is section 2 (1069bp) amplified from infectious clone plasmid and Lane 5 is positive control for section 2. Lane 6 is section 3 (867bp) amplified from infectious clone plasmid and Lane 7 is positive control for section 3, Lane 8 is section 4 (1442bp) amplified from infectious clone plasmid and Lane 9 is positive control for section 4. Lane 10 is section 5 (1044bp) amplified from infectious clone plasmid and Lane 11 is positive control for section 5, Lane 12 is section 6 (1500bp) amplified from infectious clone plasmid and Lane 13 is positive control for section 6, Lane 14 is section 7 (1400bp) amplified from infectious clone plasmid and Lane 15 is positive control for section 7, Lane 16 section 8 (922bp) and lane 17 is positive control for section 8.

4.2.7.2 Restriction analysis to confirm integrity

The recombinant extracted from *E. coli* containing the complete UCBSV genome cloned in the backbone pYES2.1 vector was approximately 14.9 kb (Figure 4.7). This was subjected to restriction analysis using *Kpn*1/*Not*1 and *Xho*1 enzymes. As predicted, the *Kpn* 1/*Not* 1 restriction digestion released a 10 kb fragment (full-length UCBSV genome section) and a 6 kb fragment (pYES2.1 vector backbone) (Figure 4.7A). The *Xho*1 restriction digestion on the recombinant plasmid gave the expected sizes of 10.8Kb, 2345bp, 1131bp and 589bp as is predicted by the *in-silico* constructed plasmid (Figure 4.7B).

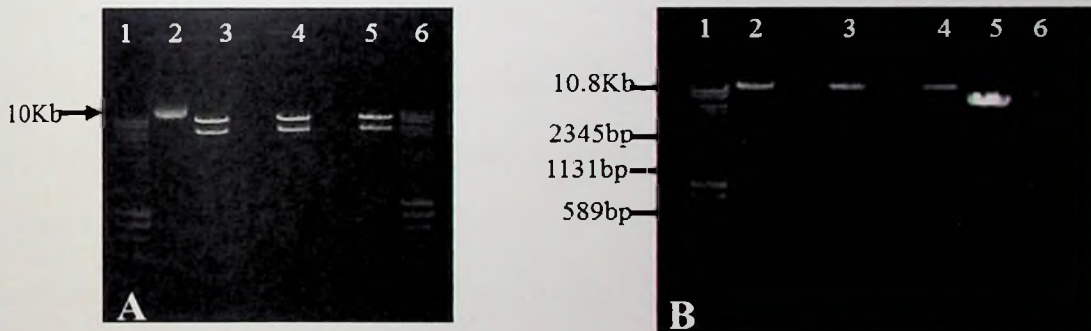


Figure 4.7: Restriction analysis to check the integrity of the different clones of constructed UCBSV infectious clone. **A:** *Kpn*1/*Not*1 restriction, Lane 1 and 6 contain 5 μ l of HyperLadder 1. Lane 2 contains 5 μ l of undigested infectious clone plasmid. Lanes 3, 4 and 5 contain three different clones showing the 10 kb UCBSV genome and the 6 kb pYES2.1 backbone vector released from the *Kpn*1/*Not*1 restriction digestion on the infectious clone plasmid. **B.** *Xho*1 restriction. Lane 1 and 6 contain 5 μ l of HyperLadder 1, Lanes 2, 3 and 4 contain three different clones of *Xho*1 restriction on different clones of UCBSV infectious clone showing the expected sizes. Lane 5 contains pYES2.1 backbone plasmid.

4.2.7.3 Sequencing of the UCBSV infectious clone genome

From the three different clones analysed by restriction analysis above (Figure 4.7A and B), one clone was chosen for further analysis and sequencing and this clone formed the basis of all subsequent work with the UCBSV Kikombe infectious clone. The recombinant infectious clone plasmid (one clone) containing the full-length UCBSV genome cloned inside the pYES2.1 vector was sequenced to confirm that there were no insertions or deletions in the sequence as these would lead to frame shifts. Sequencing was performed using the specific primers that had been used to generate the sections that were contained in the infectious clone. Sequencing was done in either the forward or reverse directions for each section (Figure 4.8). In the analysis of the sequence data, at position 5731, the coverage of the Sanger sequencing data for the infectious clone seemed to have a “T” deletion causing a frameshift in the polyprotein. On further inspection of the contig covering this region, the reads with the deletion were of low quality and thus *in-silico* analysis using Clone Manager Software Suite 7 was performed to confirm by restriction analysis if the “T” nucleotide was absent or present. The *Xho*I restriction enzyme pattern observed with a missing “T” nucleotide was different from the *Xho*I restriction pattern observed with the “T” nucleotide present at position 5731. Somewhat fortuitously, if there was indeed a T missing at this position, it would generate a target site for *Xho*I, so presence /absence of this *Xho*I site could be used to help confirm the sequence at this position. The observed restriction pattern obtained *in-vitro* confirmed that the “T” nucleotide was present at position 5731 in the genome. The infectious clone genome was then therefore correctly assembled (Appendix 2). The assembled infectious clone genome was translated and the open reading frame contained the expected 2902 amino acids. The assembled genome of the UCBSV Kikombe infectious clone has been submitted to the NCBI GenBank under accession number KX753357.



Figure 4.8: Schematic view of the UCBSV infectious clone genome assembly in Sequencher. Each region of the pYES2.1 containing the UCBSV genome was sequenced 4-5 times and the overall consensus sequence assembled. The assembled infectious clone genome has been submitted to NCBI under GenBank accession KX753357.

4.2.8 Comparison of the genome of the assembled infectious clone and wildtype UCBSV genome

The assembled genome of the constructed infectious clone (UCBSV I.C) was aligned with the genome of the original UCBSV wildtype genome (UCBSV WT) using Clone Manager Software Version 5.0 (Appendix 3). In the alignment, it was observed that 10 single nucleotide polymorphisms (SNPs) were recognized in the sequences between sequenced UCBSV wildtype genome and assembled UCBSV infectious clone genome. A comparison with other UCBSV isolates in the NCBI database confirmed that most of these polymorphisms occur naturally in the population (Table 4.2). The overall percentage similarity between all the bases in the infectious clone and wildtype UCBSV genome was 99.9%. The amino acid changes due to the SNPs were compared between the UCBSV wildtype and the assembled UCBSV infectious clone (Table 4.3). Most of the amino acid substitutions were with amino acids with similar functional groups (Appendix 4) meaning that the functional properties at these positions were probably not affected. It was also observed that most of the SNPs occurred in the CI protein of the UCBSV genome. There is also a possibility that the observed SNP's between the wildtype UCBSV genome and the assembled infectious clone genome were due to PCR errors as 8/10 of the SNP's were non-synonymous.

Table 4.2: Single nucleotide polymorphisms (SNP's): The positions of single nucleotide polymorphisms observed in the genomes of the infectious clone (UCBSV I.C) (GenBank accession KX753357) and the wildtype (UCBSV WT) (GenBank accession KX753356) genome in comparison with selected UCBSV isolates in NCBI database listed by accession numbers.

Isolate Accession	Base position of SNP in UCBSV genomes									
	1155	1156	1442	2405	2510	2672	3255	4040	4049	4076
KX753357	A	G	C	C	G	T	A	G	C	C
KX753356	C	A	T	T	A	A	G	A	T	T
FN433931	A	G	C	T	A	T	A	A	C	T
FN433930	A	G	C	C	A	T	A	A	C	T
FN433932	A	G	T	C	A	T	A	A	C	T
FN433933	A	G	C	C	A	T	A	G	C	T
FJ039520	G	A	T	T	A	A	A	A	C	T
FN434109	A	G	C	C	A	T	A	A	C	T

Table 4.3: Amino acid changes at the positions with single nucleotide polymorphisms (SNPs) in the wildtype UCBSV (GenBank accession KX753356) genome and constructed UCBSV infectious clone (GenBank accession KX753357).

Base Position	Protein	UCBSV wildtype	UCBSV infectious clone
1156	P1	Isoleucine	Valine
1441	P3	Leucine	Serine
2404	CI	Leucine	Proline
2509	CI	Glutamine	Arginine
2671	CI	Glutamine	Leucine
4039	CI	Tyrosine	Cysteine
4048	CI	Isoleucine	Threonine
4075	CI	Methionine	Threonine

4.2.9 *In-vitro* transcription and infection experiments

After confirming the integrity of the assembled infectious clone genome by PCR, restriction analysis and sequencing, the SP6 promoter was introduced at the 5' end of the clone by yeast homologous recombination. The SP6 promoter was chosen because part of the consensus sequence recognised by the SP6 RNA polymerase enzyme (ATTAGGTGACACTATAGAGNG) was present at the 5' end of the UCBSV genome. Furthermore, the sequence of the SP6 promoter is small and could thus be synthesised as part of the primer used in amplifying the 5' end of the genome. In addition, a *Kpn*I restriction site was introduced at the 3' end of the clone to allow for run off transcripts to be produced in the upstream *in-vitro* transcription reactions.

A plasmid concentration of 1 mg/ml was required for each *in-vitro* transcription reaction so the plasmid yield had to be scaled up. In order to scale up the plasmid yield for the *in-vitro* transcription reactions, an experiment was done using different *E. coli* strains to determine which of the cells gave high yields for the plasmid. The cells tested included Sure 2, CD4, Able K, XL-Gold, DH5- α , and α -select cells. It was realized that high plasmid yield was recovered

from DH5- α and α -select cells. However, high plasmid yield could only be obtained if freshly transformed cells were used and the plasmid recovered immediately from fresh colonies. High amounts of plasmids could also be obtained by performing large scale plasmid extraction using the Thermo Scientific pJET Midi Prep Plasmid Extraction Kit.

4.2.9.1 *In-vitro* transcription to generate transcripts

The infectious clone plasmid (pYES2.1_UCBSV complete genome) was linearized by *Kpn*I digestion prior to transcription, all enzyme and associated buffers were removed using the GeneJET PCR Purification Kit (Thermo Scientific). *In-vitro* transcripts were generated using a Riboprobe SP6 *in-vitro* Transcription system (Promega) in conjunction with a Ribo m7G Analog (Promega) according to the manufacturer's guidelines. The RNA transcripts were successfully generated as evidenced in (Figure 4.9).

The RNA transcripts obtained from the *in-vitro* transcription was inoculated onto *N. benthamiana* and *N. clevelandii* plants. The inoculated plants were maintained in a growth cabinet at 28°C. Two weeks post infection, the infected plants showed mild mosaic like symptoms characteristic of the virus infection (Figure 4.10).

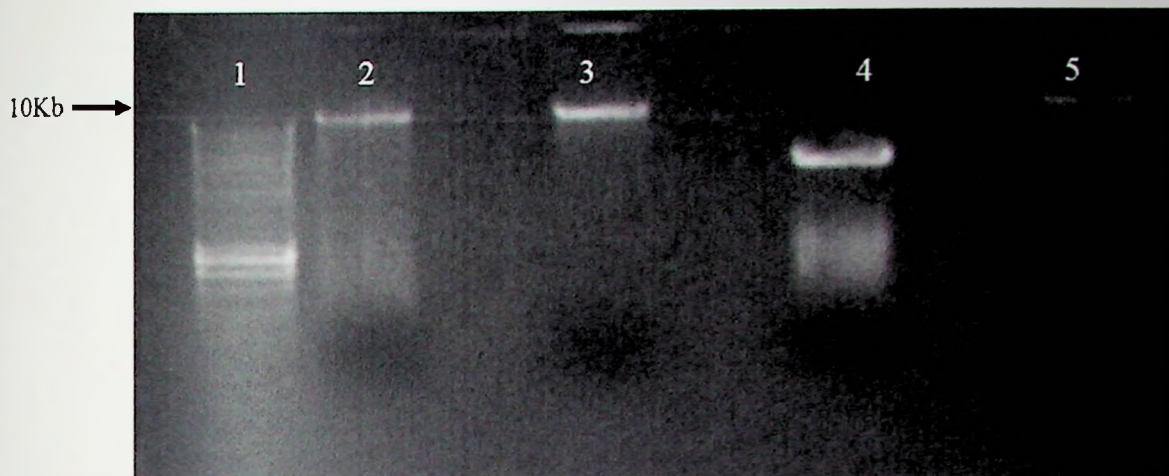


Figure 4.9: *In-vitro* transcription reactions inoculated on plants. Lane 1 contains 5 μ l of Hyper Ladder Lane 2 and 3 contain 10 μ l of *in-vitro* reactions containing the RNA transcripts. Lane 4 is the positive control *in-vitro* reaction for the Promega Kit. Lanes 5 is the *Kpn*1 digested backbone template used in the *in-vitro* transcription reactions.

4.2.9.2 Inoculation of plants with infectious clone

The *in-vitro* transcription reactions (containing RNA transcripts) were used as a source of inoculum to infect indicator plants. For a positive control, cassava leaves infected with wild type UCBSV Kikombe showing CBSD symptoms were ground in liquid nitrogen using a pestle and motor and the infection buffer was added during the grinding at a ratio of 1:6 (w/v) tissue to buffer concentration and the inoculum was rubbed onto plants. For a negative control (mock inoculation), 0.01M potassium phosphate ($K_2HPO_4 + KH_2PO_4$) pH 7.0 buffer containing 0.2% sodium sulphite and 0.01M β -Mercaptoethanol was rubbed on plants.

After 2 weeks, systematically infected leaves of *N. benthamiana* showing mild mosaic like symptoms (Figure 4.10) were collected and RNA extracted for cDNA synthesis and subsequent virus detection by RT-PCR and sequencing. The *N. clevelandii* plants did not show any symptoms however, systemic leaves were collected for detection of the virus by RT-PCR.

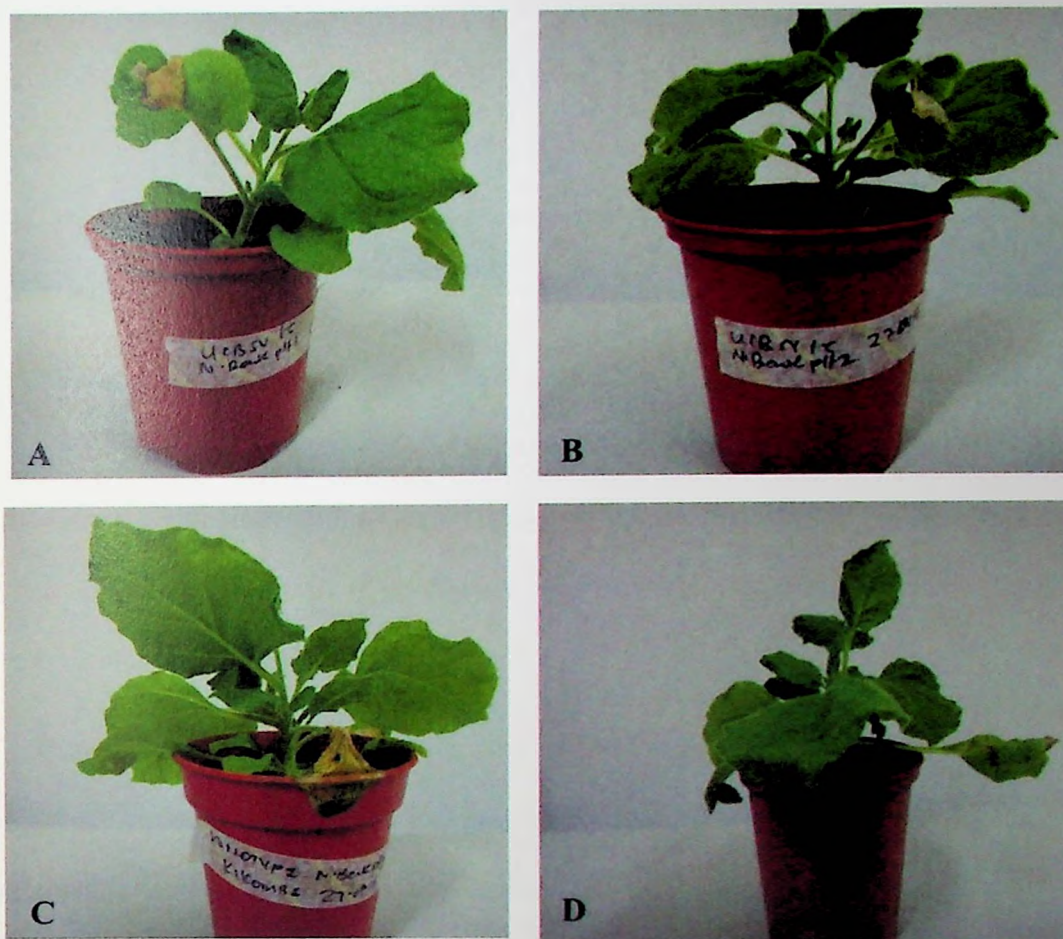


Figure 4.10: Symptoms on systemic leaves of the infectious clone infected plants: **A:** *N. benthamiana* inoculated with *in-vitro* RNA transcripts (Infectious clone virus) showing mild mosaic symptoms, **B:** *N. benthamiana* infected with *in-vitro* RNA transcripts (Infectious clone virus) with mild mosaic symptoms, **C:** *N. benthamiana* plant inoculated with wildtype UCBSV inoculum (Positive control) **D:** *N. benthamiana* plant rubbed with potassium phosphate buffer (Negative control).

4.2.9.3 Confirmation of UCBSV infection

To confirm that the mild mosaic symptoms observed were caused by UCBSV, systemically infected leaves were collected and RNA extracted for cDNA synthesis and subsequent virus detection by RT-PCR and sequencing. RNA extraction was as explained in the protocol in (section 2.6.1)

PCR was performed using UCBSV specific primers 3'867F and 3'867R (Table 4.1) to confirm that the symptoms observed were indeed caused by UCBSV. Section 3 (position 3195 to 4062 of UCBSV genome) was targeted in the RT-PCR amplification and the expected band size of 867bp (Figure 4.11) was obtained. This band was excised from the gel, cloned, and sequenced to further confirm the infection.

Quantitative qPCR was performed to compare the viral titres in the plants infected with the UCBSV infectious clone and UCBSV wildtype virus and the average C_T values for the plants infected with the UCBSV infectious clone was 30.42 whilst plants infected with the wildtype UCBSV virus had an average C_T of 30.74 confirming that the virus was accumulating at the same levels in the plants.

To confirm reproducibility of the results, infection was repeated two more times. In the second repetition, five plants (three *N. benthamiana* and two *N. clevelandii*) were inoculated. All the *N. benthamiana* plants showed mosaic like symptoms and RT-PCR and sequencing confirmed that they were infected. *N. clevelandii* plants did not show any symptoms but RT-PCR and Sanger sequencing confirmed that they were infected.

In the third repetition of the experiment, a total of 16 plants (ten *N. benthamiana* and six *N. clevelandii*) were inoculated. All ten *N. benthamiana* plants showed symptoms similar to those observed in the first experiment. There were no symptoms in the *N. clevelandii* plants. RT-PCR was carried out to detect UCBSV in four selected plants. PCR products were of the expected sizes and sequencing confirmed the UCBSV sequence.

Furthermore, passage experiments were performed using the leaves of *N. benthamiana* plants infected with the infectious clone of UCBSV as a source of inoculum to infect other *N. benthamiana* plants, the infected F2 plants showed symptoms. RNA was extracted for cDNA

synthesis and subsequent PCR was performed, the UCBSV infection could be detected and the sequencing results further confirmed that the infectious clone could be propagated.



Figure 4.11: RT-PCR on the infectious clone infected plants: Lane 1 and 10 contain 5 μ l of HyperLadder 1, Lane 2: RT-PCR product using cDNA from infectious clone infected *N. benthamiana* plant inoculated with *in-vitro* infectious clone, Lane 3 and 4: RT-PCR product using cDNA from *N. clelandii* plant inoculated with *in-vitro* infectious clone, Lane 5 and 6: wild type UCBSV infected *N. benthamiana* plant, Lane 7: Positive control. Lane 8 and 9: Negative control plants.

4.2.9.4 Sequencing the PCR products

The 867bp fragment obtained by PCR for the two *N. benthamiana* plants infected with the UCBSV infectious clone was cloned and sequenced. The sequences were aligned in Clone Manager Software Suite 7. Alignment was done alongside the assembled genomes of the infectious clone and the wildtype original UCBSV Kikombe genome. The sequencing results (Figure 4.12) show the infectious clone infected plants to be containing sequences similar to the assembled infectious clone although there were some transitions and transversions. Furthermore, the region targeted in the RT-PCR had a difference between the wildtype and infectious clone UCBSV sequences. The PCR clones had the (Adenine), showing they are from the infectious clone and not from accidental cross-contamination.

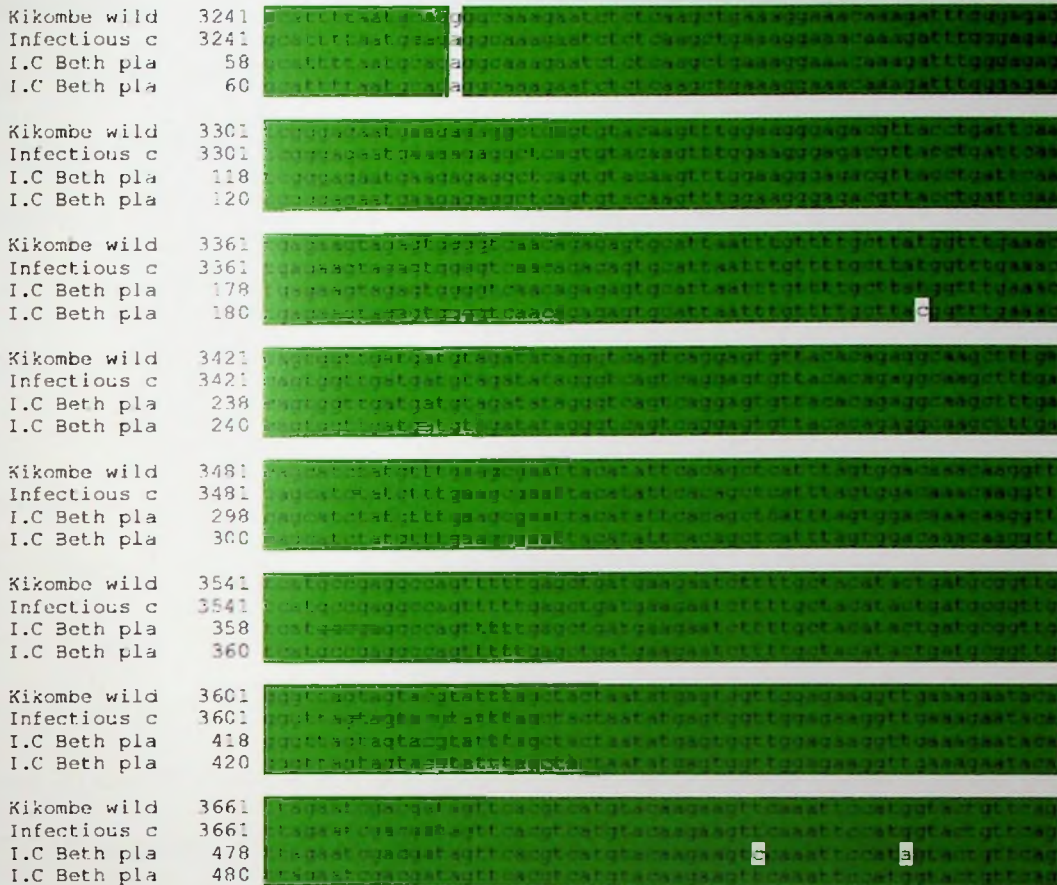


Figure 4.12: Multiple alignment of the sequences from UCBSV Infectious clone infected *N. benthamiana* plants aligned/compared against the genomes of the assembled wild type UCBSV genome and Infectious clone genome. Section 3 (867bp) of UCBSV genome was targeted. Kikombe wild is the genome of the wildtype UCBSV isolate, Infectious c is the genome of the UCBSV infectious clone. I.C Beth pla is the sequence recovered from RT-PCR and sequencing from 2 different *N. benthamiana* plants infected with the UCBSV infectious clone

4.2.10 Testing the infectious clone in cassava plants

The ultimate application of the constructed infectious clone is to test breeding lines of cassava for resistance to UCBSV. The infectious clones provide a homogenous source of virus infection for testing and breeding for resistance to CBD. It was therefore important to test and confirm the infection status of the generated infectious clones on cassava plants.

Mechanical inoculation was used to infect the cassava plants with the UCBSV infectious clone. Four cassava plants rubbed with the infectious RNA transcripts. Plants were monitored for symptom development and it was observed that plants developed symptoms after six weeks post infection (Figure 4.14). The mosaic symptoms observed in cassava plants were more conspicuous than the mild mosaics observed in *N. benthamiana* plants. RNA was extracted from four cassava plants and cDNA synthesised to confirm that the observed symptoms were due to UCBSV. Initially, there were problems with successful extraction of RNA from the cassava plants and a number of optimisations were performed to overcome this challenge. The EZNA Plant RNA extraction kit could not successfully extract RNA from the cassava leaves, this may have been probably due to the high levels of phenolics present in the cassava leaves. Finally, a modified CTAB method (section 2.6.1) allowed for good RNA extraction from the cassava leaves. To confirm that the cDNA synthesised from the cassava leaves was PCR amplifiable, a control primer set amplifying a house keeping gene ribulose –bisphosphate carboxylase (RuBisCo) gene *rbc1* 5' CTA CTA TGG TGG CTC CGT TC 3' and *rbc2* 5' CCC TTC AGT CGG AGA AAC TC 3' (Monger *et al.*, 2001) was used to check the cDNA synthesised from cassava. This primer pair generated a 619 bp product (Figure 4.13B). A 1000bp UCBSV specific PCR product was amplified from the cassava plants that were infected with the UCBSV infectious clone.

The cassava infection experiment with the UCBSV infectious clone was repeated three times. In the second repetition, four cassava plants were infected and in the third repetition, ten cassava plants were infected. In all the repetitions, reproducible results were obtained thus confirming that the UCBSV infectious clone could successfully infect cassava plants by mechanical inoculation. All the infected cassava plants showed symptoms.

Furthermore, the infected cassava leaves were used as inoculum to back inoculate ten *N. benthamiana* and ten *N. clevelandii* plants and the back inoculated plants were systemically infected. The back inoculated *N. benthamiana* plants showed mild mosaic symptoms while the *N. clevelandii* plants did not show any symptoms although the UCBSV virus could be detected by RT-PCR.

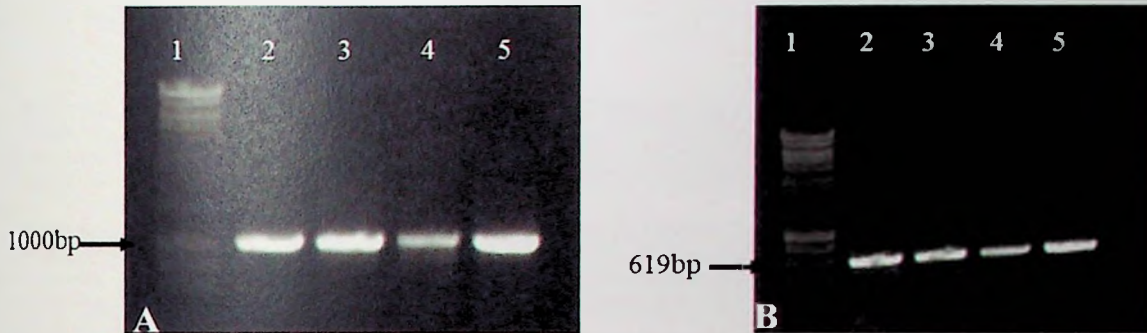


Figure 4.13: Confirming UCBSV infection in cassava plants. **A:** Amplification of section 5 (Position 4758-5798) approximately 1 kb fragment of UCBSV from cassava plants. Lane 1 contains 5 μ l of HyperLadder 1, Lane 2-5 contain 1000bp RT-PCR amplification products from the four different cassava plants infected with UCBSV infectious clone. **B:** Amplification of RuBisCo from four cassava plants amplifying a 619bp PCR product.

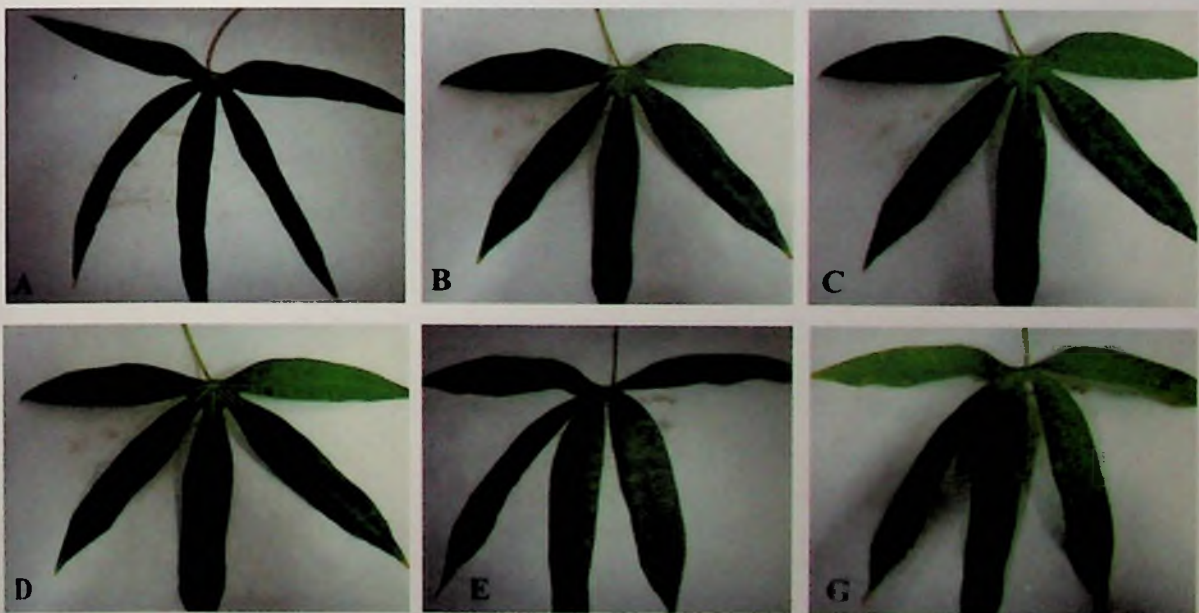


Figure 4.14: Systemic leaves of cassava plants infected with UCBSV infectious clones showing characteristic mosaic symptoms. **A:** Healthy negative control plant. **B-G:** Leaves from different cassava plants inoculated with UCBSV infectious clone.

4.2.10.1 Sequencing the PCR products to confirm UCBSV infection in cassava plants.

The PCR products obtained from the cassava plants infected with UCBSV infectious clone were cloned and sequenced to confirm that the mosaic symptoms observed were due to UCBSV infection. The sequencing results were aligned against the genomes of the wildtype UCBSV genome and the infectious clone genome (Figure 4.15).

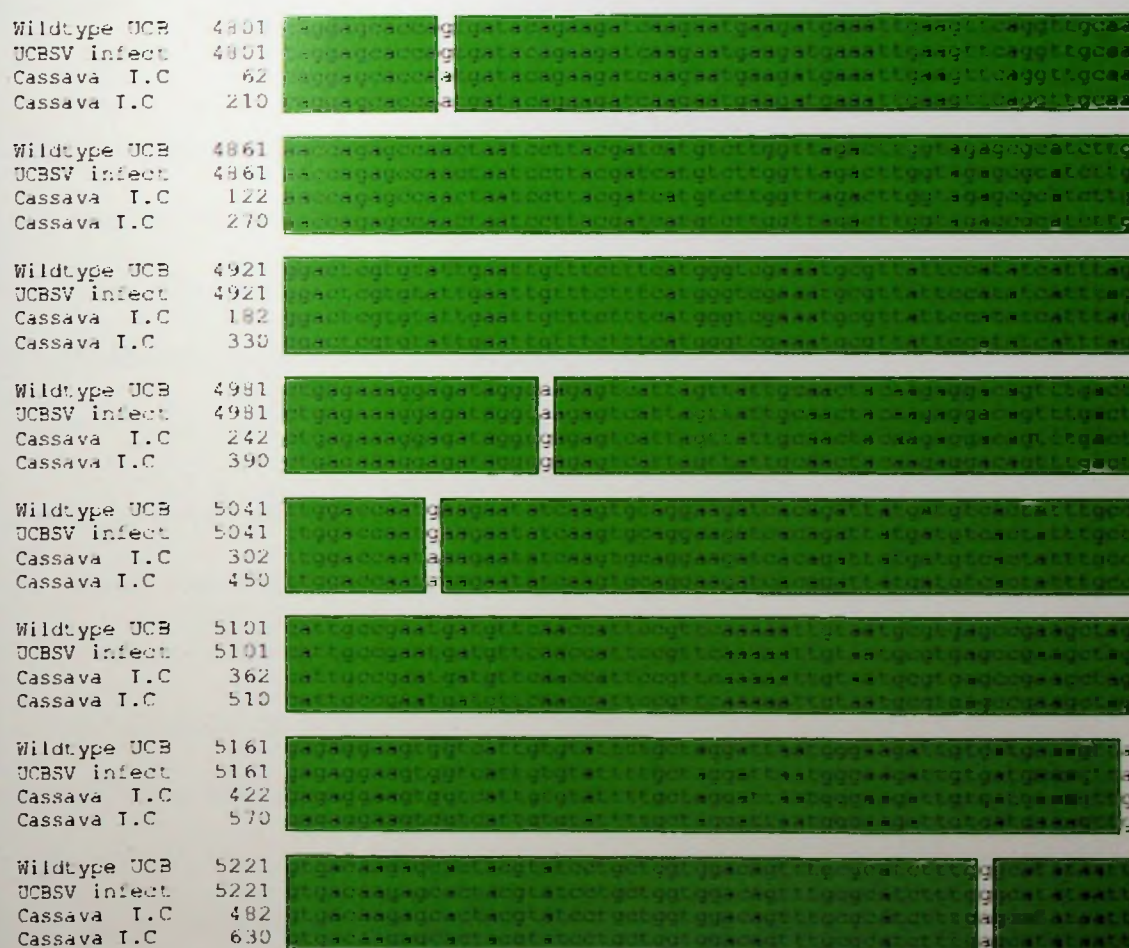


Figure 4.15: Multiple alignment of sequences from UCBSV infectious clone infected cassava plants. The 1000bp RT-PCR product amplified from two cassava plants infected with UCBSV infectious clone: The sequences aligned in the expected regions of the UCBSV wildtype and infectious clone genomes. Wildtype UCB is the genome of the wildtype UCBSV Kikombe isolate, UCBSV infect is the genome of the UCBSV infectious clone. Cassava I.C is the sequence recovered from RT-PCR of cassava plants infected with the UCBSV infectious clone.

4.2.11 Construction of *in-vivo* UCBSV Infectious Clone

It was considered important to construct an *in-vivo* UCBSV infectious clone as it has advantages over the *in-vitro* UCBSV infectious clone. One of the main advantages of *in-vivo* infectious clones over *in-vitro* infectious clones is that they can be easily delivered through *Agrobacterium* mediated agroinfiltration thus bypassing the need to purchase expensive *in-vitro* transcription kits needed for working with *in-vitro* infectious clones in addition to other advantages.

4.2.11.1 Construction of UCBSV genome in PCAMBIA 0380 backbone

Attempts to construct an *in-vivo* UCBSV infectious clone in pCAMBIA0380 backbone vector were made. The *in-vivo* clone was designed in the yeast adapted pCAMBIA0380 (pCAMBIA0380YA) (Ford *et al.*, 2015) backbone plasmid which has been modified to contain a *ura3* gene to allow for selection of transformants on uracil deficient media (Figure 4.16).

For the yeast homologous recombination in yeast experiments, full-length UCBSV genome approximately 9 kb UCBSV insert was released using *Kpn1/Not1* restriction enzymes from the pYES2.1 vector containing the full-length UCBSV genome. Alternatively PCR fragments of the 5' end of the genome (approximately 5 kb) and 3' end of the genome (approximately 5 kb) were used in homologous recombination in yeast experiments. The 35S CaMV Promoter and nopaline synthase Terminator (TNos) PCR sections were amplified from the pCAMBIA2300 vector. The pCAMBIA0380YA plasmid was linearized by digestion with *BamH1* and *Kpn1* restriction.

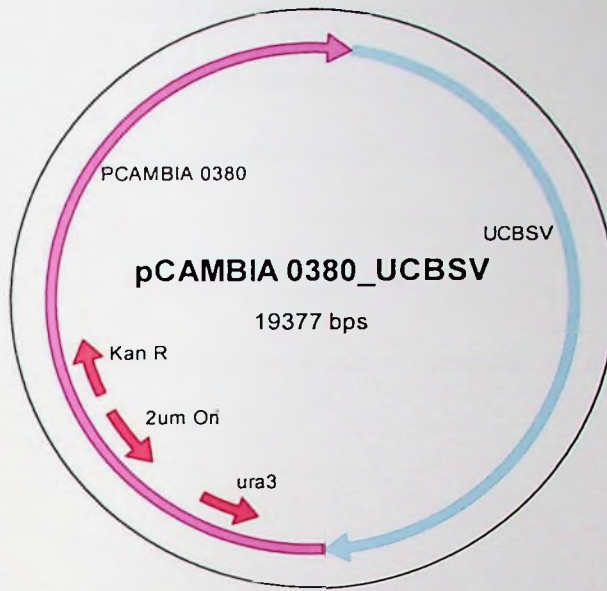


Figure 4.16: UCBSV *in-vivo* clone map: The UCBSV genome is represented by a blue arrow. The pink arrow represents the yeast adapted pCambia0380 backbone vector. The Kan resistance, 2 um Ori and *ura3* genes is a separate red arrows on the map.

The different 9 kb PCR fragments encompassing the complete genome were recombined in pCambia0380YA backbone vector and the plasmid extracted from yeast and rescued into *E. coli* cells. As a positive control, yeast cells were transformed with undigested yeast adapted pCambia0380YA backbone vector. Negative controls consisted of yeast transformation with *Bam*H1 and *Kpn*1 digested pCambia0380YA vector and water. Yeast cells transformed with undigested pCambia0380YA generated about 450 colonies confirming the viability of the yeast cells for transformation. The yeast cells transformed with *Bam*H1 and *Kpn*1 digested pCambia0380YA generated about ten colonies and there were no colonies on water transformed control plates. The experimental plates transformed with different UCBSV sections (section 1-8) generated about 300 colonies confirming successful recombination. Ten yeast colonies were individually selected and grown in SM media. Yeast plasmid was extracted and then rescued *into E. coli* cells.

However, restriction analysis on the recovered plasmids from the experimental plates using *EcoR*I and *Hind*III restriction enzymes released a 700bp section which is the expression cassette (35S Promoter and TNos Terminator) suggesting that most of the colonies contained the circularized pCAMBIA0380YA plasmid (Figure 4.17) and not the recombinant (pCAMBIA0380YA with the UCBSV genome). This suggests that most of the colonies observed on the experimental plates were a result of remnants of circular molecules of the pCAMBIA0380YA plasmid.

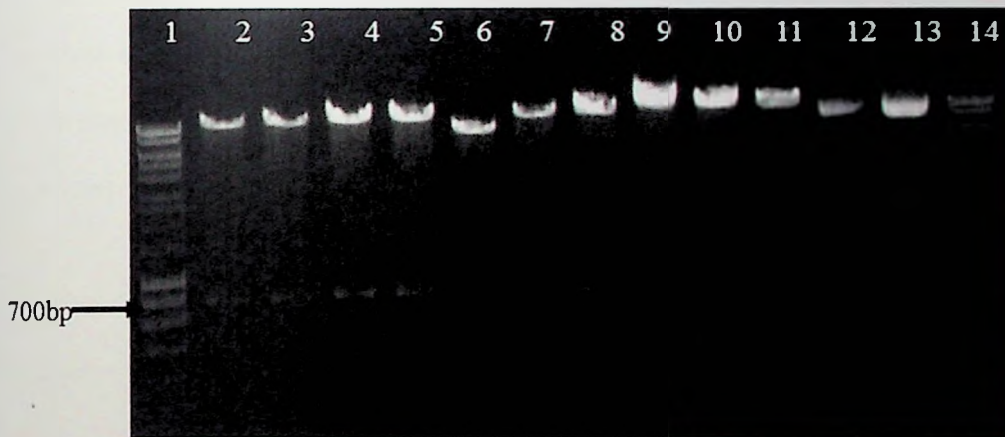


Figure 4.17: *Hind*III/*Eco*R1 restriction of extracted plasmid. Lane 1 and 14 contain 5 μ l of HyperLadder1, Lanes 2, 3, 4, 5, 7 and 8: different restrictions from different clones showing the expected 700bp. Lanes 6, 9 and 10 released an unexpected 400bp fragment. Lane 12 and 13 contain undigested circular pCAMBIA0380YA backbone plasmid.

4.2.11.2 Introducing restriction sites at the 5' end of the UCBSV genome and 3' end of the UCBSV genome

Another option to construct an *in-vivo* UCBSV infectious clone was to try and introduce restriction sites at the 5' end of the UCBSV genome and at the 3' end of the UCBSV genome. The restriction sites introduced included the *Xma*I restriction site at the 5' end of the genome and a *Kpn*I site at the 3' end of the genome. A vector backbone pCAMBIA 2300 contained

these restriction sites between the 35S Promoter and TNos Terminator and could thus be used in agroinfiltration. It was anticipated that the vector backbone and the UCBSV genome digested with similar restriction enzymes could be ligated by T4 ligase to obtain a recombinant plasmid which could be used for agroinfiltration. However, the restriction sites added extra nucleotides at the 5' end of the genome and it has been reported that the presence of extra nucleotides at the 5' end affects the infectivity of the infectious clone. The proposed restriction sites were successfully introduced at the ends of the genome (Figure 4.18) in pYES2.1 backbone vector. However, the procedure failed at the ligation step where the 10 kb UCBSV genome fragment could not be ligated *in-vitro* into the pCAMBIA2300 plasmid digested with *Xma*I and *Kpn*I restriction enzymes. This is evidenced from the fact that different *in-vitro* ligation reactions transformed into *E. coli* cells did not yield any transformants confirming that the *in-vitro* T4 DNA ligations reactions were not successful.

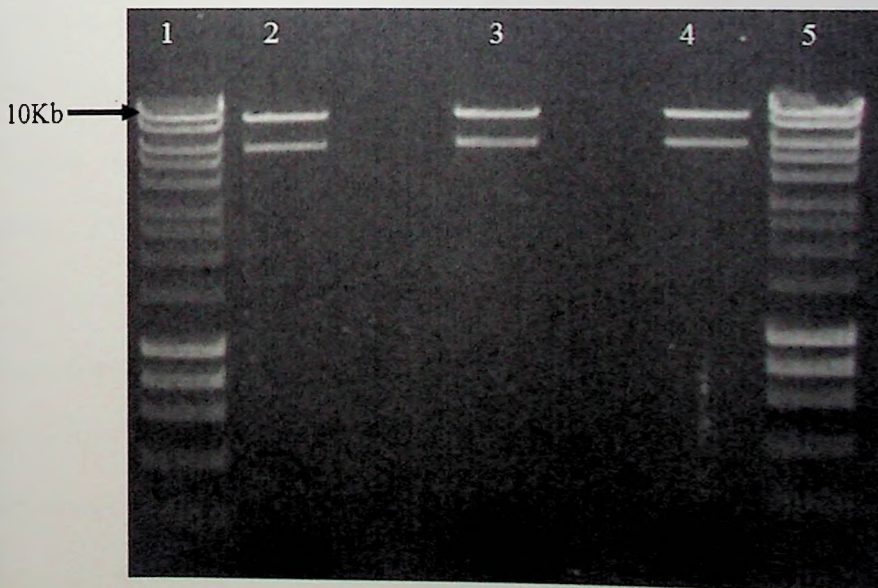


Figure 4.18: *Xma*I and *Kpn*I restriction of the pYES2.1 containing UCBSV plasmid with created *Xma*I restriction site at the 5' end of genome. Lane 1 and 5 contain 5 μ l of HyperLadder1. Lane 2, 3 and 4 contain different clones of *Xma*I and *Kpn*I restricted plasmids.

4.2.11.3 Introduction of introns into the UCBSV genome

In order to overcome the problems associated with constructing the UCBSV *in-vivo* clone, introns were introduced into the P3, CI and NIb encoding genes. The proteins were chosen as a similar CBSV Tanza infectious clone had been stabilised by intron insertion into the sequences encoding these proteins. Primers were designed to introduce introns in the UCBSV genome. Intron 1 was introduced at position 1969 in the P3 region of the UCBSV genome. Intron 2 was introduced at position 3809 in the CI region of the UCBSV genome. Intron 3 was introduced at position 5702 in the NIb region of the UCBSV genome (Figure 4.19). The PCR sections containing introns were amplified from the UCBSV pYES2.1 infectious clone (Figure 4.20). Homologous yeast recombination was then performed to ligate the intron containing UCBSV sections in the pCAMBIA0380YA backbone.

The different intron containing PCR fragments were initially recombined in pYES2.1 backbone vector and the plasmid extracted from yeast and rescued into *E. coli* cells. As a positive control, yeast cells were transformed with undigested yeast adapted pYES2.1 backbone vector. Negative controls consisted of yeast transformation with *Eco*R1 and *Sac*I digested yeast adapted pYES2.1 vector and water. Yeast cells transformed with undigested pYES2.1 generated about 500 colonies confirming the viability of the yeast cells for transformation. The yeast cells transformed with *Eco*R1 and *Sac*I digested pYES2.1 generated 15 colonies and there were no colonies on water transformed control plates. The experimental plates transformed with intron containing UCBSV PCR fragments generated 300 colonies confirming successful recombination. Ten yeast colonies were individually selected and grown in SM media. Yeast plasmid was extracted and then rescued into *E. coli* cells. The extracted plasmid was then used as a template to generate new PCR sections for yeast recombination in pCAMBIA0380YA. However, the plasmids recovered from *E. coli* cells after yeast recombination in pCAMBIA0380YA vector backbone were of unexpected sizes. It was

expected that a successful recombination would yield a plasmid of approximately 19 kb (Figure 4.16), however, the recovered plasmid were less than 1 kb in size (Figure 4.21).

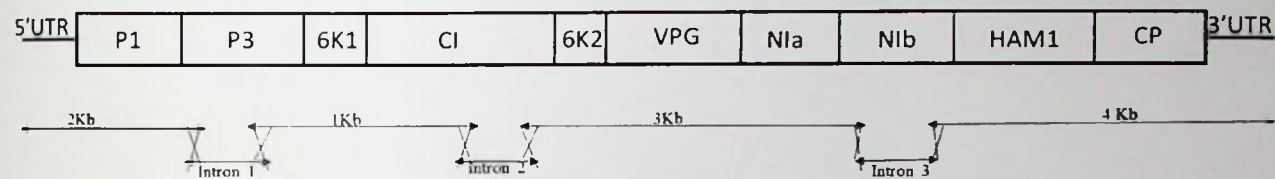


Figure 4.19: Intron insertion in the different genomic regions of UCBSV: Introns were inserted in the P3, CI and NIb regions of UCBSV genome.

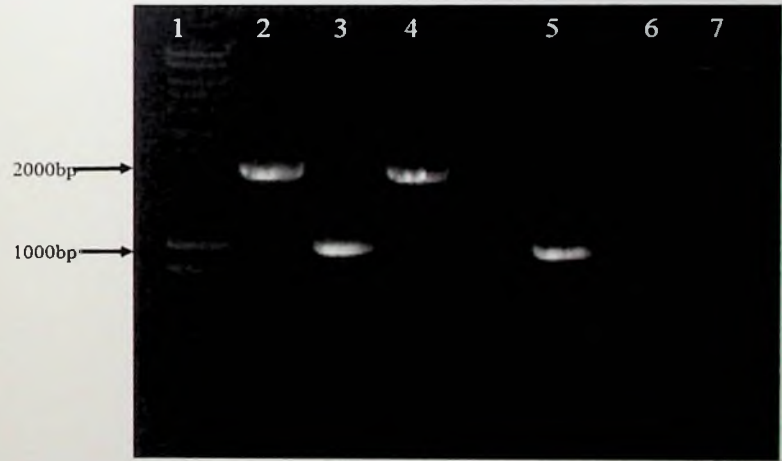


Figure 4.20: Amplification of intron containing PCR products from UCBSV infectious clone: A: Lane 1 and 7 contain 5 µl of HyperLadder1: Lane 2 ,3, 4, 5 and 6 contain intron containing UCBSV PCR products.

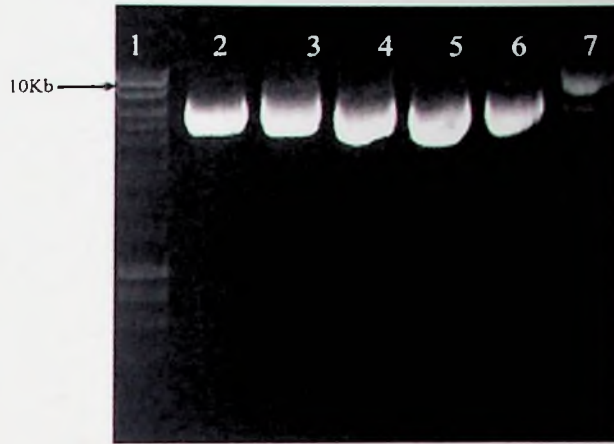


Figure 4.21: Undigested plasmids recovered from the *E. coli* cells after transferring the intron containing UCBSV sections into yeast adapted pCAMBIA0380. The plasmids recovered were of incorrect sizes (less than 10 kb) whilst the expected size of the recombinant plasmids was over 10 kb.

4.3 Discussion

In this chapter, a viable *in-vitro* infectious clone of UCBSV Kikombe was constructed and tested in tobacco and cassava plants. The UCBSV infectious clone genome was sequenced and assembled and the sequence submitted to the GeneBank under accession number KX753357. The strategy used in building the infectious clone was homologous recombination in yeast (Fernandez-Delmond *et al.*, 2004). Yeast homologous recombination is a powerful technique in the construction of infectious clones of other viral species where classical cloning methods have failed or proved to be time consuming and has been applied in the construction of infectious clones for *Zucchini yellow mosaic virus* (ZYMV) and *Papaya ringspot virus* (PRSV) (Desbiez *et al.*, 2012).

Previous reports show that *Potyvirus* sequences are toxic in *E. coli* due to the presence of toxic cryptic promoter elements (Johansen, 1996; Jakab *et al.*, 1997; Chikh *et al.*, 2011; Fakhfakh *et al.*, 1996). In most cases, introns are introduced into the genomes to interrupt the expression of toxic proteins in *E. coli*. The UCBSV isolate used in this study was stable on propagation through *E. coli* when using pYES2.1 as a backbone vector as the sequencing results confirmed that the entire open reading frame of the isolate contained no insertions or deletions that would lead to frameshifts. Stable intron-less infectious clones have earlier been constructed for other potyviruses like the *Zucchini yellow mosaic virus* (ZYMV) (Desbiez *et al.*, 2012).

However, the UCBSV genome could not be successfully propagated in pCAMBIA0380YA backbone plasmid despite several attempts including the introduction of introns in the UCBSV genome. The possible explanation could be that the UCBSV genome was toxic when placed under the strong constitutive 35S promoter present in pCAMBIA0380YA. Differences in the vector backbone have been reported to affect the stability of infectious clones as has been reported in the infectious clone of *Tobacco etch virus* (TEV) where the infectious clone

pTEV7Da was found to be unstable and stability was only obtained by generating the infectious clone in three different backbone vectors pMTEV, p35TEV and pGTEV (Bedoya & Daros 2010). In some other cases, activation of cryptic promoters in yeast may lead to expression of toxic proteins that lead to failure in propagation of infectious clones in *E. coli*. This challenge was encountered in the construction of the infectious clone of the *Potato virus Y* (PVY) and a successful infectious clone could only be generated by *in-vitro* ligation of different PCR fragments (Fakhfakh *et al.*, 1996).

A challenge was encountered in trying to scale up the concentration of the plasmid for *in-vitro* transcription reactions. In order to scale up the plasmid yield for these reactions, an experiment was done using different *E. coli* strains to determine which of the cells gave high yields for the plasmid. The cells tested included Sure 2, CD4, Able K, XL-Gold, and DH5- α , and α -select cells and it was realized that high plasmid yield was recovered from DH5- α and α -select cells. However, high plasmid yield could only be obtained if freshly transformed cells were used and the plasmid recovered immediately from fresh overnight culture. A concentration of 1mg/ml was required for each of the *in-vitro* transcription reactions. Therefore the concentration of the plasmid to be used in *in-vitro* transcription was scaled up.

In-vitro transcription has been used in making infectious clones (Domier *et al.*, 1989). It was successfully used to develop infectious *in-vitro* clone of UCBSV in the present work. However, there are challenges associated with this approach. Firstly, the actual amount of RNA transcripts that are infecting the plants seem atypical on gel and cannot be sized accurately. Secondly, the kit for *in-vitro* transcription is expensive and may pose budget constraints and thirdly, the *in-vitro* transcripts generated are sensitive to RNA degradation.

Efforts to construct an *in-vivo* infectious clone which has a 35S promoter at the 5' end of UCBSV genome and TNOS terminator at the 3' end UCBSV genome cloned inside a

pCAMBIA0380YA plasmid failed as the plasmids recovered were of incorrect sizes meaning that the UCBSV genome could stably only be stably propagated in pYES2.1 backbone vector but not the pCAMBIA0380YA vector backbone. This is confirmed by the fact that the *Xma*I restriction site was successfully introduced in pYES2.1 backbone as the correct restriction pattern was observed. A number of factors affect the stability of infectious clones among which are the vector backbone in which the infectious clone is propagated. Working with *in-vivo* infectious clones overcomes most of the challenges associated with *in-vitro* infectious clones as viral transcripts are generated in the plant via 35S promoter driven expression using *Agrobacterium* mediated infiltration into the plants (agroinfiltration).

In conclusion, although a lot of work has been published to understand genetic diversity, epidemiology, distribution, and aetiology of UCBSV. There has not been any report on generation of infectious clones of UCBSV, which is needed for studies on plant resistance and genomic functional studies. Therefore, to the best of my knowledge, this is the first work that has led to the development of the full-length infectious clones of a UCBSV isolate from East Africa.

In the next chapter, a number of modifications were performed on the generated UCBSV infectious clone which included mutating functional motifs in the P1 protein, generating knock out infectious clones of P1 and HAM to characterize the functions of these proteins in the UCBSV genome.

Chapter 5: Characterization of P1 and HAM1 proteins in UCBSV “Kikombe” isolate

5.1 Introduction

5.1.1 Functions of UCBSV proteins

The UCBSV genome is translated as a polyprotein that is autocatalytically cleaved by viral encoded proteases into 10 mature proteins namely from the (amino to carboxyl terminus) P1, P3, 6K1, CI, 6K2, VPg, NIa, NIb, HAM1 and the coat protein (Mbanzibwa *et al.*, 2009a; Winter *et al.*, 2010). The different proteins perform different roles in the viral infection process to allow effective viral spread and establishment in the host.

Understanding the pathogenicity factors in viral infection is very important in allowing the design of appropriate control strategies for a given viral pathogen (Brigneti *et al.*, 1998; Duff-Farrier *et al.*, 2015; Revers & García, 2015; Hajimorad *et al.*, 2011). However, the functions of the different genes encoded in the CBSV and UCBSV genomes have not been studied experimentally for most of the proteins although inferences on the putative roles has been predicted based on the presence of known conserved motifs present in a given protein coding region and sequence homology with characterised proteins in the family *Potyviridae* (Riechmann *et al.*, 1992; Dombrovsky *et al.*, 2014). For instance, in the P1 protein of PVY, a conserved basic LXXKA and zinc-like finger motifs have been reported to be involved in RNA silencing (Valli *et al.*, 2006; Dombrovsky *et al.*, 2014; Pruss *et al.*, 1997). A VPTRV motif is present in the potyviral CI protein and in PVY, it plays a role in the nucleotide binding activities of the protein whilst the NIa protein contains the H-34X-D-71X-GDCG motif that is important in the proteolytic activities of the protein (Adams *et al.*, 2005). The multifunctional

HC-Pro (absent from UCBSV) contains the KITC and PTK motifs that are involved in aphid transmissibility (Ala-Poikela *et al.*, 2011).

5.1.1.1 RNA silencing in viral infections

RNA silencing is an RNA degradation mechanism that occurs in a broad range of eukaryotic organisms including fungi (quelling), animals (RNA interference) and plants (post transcriptional gene silencing) (Ding & Voinnet, 2007). The RNA silencing mechanism is an antiviral defence mechanism that is triggered by double-stranded RNA (dsRNA) and requires a conserved set of gene products (Baulcombe, 2002; Lakatos *et al.*, 2004; Ratcliff, 1999., Elbashir *et al.*, 2001; Voinnet, 2001). Single-stranded RNA viruses often induce RNA silencing because the replicative form of the virus is double-stranded and can therefore activate RNA silencing. The viral double-stranded RNA is thus recognized as a pathogen associated molecular pattern (PAMP) (Ding, 2010) and is processed into small interfering RNAs (siRNAs) of 21 to 25 nucleotides by the host ribonuclease Dicer like (DCL) enzyme (an RNase III enzyme), corresponding to both sense and antisense strands of the target gene (Hamilton & Baulcombe, 1999). These siRNAs become associated with a protein complex referred to as the RNA-induced silencing complex (RISC) and argonaute (AGO) protein (an RNase H enzyme) where they serve as guides to select the target RNAs and effect their degradation (Hammond *et al.*, 2000). RNA silencing is referred to as post transcriptional gene silencing (PTGS) in plants as it reduces endogenous gene expression mainly through miRNA pathways (Jones-rhoades *et al.*, 2006) (Jones-Rhoades *et al.*, 2011) and in this way, PTGS is able to defend the plant against viruses (Voinnet, 2005).

5.1.1.2 P1 as an RNA silencing suppressor

To counteract the RNA silencing mechanism, many plant and animal viruses have evolved to acquire RNA silencing suppressor proteins that act to interfere with different of the siRNA pathways and in many cases miRNA pathways as well (Anandalakshmi *et al.*, 1998; Brigneti *et al.*, 1998; Kasschau & Carrington, 1998; Chapman *et al.*, 2004). A number of different proteins have been characterised as suppressors of RNA silencing of which the HC-Pro has been the most studied for potyviruses (Kasschau & Carrington, 2001). However, the CBSV genomes lack the HC-Pro and the RNA silencing role is known to be performed by the P1 protein (Mbanzibwa *et al.*, 2009a). In *Sweet potato mild mottle virus* (SPMMV), P1 inhibits RNA silencing by interacting with argonaute protein (AGO) thus preventing the formation of the RNA induced silencing complex (RISC) hence preventing the degradation of viral RNA (Giner *et al.*, 2010). Different motifs present in the P1 protein have been characterised to be involved in the RNA silencing suppressor activities of the P1 protein, including the WG/GW and LXKA motifs (Giner *et al.*, 2010; Dombrovsky *et al.*, 2014; Mbanzibwa *et al.*, 2009a). The motifs have been functionally characterised in most of the potyviruses by performing mutational analysis studies (Klein *et al.*, 1994; Valli *et al.*, 2008), however, similar studies have not been performed for UCBSV and CBSV and this analysis was experimented in this study.

5.1.1.3 HAM1 protein

Interestingly, the CBSV and UCBSV genomes encode the HAM1 protein incorporated between the NIa and NIb proteins of the genome (Mbanzibwa *et al.*, 2009a). This protein is absent in most other potyviruses but has been found to be present in *Euphorbia Ringspot virus*. The HAM1 gene has also been found in *Cassava torrado like virus* (CsTLV) (Carvajal-Yepes unpublished). Alignment of the available sequences of UCBSV and CBSV isolates in the NCBI shows the HAM1 protein to contain a conserved SHR motif. However, the actual function of

the HAM1 protein has not been determined although the homology of CBSV HAM1 protein sequence with other cellular Maf/HAM pyrophosphatases suggests that the HAM1 protein may reduce viral RNA mutagenesis by preventing the incorporation of non-canonical nucleotide triphosphates (Mbanzibwa *et al.*, 2009a).

The HAM1 protein is novel protein that has been found in the genomes of UCBSV and CBSV isolates but is absent in the genomes of other Ipomoviruses such as SPMNV, SqVYV and CVYV (Mbanzibwa *et al.*, 2009a). HAM1 is found incorporated between the NIb and coat protein (CP) regions in the genomes of all UCBSV and CBSV sequenced isolates. Artificial recombination of genes between the NIb and CP junction of most potyviruses has been reported and this site is normally exploited to introduce reporter genes like GFP and GUS in the potyviral genome (Arazi *et al.*, 2001; Kelloniemi *et al.*, 2008). The NIa enzyme cleavage sites have also been found to be adjacent at the ends of the N terminal and C terminal of the HAM1 coding sequence confirming that the HAM1 protein is cleaved from the polyprotein and may play a role in the virus infection cycle. HAM1 has also been found in the genome of *Euphorbia ring virus* belonging to the genus *Potyvirus*. Recently, the HAM1 gene was also found in the genomic sequence of *Cassava torrado-like virus* (CsTLV) a virus belonging to the genus *Torradovirus* in the family *Secoviridae* (Carvajal-Yepes, unpublished). BLASTP searches using the HAM1 sequence query shows the HAM1 protein to be related to the Maf/HAM1 family of proteins that have been reported to act in the reduction of mutations by reducing the incorporation of non-canonical nucleotides in the DNA or RNA (Mbanzibwa *et al.*, 2009a). Non-canonical nucleotides include dUTP and dITP arising from oxidation, deamination or other modifications of canonical nucleotides (Galperin *et al.*, 2006). The incorporation of the HAM1 gene in the CBSVs necessitates understanding the role of this

protein hence more studies are underway to understand and elucidate the function of HAM1 protein in the CBSV/UCBSV genome.

5.1.2 Aims of the Chapter

In this chapter, mutational analysis of P1 and HAM1 proteins present in the UCBSV genomes was performed. With the generation of infectious clones for UCBSV in the previous chapter, it was possible to functionally characterise the roles of different proteins by generating knockouts and mutated infectious clones. This approach was undertaken to functionally characterise the roles of P1 and HAM1 proteins in UCBSV Kikombe isolate.

Initially, a UCBSV infectious clone with a mutation created in the P1 protein was constructed. The mutation targeted the functional LRR motif which was changed to LAAA in the mutated infectious clone. Parallel infections were performed comparing the P1 mutated infectious clone and the wild type UCBSV infectious clone. In addition, a complete P1 knockout infectious clone was constructed and used to infect *N. benthamiana* plants. Previous studies have shown the P1 protein to function as an RNA silencing suppressor (Valli *et al.*, 1998; Johansen & Carrington, 2001). Mbanzibwa *et al.*, (2009a) showed that UCBSV P1 of isolate TZ: MLB3:07 (Accession number FJ039520) was an RNA silencing suppressor and multiple alignments of CBSV and UCBSV isolates revealed a conserved LRR motif which has been implicated in RNA silencing suppression (Valli *et al.*, 2008). In this study, the effects of the mutations on the LRR motifs was studied on the individual P1 protein by performing agroinfiltration assays in the *N. benthamiana* 16C line (a gift from David Baulcombe) which is constitutively expressing GFP.

In addition, the function of HAM1 protein was also studied by using knockout UCBSV infectious clones, the HAM1 full deletion and partial deletion knockout infectious clones were also infected on transgenic *N. tabacum* plants expressing CBSV HAM1 protein to study complementation.

5.2 Results

5.2.1 Identification of LRRA motif in P1 sequences of different UCBSV isolates

A MEGA6.0 multiple alignment was performed using the P1 translated sequence of UCBSV Kikombe and this was compared with the P1 sequences of selected UCBSV and CBSV isolates in the NCBI database. A conserved motif LRRR was found to be present in all the isolates (Figure 5.1). This conserved LRRR motif was thus targeted for modification to study the importance of the motif in RNA silencing properties of the P1 protein of UCBSV. Earlier studies have implicated the role of this motif in RNA silencing (Valli *et al.*, 2008; Valli *et al.*, 2007; Dombrovsky *et al.*, 2014). The suggested mutations (LRRR to LAAA) reduced the positive charge at the motif location and hence should affect the siRNA binding activities of the protein and hence reduce or eliminate its silencing suppression properties.

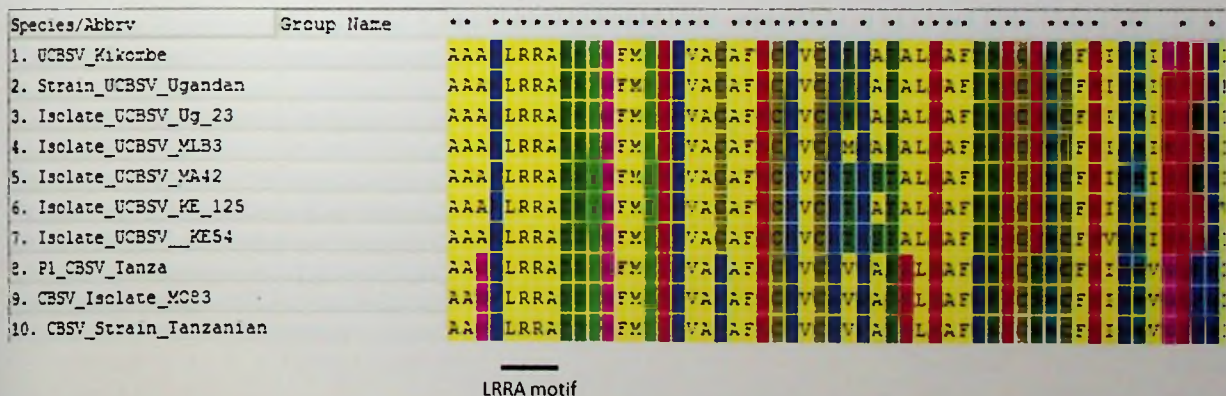


Figure 5.1: The LRRA motif in the aligned amino acid partial P1 sequence of UCBSV Kikombe and other published sequences of CBSV and UCBSV isolates. The conserved LRRA motif that was targeted for mutation in this study is underlined. Accession numbers for the isolates are; UCBSV Kikombe (KX753356), Strain Ugandan (FJ185044), Ug 23 (FN434109), MLB3 (FJ039520), Ma 42 (FN433932), Ke 125(FN433930), Ke 54 (FN 433931), Mo 83 (FN434436), Strain Tanzanian (GQ329864).

5.2.2 Primer design for site directed mutagenesis

In order to introduce a mutation in the LRRA motif in the P1 protein of the UCBSV infectious clone, mutational complementary primers were designed (Table 5.1) and used in homologous yeast recombination. The functional motif LRRA (Figure 5.1) which has been implicated to be involved in RNA silencing activities of P1 protein was targeted in the mutagenesis analysis experiment. The primers were designed in such a way as to introduce a mutation change in the LRRA motif to LAAA in the P1 protein of the UCBSV infectious clone. The amino acids chosen in the change involved converting the basic arginine amino residues (R) to the aliphatic alanine (A) and this change was anticipated to have an effect on the functional activities of the P1 protein. The primers contained mismatch sequences to the template so as to change the amino acids at the targeted position for site directed mutagenesis. A further consideration was on reducing the levels of mismatches (relative to the template) introduced in the primer sequence as this could subsequently affect the hybridisation of the mutational primers in the PCR reactions. The mutation was introduced by re-constructing the UCBSV infectious clone and generating new PCR sections for homologous yeast recombination (Figure 5.2).

Table 5.1: Oligonucleotide primers used for site directed mutagenesis of the LRRA motif to LAAA in the UCBSV P1 protein. The underlined bases represent the region targeted in the mutation analysis and the italicised amino acids represent the mutations.

Primer name	Primer Sequence	Position in UCBSV genome	Wild type amino acids	Mutated amino acids
PI Forward	TGGGCAAGAATTGAAGGCTGCCGCAAGG <u>CTAAGAA</u> <u>GGGCAACAAGCTATGGATT</u> CATGTATGATCCAGT	431-499	Leucine (L)	Not changed
PIΔLRRA Forward	TGGGCAAGAATTGAAGGCTGCCGCAAGGCTA <u>gcAgc</u> GGCA ACAAGCTATGGATT <u>CATGTATGATCCAGT</u>	431-499	Arginine (R)	Alanine (A)
PIΔLRRA Reverse	ACTGGATCATACATGAATCCATAGCTTGT <u>TGCCgcTgc</u> TAG CCTTGCGGCAGCCTTCAATTCTTGCCCA	431-499	Arginine (R)	Alanine (A)
PI Reverse	ACTGGATCATACATGAATCCATAGCTTGT <u>TGCCCTT</u> <u>CTTAGCCTTGCGGCAGCCTTCAATTCTTGCCCA</u>	431-499	Alanine (A)	Not changed

5.2.3 Construction of mutated P1 UCBSV infectious clone by yeast recombination

In the original construction of the UCBSV infectious clone (Chapter 4), the UCBSV genome had been stably reconstructed as separate halves of approximately 5 kb each representing the 5' end and 3' end of the UCBSV genome. These separate plasmids were used to introduce a mutation in the P1 protein by PCR and yeast recombination. The P1 protein in UCBSV Kikombe isolate was annotated to positions 135-1220 of the UCBSV genome so in order to introduce a mutation in the P1 protein, new PCR sections were generated for the section 1 of the UCBSV genome using the complementary primers that mutated the LRRA motif to LAAA. Hence section 1 was divided into section 1a (mutated P1) and section 1b (mutated P1) to generate new sections for introducing the mutation in the infectious clone (Table 5.2). As before, the 5' end section had an SP6 promoter sequence incorporated within the primer sequence to allow for *in-vitro* transcription while the 3' end section had a *Kpn1* restriction site for plasmid linearization to generate run off transcripts. All the sections were amplified from the original UCBSV infectious clone as template and fragments of expected sizes were obtained (Figure 5.2). All PCR amplifications were performed using Phusion High Fidelity polymerase in order to reduce on the introduction of errors in the generated PCR products.

Table 5.2: PCR sections generated in constructing UCBSV infectious clone with a mutation change in P1 protein. Sequences for primers are in Table 4.1 and Table 5.1.

Section name	Primer pair	Position in UCBSV genome	Expected product size (bp)
Section 1a (mutated P1)	VecSP62 and P1 Δ LRRA Reverse	1 to 499	459
Section 1b (mutated P1)	P1 Δ LRRA Forward and UCBSV section 1 Rev	431 to 2685	2254
Section 2,3 and 4	5'IKb FWD and 5'1442R	2382 to 4879	2497
Section 5,6,7 and 8	Mid FF1KB and 3'Final C	4758 to Poly (A) tail	4312

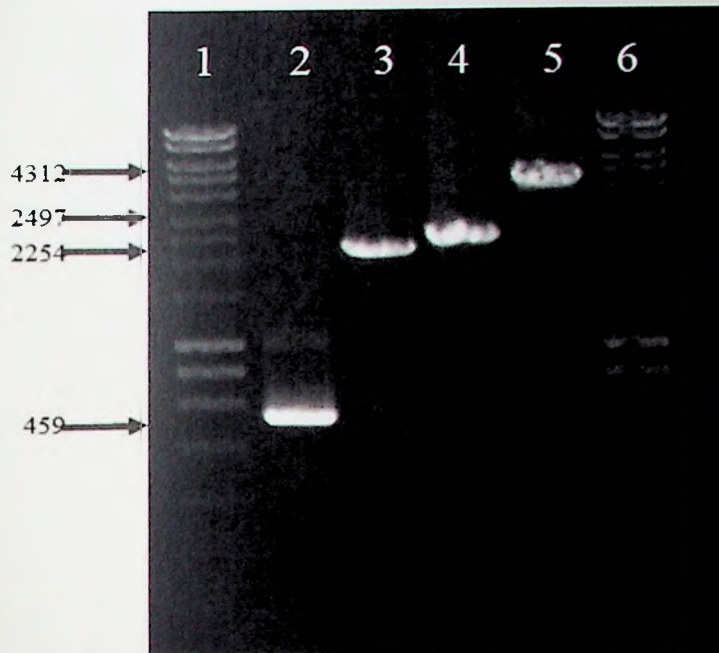


Figure 5.2: Amplification of PCR fragments from UCBSV infectious clone to generate new modified sections for introducing LRRA to LAAA mutation. Lane 1 and 6 contain 5 μ l of Hyperladder 1. Lane 2 contains 5 μ l of 459bp Section 1a (mutated P1). Lane 3 contains 5 μ l of 2254bp Section 1b (mutated P1). Lane 4 contains 5 μ l of 2497bp Section 2, 3 and 4. Lane 5 contains 5 μ l of 4312bp Section 5, 6, 7 and 8.

The generated PCR sections were used in homologous yeast recombination to re-construct a new UCBSV infectious clone with a mutation generated in the P1 protein. As a positive control, yeast cells were transformed with undigested pYES2.1 backbone vector. Negative controls consisted of yeast transformation with *Eco*R1/*Sac*I digested pYES2.1 vector and water. Yeast cells transformed with undigested pYES2.1 generated about 400 colonies confirming the viability of the yeast cells for transformation. The yeast cells transformed with *Eco*R1/*Sac*I digested pYES 2.1 generated just 9 colonies and there were no colonies on water transformed control plates. The experimental plates transformed with modified UCBSV sections generated >300 colonies confirming successful recombination. Twenty yeast colonies were individually selected and grown in SM media. Yeast plasmid was extracted and then rescued into *E. coli*

cells. Plasmid was then extracted from *E. coli* cells and analysed by restriction analysis and sequencing.

Challenges were encountered in recovering the correct plasmid from the *E. coli* cells, the initial colonies obtained did not show the correct predicted restriction pattern. The restriction pattern with *Xho*I restriction enzyme on the UCBSV infectious clone was expected to give four fragments of sizes 10,846bp, 2345bp, 1131bp and 589bp. Similar challenges had been encountered in the initial screens on colonies when the UCBSV infectious clone was being constructed. Homologous recombination in yeast was thus repeated several times and more recombinant colonies were screened for the presence of the correct plasmid and the recovered plasmids were analysed by PCR and sequencing. A single colony was identified and screened by sequencing and was confirmed to have incorporated the desired mutation. This mutated clone was used in the subsequent experiments to study the function of P1 in UCBSV.

5.2.4 Experiments of mutated P1 infectious clone infections in *N. benthamiana*

It had earlier been observed that the LRRA change to LAAA mutation in the P1 of UCBSV affected the RNA silencing function of this protein. It was imperative to study if the same mutation affected the overall functionality of the full-length UCBSV infectious clone. Therefore, an experiment was set up to study the effect of this mutation in the infectious clone on virus titres in the indicator plants *N. benthamiana*.

In this experiment, the original wildtype UCBSV infectious clone was inoculated in-parallel to the P1 mutated UCBSV infectious clone by setting up *in-vitro* transcription reactions using the different plasmids as templates in the reactions.

The generated *in-vitro* run-off transcripts were then used to infect *N. benthamiana* plants. The infected plants were monitored for the development of symptoms over a 6 week period. It was

observed that there were mild mosaic symptoms observed at two weeks post-infection in the plants infected with the wildtype UCBSV infectious clone. Surprisingly, the mutated infectious clone infected plants also showed the same mild symptoms suggesting the mutation did not alter virulence. ELISA was performed using UCBSV specific antibodies to compare the level of the titres of UCBSV virus in the wildtype infected plants as compared to the mutated infectious clone infected plants. The ELISA data was analysed using students T-test in GraphPad Prism 7 Software. It was shown that there was a gradual increase in the viral protein titres in both the wildtype UCBSV infectious clone infected plants and mutated UCBSV infectious clone plants with time although there was a statistical difference ($p < 0.001$) between the viral protein titres for the wildtype infectious clone infected plants and mutated infectious clone plants with the latter having higher viral titres (Figure 5.3). The experiment was repeated three times and the results obtained were reproducible. A similar experiment was also performed by infecting *N. clelandii* and *N. benthamiana* and the mutated virus could be detected in the infected plants by RT-PCR.

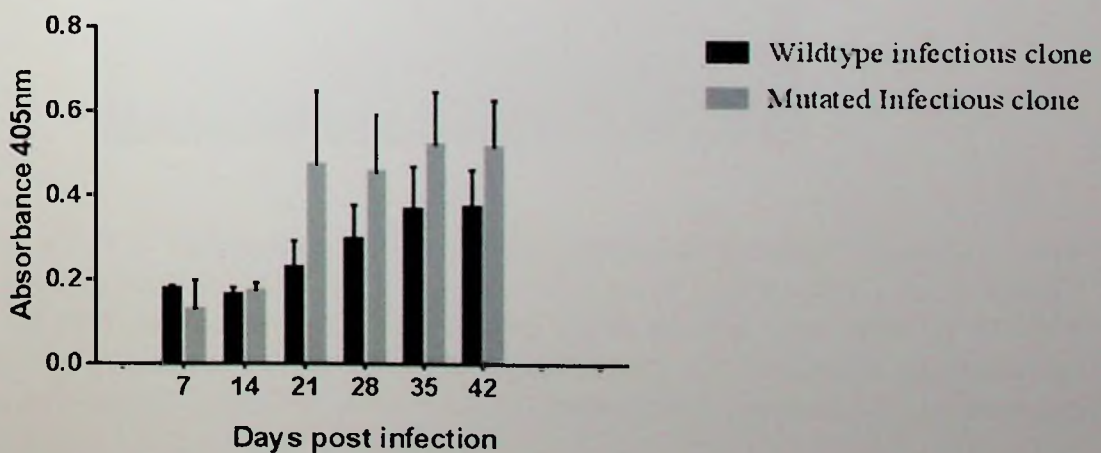


Figure 5.3: ELISA analysis of viral titres between the plants infected with LAAA P1 mutated UCBSV infectious clone and wild type infectious clone showing the levels of UCBSV viral protein titres with time.

5.2.4.1 RT-PCR to confirm UCBSV infection in the plants

RNA was extracted from systemic leaves of *N. benthamiana* plants infected with the mutated UCBSV infectious clone and the plants infected with the wildtype UCBSV infectious clone. The region spanning the P1 protein was amplified for cloning and sequencing. The region was amplified using primers P1 UCBSV Forward and P1 UCBSV Reverse that amplify a 1000bp PCR product encompassing the P1 protein and the expected product was obtained (Figure 5.4). The PCR sections were cloned and sequenced. The sequencing results confirmed that the mutated virus was replicating with the mutation (Figure 5.5).



Figure 5.4: RT-PCR amplification of P1 protein from plants infected with P1 mutated UCBSV I.C and wildtype UCBSV I.C. Lane 1 and 7 contain 5 μ l of HyperLadder 1, Lane 2 and 3 1000bp RT-PCR product from *N. benthamiana* plants infected with wildtype UCBSV infectious clone. Lane 4 and 5 1000bp RT-PCR product from *N. benthamiana* plants infected with mutated UCBSV infectious clone. Lane 6: Healthy mock inoculated *N. benthamiana* plant.

LAAA mutation



Figure 5.5: Alignment of the sequences from RT-PCR of plants infected with wildtype P1 UCBSV infectious against those infected with the LAAA mutated P1 UCBSV infectious clone. The mutated UCBSV infectious clone still has the mutation.

5.2.5 Experiments with the UCBSV P1 complete knockout infectious clone

5.2.5.1 Construction of P1 complete knockout infectious clone

A complete P1 UCBSV infectious clone knockout plasmid was constructed by yeast recombination. The entire coding sequence of the P1 protein was deleted from the genome and a new UCBSV infectious clone plasmid constructed with the 5'UTR being followed immediately by the P3 protein (Figure 5.6). The P1 knockout primers (Table 5.3) were designed.

Table 5.3: UCSBV P1 knockout primers

Primer name	Primer sequence	Position in UCBSV genome
UCBSV P1 knockout F	5'TTATATATTCATAGCACTGAGTACCTGCAGATGTCTTTC AATCTTAATTGGGCGAAATCGAAA3'	104-167
UCBSV P1 knockout R	5'TTTCGATTTGCGCCCAATTAAGATTGAAAGACATCTGCAGGTA CTCAGTGCTATGAATATATAA3'.	104-167

The P1 knockout primers were complementary to each other and encompass the end of the 5'UTR and the beginning of the P3 protein with an ATG added before the beginning of the P3 protein. The primers also contain a 30 bp overlap on either ends to allow for recombination in

yeast. An ATG start codon was added to allow for translation in the host, the ATG sequence was incorporated in the primer that was designed to generate the complete P1 knockout plasmid.

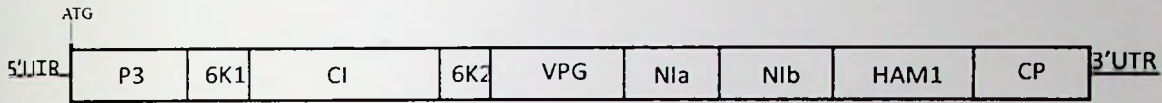


Figure 5.6: Schematic representation of genome of UCBSV P1 knockout virus: The complete P1 coding sequence was deleted from the genome.

PCR sections were generated for use in yeast recombination to construct a new UCBSV infectious clone with the P1 deletion and the expected sizes were amplified from the original UCBSV infectious clone. These different sections were recombined in pYES2.1 plasmid by homologous recombination in yeast. As a positive control, yeast cells were transformed with undigested pYES2.1 backbone vector. Negative controls consisted of yeast transformation with *EcoR1/Sac1* digested pYES2.1 vector and water. Yeast cells transformed with undigested pYES2.1 generated about 500 colonies confirming the viability of the yeast cells for transformation. The yeast cells transformed with *EcoR1/Sac1* digested pYES 2.1 generated about 20 colonies and there were no colonies on water transformed control plates. The experimental plates transformed with different UCBSV sections generated about 200 colonies confirming successful recombination. Ten yeast colonies were individually selected and grown in SM media. Yeast plasmid was extracted and then rescued into *E. coli* cells. Plasmid was then extracted from *E. coli* cells and analysed by restriction analysis and sequencing for their integrity. Confirmation of successful deletion of the P1 protein from the infectious clone was through PCR comparing the product sizes for section 1 (2700bp) in the P1 containing infectious

clone as compared to a 1700bp PCR product obtained after deletion of the P1 protein from the infectious clone (Figure 5.7). A reduction in the product size by about 1000bp was a confirmation that the entire P1 protein which is about 1000bp was deleted from the P1 knockout infectious clone.

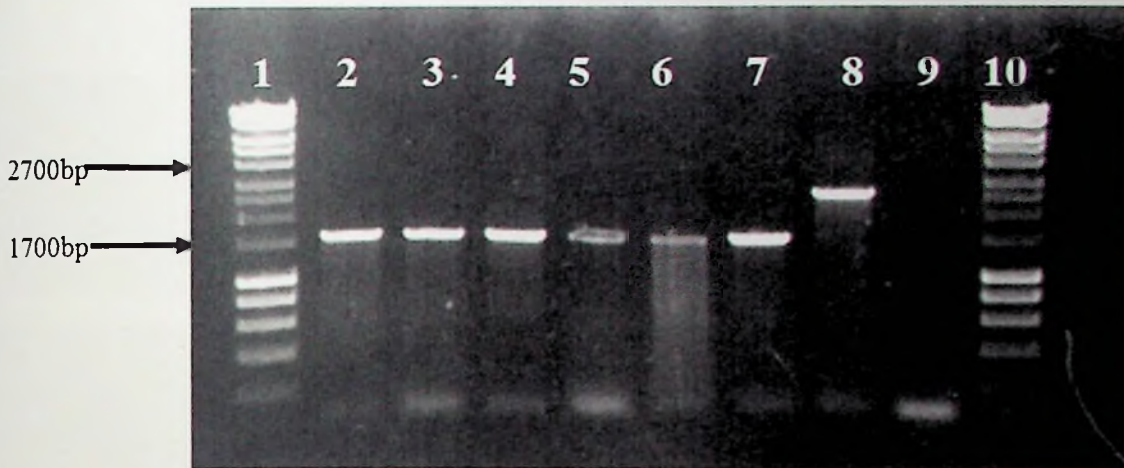


Figure 5.7: PCR confirmation of deletion of P1 protein in the UCBSV infectious clone: Lane 1 and 10 contains 5 μ l Hyperladder 1: Lane 2, 3, 4, 5, 6 and 7: 1700bp PCR products generated from different screened plasmids recovered from UCBSV P1 knockout plasmid extractions. Lane 8: 2700 bp PCR product obtained from the P1 containing UCBSV infectious clone. Lane 9 negative control with no amplification.

5.2.5.2 Infection of P1 Knockout infectious clone plasmid in *N. benthamiana*, *N. tabacum* and *N. clevelandii*.

The generated P1 knockout infectious clone was used in *in-vitro* transcriptions. As a positive control, the full-length UCBSV infectious clone was infected alongside the P1 knockout infectious clone. Plants were infected with RNA transcripts generated with a cap analog. The plants were kept at 28°C and monitored for symptom development over a 4 week period. The plants for both the wildtype UCBSV infectious clone and P1 complete knockout infectious clone showed mild mosaic symptoms in *N. benthamiana* plants as has been observed for UCBSV infections. There were no observable symptoms in *N. clevelandii* and *N. tabacum* plants infected with the P1 knockout as well as the wildtype UCBSV infectious clone infected

plants (Figure 5.8). RNA was extracted from the plants for cDNA synthesis and PCR analysis. A very faint PCR product was obtained on amplifying a 867bp fragment from the infected plants with the UCBSV P1 knockout plasmid (Figure 5.9). This experiment was repeated two times and the results obtained were reproducible.

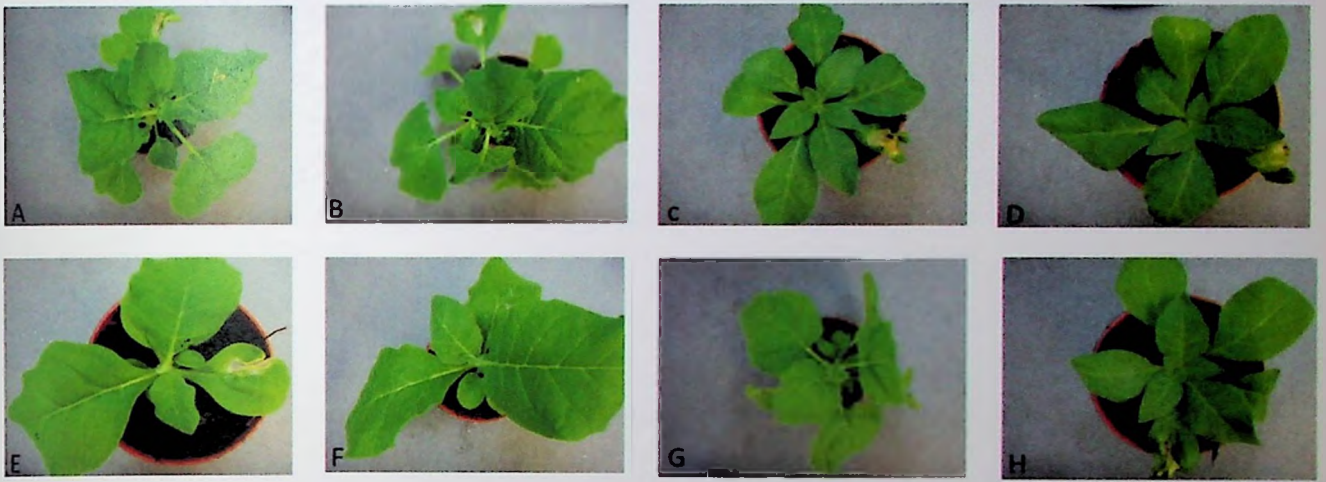


Figure 5.8: Symptoms in different indicator plants infected with UCBSV P1 knockout I.C: A: *N. benthamiana* plant infected with P1 UCBSV Knockout infectious clone B: *N. benthamiana* plant infected with full-length UCBSV infectious clone, C: *N. clevelandii* plant infected with P1 UCBSV knockout infectious clone D: *N. clevelandii* plant infected with full-length UCBSV infectious clone. E: *N. tabacum* plant infected with P1 knockout infectious clone. F: *N. tabacum* plant infected with full-length UCBSV infectious clone. G: Healthy mock inoculated *N. benthamiana* plant, H, Healthy *N. clevelandii* plant mock inoculated with buffer.

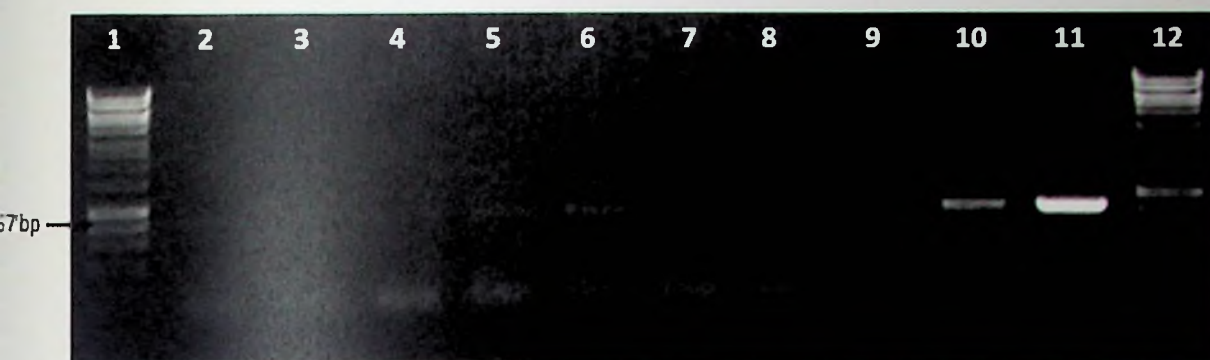


Figure 5.9: Amplification of PCR products from P1 UCBSV knockout I.C infected plants: Lane 1 and 12 contain HyperLadder 1, Lane 2, 3, 4 and 5 contain PCR products from *N. benthamiana* plants infected with P1 knockout plasmid. Lane 6 and 7 contain PCR products from *N. clevelandii* infected plants infected with P1 knockout plasmid. Lane 8 and 9 contain PCR products from *N. tabacum* plants infected with P1 knockout plasmid. Lane 10 contains the PCR from *N. benthamiana* plant infected with wildtype UCBSV infectious clone. Lane 11 is the positive control plasmid.

5.2.6 Construction of P1 agroinfiltration vectors for RNA silencing suppression studies

Different vectors were constructed in the experiment to study the effect of the LRRA mutation change to LAAA on the RNA silencing functions of P1 (Figure 5.10 B). Proteins were cloned into pCAMBIA2300 plasmid as backbone vector. In order to allow for the expression of the proteins in the backbone vector, a *Cauliflower mosaic virus* (CaMV) 35S promoter was introduced upstream of each protein coding sequence and expression was terminated by the agrobacterium TNos terminator (Figure 5.10A). The 35S promoter was chosen as it is a strong constitutive promoter in most tissues of plants.

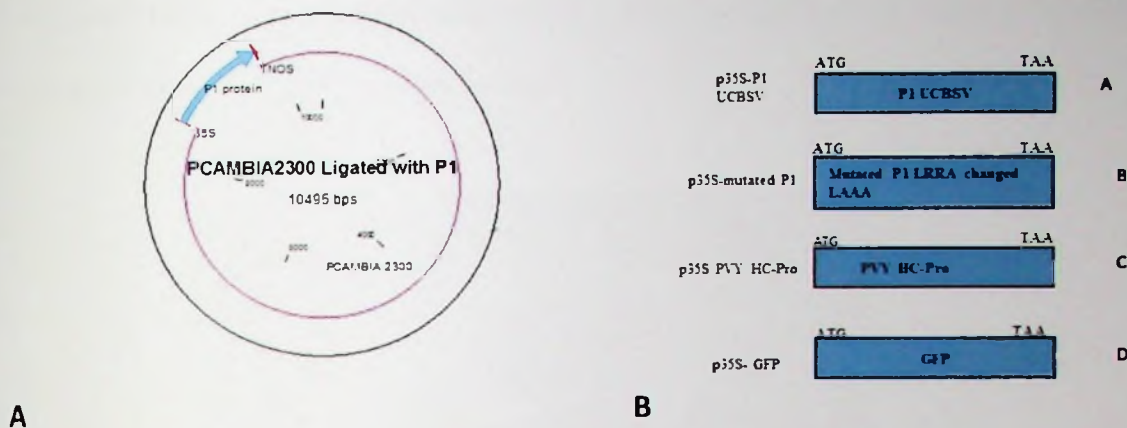


Figure 5.10: (A) Constructed recombinant pCambia2300 with P1 protein for agroinfiltration assays. (B) Schematic representation of the different constructs used in the agroinfiltration assays: A: pCambia2300 cloned with the wildtype P1 UCBSV, B: pCambia2300 cloned with LAAA mutated P1 UCBSV, C: pCambia2300 cloned with positive control PVY HC-Pro, D: pCambia2300 cloned with GFP for co-infiltration.

The LAAA mutated P1 protein was amplified from the earlier constructed UCBSV infectious clone template containing the mutated P1 protein sequence using primers that incorporated a *Bam*HI restriction site at the 5' end and a *Kpn*I restriction site at the 3' end. The enzyme restriction sites were added to allow for cloning into the pCambia2300 vector. The PCR products were amplified from different templates i.e original UCBSV infectious clone for the wildtype P1 protein and the LAAA mutated P1 infectious clone for the mutated P1 protein. As a positive control, the P1 protein of UCBSV that had previously been shown to be involved in RNA silencing was constructed alongside the mutated P1 (Luis Pablo Rodriguez, unpublished). The PCR products were cloned into pJET cloning vector and then restrict digested from pJET using *Bam*HI and *Kpn*I restriction enzymes and then ligated into pCambia2300 vector using T4 ligase enzyme. The recombinant plasmids were transformed into *E. coli* cells and the plasmids extracted from two different colonies were screened by restriction analysis using *Bam*HI and *Kpn*I (Figure 5.11) and sequencing to confirm their integrity before using the

recombinant vector in the agroinfiltration assays to study the effect of the mutation on the RNA silencing function of the P1 protein in UCBSV.

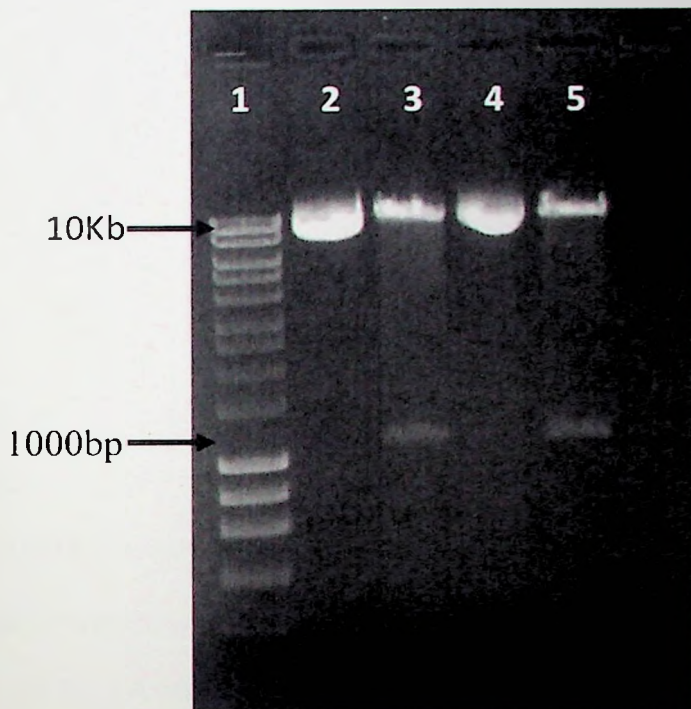


Figure 5.11: Restriction analysis of recombinant colonies digested with *Bam*H1 and *Kpn*1; Lane 1 contains 5 μ l of HyperLadder1. Lane 2 and 4 contain undigested recombinant pCambia 2300 cloned with LAAA mutated P1. Lane 3 and 5 contain *Bam*H1/*Kpn*1 restrict digested recombinant pCambia 2300 cloned with LAAA mutated P1, a 1000bp product was released on restriction digestion of recombinant pCambia2300 with *Bam*H1 and *Kpn*1 restriction enzymes.

5.2.6.1 Agroinfiltration assays to assess RNA silencing suppression

A. tumefaciens strain LB4404 was transformed with the different constructs (Figure 5.10B) by electroporation. A single colony was grown in 5 ml LB containing acetosyringone to activate T-DNA transfer for agroinfiltration. *Agrobacterium* cultures were mixed (p35S-GFP with the respective vector containing different proteins) after induction with acetosyringone and approximately 500 μ l of the mixture (optical density of 0.5 at 600nm) was used in the agroinfiltration. Two week old *N. benthamiana* 16C line which constitutively expresses GFP

protein (Gammelgard *et al.*, 2007., Valli *et al.*, 2006., Mbanzibwa *et al.*, 2009a) were infiltrated with *A. tumefaciens* strain LB4404 containing different constructs. HC-Pro from PVY was used as a positive control and the p35S-GFP construct was co-infiltrated along with each of the constructs as a silencing reporter. The infiltration was done on the underside of the plants using a syringe. Wildtype *N. benthamiana* plants were used as a negative control. In addition, the P1 protein of CBSV Nampula isolate which had previously been shown to lack RNA silencing suppression (Luis Pablo Rodriguez, unpublished) was added as an internal negative control for the experiment.

The results obtained showed that the leaves infiltrated with the construct carrying the P1 of UCBSV showed bright GFP fluorescence at 5 days post the agroinfiltration. This shows that the silencing of GFP has been impaired by the expression of P1, confirming that the unmodified UCBSV P1 protein is an effective suppressor of silencing. However, it was observed that leaves infiltrated with the LAAA modified P1 did not give clear GFP fluorescence, showing RNA silencing suppression was not occurring (Figure 5.12). This observation confirms that the mutation in the basic motif affected the functionality of the P1 protein. The experiment was repeated three times and similar observations were obtained thus confirming reproducibility of the results.

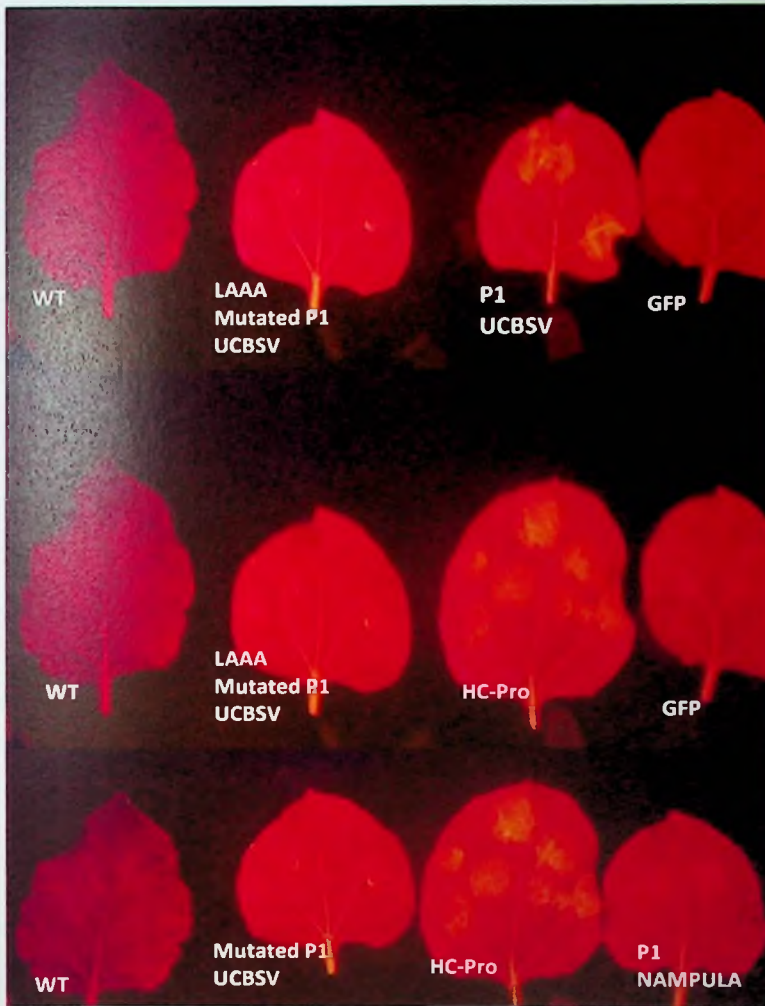


Figure 5.12: Silencing suppression of wildtype UCBSV P1 and LAAA mutated P1: WT: Wildtype *N. benthamiana* plants. LAAA mutated P1 UCBSV; No GFP green fluorescence observed with the mutated P1 infiltrated leaves, P1 UCBSV; GFP green fluorescence observed in the wildtype UCBSV P1 protein infiltrated leaves, HC-Pro; Positive control, GFP; Negative control, P1 Nampula; No green fluorescence from P1 of CBSV Nampula isolate (internal negative control).

5.2.7 HAM1 UCBSV knockout infectious clone experiments

The HAM1 gene is a novel protein that has been identified recombined in the genomes of UCBSV and CBSV isolates but is absent in the genomes of other *Ipomoviruses*. The role of HAM1 in CBSVs has not been elucidated and hence reverse genetics using UCBSV HAM1 knockout infectious clones was used to study the role of HAM1.

The UCBSV HAM1 knockout infectious clones were provided by Bunawan *et al.*, (unpublished). Two forms of the knockouts were constructed with one construct containing a full HAM1 deletion (Figure 5.13 B) and the other containing a truncated HAM1 deletion (Figure 5.13C).

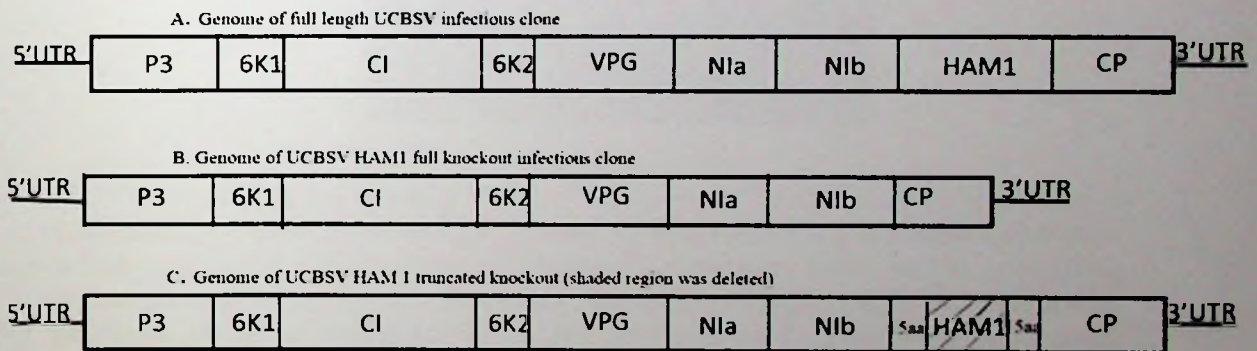


Figure 5.13: Schematic presentation of the genomes of the different UCBSV HAM1 knock out constructs. A: Genome of full-length UCBSV infectious clone, B: Genome of HAM1 full deletion UCBSV infectious clone. C: Genome of HAM1 partial deletion UCBSV infectious clone.

5.2.7.1 Verification of the integrity of the HAM1 constructs

The different constructs were verified by PCR and sequencing to confirm that they were in the correct reading frame without errors. The constructs were checked with primers 3'1.4KBFS 5' and 3' 1.4KB RS (Table 4.1) that amplify a 1400bp PCR product with a HAM1 containing UCBSV infectious clone and a 700bp product with a HAM1 knockout infectious clone (the

primers target section 7). PCR confirmed that the constructs were accurate as the expected product sizes obtained were representative of the genomes (Figure 5.14).

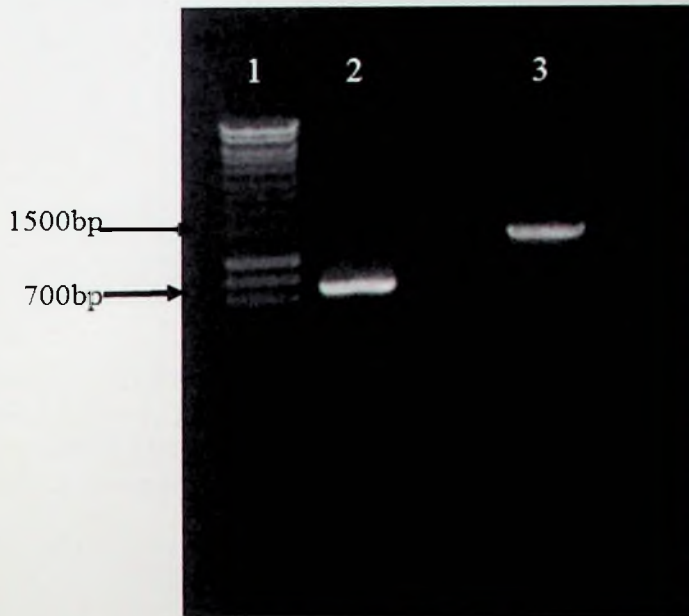


Figure 5.14: PCR to confirm the integrity of the UCBSV HAM1 knockout constructs: Lane 1: 5 μ l of HyperLadder 1. Lane 2: 700bp PCR product from the HAM1 full deletion construct confirms actual deletion of HAM1. Lane 3: 1500bp PCR product from HAM1 containing UCBSV infectious clone. A reduction in the product size from 1500bp in the HAM1 containing UCBSV infectious clone to 700bp in the HAM1 full-deletion UCBSV infectious clone confirms deletion of HAM1.

5.2.7.2 *In-vitro* transcription with the HAM1 knockout constructs

The plasmids recovered from the cells were around 10 kb as was expected of the plasmids containing the genomes (Figure 5.15). The plasmids were linearized by digestion with *Kpn*I restriction enzyme to allow for the production of run-off transcripts. *In-vitro* transcription was performed using the SP6 Riboprobe *in-vitro* transcription according to the manufacturer's instructions. The generated transcripts were diluted in infection buffer (40 μ l of *in-vitro*

reaction diluted with 200 μ l of infection buffer) and then used to infect the *N. benthamiana* plants.

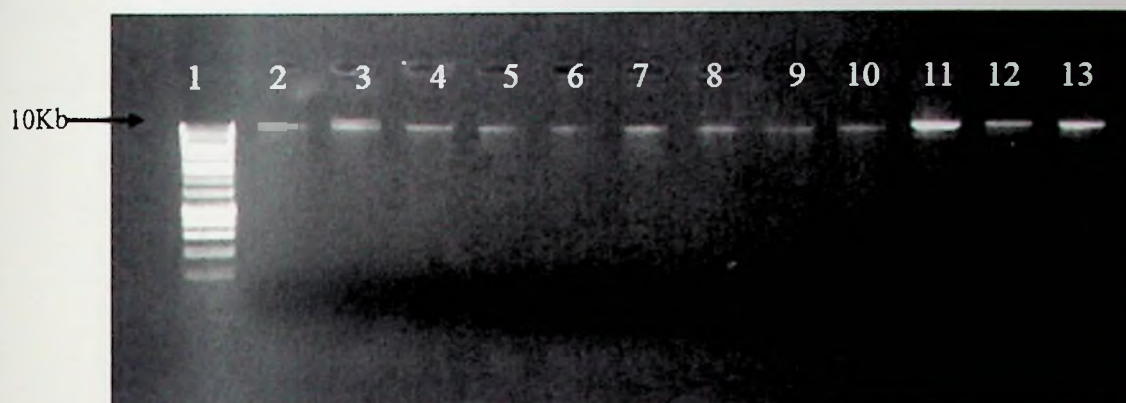


Figure 5.15: Plasmids recovered from *E. coli* cells for the different infectious clone constructs: Lane 1: 5 μ l of HyperLadder 1, Lanes 2, 3, 4 and 5 UCBSV HAM1 full deletion plasmids. Lanes 6, 7, 8 and 9 UCBSV HAM1 partial deletion plasmid. Lanes 10, 11, 12 and 13 full-length UCBSV infectious clone plasmid.

5.2.7.3 Infection of *N. benthamiana*, *N. clevelandii* and *N. tabacum* with UCBSV HAM1 knockout constructs

Plants were infected with the RNA transcripts generated from the *in-vitro* transcriptions with the UCBSV HAM1 knockout transcripts. The full-length UCBSV infectious clone was included as a positive control in the infection experiment. The plants were monitored for the development of any symptoms over a four week period. It was observed that the HAM1 knockout infected plants presented yellow chlorosis in *N. benthamiana* plants at two weeks post infection (Figure 5.16 A and B). Similar chlorosis was not observed in the *N. benthamiana* plants that were infected with the full-length UCBSV infectious clone although mild mosaics were observed (Figure 5.16C). However, the *N. clevelandii* and *N. tabacum* plants did not show any symptoms for plants infected with the HAM1 knockout and UCBSV full length infectious clones. RNA was extracted for RT-PCR analysis and the results showed that UCBSV infection

could be detected using primers that amplify a 1000bp section in UCBSV (section 5) in plants infected with the full-length UCBSV infectious clone but not in the plants infected with the HAM1 UCSBV knockout clones (Figure 5.17). The experiments were repeated two times and the results were reproducible.

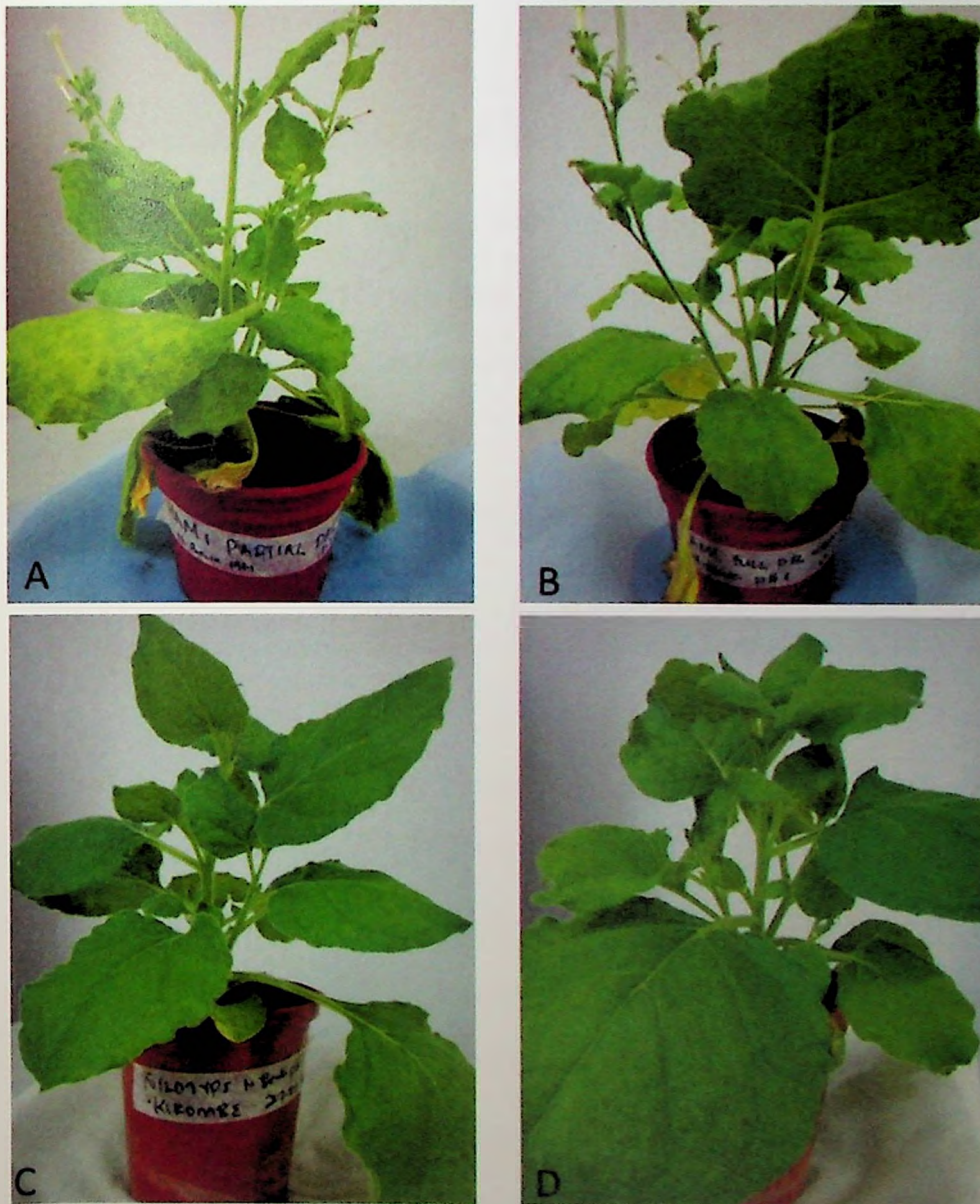


Figure 5.16: Symptoms on plants infected with UCBSV HAM1 knockout infectious clones: A: HAM1 partial deletion *N. benthamiana* infected plant. B: HAM1 full deletion *N. benthamiana* infected plant. C: UCBSV full-length infectious clone infected *N. benthamiana* plant. D: Healthy mock inoculated *N. benthamiana*.

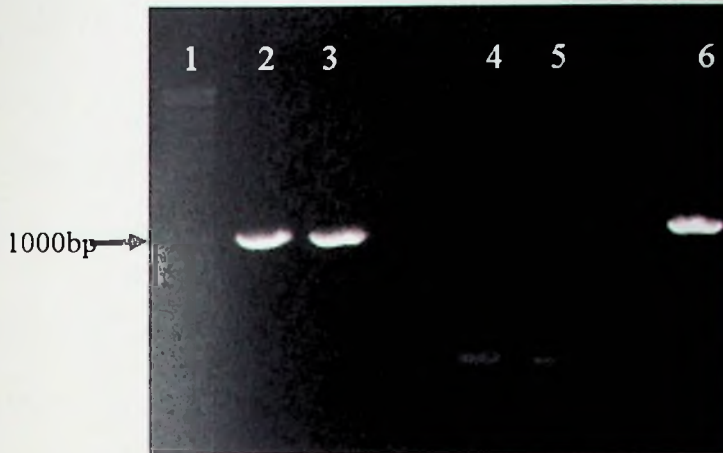


Figure 5.17: RT-PCR on HAM1 knockout infected plants. Lane 1: Contains 5 μ l of HyperLadder 1. Lane 2 and 3 1000bp RT-PCR product from UCBSV full length infected *N. benthamiana* plants. Lane 4 is *N. benthamiana* plant infected with HAM1 full deletion UCBSV infectious clone. Lane 5 *N. benthamiana* plant infected with HAM1 partial deletion UCBSV infectious clone. Lane 6 is the positive control plasmid.

5.2.7.4 Infecting the HAM1 deletions in transgenic HAM1 *N. tabacum* plants to study complementation

Transgenic *N. tabacum* lines expressing the full-length HAM1 protein from CBSV Nampula were provided by Luis Pablo- Rodriguez (unpublished). The transgenic lines had been generated by *Agrobacterium tumefaciens* mediated transformation of *N. tabacum* with the pCAMBIA2300 vector expressing the full HAM1 sequence from CBSV Nampula under the expression of the CaMV 35S promoter. The different expression levels of the HAM1 protein was confirmed in the transgenic lines by quantitative PCR (q PCR). Furthermore, the transgenic lines were screened by PCR and the presence of the HAM1 gene was confirmed. An initial sequence comparison between the HAM1 sequence from UCBSV Kikombe and CBSV Nampula showed that the HAM1 in both isolates was 678bp in length with a nucleotide identity of 55% and amino acid identity of 45% (Figure 5.18). For most trans-complementation studies, it is normally considered important that the expressed transgene is identical to the mutated viral protein (Li & Carrington, 1995). Despite the low percentages in the identities of the

HAM1 from UCBSV Kikombe and from CBSV Nampula, it was imperative to find the possibility of complementation between the transgenic plants expressing CBSV Nampula HAM1 with the UCBSV HAM1 knockout infectious clones. Furthermore, it was hypothesized that the HAM1 gene was acquired by CBSV and UCBSV to serve the same function.

The UCBSV HAM1 full deletion and HAM1 half deletion infectious clone constructs were used in *in-vitro* transcription to generate transcripts that were used to infect CBSV HAM1 transgenic lines. As a control, wildtype *N. tabacum* plants were infected alongside the transgenic lines. The full-length UCBSV infectious clone was also infected alongside the HAM1 knockout infectious clones. The plants were kept in a growth chamber at 28°C at 14 h light and 8 h dark pattern. Plants were monitored for the development of symptoms, there were no observable symptoms in the transgenic *N. tabacum* lines and wildtype *N. tabacum* plants (Figure 5.19). This is consistent with the earlier observations of UCBSV infections in *N. tabacum* plants.

HAM1 CBSV Na	1	vv	rpqslnvak	e	evtskfr	giea	i	vt	naq	k	q	f	ptipivsrkv	dp		
HAM1 UCBSV K	1	tk	lrgrekpel	i	shdgvpg	qmkf	v	lg	a	s	l	iasdviakni	dp			
HAM1 CBSV Na	181	s	ie	ike	rv	aelv	gg	iv	ea	s	dd	lg	ip	pykwlegi	e	lykl
HAM1 UCBSV K	181	v	pv	vrk	ql	vkmt	ns	lv	ea	s	dd	lg	ip	pykwlegi	e	lykl
HAM1 CBSV Na	361	v	epyqnm	s	l	v	f	nkvgd	i	i	lk	m	pss	q	lnwkk	t
HAM1 UCBSV K	361	l	safgdk	s	y	t	y	hnel	s	v	v	n	g	g	kdgc	d
HAM1 CBSV Na	541	m	tee	m	ish	fr	s	l	rdf	k	sssyfsf	a	gl	dr	dif	
HAM1 UCBSV K	541	s	sn	i	df	r	k	ek	kly	dn	lvk	qee	k	ak	valt	

Figure 5.18: Comparison of amino acid sequence between UCBSV Kikombe HAM1 and CBSV Nampula HAM1. The overall amino acid identity between the two isolates was 45%

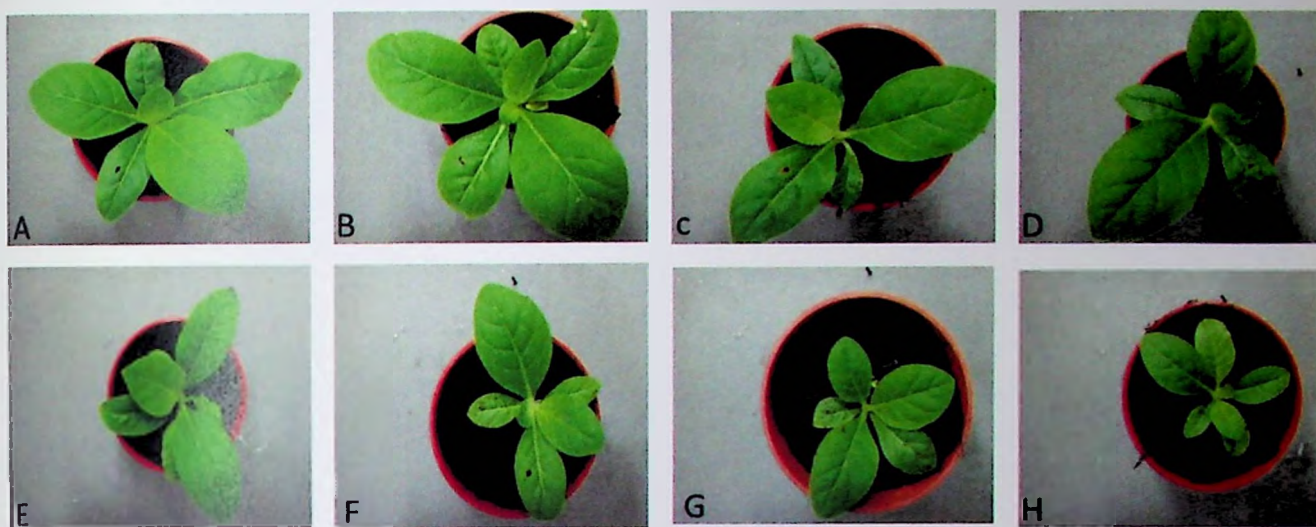


Figure 5.19: Transgenic CBSV HAM1 lines of *N. tabacum* infected with different constructs: A and B: HAM1 full deletion UCBSV infectious clone infected plants, C and D: HAM1 partial deletion UCBSV infectious clone infected plants, E and F: Full length UCBSV infectious clone infected plants. G and H: Mock infected plants

5.2.7.5 RT-PCR to detect UCBSV infection in Transgenic HAM1 plants

Despite the absence of observable symptoms in the HAM1 transgenic lines, RNA was extracted for cDNA synthesis and RT-PCR to detect UCBSV infection. The RNA was extracted from the systemic leaves at 4 weeks post infection. The PCR results showed that there was complementation between one of the transgenic lines expressing CBSV HAM1 protein with the UCBSV full deletion knockout infectious clone as was evidenced by the amplification of a very faint 867bp PCR product in the infected transgenic line (Figure 5.20). An earlier experiment in which the HAM1 mutant infectious clone was used in infecting non-transgenic plants showed that the virus could not be detected in the systemic leaves and thus complementation could be the possible explanations for this result. There was no amplification in the HAM1 partial deletion UCBSV infectious clone infected plants meaning that the partial deletion construct was not complemented. One of the HAM1 transgenic lines infected with the UCBSV full-length infectious clone showed amplification of the UCBSV virus (Figure 5.20).

Another HAM1 transgenic line infected with the UCBSV full-length infectious clone showed no amplification of the virus in the systemic leaves meaning that transgenic HAM1 expression cross-protected this line from UCBSV infection. However, the difference in the expression levels of HAM1 in the transgenic lines may be a possible explanation for the obtained results.



Figure 5.20: RT-PCR to detect UCBSV infection in transgenic HAM1 lines: Lane 1 and 10: HyperLadder 1. Lane 2: 867bp PCR product from transgenic line CBSV HAM1 *N. tabacum* infected with UCBSV HAM1 full deletion infectious clone. Lane 3: Transgenic line CBSV HAM1 infected with UCBSV HAM1 full deletion infectious clone. Lane 4 and 5: HAM1 transgenic lines infected with UCBSV HAM1 partial deletion infectious clone. Lane 6 and 7: HAM1 transgenic lines infected with full length UCBSV infectious clone. Lane 8: *N. tabacum* non transgenic plant infected with full-length UCBSV infectious clone. Lane 10: Positive control plasmid.

5.3 Discussion

Characterization of protein functions in viral genomes is very important for development of new resistance strategies (Frédéric Revers & García, 2015). Most potyviral genomes encode ten proteins that serve different roles in the viral infection processes including replication and translation to ensure viral spread and establishment with-in the host (Kekarainen *et al.*, 2002; Revers & García, 2015). Functions for most proteins are usually inferred from comparison between sequences of other characterized proteins in other potyviruses or on the basis of presence of functional motifs in protein sequences.

For most of the positive stranded RNA viruses, it is not possible to manipulate the RNA as most biotechnological tools are geared towards manipulating DNA genomes thus the RNA of positive strand RNA viruses needs to be converted into cDNA for ease of manipulation. Hence, to study the roles of different viral proteins in RNA viruses, it is important to generate infectious clones to allow deletion and mutational analysis to be performed and hence characterize the functions. In the context of the work described in this chapter, infectious clones of UCBSV Kikombe were used to functionally characterize the roles of P1 and HAM1 proteins in UCBSV genome.

Because the genomes of UCBSV and CBSV lack the multifunctional HC-Pro protein that is known to be involved in RNA silencing suppression (Soitamo *et al.*, 2011; Mbanzibwa *et al.*, 2009a), it has been proven that the P1 protein has evolved to serve the role of suppression of RNA silencing (Mbanzibwa *et al.*, 2009a; Winter *et al.*, 2010). This function has been characterized in the P1b protein of the Ipomovirus *Cucumber vein yellowing virus* (CVYV) (Valli *et al.*, 2008). A functional motif LXKA (Dombrovsky *et al.*, 2014) present in the P1 protein has been reported to be involved in the suppression of RNA silencing and it was targeted in this work. In UCBSV genomes, the LXKA motif is present as an LRRRA motif and to characterize its role in suppression of RNA silencing, this motif was mutated to LAAA and the effect of the mutation studied in agroinfiltration assays. The *Agrobacterium tumefaciens* assay

is a commonly used procedure for the identification of suppressors of RNA silencing in viruses and was applied in this study to determine the role of UCBSV P1. It was shown that a mutation in the functional LRRA motif of the UCBSV protein led to loss in the RNA silencing suppression properties of UCBSV-P1. A possible explanation for this observation may be that the change in the amino acid residues from the basic residues to the aliphatic amino acid affects the interaction of the functional motif with the proteins involved in RNA silencing like the argonaute protein and may thus further lead to loss of function. A similar mutation in the P1 of the full-length UCBSV infectious clone in comparison with the wildtype UCBSV infectious clone showed that there were higher titres of virus in the plants infected with the P1 mutated infectious clone as compared to the wildtype infectious clone. This result was an unexpected result which is to be repeated to confirm the reproducibility of this finding although a possible explanation for this finding may be that the LRRA to LAAA mutation could have interfered with the antiviral system in the plant thus permitting the virus to accumulate to high levels (Anandalakshmi *et al.*, 1998). Mutations in functional motifs of viral proteins has been shown to have an effect on viral replication and was reported by Deng *et al.*, (2015) where a motif change in the GDD motif of the NIb protein which acts as an RNA dependent RNA polymerase affected the replication of *Turnip mosaic virus* (TuMV). An increase in viral titres have been reported in the infectious clone *Tobacco etch virus* (TEV) mutated in the HC-Pro (Torres-Barcel *et al.*, 2008; Torres-Barcelo *et al.*, 2010) although this study also reported that there is no a perfect match between suppression and virulence.

The P1 full-deletion UCBSV infectious clone was able to move systemically in the *N. benthamiana* plants as was evidenced by the detection of the UCBSV virus in systemic leaves by RT-PCR although the intensity of the PCR product was faint which may be indicative of low viral titres in the plants. Similar findings have been observed with P1 knockout clones of PPV and TEV (Pasin *et al.*, 2014; Verchot & Carrington, 1995). The lack of the P1 sequence did not

compromise the ability of the P1 knockout infectious clone to move systemically although it affected the viral levels suggesting that the P1 protein in UCBSV genome may be involved in genome amplification.

When transgenic *N. tabacum* plants were infected with the UCBSV HAM1 knockout infectious clones, and it was realised that the UCBSV virus could be detected in one of the transgenic lines expressing CBSV HAM1, although the sequence identities between the HAM1 of UCBSV and HAM1 CBSV were as low as 45% amino acid identity. Furthermore, there was a possibility of cross protection observed between one of the transgenic HAM1 lines evidenced by the absence of UCBSV infection in transgenic one of HAM1 lines infected with full-length UCBSV infectious clone, similar observations were reported by Vanderschuren *et al.*, (2012) using hairpin constructs to the coat protein of UCBSV and CBSV demonstrated disease resistance in these lines. Similar observations have been reported for transiently transformed plants with RNAi constructs targeting different regions of the CP of UCBSV (Patil *et al.*, 2011). Transgenic protection has also been observed in transgenic plants that express *Tobacco mosaic virus* (TMV) coat protein (Abel *et al.*, 1986; Lindbo *et al.*, 1993). The suggested model for transgenic protection is hypothesised to be related to the elevated RNA levels which are targeted, suppressed and eventually eliminated from the cell (Lindbo *et al.*, 1993). Furthermore, the expression of transgenes in plants is known to alter the expression of a number of genes among which are the stress related genes which may affect the outcome of viral infection (Soitamo *et al.*, 2011).

Transgenic cross-protection using viral genes also referred to as pathogen-derived resistance (PDR) is hypothesised to be important in controlling viral diseases by preventing viral infections (Goldbach *et al.*, 2003; Sanford & Johnston, 1985) although caution has to be considered in situations where the presence of the viral transgene may act synergistically to cause more disease

infection in related or unrelated viruses (Abel *et al.*, 1986; Pruss *et al.*, 1997; Fondong *et al.*, 2000; Haas *et al.*, 2002). The HAM1 gene of UCBSV can thus be considered as a target for pathogen-derived resistance for development of resistance to CBSD. Similar observations have been reported for transiently transformed plants with RNAi constructs targeting different regions of the CP of UCBSV (Patil *et al.*, 2011). Transgenic protection has also been observed in transgenic plants that express *Tobacco mosaic virus* (TMV) coat protein (Abel *et al.*, 1986; Lindbo *et al.*, 1993). The suggested model for transgenic protection is hypothesised to be related to the elevated RNA levels which are targeted, suppressed and eventually eliminated from the cell (Lindbo *et al.*, 1993). Furthermore, the expression of transgenes in plants is known to alter the expression of a number of genes among which are the stress related genes which may affect the outcome of viral infection (Soitamo *et al.*, 2011).

Chapter 6: General Discussion

6.1 General Discussion

Cassava is one of the most important staple food to most of the African populations with over 800 million people depending on cassava as the main source of starch (Hillocks, 2002; Nassar & Ortiz, 2007; Jarvis *et al.*, 2012). In addition, cassava can be used for other purposes which include industrial production of starch, bio-fuel and animal feed. However, cassava production is faced by a number of challenges among which are pests and diseases. Of the viral diseases affecting cassava production, cassava mosaic disease (CMD) and cassava brown streak disease (CBSD) are considered as the diseases of economic importance due to their effects on yields (Thresh *et al.*, 2009; Vanderschuren *et al.*, 2012; Kizito *et al.*, 2005).

Cassava brown streak disease has been reported to impact on the productivity of cassava by reducing the overall yield of the crop by over 70% (Patil *et al.*, 2015; Hillocks *et al.*, 2001; Campo *et al.*, 2011). Cassava brown streak disease is caused by *Cassava brown streak virus* (CBSV) and *Ugandan cassava brown streak virus* (UCBSV). Different control methods for CBSD have been tried in order to control the impact of the disease on crop yield which include planting disease free stems, roguing of infected plants (McQuaid *et al.*, 2016) and breeding for disease resistant crops (Ogwok *et al.*, 2012; Taylor *et al.*, 2004; Wagaba *et al.*, 2016). Some other control strategies have relied on controlling the *B. tabaci* whitefly vector responsible for transmission of the virus (Legg *et al.*, 2014; Jeremiah *et al.*, 2015). Most of the breeding programs have focused on generating disease resistant plants (Ogwok *et al.*, 2012; Yadav *et al.*, 2011). The disease resistant cassava lines are obtained by conventional breeding or generating transgenic lines (Nassar & Ortiz, 2007; Taylor *et al.*, 2004). One of the challenges of testing breeding lines for resistance to CBSD has been the lack of efficient method to

transmit the viruses under laboratory conditions with the whitefly vector transmission being slow and unreliable (Maruthi *et al.*, 2005; Wagaba *et al.*, 2013). Grafting has also been used to transmit the virus (Mohammed *et al.*, 2012). In this case, one of the alternatives would be a need to construct infectious clones of CBSD viruses which provide a uniform and homogeneous source for testing breeding lines. Infectious clones further allow for understanding the pathogenesis because they can be genetically altered with precision to make desired changes which can guide in developing anti-viral strategies to control disease (Boyer and Haenni, 1994; Youssef *et al.*, 2011). With the availability of infectious clones of UCBSV, a number of studies can be undertaken to study the pathogenicity of this virus.

6.1.1 Genome sequencing and characterization of UCBSV Kikombe isolate

The UCBSV Kikombe isolate used in this study was not characterised at sequence level and hence initial work involved sequencing the complete genome of the UCBSV Kikombe. A genome overlapping strategy was used and different overlapping regions (herein called sections) were amplified by RT-PCR and cloned into a pJET cloning vector for sequencing. Cassava leaves confirmed to be infected with UCBSV was used as the source of nucleic acid that was used as template in RT-PCR. The genome was successfully sequenced and assembled and found to consist of 9070 nucleotides without including the polyadenylated tract at the 3'terminus. The 5' UTR consisted of 134 nucleotides followed by a start codon, ATG. The 3' UTR consisted of 226 nucleotides. The genome translated into a polyprotein of 2902 amino acids, this was similar to the earlier assembled genomes of UCBSV isolates in the NCBI database (Mbanzibwa *et al.*, 2009a; Winter *et al.*, 2010). All the predicted NIa protease cleavage sites reported for potyviruses (Adams *et al.*, 2005) were present in the assembled UCBSV genome and ten proteins were annotated namely; P1, P3, 6K1, CI, 6K2, VPg, NIa-Pro,

Nlb, HAM1 and CP. Similar to the Ipomoviruses *Cucumber vein yellow virus* (CVYV) and *Squash vein yellowing virus* (SqVYV), the UCBSV Kikombe genome lacked the HC-Pro coding sequence (Janssen *et al.*, 2005; Li *et al.*, 2008). The HAM1 gene was found incorporated between the Nlb and CP coding regions, the HAM1 protein has also been reported to be present in the genomes of other UCBSV and CBSV isolates (Mbanzibwa *et al.*, 2009a; Winter *et al.*, 2010). Phylogenetic analysis of the assembled UCBSV isolate genome with other sequences of UCBSV and CBSV isolates in the NCBI database clustered the isolate with the known UCBSV isolates confirming the taxonomic status of the isolate. The UCBSV Kikombe genome has been submitted to the GenBank under the accession number KX753356.

An appropriate source of virus template is important in characterising a virus, in most cases, the virus particles can be purified from infected plant material or nucleic acid can be used as a template (Youssef *et al.*, 2011; Olsen & Johansen 2001). In this study, total nucleic acid was extracted from infected cassava leaves and used as a template. However, problems were encountered in trying to extract good quality RNA from cassava leaves using the commercial Plant RNA Mini Kit (Omega Bio-tek), despite several attempts, it seemed not possible to extract RNA with the kit. Alternative optimizations were done to overcome this problem which included using the modified CTAB buffer for the RNA extraction. The reasons why the RNA extractions could have failed with the commercial kit may be because the cassava leaves contain a lot of polysaccharides and polyphenols (Essers, 1995). Other reasons for failed RNA extractions could be due to insoluble RNA pellets and high degree of RNA degradation. The CTAB buffer used contains polyvinylpyrrolidone (PVP) which reduces on the amount of phenols in the RNA extraction. To reduce degradation of RNA extracted, the RNA samples were reconstituted in DEPC treated water. The yield and purity (A_{280}/A_{260}) of the extracted RNA was measured using a Nanodrop and the readings ranged from 1.8-2.0. Once good quality RNA was obtained, cDNA synthesis was performed for RT-PCR. The synthesized cDNA was

checked for the amplification of the house keeping cassava ribulose 1-5 biphosphate carboxylase gene *rbc* control primers *rbc1* 5'CTACTATGGTGGCTCCGTTTC3' and *rbc2* 5'CCGTTTCAGTCGGAGAAACTC3' which amplify a 619bp PCR product with cassava cDNA (Monger *et al.*, 2001).

The UCBSV Kikombe isolate was also confirmed to infect a number of tested herbaceous hosts which included *N. benthamiana*, *N. tabacum*, *N. glutinosa* and *N. clevelandii*. Earlier reports have confirmed UCBSV isolates to cause mild symptoms in *N. benthamiana* plants and similar observations were obtained for the UCBSV Kikombe isolate (Mohammed *et al.*, 2012). This is in contrast to the CBSV isolates which cause necrotic local lesions and eventual death in *N. benthamiana*. However, both UCBSV and CBSV isolates cause similar yellow mosaic symptoms in cassava (Winter *et al.*, 2010). A similar observation was obtained in this study with the UCBSV Kikombe isolate infectious clone which caused very mild mosaics in *N. benthamiana* plants but yellow mosaic symptoms were observed in cassava infections. With the wide occurrence of different strains and variants of cassava brown streak viruses, there is an urgent need to study and understand the pathogenicity determinants of the different viral strains, information about these variants may help to explain and understand the variation in the symptoms observed between UCBSV and CBSV viruses. In addition, the presence of two different forms of the virus possess serious challenge to breeding programs which need to generate improved germplasm with resistance to all the various CBSD causing viruses (Patil *et al.*, 2015).

6.1.2 UCBSV Kikombe isolate infectious clone construction

In this PhD project, a viable UCBSV Kikombe *in-vitro* infectious clone was constructed and tested in tobacco and cassava plants. Infectious clones allow for the generation of a cDNA

stage for RNA viruses which allow for further manipulation in order to study and understand viral pathogenesis (Dawson *et al.*, 1986; Nagyová & Šubr 2007; Yang *et al.*, 1998). The technique used in the construction of the infectious clones was homologous recombination in yeast. Homologous recombination in yeast is a gap repair mechanism that joins overlapping DNA sequences that contain homologous regions (Ma *et al.*, 1987). Homologous recombination has been utilised in the construction of infectious clones for a number of different RNA viruses (Desbiez *et al.*, 2012; Nagyová & Šubr 2007; Liang *et al.*, 2004; Youssef *et al.*, 2011). In the construction of the UCBSV infectious clone, the genome was initially constructed as two separate halves of approximately 5 kb each, this was to check for the stability of propagating the genome in *E. coli* before constructing the complete UCBSV infectious clone. The UCBSV infectious clone plasmid pYES2.1 containing the complete UCBSV genome without introns was stable on propagation through *E. coli*, this was unlike observations of most potyviral infectious clones that have been reported to be toxic on propagation through *E. coli* cells (Johansen 1996; López-Moya & García 2000; Desbiez *et al.*, 2012). The constructed UCBSV Kikombe infectious clone was sequenced and the genome of the assembled infectious clone compared with the genome of the wildtype UCBSV genome. The genomes were 99.9% identical although single nucleotide polymorphisms (SNPs) were identified. Each of these SNPs were found to be present in the natural populations of other UCBSV isolates, however, differences between sequences of infectious clones and wildtype genomes have been reported to affect the overall biological activity of the infectious clones and in some cases, the infectious clone may be more aggressive than the wildtype virus (Boyer and Haenni, 1994).

Mechanical inoculation was used to infect the susceptible plants with the *in-vitro* UCBSV infectious clone. It involved the application of carborandum powder which acts as an abrasive and allows entry of the infectious transcripts into the leaves. Different methods have been used

to introduce infectious clones among which are the use of biolistics (Fakhfakh *et al.*, 1996; Gal-On *et al.*, 1997), all the different methods available for use have advantages and disadvantages as regards the efficiency with which the infectious clone causes infection in the host plant. The infectious clones were confirmed to infect tobacco and cassava plants and the infections caused by the UCBSV infectious clone were found to be similar to the infections caused by the wildtype virus. The RNA of the viruses recovered from the infectious clone infected plants was amplified by PCR cloned and sequenced and results confirmed successful infection of the plants. The RNA transcripts were capped with the 5' Cap analog to ensure infectivity, capped infectious clones have been used for a number of infectious viruses including *Plum pox virus* (PPV) (Riechmann *et al.*, 1990), *Brome mosaic virus* (BMV) (Ahlquist & Janda, 1984).

Attempts to construct an *in-vivo* UCBSV infectious clone failed despite the different methods that were tried. Construction of infectious clones of potyviruses is very unpredictable as some forms of the virus may be found to be stable in *E. coli* whilst other forms are unstable (Johansen, 1996). The possible reason for the failure to construct the *in-vivo* UCBSV infectious clone was thought to be due to instability of the infectious clone in the yeast adapted pCambia0380 plasmid backbone. The UCBSV genome could have been toxic in *E. coli* when placed under a strong constitutive 35S promoter, hence introns were introduced into the UCBSV genome. Three different introns were introduced in the P3, CI and NIb coding regions of the UCBSV genome and the positions selected for the intron insertion contained the consensus sequences both at the beginning and the end of the intron insertion points to allow for subsequent splicing of the introns *in-planta* (Gao *et al.*, 2012). The recommended consensus sequences for intron insertions are GT-AG at the beginning and the end of the intron insertion site. It has been reported that the number and location of introns needed to stabilise a given infectious clone may vary for different viruses (Johansen, 1996). The P3 and CI proteins

have been implicated in the expression of toxic proteins thus preventing the propagation of potyviral infectious clones in the *E. coli* cells (Subr *et al.*, 2000; Johansen 1996). Intron insertion in proteins thought to cause toxicity in *E. coli* has been reported to stabilise infectious clones (Johansen, 1996), for instance, the *Plum pox virus* (PPV) infectious clone could be stabilised by the insertion of an intron in the P3 coding region (López-Moya & García, 2000). In this study, three different introns were inserted in the UCBSV genome, however, it was observed that even with the introduction of introns in the genome, a stable UCBSV *in-vivo* clone recombined in pCAMBIA0380YA could not be recovered from the *E. coli* cells. It is possible that the positions in the selected proteins (or even the targeted proteins) where the introns were placed or the number of introns required to stabilize the UCBSV *in-vivo* clone was not appropriate (Johansen, 1996).

In-vitro ligation was also used in an attempt to construct a UCBSV *in-vivo* infectious clone. This involved the introduction of appropriate restriction sites at the 5' and 3' ends of the UCBSV genome. The restriction sites were successfully added to the ends of the genome however, *in-vitro* ligation into the pCAMBIA 2300 backbone vector was not possible. A lot of inconsistencies have been reported about the possibility of utilising *in-vitro* ligation of genomes for generating infectious clones.

It remains unclear if infectious clones can be obtained for other UCBSV isolates without the need to introduce introns as was the case for the UCBSV Kikombe isolate. However, high genetic variability has been reported between isolates of both CBSV and UCBSV (Ndunguru *et al.*, 2015; Mbanzibwa *et al.*, 2011b). This implies that the construction of infectious clones for some of the isolates of UCBSV may present challenges and require introns to stabilise them. Nevertheless, in this study *in-vitro* UCBSV infectious clones were generated and shown to infect both cassava and alternative hosts such as *N. benthamiana*. The infectious clone will be

useful in a number of ways: 1) Functional analysis as exemplified in this study; 2) Use in breeding programs; 3) The infectious clone can be engineered to understand the pathology of the UCBSV virus.

However, caution has to be considered as regards the biosafety of working with the infectious clones to avoid accidental release into the environment by ensuring that appropriate disposal methods are followed with experimental plants (Fuchs & Gonsalves, 2007). This is because infectious clones are biologically amplifiable synthetic virus molecules so their use should be restricted to confined and controlled biological environments.

6.1.3 Characterization of P1 and HAM1 proteins in the UCBSV genome

With the generation of viable *in-vitro* UCBSV infectious clones, it was possible to use reverse genetics to study and characterise the functions of P1 and HAM1 proteins in the UCBSV genome.

The P1 protein encoded in the UCBSV genome has been previously reported as a pathogenicity factor that is involved in suppression of RNA silencing (Mbanzibwa *et al.*, 2009a). RNA silencing is one of the most common means through which plants defend themselves against viruses (Wang *et al.*, 2012; Ding, 2010; Baulcombe, 2002; Soosaar *et al.*, 2005). As a counter-defensive strategy, most viruses encode RNA silencing suppressors (Valli *et al.*, 2006; Sharp, 2001) in addition to other counter-defense mechanisms (Schwartz *et al.*, 2002; Taliansky *et al.*, 2003). In this work, the P1 of UCBSV Kikombe isolate was determined to be involved in RNA silencing and similar findings were observed in the UCBSV MLB3 isolate (GenBank accession FJ039520) (Mbanzibwa *et al.*, 2009a). Mutating the functional motif LXKA in the P1 protein affected the RNA silencing suppression functionality of the protein (loss of RSS activity) in

addition to the increasing viral titres in *N. benthamiana* plants infected with the LAAA mutated P1 full-length infectious clone. Both the LXKA motif and putative zinc finger have been implicated to be involved in RNA silencing suppression (Valli *et al.*, 2008), however, this study showed that a mutated LXKA motif and a functional zinc finger motif was sufficient to cause loss of this function in the P1 protein of UCBSV. Mutational analysis has been applied to study the functions of different viral proteins, a number of different motifs have been targeted and mutated in order to elucidate their role in a given viral infection process, for instance, the DAG motif in the N-terminus of the CP of *Tobacco vein mottling virus* (TVMV) has been implicated to be involved in aphid transmissibility of the virus (López-Moya *et al.*, 1999; Atreya *et al.*, 1990). Mutational analysis on the LXKA motif of the ipomoviral *Cucumber vein yellowing virus* (CVYV) P1b protein led to loss of function in the RNA silencing suppression roles of the protein (Valli *et al.*, 2008). Valli *et al.*, (2008) mutated both the LXKA and putative zinc finger motif (CX₂CX_nCX₂) and reported loss of function. In this study, only the LXKA and not the zinc finger was mutated and this motif alone was important in RNA silencing suppression.

Reverse genetics was used to further study the roles of the UCBSV proteins hence, different deletion strains of UCBSV were constructed and evaluated to study the roles of P1 and HAM1 proteins in the UCBSV genome. It was observed that the P1 deletion infectious clone was able to move systemically within the plant although the RT-PCR showed low levels of the virus in the plant implying that the P1 protein may probably be involved in genome amplification. A similar observation has been reported in the P1 of the *Tobacco etch virus* (TEV) (Verchot & Carrington, 1995). A complete deletion of P1 in TEV affected genome amplification in the mutants in the plant.

A mutation was created in the LXKA motif of the P1 protein of the UCBSV infectious clone. It was observed that the mutated infectious clone was still infectious and the viral protein titres

were higher than the wildtype infected plants. RNA was extracted to confirm that the mutated infectious clone was still replicating with the mutation and the sequencing results confirmed that the mutation was still present in the viral progeny recovered from the plants. Similar observations have been reported in the mutational analysis of *Tobacco vein mottling virus* (TVMV) (Klein *et al.*, 1994). Other studies involving the mutation of the multifunctional HC-Pro protein involved in RNA silencing functions of different potyviruses showed contrary results in which the mutations in HC-Pro are reported to reduce viral titers as has been found in *Clover yellow vein virus* (Yambao *et al.*, 2008), *Plum pox virus* (González-Jara *et al.*, 2005) and *Zucchini yellow mosaic virus* (Lin *et al.*, 2007).

The HAM1 UCBSV knockout mutant caused yellow mosaic symptoms on *N. benthamiana* plants, however, RT-PCR could not detect the UCBSV virus in these plants although the virus could be detected in the plants infected with the wildtype UCBSV infectious clone. This observation was similar to the finding of (Hamidun Bunawan, PhD Thesis). The possible explanation for this finding may be that the levels of viral RNA in the plants may be too low for accurate detection by RT-PCR although PCR is by far the most sensitive method for detection of plant pathogens (Boonham *et al.*, 2008). Other diagnostic methods like performing RNA northern blotting and using UCBSV specific labelled probes for the detection of the virus in the infected plants could be tried although challenges in diagnostics of plant pathogens have been reported (Smith *et al.*, 2008).

In *Saccharomyces cerevisiae*, it has been reported that overexpression of HAM1 gene confers resistance to fluorouracil a pyrimidine analog (Carlsson *et al.*, 2013) however, overexpression of HAM1 in UCBSV did not show similar findings (Hamidun Bunawan, PhD Thesis). HAM1 could also be thought to serve the function of the missing HC-Pro protein in the UCBSV/CBSV genomes by functioning as suppressor of RNA silencing, however, agroinfiltration assays for

RNA silencing suppression with HAM1 protein from UCBSV MLB3 isolate proved otherwise (Mbanzibwa *et al.*, 2009a). Depending on observations of symptoms on the UCBSV HAM1 deletion mutant, the role of HAM1 in UCBSV may be hypothesised to be involved in symptom expression in UCBSV.

One of the transgenic lines expressing CBSV Nampula HAM1 protein was found to complement the UCBCV HAM1 full deletion mutant infectious clone despite the differences in the percentage amino acid identities between the two isolates (45%). Previously, the HAM1 full deletion mutant could not be detected in the non-transgenic *N. tabacum* plants meaning the observed result was due to the complementation by the HAM1 expressed in the transgenic lines. Similar observations have been reported in the complementation experiments with the Nlb mutant infectious clone of *Tobacco etch virus* (TEV) which was complemented in Nlb expressing *N. tabacum* transgenic lines (Li & Carrington, 1995).

6.2 Recommendations

To further understand UCBSV pathogenicity and infection process, the now available UCBSV *in-vitro* infectious clone can have more modifications performed and studied which include;

- Tagging the UCBSV *in-vitro* infectious clones with reporter genes like GFP or GUS to understand and monitor the movement of the UCBSV viruses in host plants.
- Construct chimeras between the proteins of UCBSV and CBSV isolates to try and understand and elucidate the mechanisms behind the differences in the symptoms presented by the two viruses.
- Study vector transmissibility of the infectious clone virus by the *B. tabaci* whitefly vector.
- Further studies are needed to investigate appropriate positions for insertions of introns for construction of *in-vivo* UCBSV infectious clones

6.3 Conclusions

From the observations obtained in this study, the following conclusions can be drawn;

- The UCBSV Kikombe wildtype isolate studied in this thesis belongs to the genus *Ipomovirus* and causes very mild mosaic symptoms on *N. benthamiana* host plants. It causes no observable symptoms on *N. clevelandii*, *N. tabacum* and *N. glutinosa* plants although the isolate can infect all the different host plants. The UCBSV infectious clone constructed causes similar symptoms like the wildtype virus in indicator plants, however, yellow mosaic symptoms are observed in cassava plants.
- The UCBSV Kikombe *in-vitro* infectious clone plasmid could be stably propagated in *E. coli* cells using pYES2.1 as a backbone vector. However, instability was observed in the construction of the *in-vivo* UCBSV infectious clone in the yeast adapted pCAMBIA0380 backbone vector.
- P1 protein in UCBSV Kikombe isolate is a suppressor of RNA silencing. The P1 protein may also be involved in genome amplification of the virus in the plant.
- HAM1 protein in UCBSV may be involved in symptom development in the plants.

References

- Abarshi, M. M., Mohammed, I. U., Jeremiah, S. C., Legg, J. P., Kumar, P. L., Hillocks, R. J., & Maruthi, M. N. (2012). Multiplex RT-PCR assays for the simultaneous detection of both RNA and DNA viruses infecting cassava and the common occurrence of mixed infections by two cassava brown streak viruses in East Africa. *Journal of Virological Methods*, 179(1), 176–184. doi:10.1016/j.jviromet.2011.10.020
- Abel, P. P., Nelson, R. S., De, B., Hoffmann, N., Rogers G., S., Fraley, R. T., & Beachy, R. N. (1986). Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene. *Science*, 232, 738–743.
- Adams, I. P., Abidrabo, P., Miano, D. W., Alicai, T., Kinyua, Z. M., Clarke, J., ... Smith, J. (2013). High throughput real-time RT-PCR assays for specific detection of cassava brown streak disease causal viruses, and their application to testing of planting material. *Plant Pathology*, 62, 233–242. doi:10.1111/j.1365-3059.2012.02622.x
- Adams, M. J., Antoniw, J. F., & Beaudoin, F. (2005). Overview and analysis of the polyprotein cleavage sites in the family Potyviridae. *Molecular Plant Pathology*, 6(4), 471–487. doi:10.1111/j.1364-3703.2005.00296.x
- Adams, M. J., Antoniw, J. F., & Fauquet, C. M. (2005). Molecular criteria for genus and species discrimination within the family Potyviridae. *Archives of Virology*, 150, 459–479. doi:10.1007/s00705-004-0440-6
- Adkins, S., Webb, S. E., Baker, C. A., & Kousik, C. S. (2008). Squash vein yellowing virus Detection Using Nested Polymerase Chain Reaction Demonstrates that the Cucurbit Weed *Momordica charantia* Is a Reservoir Host. *Plant Disease*, 92(7), 1119–1123. doi:10.1094/PDIS-92-7-1119
- Agrios, G. N. (2005). *Plant Pathology* (5th ed.). Elsevier Academic Press, Burlington, Mass.
- Ahlquist, P., & Janda, M. (1984). cDNA cloning and in vitro transcription of the complete brome mosaic virus genome. *Molecular and Cellular Biology*, 4(12), 2876–2882. doi:10.1128/MCB.4.12.2876
- Ala-Poikela, M., Goytia, E., Haikonen, T., Rajamäki, M.-L., & Valkonen, J. P. T. (2011). Helper component proteinase of the genus Potyvirus is an interaction partner of translation initiation factors eIF(iso)4E and eIF4E and contains a 4E binding motif. *Journal of Virology*, 85(13), 6784–6794. doi:10.1128/JVI.00485-11
- Alfano, J. R., & Collmer, A. (1996). Bacterial Pathogens in Plants: Life up against the Wall. *The Plant Cell*, 8(10), 1683–1698. doi:10.1105/tpc.8.10.1683
- Alicai, T., Ndunguru, J., Sseruwagi, P., Tairo, F., Okao-Okuja, G., Nanvubya, R., ... Boykin, L. M. (2016). Characterization by Next Generation Sequencing Reveals the Molecular

Mechanisms Driving the Faster Evolutionary rate of Cassava brown streak virus Compared with Ugandan cassava brown streak virus (Preprint).

- Alicai, T., Omongo, C. A., Maruthi, M. N., Hillocks, R., Baguma, Y., Kawuki, R., ... Colvin, J. (2007). Re-emergence of Cassava Brown Streak Disease in Uganda. *Plant Disease*, 91(1), 24–29. doi:10.1094/PD-91-0024
- Allem, A. C. (2002). The origins and taxonomy of cassava. Cassava biology, production and utilization (pp. 1–16).
- Al-Musa, A. M., Qusus, S. J., & Mansour, A. N. (1985). Cucumber vein yellowing virus on cucumber in Jordan. *Plant Disease*, 69(4).
- Anandalakshmi, R., Pruss, G. J., Ge, X., Marathe, R., Mallory, A. C., Smith, T. H., & Vance, V. B. (1998). A viral suppressor of gene silencing in plants. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 13079–13084. doi:10.1073/pnas.95.22.13079
- Arazi, T., Slutsky, S. G., Shibolet, Y. M., Wang, Y., Rubinstein, M., Barak, S., ... Gal-On, A. (2001). Engineering zucchini yellow mosaic potyvirus as a non-pathogenic vector for expression of heterologous proteins in cucurbits. *Journal of Biotechnology*, 87(1), 67–82. doi:10.1016/S0168-1656(01)00229-2
- Atreya, C. D., Raccach, B., & Pirone, T. D. (1990). A point mutation in the coat protein abolish aphid transmissibility of a potyvirus. *Virology*, 178(1), 161–165.
- Baltimore, D. (1971). Expression of animal virus genomes. *Bacteriological Reviews*, 35(3), 235–241.
- Baulcombe, D. (2002). RNA silencing. *Current Biology*, 12(3), R82–R84. doi:10.1042/BJ20100932
- Baulcombe, D. (2004). RNA silencing in plants. *Nature*, 431(7006), 356–363. doi:10.1038/nature02874
- Bedoya, L. C., & Daros, J. A. (2010). Stability of Tobacco etch virus infectious clones in plasmid vectors. *Virus Research*, 149(2), 234–240. doi:10.1016/j.virusres.2010.02.004
- Beijerinck, M. W. (1898). Over een contagium vivum fluidum als oorzaak van de vlekziekte der tabaksbladen. *Ver Sl Gewone Vergad Wis- Natuurkd Afd, K Akad Wet Amsterdam*, 7, 229–235.
- Beijerinck, M. W. (1898). Ueber ein contagium vivum fluidum als Ursache der Fleckenkrankheit der Tabaksblätter. *Verhandelingen Der Koninklijke Akademie van Wetenschappen Te Amsterdam*, (2), 3–21.
- Bendahmane, M., Szecsi, J., Chen, I., Berg, R. H., & Beachy, R. N. (2002). Characterization of mutant tobacco mosaic virus coat protein that interferes with virus cell-to-cell

- movement. *Proceedings of the National Academy of Sciences of the United States of America*, 99(6), 3645–3650. doi:10.1073/pnas.062041499
- Bigirimana, S., Barambanze, P., Ndayihanzamaso, P., Shirma, R., & Legg, J. P. (2011). First report of cassava brown streak disease and associated Ugandan cassava brown streak virus in Burundi. *New Disease Reports*, 24(26), 2044–0588.
- Bock, K. R., & Woods, R. D. (1983). Etiology of African cassava mosaic disease. *Plant Disease*, 67(9), 994–995.
- Boonham, N., Glover, R., Tomlinson, J., & Mumford, R. (2008). Exploiting generic platform technologies for the detection and identification of plant pathogens. *European Journal of Plant Pathology*, 121(3), 355–363. doi:10.1007/s10658-008-9284-3
- Boyer, J.-C., & Haenni, A.-L. (1994). Infectious Transcripts and cDNA Clones of RNA Viruses. *Virology*, 198(2), 415–426. doi:10.1006/viro.1994.1053
- Bredeson, J. V., Lyons, J. B., Prochnik, S. E., Wu, G. A., Ha, C. M., Edsinger-Gonzales, E., ... Rokhsar, D. S. (2016). Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. *Nature Biotechnology*, 34(5), 562–570. doi:10.1038/nbt.3535
- Brigneti, G., Voinnet, O., Li, W., & Ji, L. (1998). Viral pathogenicity determinants are suppressors of transgene silencing in *Nicotiana Benthamiana*. *EMBO Journal*, 17(22), 6739–6746.
- Brunt, A. A., Crabtree, K., Dallwitz, M. J., Gibbs, A. J., & Watson, L. (1996). Viruses of plants: description and list from the VIDE database. *CAB International*.
- Butler, E. T., & Chamberlin, M. J. (1982). Bacteriophage SP6-specific RNA polymerase. I. Isolation and characterization of the enzyme. *Journal of Biological Chemistry*, 257, 5772–5778.
- Calvert, L. A., & Thresh, J. M. (2002). The Viruses and Virus diseases of Cassava. *Cassava: Biology, Production and Utilization*, 237–260. doi:10.1079/9780851995243.0000
- Campo, B. V. H., Hyman, G., & Bellotti, A. (2011). Threats to cassava production: Known and potential geographic distribution of four key biotic constraints. *Food Security*, 3(3), 329–345. doi:10.1007/s12571-011-0141-4
- Carlsson, M., Gustavsson, M., Hu, G. Z., Muren, E., & Ronne, H. (2013). A Ham1p-Dependent Mechanism and Modulation of the Pyrimidine Biosynthetic Pathway can both Confer Resistance to 5-Fluorouracil in Yeast. *PLoS ONE*, 8(10), e52094. doi:10.1371/journal.pone.0052094
- Carrington, J. C., Kasschau, K. D., Mahajan, S. K., & Schaad, M. C. (1996). Cell-to-Cell and long-distance transport of viruses in plants. *The Plant Cell*, 8, 1669–1681. doi:10.1105/tpc.8.10.1669

- Cheng, S., Chang, S. Y., Gravitt, P., & Respass, R. (1994). Long PCR. *Nature*, 369(6482), 684–685. doi:10.1038/369684a0
- Chikh Ali, M., Said Omar, A., & Natsuaki, T. (2011). An infectious full-length cDNA clone of potato virus Y(NTN-NW), a recently reported strain of PVY that causes potato tuber necrotic ringspot disease. *Archives of Virology*, 156(11), 2039–2043. doi:10.1007/s00705-011-1062-4
- Chung, B. Y.-W., Miller, W. A., Atkins, J. F., & Firth, A. E. (2008). An overlapping essential gene in the Potyviridae. *Proceedings of the National Academy of Sciences*, 105(15), 5897–5902. doi:10.1073/pnas.0800468105
- Colinet, D., Kummert, J., & Lepoivre, P. (1998). The nucleotide sequence and genome organisation of the whitefly transmitted sweetpotato mild mottle virus: a close relationship with members of the family Potyviridae. *Virus Research*, 53, 187–196.
- Dagert, M., & Ehrlich, S. D. (1979). Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. *Gene*, 6(1), 23–38. doi:10.1016/0378-1119(79)90082-9
- Daròs, J. A., Schaad, M. C., & Carrington, J. C. (1999). Functional analysis of the interaction between VPg-proteinase (Nla) and RNA polymerase (Nlb) of tobacco etch potyvirus, using conditional and suppressor mutants. *Journal of Virology*, 73(10), 8732–8740.
- Davis, N. L., Willis, L. V., Smith, J. F., & Johnston, R. E. (1989). In vitro synthesis of infectious Venezuelan equine encephalitis virus RNA from a cDNA clone: analysis of a viable deletion mutant. *Virology*, 171, 189–204.
- Dawson, W. O., Beck, D. L., Knorr, D. A., & Grantham, G. L. (1986). cDNA cloning of the complete genome of tobacco mosaic virus and production of infectious transcripts. *Proceedings of the National Academy of Sciences*, 83(6), 1832–1836. doi:10.1073/pnas.83.6.1832
- Deng, P., Wu, Z., & Wang, A. (2015). The multifunctional protein CI of potyviruses plays interlinked and distinct roles in viral genome replication and intercellular movement. *Virology Journal*, 12(1), 141. doi:10.1186/s12985-015-0369-2
- Desbiez, C., Chandeysson, C., Lecoq, H., & Moury, B. (2012). A simple, rapid and efficient way to obtain infectious clones of potyviruses. *Journal of Virological Methods*, 183(1), 94–97. doi:10.1016/j.jviromet.2012.03.035
- Ding, S. W., & Voinnet, O. (2007). Antiviral immunity directed by small RNAs. *Cell*, 130(3), 413–426. doi:10.1016/j.cell.2007.07.039
- Ding, S. W. (2010). RNA-based antiviral immunity. *Nature Reviews Immunology*, 10(9), 632–644. doi:10.1038/nri2824

- Doija, V. V., Haldeman, R., Robertson, N. L., Dougherty, W. G., & Carrington, J. C. (1994). Distinct functions of capsid protein in assembly and movement of tobacco etch potyvirus in plants. *The EMBO Journal*, 13(6), 1482–1491.
- Dombrovsky, A., Reingold, V., & Antignus, Y. (2014). Ipomovirus--an atypical genus in the family Potyviridae transmitted by whiteflies. *Pest Management Science*, 70(10), 1553–1567. doi:10.1002/ps.3735
- Domier, L. L., Franklin, K. M., Hunt, A. G., Rhoads, R. E., & Shaw, J. G. (1989). Infectious *in vitro* transcripts from cloned cDNA of a potyvirus, tobacco vein mottling virus. *Proceedings of the National Academy of Sciences*, 86(10), 3509–3513. doi:10.1073/pnas.86.10.3509
- Domier, L. L., Shaw, J. G., & Rhoads, R. E. (1987). Potyviral Proteins Share Amino Acid Sequence Homology With Picorna-, Como-, and Caulimoviral Proteins. *Virology*, 158, 20–27.
- Dougherty, W. G., & Carrington, J. C. (1988). Expression and Function of Potyviral Gene Products. *Annual Review of Phytopathology*, 26(1), 123–143. doi:10.1146/annurev.py.26.090188.001011
- Duff-Farrier, C. R. A., Bailey, A. M., Boonham, N., & Foster, G. D. (2015). A pathogenicity determinant maps to the N-terminal coat protein region of the Pepino mosaic virus genome. *Molecular Plant Pathology*, 16(3), 308–15. doi:10.1111/mpp.12184
- Dunoyer, P., Lecellier, C.-H., Parizotto, E. A., Himber, C., & Voinnet, O. (2004). Probing the microRNA and small interfering RNA pathways with virus-encoded suppressors of RNA silencing. *The Plant Cell*, 16(5), 1235–1250. doi:10.1105/tpc.020719
- Elbashir, S. M., Lendeckel, W., & Tuschl, T. (2001). RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes and Development*, 15, 188–200. doi:10.1101/gad.862301.vents
- El-Sharkawy, M. A. (2003). Cassava biology and physiology. *Plant Molecular Biology*, 53(5), 621–641.
- Essers, S. A. J. A. (1995). Removal of cyanogens from cassava roots: Studies on domestic sun-drying and solid-substrate fermentation in rural Africa.
- Estevan, J., Maréna, A., Callot, C., Lacombe, S., Moretti, A., Caranta, C., & Gallois, J.-L. (2014). Specific requirement for translation initiation factor 4E or its isoform drives plant host susceptibility to Tobacco etch virus. *BMC Plant Biology*, 14(1), 67. doi:10.1186/1471-2229-14-67
- Fakhfakh, H., Vilaine, F., Makni, M., & Robaglia, C. (1996). Cell-free cloning and biolistic inoculation of an infectious cDNA of potato virus Y. *Journal of General Virology*, 77(3), 519–523. doi:10.1099/0022-1317-77-3-519

- Fauquet, C. M. D., & Fargette, D. (1990). African Cassava Mosaic Virus : Etiology, Epidemiology and Control. *Plant Disease*, 74(6), 404–411.
- Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U., & Ball, L. A. (2005). Virus Taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses. Retrieved August 30, 2015, from http://samples.sainsburysebooks.co.uk/9780080886909_sample_799973.pdf
- Fernández, A., Laín, S., & García, J. A. (1995). RNA helicase activity of the plum pox potyvirus CI protein expressed in *Escherichia coli*. Mapping of an RNA binding domain. *Nucleic Acids Research*, 23(8), 1327–1332.
- Fernandez-Delmond, I., Pierrugues, O., Wispelaere, M. De, Guilbaud, L., Gaubert, S., Divéki, Z., ... Jacquemond, M. (2004). A novel strategy for creating recombinant infectious RNA virus genomes. *Journal of Virological Methods*, 121(2), 247–257. doi:10.1016/j.jviromet.2004.07.003
- Fondong, V. N., Pita, J. S., Rey, M. E. C., Kochko, A., Beachy, R. N. and, & Fauquet, C. M. (2000). Evidence of synergism between African cassava mosaic virus and new double-recombination geminivirus infecting cassava in Cameroon. *Journal of General Virology*, 81, 287–297.
- Food and Agriculture Organisation Statistics Database (2014). <http://faostat3.fao.org/faostat-gateway/go/to/home>
- Ford, K. L., Baumgartner, K., Henricot, B., Bailey, A. M., & Foster, G. D. (2015). A reliable in vitro fruiting system for *Armillaria mellea* for evaluation of *Agrobacterium tumefaciens* transformation vectors. *Fungal Biology*, 119(10), 859–869. doi:10.1016/j.funbio.2015.06.007
- Fuchs, M., & Gonsalves, D. (2007). Safety of virus-resistant transgenic plants two decades after their introduction: lessons from realistic field risk assessment studies. *Annual Review of Phytopathology*, 45, 173–202. doi:10.1146/annurev.phyto.45.062806.094434
- Galipienso, L., Rubio, L., Aramburu, J., Velasco, L., & Janssen, D. (2012). Complete nucleotide sequence of a severe isolate of cucumber vein yellowing virus from Jordan. *Archives of Virology*, 157(6), 1189–92. doi:10.1007/s00705-012-1264-4
- Gallie, D. R., & Kobayashi, M. (1994). The role of the 3'-untranslated region of non-polyadenylated plant viral mRNAs in regulating translational efficiency. *Gene*, 142(2), 159–165. doi:10.1016/0378-1119(94)90256-9
- Gallie, D. R., Tanguay, R. L., & Leathers, V. (1995). The tobacco etch viral 5' leader and poly(A) tail are functionally synergistic regulators of translation. *Gene*, 165(2), 233–238. doi:10.1016/0378-1119(95)00521-7
- Gal-On, A. (2000). A Point Mutation in the FRNK Motif of the Potyvirus Helper Component-Protease Gene Alters Symptom Expression in Cucurbits and Elicits

- Protection Against the Severe Homologous Virus. *Phytopathology*, 90(5), 467–473. doi:10.1094/PHYTO.2000.90.5.467
- Gal-On, A., Antignus, Y., Rosner, A., & Raccach, B. (1991). Infectious in vitro RNA transcripts derived from cloned cDNA of the cucurbit potyvirus, zucchini yellow mosaic virus. *Journal of General Virology*, 72(11), 2639–2643. doi:10.1099/0022-1317-72-11-2639
- Gal-On, A., Meiri, E., Elman, C., Gray, D. J., & Gaba, V. (1997). Simple hand-held devices for the efficient infection of plants with viral-encoding constructs by particle bombardment. *Journal of Virological Methods*, 64(1), 103–110. doi:10.1016/S0166-0934(96)02146-5
- Galperin, M. Y., Moroz, O. V., Wilson, K. S., & Murzin, A. G. (2006). House cleaning, a part of good house keeping. *Molecular Microbiology*, 59(1), 5–19.
- Gammelgard, E., Mohan, M., & Valkonen, J. P. T. (2007). Potyvirus-induced gene silencing: the dynamic process of systemic silencing and silencing suppression. *Journal of General Virology*, 88(8), 2337–2346. doi:10.1099/vir.0.82928-0
- Gao, R., Tian, Y. P., Wang, J., Yin, X., Li, X. D., & Valkonen, J. P. T. (2012). Construction of an infectious cDNA clone and gene expression vector of Tobacco vein banding mosaic virus (genus Potyvirus). *Virus Research*, 169(1), 276–281. doi:10.1016/j.virusres.2012.07.010
- Garcia-Arenal, F., Fraile, A., & Malpica, M. (2001). Variability and Genetic structure of Plant Virus Populations. *Annual Review of Phytopathology*, 39, 157–86.
- Garcia-Ruiz, H., Takeda, A., Chapman, E. J., Sullivan, C. M., Fahlgren, N., Bremmelis, K. J., & Carrington, J. C. (2010). Arabidopsis RNA-dependent RNA polymerases and dicer-like proteins in antiviral defense and small interfering RNA biogenesis during Turnip Mosaic Virus infection. *The Plant Cell*, 22(2), 481–96. doi:10.1105/tpc.109.073056
- Gietz, D. R., and Woods, R. A. (2002). Transformation of yeast by lithium acetate/single stranded carrier DNA/ Polyethylene Glycol method. *Methods in Enzymology*, 350, 87–96.
- Giner, A., Lakatos, L., García-Chapa, M., López-Moya, J. J., & Burguán, J. (2010). Viral protein inhibits RISC activity by argonaute binding through conserved WG/GW motifs. *PLoS Pathogens*, 6(7), e1000996. doi:10.1371/journal.ppat.1000996
- Goldbach, R., Bucher, E., & Prins, M. (2003). Resistance mechanisms to plant viruses: An overview. *Virus Research*, 92(2), 207–212. doi:10.1016/S0168-1702(02)00353-2
- González-Jara, P., Atencio, F. A., Martínez-García, B., Barajas, D., Tenllado, F., & Díaz-Ruiz, J. R. (2005). A Single Amino Acid Mutation in the Plum pox virus Helper Component-Proteinase Gene Abolishes Both Synergistic and RNA Silencing Suppression Activities. *Phytopathology*, 95(8), 894–901. doi:10.1094/PHYTO-95-0894

- Haas, M., Bureau, M., Geldreich, A., Yot, P., & Keller, M. (2002). Cauliflower mosaic virus : still in the news. *Molecular Plant Pathology*, 3, 419–429.
- Hamilton, A. J., & Baulcombe, D. C. (1999). A Species of Small Antisense RNA in Posttranscriptional Gene Silencing in Plants. *Science*, 286(5441), 950–952. doi:10.1126/science.286.5441.950
- Hamilton, W. D. O., Bisaro, D. M., Coutts, R. H. A., & Buck, K. W. (1983). Demonstration of the bipartite nature of the genome of a single-stranded DNA plant virus by infection with the cloned DNA components. *Nucleic Acids Research*, 11(21), 7387–7396.
- Hammond, S. M., Bernstein, E., Beach, D., & Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature*, 404(6775), 293–296. doi:10.1038/35005107
- Harpaz, I., & Cohen, S. (1965). Semipersistent Relationship between Cucumber Vein Yellowing Virus (CVYV) and its Vector, the Tobacco Whitefly (*Bemisia tabaci* Gennadi us). *Journal of Phytopathology*, 54(3), 240–248. doi:10.1111/j.1439-0434.1965.tb04096.x
- Harrison, B. D., Swanson, M. M., & Robinson, D. J. (1995). Cassava viruses in the old world. Retrieved February 18, 2016, from <http://www.sidalc.net/cgi-bin/wxis.exe/?IsisScript=catalco.xis&method=post&formato=2&cantidad=1&expresion=mfn=015322>
- Hillocks, R. J. (2002). Cassava in Africa (pp. 41–54).
- Hillocks, R. J. (1997). Cassava virus diseases and their control with special reference to southern Tanzania. *Integrated Pest Management Reviews*, 2(3), 125–138.
- Hillocks, R. J., Raya, M. D., Mtunda, K., & Kiozia, H. (2001). Effects of brown streak virus disease on yield and quality of cassava in Tanzania. *Journal of Phytopathology*, 149, 389–394. doi:10.1046/j.1439-0434.2001.00641.x
- Hillocks, R. J., Raya, M., & Thresh, J. M. (1996). The association between root necrosis and above-ground symptoms of brown streak virus infection of cassava in southern Tanzania. *International Journal of Pest Management*, 42(4), 285–289. doi:10.1080/09670879609372008
- Hillocks, R. J., & Thresh, J. M. (2000). Cassava mosaic and cassava brown streak virus diseases in Africa. A comparative guide to symptoms and aetiologies. *Roots*, 7(1), 1–8.
- Hillocks, R., & Jennings, D. (2003). Cassava brown streak disease: A review of present knowledge and research needs. *International Journal of Pest Management*, 49(3), 225–234. doi:10.1080/0967087031000101061
- Hillocks, R.J., Raya, M.D., Mtunda, K., and Kiozia, H. (2001). Effects of brown streak virus disease on yield and quality of cassava in Tanzania. *Journal of Phytopathology*, 149(1-8), 389–394.

- Hong, Y., & Hunt, A. G. (1996). RNA polymerase activity catalyzed by a potyvirus-encoded RNA-dependent RNA polymerase. *Virology*, 226(1), 146–151. doi:10.1006/viro.1996.0639
- Hull, R. (2014). *Plant Virology* (Fifth Edit.). Elsevier Inc.
- Inaba, J., Kim, B. M., Shimura, H., & Masuta, C. (2011). Virus-induced necrosis is a consequence of direct protein-protein interaction between a viral RNA-silencing suppressor and a host catalase. *Plant Physiology*, 156(4), 2026–2036. doi:10.1104/pp.111.180042
- Ivanov, K. I., Eskelin, K., Lõhmus, A., & Mäkinen, K. (2014). Molecular and cellular mechanisms underlying potyvirus infection. *Journal of General Virology*, 95(7), 1415–1429. doi:10.1099/vir.0.064220-0
- Jakab, G., Droz, E., Brigneti, G., Baulcombe, D., & Malnoš, P. (1997). Infectious in vivo and in vitro transcripts from a full-length cDNA clone of PVY-N605, a Swiss necrotic isolate of potato virus Y. *Journal of General Virology*, 78(12), 3141–3145.
- Janssen, D., Martín, G., Velasco, L., Gómez, P., Segundo, E., Ruiz, L., & Cuadrado, I. M. (2005). Absence of a coding region for the helper component-proteinase in the genome of cucumber vein yellowing virus, a whitefly-transmitted member of the Potyviridae. *Archives of Virology*, 150(7), 1439–1447. doi:10.1007/s00705-005-0515-z
- Jarvis, A., Ramirez-Villegas, J., Herrera Campo, B. V. H., & Navarro-Racines, C. (2012). Is Cassava the Answer to African Climate Change Adaptation? *Tropical Plant Biology*, 5(1), 9–29. doi:10.1007/s12042-012-9096-7
- Jennings, D. L. (1994). Breeding for resistance to African cassava mosaic geminivirus in East Africa. *Tropical Science*, 34(1), 110–122.
- Jeremiah, S. C., Ndyetabula, I. L., Mkamilo, G. S., Haji, S., Muhanna, M. M., Chuwa, C., ... Legg, J. P. (2015). The Dynamics and Environmental Influence on Interactions Between Cassava Brown Streak Disease and the Whitefly, *Bemisia tabaci*. *Phytopathology*, 105(5), 646–655. doi:10.1094/PHYTO-05-14-0146-R
- Jiang, L., Wei, C., & Li, Y. (2012). Viral suppression of RNA silencing. *Science China Life Sciences*, 55(2), 109–18. doi:10.1007/s11427-012-4279-x
- Johansen, I. E. (1996). Intron insertion facilitates amplification of cloned virus cDNA in *Escherichia coli* while biological activity is reestablished after transcription in vivo. *Proceedings of the National Academy of Sciences*, 93(22), 12400–12405.
- Johansen, L. K., & Carrington, J. C. (2001). Silencing on the spot. Induction and suppression of RNA silencing in the *Agrobacterium*-mediated transient expression system. *Plant Physiology*, 126(3), 930–938. doi:10.1104/pp.126.3.930
- Johnston, M., Flick, J. S., & Pexton, T. (1994). Multiple mechanisms provide rapid and stringent glucose repression of GAL gene expression in *Saccharomyces cerevisiae*.

- Molecular and Cellular Biology*, 14(6), 3834–3841.
doi:10.1128/MCB.14.6.3834.Updated
- Jones, W. O. (1959). Manioc in Africa. In *Stanford University Press, Stanford, CA USA*.
- Jones-rhoades, M. W., Bartel, D. P., & Bartel, B. (2006). MicroRNAs and their regulatory roles in plants. *The Annual Review of Plant Biology*, 57, 19–53.
doi:10.1146/annurev.arplant.57.032905.105218
- Kanju, E. E., Mtunda, K. J., Muhana, M., Raya, M. D., & Mahungu, N. M. (2002). Management of Cassava brown streak virus disease in Tanzania. In *Cassava brown streak disease Past, Present and the Future* (pp. 66–69).
- Kasschau, K. D., & Carrington, J. C. (1998). A Counterdefensive Strategy of Plant Viruses: Suppression of Posttranscriptional Gene Silencing. *Cell*, 95(4), 461–470.
doi:10.1016/S0092-8674(00)81614-1
- Kasschau, K. D., & Carrington, J. C. (2001). Long-distance movement and replication maintenance functions correlate with silencing suppression activity of potyviral HC-Pro. *Virology*, 285, 71–81. doi:10.1006/viro.2001.0901
- Keating, B., Wilson, G., & Evenson, J. (1988). Effects of length, thickness, orientation, and planting density of cassava (*Manihot esculenta* Crantz) planting material on subsequent establishment, growth and yield. *East African Agricultural and Forestry*, 53, 145–149.
- Kekarainen, T., Savilahti, H., & Valkonen, J. P. T. (2002). Functional Genomics on Potato Virus A : Virus Genome-Wide Map of Sites Essential for Virus Propagation. *Genome Research*, 12(4), 584–594. doi:10.1101/gr.220702.Article
- Kelloniemi, J., Mäkinen, K., & Valkonen, J. P. T. (2008). Three heterologous proteins simultaneously expressed from a chimeric potyvirus: Infectivity, stability and the correlation of genome and virion lengths. *Virus Research*, 135(2), 282–291.
doi:10.1016/j.virusres.2008.04.006
- Kizito, E. B., Bua, A., Fregene, M., Egwang, T., Gullberg, U., & Westerbergh, A. (2005). The effect of cassava mosaic disease on the genetic diversity of cassava in Uganda. *Euphytica*, 146(1-2), 45–54. doi:10.1007/s10681-005-2959-3
- Klein, P. G., Klein, R. R., Rodrigues-Cerezo, E., Hunt, A. G. and, & Shaw, J. G. (1994). Mutational analysis of the Tobacco Vein Mottling Virus Genome. *Virology*, 204, 759–769.
- Lain, S., Riechmann, J. L., & García, J. A. (1990). RNA helicase: a novel activity associated with a protein encoded by a positive strand RNA virus. *Nucleic Acids Research*, 18(23), 7003–7006. doi:10.1093/nar/18.23.7003
- Lakatos, L., Szittyá, G., Silhavy, D., & Burgyán, J. (2004). Molecular mechanism of RNA silencing suppression mediated by p19 protein of tombusviruses. *The EMBO Journal*, 23(4), 876–884. doi:10.1038/sj.emboj.7600096

- Lecellier, C. H., & Voinnet, O. (2004). RNA silencing: No mercy for viruses? *Immunological Reviews*, 198, 285–303. doi:10.1111/j.0105-2896.2004.00128.x
- Lecoq, H., Desbiez, C., Delécolle, B., Cohen, S., & Mansour, A. (2000). Cytological and molecular evidence that the whitefly-transmitted Cucumber vein yellowing virus is a tentative member of the family Potyviridae. *Journal of General Virology*, 81(9), 2289–2293. doi:10.1099/0022-1317-81-9-2289
- Legg, J. P., & Fauquet, C. M. (2004). Cassava mosaic geminiviruses in Africa. *Plant Molecular Biology*, 56(4), 585–599. doi:10.1007/s11103-004-1651-7
- Legg, J. P., Jeremiah, S. C., Obiero, H. M., Maruthi, M. N., Ndyetabula, I., Okao-Okuja, G., ... Lava Kumar, P. (2011). Comparing the regional epidemiology of cassava mosaic and cassava brown streak virus pandemics in Africa. *Virus Research*, 159(2), 161–170. doi:10.1016/j.virusres.2011.04.018
- Legg, J. P., & Raya, M. D. (1998). Survey of cassava virus diseases in Tanzania. *International Journal of Pest Management*, 44(1), 17–23. doi:10.1080/096708798228473
- Legg, J. P., Shirima, R., Tajebe, L. S., Guastella, D., Boniface, S., Jeremiah, S., ... Rapisarda, C. (2014). Biology and management of Bemisia whitefly vectors of cassava virus pandemics in Africa. *Pest Management Science*, 70(10), 1446–1453. doi:10.1002/ps.3793
- Legg, J. P., Lava Kumar, P., Makeshkumar, T., Tripathi, L., Ferguson, M., Kanju, E., Ntawuruhunga, P., W. C. (2015). Cassava virus diseases: Biology, Epidemiology and Management. *Advances in Virus Research*, 91, 85–142.
- Leihner, D. (2002). *Agronomy and cropping systems. Cassava: production, biology and utilization*. CAB international.
- Lennon, A. (1985). *Cassava viruses from Africa: Report of the Scottish Crop Research Institute*.
- Li, W., Hilf, M. E., Webb, S. E., Baker, C. a., & Adkins, S. (2008). Presence of P1b and absence of HC-Pro in Squash vein yellowing virus suggests a general feature of the genus Ipomovirus in the family Potyviridae. *Virus Research*, 135(2), 213–219. doi:10.1016/j.virusres.2008.03.015
- Li, X. H., & Carrington, J. C. (1995). Complementation of tobacco etch potyvirus mutants by active RNA polymerase expressed in transgenic cells. *Proceedings of the National Academy of Sciences*, 92(2), 457–61. doi:10.1073/pnas.92.2.457
- Liang, D., Gray, S. M., Kaplan, I., & Palukaitis, P. (2004). Site-directed mutagenesis and generation of chimeric viruses by homologous recombination in yeast to facilitate analysis of plant-virus interactions. *Molecular Plant-Microbe Interactions*, 17(6), 571–6. doi:10.1094/MPMI.2004.17.6.571

- Lin, S.-S., Wu, H.-W., Jan, F.-J., Hou, R. F., & Yeh, S.-D. (2007). Modifications of the Helper Component-Protease of Zucchini yellow mosaic virus for Generation of Attenuated Mutants for Cross Protection Against Severe Infection. *Phytopathology*, 97(3), 287–296. doi:10.1094/PHYTO-97-3-0287
- Lindbo, J. A., Silva-Rosales, L., Proebsting, W. M., & Dougherty, W. G. (1993). Induction of a Highly Specific Antiviral State in Transgenic Plants: Implications for Regulation of Gene Expression and Virus Resistance. *The Plant Cell*, 5(12), 1749–1759. doi:10.1105/tpc.5.12.1749
- Lister, R. M. (1959). Mechanical Transmission of Cassava Brown Streak Virus. *Nature*, 182, 1588–1589.
- Llave, C. (2010). Virus-derived small interfering RNAs at the core of plant-virus interactions. *Trends in Plant Science*, 15(12), 701–707. doi:10.1016/j.tplants.2010.09.001
- Loebenstein G and Thottapilly G. (2013). *Virus and Virus-like Diseases of Major Crops in Developing Countries*. Springer Science & Business Media.
- López-Moya, J. J., & García, J. A. (2000). Construction of a stable and highly infectious intron-containing cDNA clone of plum pox potyvirus and its use to infect plants by particle bombardment. *Virus Research*, 68(2), 99–107. doi:10.1016/S0168-1702(00)00161-1
- López-Moya, J. J., Wang, R. Y., & Pirone, T. P. (1999). Context of the coat protein DAG motif affects potyvirus transmissibility by aphids. *Journal of General Virology*, 80(12), 3281–3288.
- Ma, H., Kunes, S., Schatz, P. J., & Botstein, D. (1987). Plasmid construction by homologous recombination in yeast. *Gene*, 58(2), 201–216. doi:10.1016/0378-1119(87)90376-3
- Mahungu, N. M., Bidiaka, M., Tata, H., Lukombo, S., & N'luta, S. (2003). Cassava brown streak disease-like symptoms in Democratic Republic of Congo. *Roots*, 8(2), 8–10.
- Maia, I. G., Haenni, A., & Bernardi, F. (1996). Potyviral HC-Pro : A multifunctional. *Journal of General Virology*, 77(1996), 1335–1341. doi:10.1099/0022-1317-77-7-1335
- Mallory, A. C., Reinhart, B. J., Bartel, D., Vance, V. B., & Bowman, L. H. (2002). A viral suppressor of RNA silencing differentially regulates the accumulation of short interfering RNAs and micro-RNAs in tobacco. *Proceedings of the National Academy of Sciences*, 99(23), 15228–15233. doi:10.1073/pnas.232434999
- Mallory, A. C., Reinhart, B. J., Bartel, D., Vance, V. B., & Bowman, L. H. (2002). A viral suppressor of RNA silencing differentially regulates the accumulation of short interfering RNAs and micro-RNAs in tobacco. *Proceedings of the National Academy of Sciences of the United States of America*, 99(23), 15228–33. doi:10.1073/pnas.232434999

- Mansour, A., & Al-Musa, A. (1993). Cucumber Vein Yellowing Virus; Host Range and Virus Vector Relationships. *Journal of Phytopathology*, 137(1), 73–78. doi:10.1111/j.1439-0434.1993.tb01327.x
- Manyong, V. ., Maeda, C., Kanju, E., & Legg, J. P. (2010). Economic damage of cassava brown streak disease in sub-Saharan Africa. In *Tropical Root and Tuber Crops and the Challenges of Globalization and Climate Change* (pp. 4–8).
- Maruthi, M. N., Bouvaine, S., Tufan, H. A., Mohammed, I. U., & Hillocks, R. J. (2014). Transcriptional response of virus-infected cassava and identification of putative sources of resistance for cassava brown streak disease. *PLoS ONE*, 9(5). doi:10.1371/journal.pone.0096642
- Maruthi, M. N., Hillocks, R. J., Mtunda, K., Raya, M. D., Muhanna, M., Kiozia, H., ... Thresh, J. M. (2005). Transmission of Cassava brown streak virus by *Bemisia tabaci* (Gennadius). *Journal of Phytopathology*, 153(5), 307–312. doi:10.1111/j.1439-0434.2005.00974.x
- Mathur, C., Jimsheena, V. K., Banerjee, S., Makinen, K., Gowda, L. R., & Savithri, H. S. (2012). Functional regulation of PVBV Nuclear Inclusion protein-a protease activity upon interaction with Viral Protein genome-linked and phosphorylation. *Virology*, 422(2), 254–264. doi:10.1016/j.virol.2011.10.009
- Mbanzibwa, D. R., Tian, Y., Mukasa, S. B., & Valkonen, J. P. T. (2009a). Cassava brown streak virus (Potyviridae) encodes a putative Maf/HAM1 pyrophosphatase implicated in reduction of mutations and a P1 proteinase that suppresses RNA silencing but contains no HC-Pro. *Journal of Virology*, 83, 6934–40. doi:10.1128/JVI.00537-09
- Mbanzibwa, D. R., Tian, Y. P., Tugume, A. K., Mukasa, S. B., Tairo, F., Kyamanywa, S., ... Valkonen, J. P. T. (2009b). Genetically distinct strains of Cassava brown streak virus in the Lake Victoria basin and the Indian Ocean coastal area of East Africa. *Archives of Virology*, 154, 353–359. doi:10.1007/s00705-008-0301-9
- Mbanzibwa, D. R., Tian, Y. P., Tugume, A. K., Mukasa, S. B., Tairo, F., Kyamanywa, S., ... Valkonen, J. P. T. (2011a). Simultaneous virus-specific detection of the two cassava brown streak-associated viruses by RT-PCR reveals wide distribution in East Africa, mixed infections, and infections in *Manihot glaziovii*. *Journal of Virological Methods*, 171(2), 394–400. doi:10.1016/j.jviromet.2010.09.024
- Mbanzibwa, D. R., Tian, Y. P., Tugume, A. K., Patil, B. L., Yadav, J. S., Bagewadi, B., ... Valkonen, J. P. T. (2011b). Evolution of cassava brown streak disease-associated viruses. *Journal of General Virology*, 92(4), 974–987. doi:10.1099/vir.0.026922-0
- Mbewe, W., Kumar, P. L., Changadeya, W., Ntawuruhunga, P., & Legg, J. (2015). Diversity, Distribution and Effects on Cassava Cultivars of Cassava Brown Streak Viruses in Malawi. *Journal of Phytopathology*, 163(6), 433–443. doi:10.1111/jph.12339

- McQuaid, C. F., Sseruwagi, P., Pariyo, A., & van den Bosch, F. (2016). Cassava brown streak disease and the sustainability of a clean seed system. *Plant Pathology*, 65(2), 299–309. doi:10.1111/ppa.12453
- Melrose, J., Perroy, R., & Careas, S. (2015). *World population prospects. United Nations, Department of Economic and Social Affairs* (Vol. ESA/P/WP.2). doi:10.1017/CBO9781107415324.004
- Merits, A., Guo, D., & Saarma, M. (1998). VPg, coat protein and five non-structural proteins of potato A potyvirus bind RNA in a sequence-unspecific manner. *Journal of General Virology*, 79, 3123–3127.
- Meshi, T., Ishikawa, M., Motoyoshi, F., Semba, K., & Okada, Y. (1986). *In-vitro* transcription of infectious RNAs from full-length cDNAs of tobacco mosaic virus. *Proceedings of the National Academy of Sciences*, 83(14), 5043–5047.
- Mingot, A., Valli, A., Rodamilans, B., San León, D., Baulcombe, D. C., García, J. A., & López-Moya, J. J. (2016). The P1N-PISPO Trans-frame Gene of Sweet Potato Feathery Mottle Potyvirus is Produced During Virus Infection and Functions as RNA Silencing Suppressor. *Journal of Virology*, 90(7), 3543–3557. doi:10.1128/JVI.02360-15
- Mohammed, I. U., Abarshi, M. M., Muli, B., Hillocks, R. J., & Maruthi, M. N. (2012). The symptom and genetic diversity of cassava brown streak viruses infecting cassava in East Africa. *Advances in Virology*, 2012, 1–10. doi:10.1155/2012/795697
- Monger, W. A., Alicai, T., Ndunguru, J., Kinyua, Z. M., Potts, M., Reeder, R. H., ... Smith, J. (2010). The complete genome sequence of the Tanzanian strain of Cassava brown streak virus and comparison with the Ugandan strain sequence. *Archives of Virology*, 155, 429–433. doi:10.1007/s00705-009-0581-8
- Monger, W. A., Seal, S., Cotton, S., & Foster, G. D. (2001). Identification of different isolates of Cassava brown streak virus and development of a diagnostic test. *Plant Pathology*, 50(6), 768–775. doi:10.1046/j.1365-3059.2001.00647.x
- Monger, W. A., Seal, S., Isaac, A. M., & Foster, G. D. (2001). Molecular characterization of the Cassava brown streak virus coat protein. *Plant Pathology*, 50, 527–534. doi:10.1046/J.1365-3059.2001.00589.X
- Nagyová, A., & Šubr, Z. (2007). Infectious full-length clones of plant viruses and their use for construction of viral vectors. *Acta Virologica*, 51, 223–237.
- Nakahara, K. S., & Masuta, C. (2014). Interaction between viral RNA silencing suppressors and host factors in plant immunity. *Current Opinion in Plant Biology*, 20, 88–95. doi:10.1016/j.pbi.2014.05.004
- Nakahara, K.S., Nishino, K., Uyeda, I. (2015). Construction of infectious cDNA clones derived from the potyviruses clover yellow vein virus and bean yellow mosaic virus. *Methods in Molecular Biology*, 1236, 219–227.

- Nassar, N. M. A. (2002). Cassava, *Manihot esculenta* Crantz Genetic Resources : The Origin of the Crop, its Evolution and Relationships of Wild Relatives. *Genetics and Molecular Research*, 4(1), 298–305.
- Nassar, N. M. A., & Ortiz, R. (2007). Cassava improvement: challenges and impacts. *The Journal of Agricultural Science*, 145(2), 163–171. doi:10.1017/S0021859606006575
- Ndunguru, J., & Kapinga, R. (2007). Viruses and virus-like diseases affecting sweet potato subsistence farming in southern Tanzania. *African Journal of Agricultural Research*, 2(5), 232–239.
- Ndunguru, J., Sseruwagi, P., Tairo, F., Stomeo, F., Maina, S., Djinkeng, A., ... Boykin, L. M. (2015). Analyses of twelve new whole genome sequences of cassava brown streak viruses and ugandan cassava brown streak viruses from East Africa: Diversity, supercomputing and evidence for further speciation. *PLoS ONE*, 10(10), 1–18. doi:10.1371/journal.pone.0139321
- Nichols, R. F. W. (1950). The brown streak disease of Cassava. Distribution, climatic effects and diagnostic symptoms. *East African Agricultural Journal*, 15(3), 154–160.
- Odipio, J., Ogwok, E., Taylor, N. J., Halsey, M., Bua, A., Fauquet, C. M., & Alicai, T. (2014). RNAi-derived field resistance to Cassava brown streak disease persists across the vegetative cropping cycle. *GM Crops & Food*, 5(1), 16–19. doi:10.4161/gmcr.26408
- Ogwok, E., Odipio, J., Halsey, M., Gaitán-Solís, E., Bua, A., Taylor, N. J., ... Alicai, T. (2012). Transgenic RNA interference (RNAi)-derived field resistance to cassava brown streak disease. *Molecular Plant Pathology*, 13, 1019–1031. doi:10.1111/j.1364-3703.2012.00812.x
- Ogwok, E., Patil, B. L., Alicai, T., & Fauquet, C. M. (2010). Transmission studies with Cassava brown streak Uganda virus (Potyviridae: Ipomovirus) and its interaction with abiotic and biotic factors in *Nicotiana benthamiana*. *Journal of Virological Methods*, 169(2), 296–304. doi:10.1016/j.jviromet.2010.07.030
- Olsen, B. S., & Johansen, I. E. (2001). Nucleotide sequence and infectious cDNA clone of the L1 isolate of Pea seed-borne mosaic potyvirus. *Archives of Virology*, 146(1), 15–25. doi:10.1007/s007050170187
- Olsen, K. M., & Scaal, B. (1999). Evidence on the Origin of Cassava : Phylogeography of *Manihot esculenta*. *PNAS*, 96(10), 5586–5591.
- Osogo, A. ., Muoma, J., Nyamwamu, P., Omuse, C. N., & Were, H. K. (2014). Occurrence and Distribution of Cassava Brown Streak Viruses in Western Kenya. *Journal of Agri-Food and Applied Sciences*, 2(7), 184–190.
- Pallas, V., & García, J. A. (2011). How do plant viruses induce disease? Interactions and interference with host components. *Journal of General Virology*, 92(12), 2691–2705. doi:10.1099/vir.0.034603-0

- Pasin, F., Simon-Mateo, C., & Garcia, J. A. (2014). The Hypervariable Amino-Terminus of P1 Protease Modulates Potyviral Replication and Host Defense Responses. *PLoS Pathogens*, 10(3). doi:10.1371/journal.ppat.1003985
- Patil, B. L., Legg, J. P., Kanju, E., & Fauquet, C. M. (2015). Cassava brown streak disease: A threat to food security in Africa. *Journal of General Virology*, 96(5), 956–968. doi:10.1099/vir.0.000014
- Patil, B. L., Ogowok, E., Wagaba, H., Mohammed, I. U., Yadav, J. S., Bagewadi, B., ... Fauquet, C. M. (2011). RNAi-mediated resistance to diverse isolates belonging to two virus species involved in Cassava brown streak disease. *Molecular Plant Pathology*, 12, 31–41. doi:10.1111/j.1364-3703.2010.00650.x
- Powell, P. A., Sanders, P. R., Tumer, N., Fraley, R. T., & Beachy, R. N. (1990). Protection against tobacco mosaic virus infection in transgenic plants requires accumulation of coat protein rather than coat protein RNA sequences. *Virology*, 175(1), 124–130. doi:10.1016/0042-6822(90)90192-T
- Prins, M., Laimer, M., Noris, E., Schubert, J., Wassenegger, M., & Tepfer, M. (2008). Strategies for antiviral resistance in transgenic plants. *Molecular Plant Pathology*, 9(1), 73–83. doi:10.1111/j.1364-3703.2007.00447.x
- Pruss, G., Ge, X., Shi, X. M., Carrington, J. C., & Bowman Vance, V. (1997). Plant viral synergism: the potyviral genome encodes a broad-range pathogenicity enhancer that transactivates replication of heterologous viruses. *The Plant Cell*, 9, 859–868. doi:10.1105/tpc.9.6.859
- Radford, A. D., Chapman, D., Dixon, L., Chantrey, J., Darby, A. C., & Hall, N. (2012). Application of next-generation sequencing technologies in virology. *Journal of General Virology*, 93(9), 1853–1868. doi:10.1099/vir.0.043182-0
- Rashid, U. J., Hoffmann, J., Brutschy, B., Piehler, J., & Chen, J. C. H. (2008). Multiple targets for suppression of RNA interference by tomato aspermy virus protein 2B. *Biochemistry*, 47(48), 12655–12657. doi:10.1021/bi801281h
- Ratcliff, F. G., MacFarlane, S. A., & Baulcombe, D. (1999). Gene Silencing without DNA: RNA-Mediated Cross-Protection between Viruses. *The Plant Cell*, 11(7), 1207–1215. doi:10.1105/tpc.11.7.1207
- Revers, F., Gall, O. L., Candresse, T., & Maule, A. J. (1999). New advances in understanding the molecular biology of plant/potyvirus interactions. *Molecular Plant-Microbe Interactions*, 12(5), 367–376.
- Revers, F., & García, J. A. (2015). Chapter 3: Molecular biology of potyviruses. *Advances in Virus Research*, 92, 101–99. doi:10.1016/bs.aivir.2014.11.006
- Riechmann, J. L., Lain, S., & Garcia, J. A. (1990). Infectious in-vitro transcripts from a plum pox potyvirus cDNA clone. *Virology*, 177(2), 710–716. doi:10.1016/0042-6822(90)90537-2

- Riechmann, J. L., Lain, S., & Garcia, J. A. (1992). Highlights and prospects of potyvirus molecular biology. *Journal of General Virology*, 73(1), 1–16. doi:10.1099/0022-1317-73-1-1
- Robaglia, C., Durand-Tardif, M., Tronchet, M., Boudazin, G., Astier-Manifacier, S., & Casse-Delbart, F. (1989). Nucleotide sequence of potato virus Y (N Strain) genomic RNA. *Journal of General Virology*, 70(4), 935–947.
- Sanford, J. C., & Johnston, S. A. (1985). The concept of parasite-derived resistance-Deriving resistance genes from the parasite's own genome. *Journal of Theoretical Biology*, 113(2), 395–405. doi:10.1016/S0022-5193(85)80234-4
- Satheshkumar, P. S., Gayathri, P., Prasad, K., & Savithri, H. S. (2005). "Natively unfolded" VPg is essential for Sesbania mosaic virus serine protease activity. *Journal of Biological Chemistry*, 280(34), 30291–30300. doi:10.1074/jbc.M504122200
- Schaad, M. C., Jensen, P. E., & Carrington, J. C. (1997). Formation of plant RNA virus replication complexes on membranes: role of an endoplasmic reticulum-targeted viral protein. *The EMBO Journal*, 16(13), 4049–4059.
- Schaad, M. C., Lellis, A. D., & Carrington J. (1997). VPg of tobacco etch potyvirus is a host genotype-specific determinant for long-distance movement. *Journal of Virology*, 71(11), 8624–8631.
- Scholthof, K. B. G., Adkins, S., Czosnek, H., Palukaitis, P., Jacquot, E., Hohn, T., ... Foster, G. D. (2011). Top 10 plant viruses in molecular plant pathology. *Molecular Plant Pathology*, 12(9), 938–954. doi:10.1111/j.1364-3703.2011.00752.x
- Schwartz, M., Chen, J., Janda, M., Sullivan, M., Den Boon, J., & Ahlquist, P. (2002). A positive-strand RNA virus replication complex parallels form and function of retrovirus capsids. *Molecular Cell*, 9(3), 505–514. doi:10.1016/S1097-2765(02)00474-4
- Sharma, R. C., & Schimke, R. T. (1996). Preparation of electro-competent E. coli using salt-free growth medium. *BioTechniques*, 20(1), 42–44. doi:10.1007/s13239-011-0061-7
- Sharp, P. A. (2001). RNA interference. *Genes Development*, 15, 485–490. doi:10.1101/gad.880001.et
- Sherman Fred. (2002). Getting started with yeast. *Methods in Enzymology*, 350, 3–41.
- Shukla, D. D., & Ward, C. W. (1988). Amino Acid Sequence Homology of Coat Proteins as a Basis for Identification and Classification of the Potyvirus Group. *Journal of General Virology*, 69(11), 2703–2710. doi:10.1099/0022-1317-69-11-2703
- Smith, J. J., Waage, J., Woodhall, J. W., Bishop, S. J., & Spence, N. J. (2008). The challenge of providing plant pest diagnostic services for Africa. *European Journal of Plant Pathology*, 121(3), 365–375. doi:10.1007/s10658-008-9311-4

- Soitamo, A. J., Jada, B., & Lehto, K. (2011). HC-Pro silencing suppressor significantly alters the gene expression profile in tobacco leaves and flowers. *BMC Plant Biology*, 11, 68. doi:10.1186/1471-2229-11-68
- Soosaar, J. L. M., Burch-Smith, T. M., & Dinesh-Kumar, S. P. (2005). Mechanisms of plant resistance to viruses. *Nature Reviews Microbiology*, 3(10), 789–98. doi:10.1038/nrmicro1239
- Storey, H.H. (1936a). Virus diseases of East African plants. *East African Agricultural Journal*, 2, 34–39.
- Storey, H.H. (1936b). Virus diseases of East African plants. VI-A progress report on studies of diseases of cassava. *East African Agricultural Journal*, 2, 34–39.
- Subr, Z., Fomitcheva, V. W., & Ühne, T. K. (2000). The complete nucleotide sequence of RNA1 of barley mild mosaic virus (ASL1) and attempts at bacterial expression of the P3 protein. *Journal of Phytopathology*, 148(7-8), 461–467. doi:10.1046/j.1439-0434.2000.00531.x
- Tairo, F., Mukasa, S. B., Jones, R. A. C., Kullaya, A., Rubaihayo, P. R., & Valkonen, J. P. T. (2005). Unravelling the genetic diversity of the three main viruses involved in Sweet Potato Virus Disease (SPVD), and its practical implications. *Molecular Plant Pathology*, 6(2), 199–211. doi:10.1111/j.1364-3703.2005.00267.x
- Taliansky, M., Roberts, I. M., Kalinina, N., Ryabov, E. V., Raj, S. K., Robinson, D. J., & Oparka, K. J. (2003). An umbraviral protein, involved in long-distance RNA movement, binds viral RNA and forms unique, protective ribonucleoprotein complexes. *Journal of Virology*, 77(5), 3031–3040. doi:10.1128/JVI.77.5.3031-3040.2003
- Taylor, N., Chavarriaga, P., Raemakers, K., Siritunga, D., & Zhang, P. (2004). Development and application of transgenic technologies in cassava. *Plant Molecular Biology*, 56(4), 671–688. doi:10.1007/s11103-004-4872-x
- Thivierge, K., Cotton, S., Dufresne, P. J., Mathieu, I., Beauchemin, C., Ide, C., ... Laliberte, J. F. (2008). Eukaryotic elongation factor 1A interacts with Turnip mosaic virus RNA-dependent RNA polymerase and VPg-Pro in virus-induced vesicles. *Virology*, 377(1), 216–225. doi:10.1016/j.virol.2008.04.015
- Thresh, J. M. (2006). Control of Tropical Plant Virus Diseases. *Advances in Virus Research*, 67, 245–295.
- Thresh, J. M., Fargette, D., & Otim-Nape, G. (2009). The viruses and virus diseases of cassava in Africa.
- Thresh, J. M., & Mbwana M. (1998). Cassava mosaic and cassava brown streak diseases in Zanzibar. *Roots*, 5(1), 6–8.
- Tomlinson, J. A., Ostojka-Starzewska, S., Adams, I. P., Miano, D. W., Abidrabo, P., Kinyua, Z., ... Smith, J. (2013). Loop-mediated isothermal amplification for rapid detection of

the causal agents of cassava brown streak disease. *Journal of Virological Methods*, 191(2), 148–154. doi:10.1016/j.jviromet.2012.07.015

Torres-Barcelo, C., Daros, J. A., & Elena, S. F. (2010). HC-Pro hypo- and hypersuppressor mutants: Differences in viral siRNA accumulation in vivo and siRNA binding activity in vitro. *Archives of Virology*, 155(2), 251–254. doi:10.1007/s00705-009-0563-x

Torres-Barceló, C., Martín, S., Daròs, J. A., Elena, S. F., Torres-Barcel, C., Martin, S., ... Elena, S. F. (2008). From hypo- to hypersuppression: Effect of amino acid substitutions on the RNA-silencing suppressor activity of the Tobacco etch potyvirus HC-pro. *Genetics*, 180(2), 1039–1049. doi:10.1534/genetics.108.091363

Tugume, A. K., Mukasa, S. B., Kalkkinen, N., & Valkonen, J. P. T. (2010). Recombination and selection pressure in the ipomovirus sweet potato mild mottle virus (Potyviridae) in wild species and cultivated sweetpotato in the centre of evolution in East Africa. *Journal of General Virology*, 91(4), 1092–1108. doi:10.1099/vir.0.016089-0

Urcuqui-Inchima, S., Maia, I. G., Arruda, P., Haenni, a L., & Bernardi, F. (2000). Deletion mapping of the potyviral helper component-proteinase reveals two regions involved in RNA binding. *Virology*, 268(1), 104–111. doi:10.1006/viro.1999.0156

Valli, A., Dujovny, G., & García, J. A. (2008). Protease activity, self interaction, and small interfering RNA binding of the silencing suppressor p1b from cucumber vein yellowing ipomovirus. *Journal of Virology*, 82(2), 974–86. doi:10.1128/JVI.01664-07

Valli, A., Gallo, A., Calvo, M., Perez, J. D. J., & Garcia, J. A. (2014). A novel role of the potyviral Helper Component proteinase contributes to enhance the yield of viral particles. *Journal of Virology*, 88(17), 9808–9818. doi:10.1128/JVI.01010-14

Valli, A., Lopez-Moya, J. J., & Garcia, 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:10.1099/vir.0.82402-0

Valli, A., Montserrat Martín-Hernández, A., López-Moya, J. J., & García, J. A. (2006). RNA Silencing Suppression by a Second Copy of the P1 Serine Protease of Cucumber Vein Yellowing Ipomovirus, a Member of the Family Potyviridae That Lacks the Cysteine Protease HCPro. *Journal of Virology*, 80(20), 10055–10063. doi:10.1128/JVI.00985-06

Van Dun, C. M. P., Overduin, B., Van Vloten-Doting, L., & Bol, J. F. (1988). Transgenic tobacco expressing tobacco steak virus or mutated alfalfa mosaic virus coat protein does not protect against alfalfa mosaic virus infection. *Virology*, 164(2), 383–389.

Vance, B. V., Berger, P. H., Carrington, J. C., Hunt, A. G., & Ming Shi, X. (1995). 5' Proximal potyviral sequences mediate potato virus X/potyviral synergistic disease in transgenic tobacco. *Virology*, 206(1), 583–590. doi:10.1016/S0042-6822(95)80075-1

Vanderschuren, H., Moreno, I., Anjanappa, R. B., Zainuddin, I. M., & Grussem, W. (2012). Exploiting the Combination of Natural and Genetically Engineered Resistance to

- Cassava Mosaic and Cassava Brown Streak Viruses Impacting Cassava Production in Africa. *PLoS ONE*, 7(9), e45277.
- Verchot, J., & 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.
- Verchot, J., Koonin, E. V., & Carrington, J. C. (1991). The 35-kDa protein from the N-terminus of the potyviral polyprotein functions as a third virus-encoded proteinase. *Virology*, 185(2), 527–535. doi:10.1016/0042-6822(91)90522-D
- Verchot-Lubicz, J., Ye, C.-M. M., & Bamunusinghe, D. (2007). Molecular biology of potexviruses: recent advances. *Journal of General Virology*, 88(6), 1643–1655. doi:10.1099/vir.0.82667-0
- Voinnet, O. (2001). RNA silencing as a plant immune system against viruses. *Trends in Genetics*, 17(8), 449–459. doi:10.1016/S0168-9525(01)02367-8
- Voinnet, O. (2005). Induction and suppression of RNA silencing: insights from viral infections. *Nature Reviews Genetics*, 6(3), 206–220. doi:10.1038/nrg1555
- Vos, P., Jaegle, M., Wellink, J., Verver, J., Eggen, R., Van Kammen, A., & Goldbach, R. (1988). Infectious RNA transcripts derived from full-length DNA copies of the genomic RNAs of cowpea mosaic virus. *Virology*, 165(1), 33–41. doi:10.1016/0042-6822(88)90655-1
- Wagaba, H., Beyene, G., Trembley, C., Alicai, T., Fauquet, C. M., & Taylor, N. J. (2013). Efficient transmission of cassava brown streak disease viral pathogens by chip bud grafting. *BMC Research Notes*, 6(1), 1. doi:10.1186/1756-0500-6-516
- Wagaba, H., Patil, B. L., Mukasa, S., Alicai, T., Fauquet, C. M., & Taylor, N. J. (2016). Artificial microRNA-derived resistance to Cassava brown streak disease. *Journal of Virological Methods*, 231, 38–43. doi:10.1016/j.jviromet.2016.02.004
- Wang, A. (2015). Dissecting the molecular network of virus-plant interactions: the complex roles of host factors. *Annual Review of Phytopathology*, 53, 45–66. doi:10.1146/annurev-phyto-080614-120001
- Wang, A., & Krishnaswamy, S. (2012). Eukaryotic translation initiation factor 4E-mediated recessive resistance to plant viruses and its utility in crop improvement. *Molecular Plant Pathology*, 13(7), 795–803.
- Wang, M.-B., Masuta, C., Smith, N. A., & Shimura, H. (2012). RNA Silencing and Plant Viral Diseases. *Molecular Plant-Microbe Interactions*, 25(10), 1275–1285. doi:10.1094/MPMI-04-12-0093-CR
- Wang, X., Kohalmi, S. E., Svircev, A., Wang, A., Sanfaçon, H., & Tian, L. (2013). Silencing of the Host Factor eIF(iso)4E Gene Confers Plum Pox Virus Resistance in Plum. *PLoS ONE*, 8(1), e50627. doi:10.1371/journal.pone.0050627

- Webb, S., Hiebert, E., & Kucharek, T. (2003). Identity and distribution of viruses infecting cucurbits in Florida. *Phytopathology*, 93, S89.
- Winter, S., Koerbler, M., Stein, B., Pietruszka, A., Paape, M., & Butgereitt, A. (2010). Analysis of cassava brown streak viruses reveals the presence of distinct virus species causing cassava brown streak disease in East Africa. *Journal of General Virology*, 91(5), 1365–1372. doi:10.1099/vir.0.014688-0
- Yadav, J. S., Ogwok, E., Wagaba, H., Patil, B. L., Bagewadi, B., Alicai, T., ... Fauquet, C. M. (2011). RNAi-mediated resistance to Cassava brown streak Uganda virus in transgenic cassava. *Molecular Plant Pathology*, 12(7), 677–687. doi:10.1111/j.1364-3703.2010.00700.x
- Yambao, M. L. M., Yagihashi, H., Sekiguchi, H., Sekiguchi, T., Sasaki, T., Sato, M., ... Uyeda, I. (2008). Point mutations in helper component protease of clover yellow vein virus are associated with the attenuation of RNA-silencing suppression activity and symptom expression in broad bean. *Archives of Virology*, 153(1), 105–115. doi:10.1007/s00705-007-1073-3
- Yang, S. J., Revers, F., Souche, S., Lot, H., Le Gall, O., Candresse, T., & Dunez, J. (1998). Construction of full-length cDNA clones of lettuce mosaic virus (LMV) and the effects of intron-insertion on their viability in *Escherichia coli* and on their infectivity to plants. *Archives of Virology*, 143(12), 2443–2451. doi:10.1007/s007050050474
- Youssef, F., Marais, A., Faure, C., Gentit, P., & Candresse, T. (2011). Strategies to facilitate the development of uncloned or cloned infectious full-length viral cDNAs: Apple chlorotic leaf spot virus as a case study. *Virology Journal*, 8(1), 1. doi:10.1186/1743-422X-8-488
- Zhang, X., Li, H., Zhang, J., Zhang, C., Gong, P., Ziaf, K., ... Ye, Z. (2011). Expression of artificial microRNAs in tomato confers efficient and stable virus resistance in a cell-autonomous manner. *Transgenic Research*, 20(3), 569–581. doi:10.1007/s11248-010-9440-3
- Zhang, X., Yuan, Y. R., Pei, Y., Lin, S. S., Tuschl, T., Patel, D. J., & Chua, N. H. (2006). Cucumber mosaic virus-encoded 2b suppressor inhibits Arabidopsis Argonaute1 cleavage activity to counter plant defense. *Genes and Development*, 20(23), 3255–3268. doi:10.1101/gad.1495506

Appendix 1: Complete genome of UCBSV wildtype Kikombe isolate (9096 Nucleotides) GenBank Accession Number KX753356

[illegible]

Appendix 3: Alignment of the Nucleotide sequence between the UCBSV wildtype and UCBSV infectious clone genomes (99% homology)

[illegible]

3571	3572	3573	3574	3575	3576	3577	3578	3579	3580	3581	3582	3583	3584	3585	3586	3587	3588	3589	3590	3591	3592	3593	3594	3595	3596	3597	3598	3599	3600	3601	3602	3603	3604	3605	3606	3607	3608	3609	3610	3611	3612	3613	3614	3615	3616	3617	3618	3619	3620	3621	3622	3623	3624	3625	3626	3627	3628	3629	3630	3631	3632	3633	3634	3635	3636	3637	3638	3639	3640	3641	3642	3643	3644	3645	3646	3647	3648	3649	3650	3651	3652	3653	3654	3655	3656	3657	3658	3659	3660	3661	3662	3663	3664	3665	3666	3667	3668	3669	3670	3671	3672	3673	3674	3675	3676	3677	3678	3679	3680	3681	3682	3683	3684	3685	3686	3687	3688	3689	3690	3691	3692	3693	3694	3695	3696	3697	3698	3699	3700	3701	3702	3703	3704	3705	3706	3707	3708	3709	3710	3711	3712	3713	3714	3715	3716	3717	3718	3719	3720	3721	3722	3723	3724	3725	3726	3727	3728	3729	3730	3731	3732	3733	3734	3735	3736	3737	3738	3739	3740	3741	3742	3743	3744	3745	3746	3747	3748	3749	3750	3751	3752	3753	3754	3755	3756	3757	3758	3759	3760	3761	3762	3763	3764	3765	3766	3767	3768	3769	3770	3771	3772	3773	3774	3775	3776	3777	3778	3779	3780	3781	3782	3783	3784	3785	3786	3787	3788	3789	3790	3791	3792	3793	3794	3795	3796	3797	3798	3799	3800	3801	3802	3803	3804	3805	3806	3807	3808	3809	3810	3811	3812	3813	3814	3815	3816	3817	3818	3819	3820	3821	3822	3823	3824	3825	3826	3827	3828	3829	3830	3831	3832	3833	3834	3835	3836	3837	3838	3839	3840	3841	3842	3843	3844	3845	3846	3847	3848	3849	3850	3851	3852	3853	3854	3855	3856	3857	3858	3859	3860	3861	3862	3863	3864	3865	3866	3867	3868	3869	3870	3871	3872	3873	3874	3875	3876	3877	3878	3879	3880	3881	3882	3883	3884	3885	3886	3887	3888	3889	3890	3891	3892	3893	3894	3895	3896	3897	3898	3899	3900	3901	3902	3903	3904	3905	3906	3907	3908	3909	3910	3911	3912	3913	3914	3915	3916	3917	3918	3919	3920	3921	3922	3923	3924	3925	3926	3927	3928	3929	3930	3931	3932	3933	3934	3935	3936	3937	3938	3939	3940	3941	3942	3943	3944	3945	3946	3947	3948	3949	3950	3951	3952	3953	3954	3955	3956	3957	3958	3959	3960	3961	3962	3963	3964	3965	3966	3967	3968	3969	3970	3971	3972	3973	3974	3975	3976	3977	3978	3979	3980	3981	3982	3983	3984	3985	3986	3987	3988	3989	3990	3991	3992	3993	3994	3995	3996	3997	3998	3999	4000
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

[illegible]

Appendix 4: Alignment of amino acid sequence (ORF) between UCBSV Infectious clone and UCBSV wildtype

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80																				

Appendix 5: Translation of UCBSV Genome with cleavage sites highlighted

[illegible]

Appendix 6: Nucleotide alignments of P1 sequences from different UCSV and CBSV isolates

[illegible]

PL UCSBY Kik
 KR102837
 KR102838
 KR102839
 KR102840
 KR102841
 KR102842
 KR102843
 KR102844
 KR102845
 KR102846
 KR102847
 KR102848
 KR102849
 KR102850
 KR102851
 KR102852
 KR102853
 KR102854
 KR102855
 KR102856
 KR102857
 KR102858
 KR102859
 KR102860
 KR102861
 KR102862
 KR102863
 KR102864
 KR102865
 KR102866
 KR102867
 KR102868
 KR102869
 KR102870
 KR102871
 KR102872
 KR102873
 KR102874
 KR102875
 KR102876
 KR102877
 KR102878
 KR102879
 KR102880
 KR102881
 KR102882
 KR102883
 KR102884
 KR102885
 KR102886
 KR102887
 KR102888
 KR102889
 KR102890
 KR102891
 KR102892
 KR102893
 KR102894
 KR102895
 KR102896
 KR102897
 KR102898
 KR102899
 KR102900
 KR102901
 KR102902
 KR102903
 KR102904
 KR102905
 KR102906
 KR102907
 KR102908
 KR102909
 KR102910
 KR102911
 KR102912
 KR102913
 KR102914
 KR102915
 KR102916
 KR102917
 KR102918
 KR102919
 KR102920
 KR102921
 KR102922
 KR102923
 KR102924
 KR102925
 KR102926
 KR102927
 KR102928
 KR102929
 KR102930
 KR102931
 KR102932
 KR102933
 KR102934
 KR102935
 KR102936
 KR102937
 KR102938
 KR102939
 KR102940
 KR102941
 KR102942
 KR102943
 KR102944
 KR102945
 KR102946
 KR102947
 KR102948
 KR102949
 KR102950
 KR102951
 KR102952
 KR102953
 KR102954
 KR102955
 KR102956
 KR102957
 KR102958
 KR102959
 KR102960
 KR102961
 KR102962
 KR102963
 KR102964
 KR102965
 KR102966
 KR102967
 KR102968
 KR102969
 KR102970
 KR102971
 KR102972
 KR102973
 KR102974
 KR102975
 KR102976
 KR102977
 KR102978
 KR102979
 KR102980
 KR102981
 KR102982
 KR102983
 KR102984
 KR102985
 KR102986
 KR102987
 KR102988
 KR102989
 KR102990
 KR102991
 KR102992
 KR102993
 KR102994
 KR102995
 KR102996
 KR102997
 KR102998
 KR102999
 KR103000

PL UCSBY Kik
 KR102937
 KR102938
 KR102939
 KR102940
 KR102941
 KR102942
 KR102943
 KR102944
 KR102945
 KR102946
 KR102947
 KR102948
 KR102949
 KR102950
 KR102951
 KR102952
 KR102953
 KR102954
 KR102955
 KR102956
 KR102957
 KR102958
 KR102959
 KR102960
 KR102961
 KR102962
 KR102963
 KR102964
 KR102965
 KR102966
 KR102967
 KR102968
 KR102969
 KR102970
 KR102971
 KR102972
 KR102973
 KR102974
 KR102975
 KR102976
 KR102977
 KR102978
 KR102979
 KR102980
 KR102981
 KR102982
 KR102983
 KR102984
 KR102985
 KR102986
 KR102987
 KR102988
 KR102989
 KR102990
 KR102991
 KR102992
 KR102993
 KR102994
 KR102995
 KR102996
 KR102997
 KR102998
 KR102999
 KR103000

PL UCSBY Kik
 KR102937
 KR102938
 KR102939
 KR102940
 KR102941
 KR102942
 KR102943
 KR102944
 KR102945
 KR102946
 KR102947
 KR102948
 KR102949
 KR102950
 KR102951
 KR102952
 KR102953
 KR102954
 KR102955
 KR102956
 KR102957
 KR102958
 KR102959
 KR102960
 KR102961
 KR102962
 KR102963
 KR102964
 KR102965
 KR102966
 KR102967
 KR102968
 KR102969
 KR102970
 KR102971
 KR102972
 KR102973
 KR102974
 KR102975
 KR102976
 KR102977
 KR102978
 KR102979
 KR102980
 KR102981
 KR102982
 KR102983
 KR102984
 KR102985
 KR102986
 KR102987
 KR102988
 KR102989
 KR102990
 KR102991
 KR102992
 KR102993
 KR102994
 KR102995
 KR102996
 KR102997
 KR102998
 KR102999
 KR103000

Appendix 7: Amino acid alignments between P1 protein sequences of different UCSV and CBSV isolates

[illegible]

Appendix 8: Nucleotide alignments of HAM1 sequences from different UCSV and CBSV isolates

[illegible]

[illegible][illegible]

Appendix 9: Amino acid alignments between HAM1 sequences between UCSV and CBSV isolates

[illegible][illegible]