Expanding the knowledge of Genetically Modified Cassava through development of stacked traits for resistance and enhanced starch

by

Helen Ann Walsh

(474725)

Thesis

Submitted in fulfilment of the requirements for the degree

Philosophiae Doctor

in

Molecular and Cell Biology

in the Faculty of Science, University of the Witwatersrand, Johannesburg, South Africa

Supervisor: Prof. MEC Rey

August 2019

Declaration

I, Helen Ann Walsh (474725), am a student registered for the degree of Doctor of Philosophy (MCBG 9002) in the academic year 2019. I hereby declare the following

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where explicitly indicated otherwise and acknowledged.
- I have not submitted this work before for any other degree or examination at this or any other University.
- The information used in the Thesis HAS NOT been obtained by me while employed by, or working under the aegis of, any person or organisation other than the University. •
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.

Signature:  Date: 2nd of August, 2019.

Acknowledgements

- First and foremost I would like to thank my Supervisor, Prof. Chrissie Rey, without whom it would not have been possible to complete this work. I would like to thank her for guidance, her endless enthusiasm, patience and kindness.
- I would like to thank my work colleagues, both in the Cassava Biotechnology Research group and the wider School of Molecular Biology, for their support, and guidance.
- Specifically I would like to thank Angela Eni, who guided me through my first few months of research and was endlessly patient with me. Maabo Moralo, who has forgotten more about growing and transforming cassava than I will ever know and Imah Mwaba who was always willing to share her vast knowledge.
- I would like to thank Frances Ayres, for always being willing to go the extra mile to make sure the lab runs smoothly.
- I would like to thank the National research foundation (NRF), University of the Witwatersrand, Technology innovation agency (TIA) and the SA-Swiss Bilateral Grant for financial support.
- Lastly I would like to thank my family; especially my parents who have been unendingly supportive, generous and kind. I could have never have achieved this without them. And Joe Penn, who has endured all this too, and still brought me tea.

Dedication

This thesis is dedicated to my granddad, Ernest Benfield.

Research Outputs

H.A Walsh and M.E.C Rey: (Presentation) Development of a cassava mosaic disease resistant cassava through enhanced post-transcriptional gene silencing. South African society of Microbiology Conference, Durban, 2016

H.A Walsh and M.E.C Rey: (Poster) Expanding the knowledge of Genetically Modified Cassava through development of stacked traits for resistance and enhanced starch, 49th Conference of South African Society of Plant Pathology, Bloemfontein. 2016

H.A Walsh and M.E.C Rey: (Poster and Presentation) Toxic effects on transformed friable embryogenic callus by a triple-stacked hpRNA- silencing construct derived from three cassava geminiviruses, 8th International Geminivirus Symposium and the 6th International ssDNA Comparative Virology Workshop, India. 2017.

H.A. Walsh, H. Vanderschuren and M.E.C Rey: **(**Paper**)** RNA silencing of *South African cassava mosaic virus* in AC1/AC4-expressing RNA transgenic cassava cv.60444 induces tolerance, submitted to New Biotechnology.

Abstract

Cassava (*Manihot esculenta* Crantz) is ranked as the world's fourth most important food crop (FAO, 2016). Cassava mosaic disease (CMD), caused by 9 species of cassava mosaic geminiviruses including *South African cassava mosaic virus*, *African cassava mosaic virus* and *East African cassava mosaic Cameroon virus* (Brown *et al*, 2015), is one of the greatest hurdles to cassava cultivation in sub-Saharan Africa. The most promising method of improving cassava is through transgenic technology. The over-all aim of this project was the improvement of cassava through; improved disease resistance to *South African cassava mosaic virus,* and improved starch yield though the down-regulation of two genes involved in starch synthesis: plastidial *Adenylate kinase* and *Uridine-5'-monophosphate synthase*. In addition to these aims, another study conducted here investigated the off-target effects caused by a triple stacked construct targeting African cassava mosaic virus (ACMV), East African cassava mosaic virus (EACMV) and South African cassava mosaic virus (SACMV) which was found to be toxic to cassava friable embryogenic callus (FECs).

Previously, in order to increase CMD-susceptible cassava model cultivar 60444 resistance to SACMV, a RNA hairpin construct derived from the overlapping region of the replication associated protein (AC1)/ putative virus suppressor (AC4) of SACMV was used to transform cv.60444 (Taylor, 2009). This study reports on the screening of transformed lines for SACMV resistance. Three independent lines: O13-5, O13-8 and O12-2 showed tolerance to SACMV, with decreased symptom severity and viral loads compared to untransformed cv.60444 control. One line, O13-8, also showed a recovery phenotype associated with the resistant TME3 cultivar. This is the first time tolerance to SACMV has been reported in cassava cv.60444.

The second aim of project was further elucidate the mechanism by which a triple stacked construct (pC-AES) targeting the AC1/AC4 overlap of ACMV, as well as the AC1/IR of SACMV and EACMV, induced 'off-target effects' in cv.60444, in order to improve virus target selection for future studies. This construct produced 'off-target effects' including high levels of mortality and very low transformation efficiency in FECs and regenerated lines also showed aberrant phenotypes including stunted growth and misshaped leaves. In order to confirm that the observed off-target effects were due to the stacked construct, possible siRNAs associated with the triple

stacked construct were predicted bioinformatically, and siRNAs were compared to the cassava genome. Ten predicted gene targets which had partial homology to the predicted siRNA were identified, seven of which showed differential gene expression in FECs transformed with the pC-AES triple construct. The differentially regulated genes are involved in plant development, as well as defence response. These results show the limitations of PTGS, where the expression viral targets could interfere with the host biology.

The third aim of this project was to increase starch yields through the down regulation of two genes plastidial *Adenylate kinase* (ADK) and *Uridine-5'-monophosphate* (UMP) *synthase,* which are involved in starch synthesis. Hairpin-RNA constructs targeting *ADK* and *UMP synthase*, inserted into pCambia 1305.1 transformation vectors were used to transform cassava cv.60444 FECs. Ten independent UMPS lines and eight independent ADK lines were produced and were assessed over a 10 month period for tuber development. Two ADK lines, (1 and 5) and four UMPS lines (1, 2, 13 and 17) as well as control cv.60444 produced storage roots. Two of the UMPS lines storage roots (13 and 17) were significantly greater in size and weight than the untransformed regenerated control cv.60444 and three of the lines (1, 13 and 17) had a dry weight percentage mass higher than 25%, indicating high starch content. Analysis of the relative expression of *UMP synthase* in the 4 lines showed that there was a correlation between the increased storage root weight and the decreased expression of *UMP synthase*. The ADK lines did not produce any storage roots that were significantly greater than the control lines, although ADK 5 showed down-regulation of the *ADK*. Due to time constraints, no further testing of the starch content of these lines was performed, but this will be done in future studies. This is the first report of increased in storage root yields in cassava cv.60444 transformed with hp-RNA targeting *UMP synthase*.

This study contributes to improvement of cassava cv.60444, through the use of RNAi technology. Although several cassava lines have been developed which show increased resistance to CMG including ACMV and Sri Lankan cassava mosaic virus, this is the first report of tolerance to SACMV. Further, the improved storage root yield shown by UMPS lines increases the commercial value of cassava. These results could also contribute to further improvement in cassava, as the transgenes could be

stacked to further improve yield. Further understanding of the production of 'off-targets' can also be used in improve transgene selection.

Table of Figures:

Table of Tables:

Table of Appendices

List of Abbreviations

ABA- Abscisic acid

AC1- Replication-associated protein

AC2- Transcriptional activator protein

AC3- Replication enhancer protein

AC4- Putative viral suppressor protein

AC5- Putative silencing suppressor protein

ACMV-CM- African cassava mosaic virus - Cameroon

ACMV-NOg- African cassava mosaic virus - Nigeria

ADK- Adenylate kinase

ADP- Adenosine diphosphate

AGO- Argonaut protein

AGPase- ADP-glucose pyrophosphorylase

AMIR- artificial miRNA induced resistance

AMP-Adenosine monophosphate

AS2 - ASSYMETRICAL LEAVES 2

AS-PTGS- Anti-sense post-transcriptional gene silencing

ATP- Adenosine triphosphate

AV1- Coat Protein

AV2- Pathogenicity determinant

B33- Patatin class I promoter

BAP- Benzylaminopurine

BC1 - Movement protein

BV1- Nuclear shuttle protein

CaMV- Cauliflower mosaic virus

CBM- Cassava basic medium

CBSV- Cassava brown streak virus

CEM- Cassava elongation medium

CMD- Cassava mosaic disease

CMG-Cassava mosaic geminivirus

CMT3- CHROMOMETHYLASE 3

CMV- Cucumber mosaic virus

CRISPR- Clustered Regularly Interspaced Short Palindromic Repeats

Cas9- CRISPR-associated 9 protein

DCP- Decapping protein

DCL- DICER-like proteins

DPI- Days post inoculation

DRB- DOUBLE-STRANDED RNA BINDING

EACMCV- East African cassava mosaic Cameroon virus

EACMV- East African cassava mosaic virus

FAC1- EMBRYONIC FACTOR 1

FEC- Friable embryogenic callus

FFECs- Fresh FECs

FPS- Farnesyl pyrophosphate synthase

GBSS- GRANULE BOUND STARCH SYNTHASE

GD- Gresshoff and Doy

GEM- GL2-EXPRESSION MODULATOR

GFP- Green fluorescent protein

GRAM - Glycosyltransferases, Rab-like GTPase Activators, Myotubularins

gRNA- Guide RNA strand

HDR- homology-directed repair

HEN- HUA-Enhancer

HST- HASTY protein

Het-siRNA- Heterochromatic siRNAs

IAA- Indole-3-acetic acid

ICMV- Indian cassava mosaic virus

IR- Intergenic region

LB- Luria broth

LEC- LEAFY COTYLEDON

MET- METHYLTRANSFERASE

MP- Movement protein

MS- Murashige and Skoog

MYMV- Mungbean yellow mosaic virus

mRNA- Messenger RNA

MicroRNA- miRNA

Mt-Metric tons

NAA- Naphthaleneacetic acid

NHEJ- non-homologous end joining

NOS- Nopaline synthase

NSP- Nuclear shuttle protein

OFECs- Old FECs

ORF- Open Reading Frame

PAM -protospacer motif

PCNA- proliferating cell nuclear antigen

PCR- Polymerase chain reaction

PDR- Pathogen-derived resistance

PSY- Phytoene synthase gene

PTGS- Post-transcriptional gene silencing

PTST1- PROTEIN TARGETING TO STARCH

PolII- RNA polymerase II

Pre-miRNA- Precursor miRNA

Pri-miRNA- Primary miRNA

QTL- Quantitative trait locus

RCA- Rolling-circle amplification

RDR- RNA-dependent RNA polymerase

REn- Replication enhancer protein

RISC- RNA-induced silencing complex

RITS- RNA-induced transcriptional silencing complex

RNAi- RNA interference

ROS- Reactive oxygen species

RT-PCR- Reverse transcriptase PCR

Rep- Replication associated protein

S-PTGS- Sense post-transcriptional gene silencing

SACMV- South African cassava mosaic virus

SDE- SILENCING DEFECTIVE

SE- Somatic embryo

SEGS- Sequences enhancing geminivirus symptoms

SGS- Suppressor of GENE SILENCINIG

SLCMV- Sri-Lankan cassava mosaic virus

SNP- Single nucleotide polymorphism

SS- Starch synthase

SSS-Soluble starch synthase

SUVH4/KYP- Su(var)3-9 homolog 4/Kryptonite

TALENs- transcription activator-like effector nucleases

TCP-TEOSINTE BRANCHED, CYCLOIDEA, and PROLIFERATING CELL NUCLEAR ANTIGEN BINDING FACTOR

TEV- Tobacco etch virus

TGMV- Tomato Golden mosaic virus

TGS- Transcriptional gene silencing

TME-Tropical *Manihot esculenta*

TMS-Tropical Manioc Selection

ToRSV - Tomato ringspot virus

TrAP- Transcriptional activator protein

TYLCV- Tomato Yellow Leaf Curl Virus

ToLCNDV- Tomato leaf curl New Delhi virus

U- units

UBQ10 - Ubiquitin 10

UTR- Untranslated region

VSR- Viral silencing suppressor

WRKY 14- WRKY transcription factor 14

XET- Xyloglucan endotransglucosylase

YEP- Yeast extract-phosphate

cv- Cultivar

dsDNA- Double-stranded DNA

dsRNA- Double-stranded RNA

ea-siRNA- epigenetically-activated siRNA

g- Gram

h- Hour

hp-RNA- Hairpin RNA

lncRNA- Long non-coding RNA

mM- Micromolar

mRNA- messenger RNA

min- Minute

nat-siRNA- Natural anti-sense short interfering RNA

nt- nucleotide

pRb- Retinoblastoma protein

pha-siRNA - phased siRNAs

prxs- Peroxidase

qPCR- Quantitative PCR

rpm- revolutions per minute

sRNA- Small RNA

siRNA- Small interfering RNA

ssDNA- Single stranded DNA

ssRNA- Single stranded RNA

ta-siRNA- trans-acting small interfering RNA

µg- Microgram

vsRNA- Viral small RNA

vsiRNA- Virus-derived small interfering RNA

µM- Micromolar

µl- Microlitre

Chapter 1: Literature Review

1.1 Cassava

Cassava (*Manihot esculenta* Crantz) is a perennial New World shrub, grown in tropical and sub-tropical regions of the world for its starchy, tuberous roots (El-Sharkawy, 2004). Cassava is mainly grown by subsistence farmers as a food crop (Patil and Fauquet, 2009) however; the large range of industrial products that can be made from processed cassava storage roots means that this status is changing and cassava is considered one of the fastest expanding food crops (FAO, 2016).

Cassava was first cultivated as a crop in South America, and is thought to have been bred from a wild progenitor (*Manihot esculenta* subspecies *flabellifolia*) in the Brazil (Allem and Postal, 1994; Olsen and Schaal, 1999). Breeding of the domestic crop was achieved through interspecific hybridisation with interogression between cassava and the wild species (Boster, 1985; Rogers, 1963). Cassava is a shrub measuring between 1-4m in height (in the field). There are two distinct plant types: spreading or, erect with or without branching tops. Due to the high degree of interspecific hybridisation cassava morphological characteristics vary widely (Alves, 2002) with over 6400 different cultivars held at CIAT and EMBRAPAS alone (Alves, 2002).

Cassava can be propagated by seed or by stem cuttings, however stem cuttings is the most common method used (Alves, 2002), because cassava production using seeds increases time to production. Cassava is mainly grown for its tuberous roots which are the main storage organ of cassava. These roots develop from the primary tap root (in the case of cassava borne from seed) and adventitious roots (from stem cuttings). Storage roots develop only after the primary roots have developed to form the fibrous roots (Alves, 2002). The cassava root is not a tuberous root and therefore cannot be used for vegetative propagation. It is made up of three layers: periderm, cortex and parenchyma which is the edible portion of the root where the majority of the starch is stored (Alves, 2002; El-Sharkawy, 2004).

Cassava is grown widely in tropical and sub-tropical areas of the world, where it is used mainly as a staple food crop and as a source of animal feed (El-Sharkawy,

2004). Globally there has been an increase in the demand for cassava (2.6% between 2011- 2016) with 253.4 million tons produced in 2018 (FAO, 2018). Cassava roots and leaves can be consumed, with the roots providing a rich source of carbohydrates and the leaves providing a source of proteins and minerals. Although the starch content can vary between cultivars, typically cassava roots are composed of 24% starch and 1% protein (FAO, 2000a). In addition to its use as a food stuff, cassava is used globally as an animal feed, with up to one fourth of all cassava produced globally used an animal feed for pigs, poultry fish and cattle (Bellotti *et al*., 1999) (FAO, 2018), however this varied based on geographical regions. Cassava roots are high in starch and as such also have a wide range of industrial applications including paper, glue, cardboard and the production of bioethanol (the production of ethanol from starch) (Westerbergh *et al*, 2009; Zeeman *et al*, 2010).

Cassava is drought tolerant, and can be grown in nutrient poor soils, which makes it ideal for growing in areas where other crops cannot survive. However the production of cassava is affected by a number of factors including socioeconomic factors (Balat and Balat, 2009) abiotic factors and biotic factors such as pests and diseases (Patil and Fauquet, 2009). Globally cassava production is threatened by a number of emerging and reoccurring pests and problems (Campo *et al*., 2011). In Africa the spread of cassava brown streak disease, caused by cassava brown streak virus (CBSV) and Ugandan cassava brown streak virus (UCBSV) (Mbanzibwa *et al*., 2011; Monger *et al*., 2001) has resulted in crop losses of up to 80% in East Africa (FAO, 2000b). Cassava mosaic disease (CMD) which is caused by a number of cassava mosaic geminiviruses (CMGs) and is transmitted by *Bemisia tabaci* (whitefly), is the most important disease of cassava in Africa (Legg *et al*, 2015; Rey and Vanderschuren, 2017) and causes large yield loses (Legg and Fauquet, 2004; Owor and Legg, 2004). Cassava is a single species in the *Euphorbiaceae* family and cultivars are classified according the content of glycosides in the roots and plant morphology, including leaf shape and size, stem colour, plant height, inflorescence and storage root shape and colour (Nassar and Ortiz, 2006). Cassava is an outbreeding species with 2n=36 chromosomes. It is considered to be a sequential allopolyploid or amphidiploid (El-Sharkawy, 2004).

Cassava is drought tolerant, and can be grown in nutrient poor soils with good yields. Globally, climate change, increased demands for foods security and increased cassava starch industrialisation opportunities for rural farmers, have accelerated the demand for, and production of, cassava varieties with increased nutrition, starch content and stress resistance (Zhang *et al*, 2017). Although several breeding programs aimed at cassava crop improvement exist (FAO, 2013), breeding for specific traits in cassava is difficult due to its heterozygous nature. Using transgenic technologies, several cassava lines have been developed with a number of improved traits including disease resistance, and abiotic stress tolerance as well is improve nutrition and starch (Bull *et al*, 2018b; Leyva-Guerrero *et al*, 2012; Zhang *et al*, 2017).

1.1.1 Transformation of cassava

Transformation of plants involves the incorporation and stable expression of genes into a plant with the aim of giving it new traits (Birch, 1997). This process has become widely adopted as it allows for the improvement of crops more rapidly than traditional breeding. Due to the allopolyploid nature of cassava and its naturally low fertility, breeding for improved traits in cassava is difficult (Woodward and Puonti-Kaerlas, 2001). Therefore many cassava improvement schemes have taken advantage of the advent of genetic engineering to improve cassava. In order to improve cassava through genetic engineering, there must be a stable efficient means of transforming and regenerating cassava material.

Stamp and Henshaw first reported on the production of somatic embryos (SE) from zygotic cotyledons (Stamp and Henshaw, 1982) and leaf lobes (Stamp and Henshaw, 1987). These primary embryos or SE can also be produced from floral tissue (Mukharjee, 1995). Somatic embryos are then sub-cultured on media containing auxin to induce secondary somatic embryos (Stamp and Henshaw, 1987). Embryogenesis, characterised by the direct production of propagules, is referred to as direct somatic embryogenesis. In contrast to this Taylor *et al* (1996) developed 'indirect embryogenesis' where somatic embryos are exposed to continuous selection on Gresshof and Doy media supplemented with picloram, which gradually converts the organised structures associated with SE to less organised friable callus-like pro-embryos (Raemakers *et al*., 1997; Schopke *et al*., 1996; Taylor *et al*., 1996).

The first genetic transformation system which relied on the production of friable embryogenic callus (FEC) were established for cassava in 1996 and several transformation protocols have been developed which allow for the transformation of many cassava cultivars including TMS (Tropical manioc selection) 60444 (cv.60444) (Bull *et al*, 2009), T200 (Chetty *et al*, 2013) as well as many of the TME (Tropical *Manihot esculenta*) cultivars (Nyaboga *et al*, 2015) and landraces (Zainuddin *et al*., 2012). The production of transgenic cassava from this unorganised FEC material also lowers the risk of regenerating chimeric plants, associated with organogenesis from organized tissue (Gonzalez *et al*., 1998). In 1996 protocols for the transformation of cassava were reported by a number of different groups. Li *et al* (1996) used *Agrobacterium*-mediated transformation to transform somatic embryos (SE) which could then be used to generate transgenic shoots, while Schopke *et al* (1996) employed micro-particle bombardment of embryonic suspension-derived tissues which could then be used to regenerate transgenic plants. The combined method where *Agrobacterium*-mediated transformation is used to transform FECs has emerged as the most efficient means of producing transgenic cassava (Bull *et al*, 2009; Raemakers *et al*, 1997), even though the transformation of cassava remains recalcitrant and genotype dependent.

The production of genetically modified cassava, using FECs is complicated and time-consuming (Nyaboga *et al*, 2015). The process involves a number of steps: SE induction, FEC induction and multiplication, *Agrobacterium*-mediated transformation and hormone-induced regeneration (Bull *et al*, 2009). SEs are produced by incubating explants (leaf lobes or auxiliary buds) on media containing the auxin-like herbicide picloram (Bull *et al*, 2009). FEC clusters can then be induced from secondary SEs by continuous incubation on GD (Gresshoff and Doy) medium supplemented with at least 50µM picloram. This process is highly genotype dependent and has only been demonstrated in a limited number of cultivars (Bull *et al*, 2009; Chetty *et al*, 2013; Nyaboga *et al*, 2015). Although the exact factors involved in FEC induction are still not clear, Ma *et al* (2015) showed that in cassava cv.60444 during the transition from SE to FECs, 6871 genes are differentially expressed. These genes were mainly involved in 'cell periphery', 'external encapsulating processes', glycolysis and the metabolism of glutamate, alanine and aspartate. A recent finding by Brand *et al* (2019) two *Arabidopsis* orthologs LEAFY

COTYLEDON 1 (LEC1) and LEAFY COTYLEDON 2 (LEC2) showed a rapid increase in expression during the induction of SE material, and were more highly expressed in SE tissues than in mature differentiated material. This could indicate that these genes are very important for SE development in cassava varieties. Another interesting finding was that there was reduction in the methylation of genes in old FECs (OFEC) (>9 months), which could indicate a reduction in gene function in the OFEC material, and could indicate why plants regenerated from OFECs display larger amounts of somaclonal variation and physiological defects. Once the FECs have been transformed via *Agrobacterium*-mediated transformation, the cassava plantlets are regenerated from FECs through incubation on media containing the synthetic cytokinin and auxin: 6-Benzylaminopurine (BAP), Naphthaleneacetic acid (NAA) hormones. This process, in conjunction with a selection stage, where FECs are exposed to selective media (Bull *et al*., 2009; Zhang and Puonti-Kaerlas, 2000) (Bull *et al*,2009) has been shown to be highly efficient, with a high number of transformants. This technology has allowed for the reliable development of a number of improved cassava lines that are resistant to biotic and abiotic stress factors including CMD (Chavarriaga-aguirre *et al*., 2016).

1.1.2 Cassava mosaic disease

Cassava, originally grown in South America, is thought to have been brought to Africa by the Portuguese in the 16th Century (Legg & Thresh, 2000). This relatively recent introduction has made it susceptible to many pests and diseases endemic to Africa. CMD was first reported in Tanzania in East Africa by Warburg (1894) and has since been reported in all African countries where cassava is grown (Legg & Fauquet, 2004). It is considered the most economically important disease of cassava in Africa (Legg and Thresh, 2000; Patil and Fauquet, 2009) and epidemics can result in the loss of entire crops and even in the absence of an epidemic, crop losses can range between 20 and 90% (Akano *et al*, 2002; Hahn *et al*, 1980).

Cassava mosaic disease symptoms are characterised by expression of irregular yellow or yellow green chlorotic mosaic pattern on leaves (**Figure 1.1**) and stunted root development (Legg & Thresh, 2000) but symptom expression can vary with the cassava variety and the CMG involved (Patil and Fauquet, 2015). Symptoms of the disease are first displayed by younger plants with symptoms increasing as the plants

ages, however some cassava cultivars and landraces display recovery, where after 60 days, depending on the cultivar, season and climatic conditions, the symptoms can become more moderate, lessen or will not develop on new leaves (Fauquet & Fargette 1990). Resistance also affects symptom expression, where CMD-resistant cultivars express either do not show symptoms, or the very mild symptoms remain local to the sight of infection and disappear shortly after infection (Kuria et al., 2017; Thresh and Cooter, 2005).

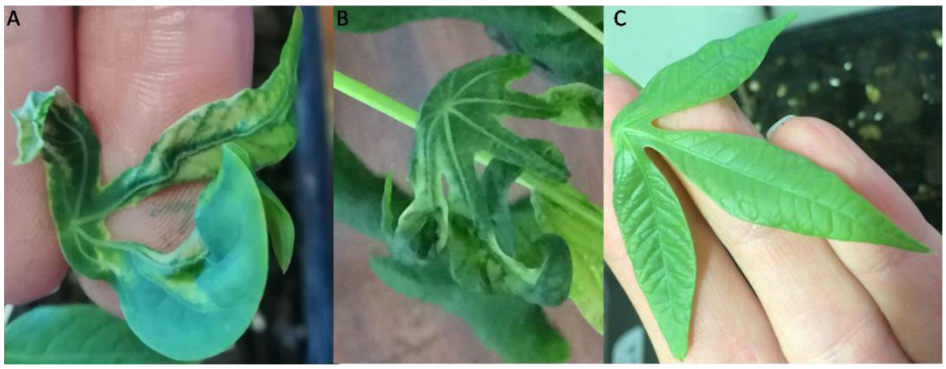


Figure 1.1: Cassava cv.60444 infected with African cassava mosaic virus- Nigeria (A) and South African cassava mosaic virus (B). Leaves display curling, mosaic and are reduced in size compared to the healthy plant (C).

Cassava mosaic disease is caused by a group of geminiviruses belonging to the genus *Begomovirus* in the family *Geminiviridae* referred to as cassava mosaic geminiviruses (Briddon and Markham, 1995). CMGs are vectored by the whitefly *Bemisia tabaci* (Genn.), and by cuttings which is the most prominent means of cassava propagation (Fondong *et al*, 2000). Globally, 11 species of CMGs have been characterised, while in Africa CMD is caused by; *African cassava mosaic virus*, *East African cassava mosaic virus*, *East African cassava mosaic Cameroon virus*, *East African cassava mosaic Kenya virus*, *East African cassava mosaic Malawi virus*, *East African cassava mosaic Zanzibar* virus and *South African cassava mosaic virus* as well as large number of strains (Brown *et al*, 2015).

1.2 *Geminiviridae*

Geminiviruses are classified into nine genera (*Becurtovirus*, *Capulavirus*, *Curtovirus*, *Eragrovirus*, *Grablovirus*, *Mastrevirus*, *Topocuvirus, Turncurtovirus* and *Begomovirus*) depending on their genome organisation, vector and host range

(Zerbini *et al*, 2017). Geminiviruses are transmitted by hemipteran insects including: whiteflies, aphids, treehoppers and leafhoppers. *Curtovirus, Becurtovirus, Turncurtovirus* and *Mastrevirus* are transmitted by leafhoppers, *Capulavirus* are transmitted by aphids and tomato pseudo-curly top virus, the only member of *Topocuvirus,* is transmitted by treehoppers. The *Eragrovirus* and *Grablovirus* vectors have yet to be identified. *Begomoviruses* (type member: *Bean golden mosaic virus*) are can be either monopartite or bipartite viruses, which infect dicotyledonous plants and are transmitted by whiteflies. Begomoviruses can further be separated into Old World and New World viruses based on their geographic distribution and genetic diversity (Zerbini *et al*, 2017).

Geminiviruses, named for their twinned icosahedral particles, are the largest group of plant DNA viruses (Briddon and Markham, 1995). Geminiviruses infect a large variety of plants including food crops, ornamentals and weeds, and infection is associated with mosaic and curling of the leaves and stunted growth of plants (Mansoor *et al*, 2003). Geminivirus infections can cause huge economic losses in crop plants (Mansoor *et al*, 2003; Patil *et al*, 2016; Patil and Fauquet, 2009). Geminiviruses are single stranded DNA (ssDNA) viruses which can be either monopartite or bipartite. Each single molecule of circular ssDNA, which varies from 2.5-3 kb in length, is packaged in an individual particle. The ssDNA particles can also be accompanied by alpha-, beta- and deta-satellites, which enhance their ability to infect the host (Bhattacharyya *et al*, 2015; Hanley-Bowdoin *et al*, 1999; Nawaz-ul-Rehman and Fauquet, 2009). Replication of the genome occurs in the nucleus of the host using both rolling circle amplification and recombination-dependent replication (Gutierrez, 1999; Haible *et al*, 2006; Hanley-Bowdoin *et al*, 1999).Genome organisation of Begomoviruses

The *Begomovirus* genus is made up of >360 species, which infect dicotyledonous plants (Brown *et al*., 2015; Zerbini *et al*., 2017). Monopartite begomoviruses are usually associated with DNA satellites (Zerbini *et al*, 2017). The genome organisation of bipartite viruses consists of two DNA components; DNA-A and DNA-B each which have a 2.6-2.8 kb genome(Gutierrez, 1999; Rojas *et al*, 2001) (**Figure 1.2**). Both DNA-A and DNA-B contain an Intergenic region (IR) with a conserved nona-nucleotide sequence which is necessary for gene replication and control of gene expression (Hanley-Bowdoin *et al*, 1999; Rojas *et al*, 2001).

DNA-A contains 6 ORFs: a replication associated protein (AC1/Rep), a transcriptional activator protein (AC2/TrAP), Replication enhancer protein (AC3/ REn), and AC4 (VIR suppressor protein) in the complementary sense; and the coat protein (AV1/CP) and pathogenicity determinant (AV2) in the virion sense (Hanley-Bowdoin *et al*, 1999) (**Figure 1.2**). More recently a seventh ORF, encoding for AC5 has been described in a number of monopartite and bipartite geminiviruses. The putative protein is located downstream of AC3, overlapping the CP, on the complementary strand of DNA-A (Ilyas *et al*, 2010). Although multiple studies have been conducted on AC5 in different viruses, its role in viral replication and infection has still not been clarified. In a study conducted in bipartite *Tomato chlorotic mottle virus* (MG-Bt) AC5 was not essential for virus Replication (Fontenelle *et al*, 2007), however a *Tomato leaf deformation virus* C5 mutant (PA10-3) was reported to produce less severe symptoms in plants (Melgarejo *et al*, 2013). This result was reiterated in a study conducted by Li *et al*(2015) who showed that AC5 played a critical role in infection of *Mungbean yellow mosaic virus* and acted as a suppressor of both post-transcription gene silencing and transcriptional gene silencing.

DNA-B is only found in bipartite viruses and contains two ORFs: the movement protein (BC1) in the complementary sense and the nuclear shuttle protein (BV1/NSP) in the virion sense (Hanley-Bowdoin *et al*, 1999; Patil and Fauquet, 2009) (**Figure 1.2)**. DNA-A is required for replication, transcription and encapsidation as well being involved in suppression of plant viral resistance, while DNA-B is required for the cell-to-cell and the long-distance movement of the virus (Gutierrez, 1999; Hanley-Bowdoin *et al*, 2004; Vanitharani *et al*, 2005). Some Old World bipartite begomoviruses are able to cause systemic infection without DNA-B, however New World begomoviruses require DNA-B (Nawaz-ul-Rehman and Fauquet, 2009; Rojas *et al*, 2005).

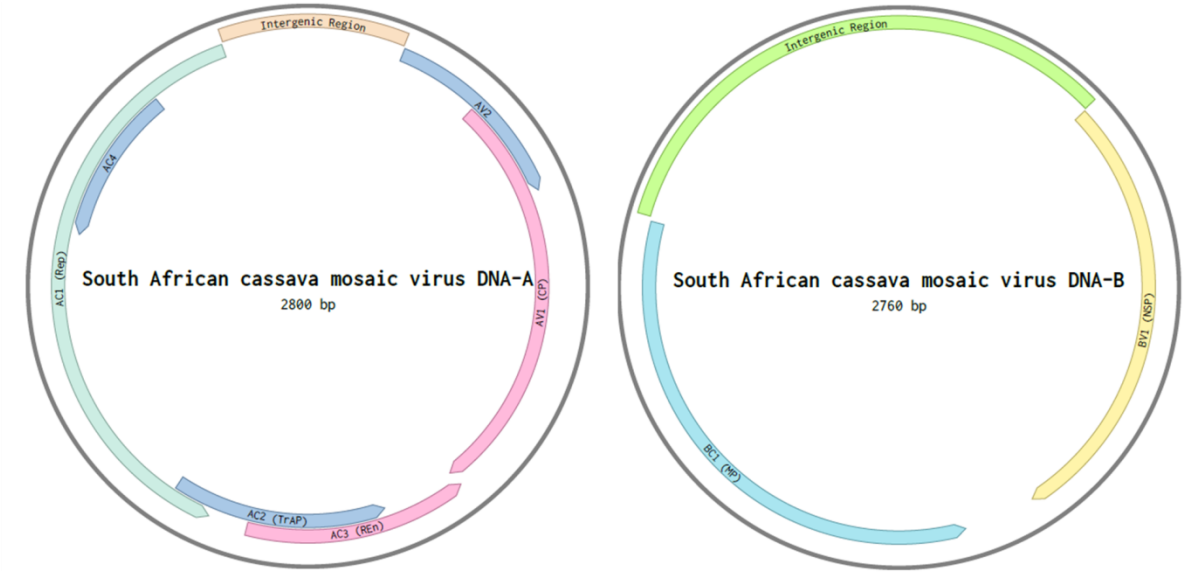


Figure 1.2: Genome organisation of the bipartite Begomovirus South African cassava mosaic virus. DNA-A consists of 6 ORFs: AC1replication associated protein (Rep), AC2 Transcriptional Activator protein (TrAP), AC3 Replication Enhancer protein (REn), AC4 (silencing suppressor/ pathogenicity determinant), AV1 Coat protein (CP) and AV2 pathogenicity determinant and DNA-B consists of two ORFs: BC1 Movement protein and BV1 nuclear shuttle protein.

The geminivirus Rep protein is essential for viral replication and is highly conserved across the *Geminiviridae* family (Rojas *et al*, 2005). Rep has a number of functions, which is evident by the number of host factors it is able to bind (Rizvi *et al*, 2014). Rep regulates vial transcription from the bidirectional promoter located in the IR of both DNA-A and B by mediating recognition of the origin and inducting DNA cleavage and ligation during rolling-circle amplification (Fontes *et al*, 1994, 1992). Geminiviruses rely on host machinery to mediate their replication and Rep is a key regulator, interacting with plant cell regulation factors to mediate viral replication (Kong and Hanley-Bowdoin, 2002). The Rep protein also mediates its own expression, as well as that of *AC2* and *AC3* genes, by the binding to the promoter in located in the IR (Eagle *et al*, 1994). Rep has also been shown to interfere with transcriptional gene silencing, preventing methylation of the viral genome by reducing the expression of DNA methyltransferases (Rodríguez-Negrete *et al*, 2013).

Transcriptional activation protein (TRaP) is a nuclear protein which functions in the trans-activation of other viral proteins including the *CP* and *BV1* genes. TRaP and AC4 also act as silencing repressors. The REn protein is involved in the up-regulation of viral replication, enhancing the accumulation of viral DNA by interacting

with Rep (Vanitharani *et al*, 2004), while both Rep and REn are required for efficient viral replication, only Rep is essential (Morilla *et al*, 2006).

1.2.1 Geminiviruses replication

Geminiviruses have relatively small genomes and use bidirectional promoters and overlapping genes in order to efficiently replicate their genomes (Rojas *et al*, 2005). Geminiviruses replicate through double stranded replicative intermediates using both recombination-dependent replication and rolling-circle mechanisms (Preiss and Jeske, 2003). Geminiviruses DNA have both a single stranded DNA (ssDNA) form which is packaged into the capsid, and replicative double stranded DNA (dsDNA) form which is found in the nucleus of the host cell (Hanley-Bowdoin *et al*, 1999; Saunders *et al*, 1992).

Begomovirus replication is initiated when the whitefly transmits the virions to the phloem-associated cells (Hanley-Bowdoin *et al*, 1999). The ssDNA is released and converted to an intermediate dsDNA using host-directed synthesis of the complementary DNA, primed by a RNA primer which binds within the conserved sequence of the IR (Saunders *et al*, 1992). The dsDNA is then transcribed by host RNA polymerase II, producing Rep which then initiates viral replication (Jeske *et al*, 2001). Once ssDNA has been produced through either recombination-dependent replication or rolling circle amplification, the ssDNA can either move out of the cell (either packaged in the presence of the CP, or using the NSP and MPs), or it can re-enter the replication cycle (Hanley-Bowdoin *et al*, 1999).

Geminiviruses (like most viruses) do not code for their own replicative machinery but rather rely on the host. Geminiviruses are very efficient at taking control of host machinery, which is demonstrated by the extent of symptoms induced by viral infection and the vast numbers of differentially expressed genes in hosts infected with geminiviruses(Allie *et al*, 2014; Hanley-Bowdoin *et al*, 1999; Kong and Hanley-Bowdoin, 2002). Geminiviruses do not code for their own DNA polymerases and rely on the host replisome to replicate their DNA. This requires the reprogramming of the tightly regulated host replisome (polymerase and host factors).

The first stage of viral replication referred to as the initiation phase involves reprogramming the cell cycle in infected cells to allow for the accumulation of host

machinery involved in DNA replication (Kong *et al*, 2000). This is achieved by a viral replisome, which consists of Rep, REn and host factors. Rep is assumed to be vital for the formation of the viral replisome, as it has been shown to interact with a number of host factors which are involved in DNA replication and repair (Fontes *et al*, 1994; Hanley-Bowdoin *et al*, 1999; Rojas *et al*, 2005). REn, enhances the accumulation of DNA in begomoviruses and is also likely to be part of the replisome (Hanley-Bowdoin *et al*, 1999). Both Rep and REn bind to a number of host factors including PCNA (proliferating cell nuclear antigen), pRb (retinoblastoma protein), Histone H3, a protein Ser/Thr kinase and kinesin (Hanley-Bowdoin *et al*, 2013, 1999; Kong and Hanley-Bowdoin, 2002; Rojas *et al*, 2001). pRb interacts with the E2F transcription factors, which are involved in the suppression host replication proteins(Kong *et al*, 2000).The binding of Histone H3 is also thought to function in prevention of transcriptional gene silencing (TGS) of the viral genome (Kong and Hanley-Bowdoin, 2002). These factors allow Rep to reprogram the cell, re-initiating the S-phase cell cycle in infected cells (Hanley-Bowdoin *et al*, 2004).

Following the reprogramming of the cells' replication machinery, Rep initiates rolling-circle amplification (RCA) by binding to the IR, in a sequence specific manner. Rep nicks the nano-nucleotide sequence of the plus-strand (Laufs *et al*, 1995), and binds to the 5' end of the nick site and where it acts as a helicase for the progression of the RCA fork allowing strand to be synthesised via rolling-circle amplification. After the strand has been completely synthesised, Rep cleaves the strand in the nano-nucleotide sequence again and mediates the ligation of the ends, forming a circular ssDNA particle (Rizvi *et al*, 2014; Ruhel and Chakraborty, 2018). The ssDNAs are either then packaged into virion particles or are transported out of the cell via the plasmodesmata channels, in the form of a nucleo-protein complex with the help of the NSP and MP encoded by the virus. In this manner the virus spreads throughout the plant (Pradhan *et al*, 2017).

1.2.2. Mixed infections and synergistic effect

Begomoviruses associated with CMD have been shown to occur as mixed infection in the same plant (Fondong *et al*, 2000; Harrison and Robinson, 1999). It has been documented that this results in an increase in viral titre of all the viruses infecting the plant which is usually linked to an increase in severity of the disease (Roth *et al*,

2004). This synergistic interaction of viruses has been shown to exist between *African cassava mosaic virus*-Cameroon and *East African cassava mosaic Cameroon virus* where a seven-fold increase in ACMV-CM DNA-A and B was observed in *Nicotiana benthamiana* which was co-infected with both viruses compared with either EACMCV or ACMV-CM alone (Pita *et al*, 2001).

Mixed infection can also lead to recombination and pseudo-recombination (Zhou *et al*, 1997). This can lead to a higher diversity in geminiviruses and has been reported to have occurred between several African strains of CMGs including the Uganda variant (UgV), which was found to be extremely similar to the *East African cassava mosaic* Tanzanian *virus*; however, the coat protein was found to be composed of both EACMTV and ACMV CPs. These recombination events can lead to unusual and severe CMD (Zhou *et al*, 1997).

Designing hybrid RNA silencing constructs targeting several viral regions and several viruses simultaneously ensures greater silencing efficiency and allows successful silencing in the event of mixed infections and in the event that virus recombination or sequence mutations have occurred.

1.3 RNA silencing

RNA silencing is a highly specific mechanism of gene regulation that plays a role in wide variety of biological processes including: development, metabolism and stress response, as well as in defence against invading nucleic acids (such as transposons and viruses) (Jones *et al*, 2001; Parent *et al*, 2015). RNA silencing is regulated by a group of small (20-40nt), non-coding RNAs which regulate gene expression at either a transcriptional or post-transcriptional level. The production of small RNA (sRNA) is triggered by the presence of double stranded RNA (dsRNA) which is processed by one of several Dicer-like (DCL) proteins to produce small single stranded RNAs (ssRNA). The small ssRNA is loaded onto Argonaut (AGO) proteins and, with other plant-specific proteins, to form either a RNA-induced silencing complex (RISC), or RNA-induced transcriptional silencing complex (RITS) which target both coding and non-coding RNAs by sequence complementarity. This process leads to either transcriptional or translational repression depending on the nature of target transcript, the DCL and AGO involved (Borges and Martienssen, 2016).

Small RNA can be broadly separated into microRNAs (miRNA) and small interfering RNA (siRNA) based on their biogenesis and precursor structures. All sRNA classes play roles in both epigenetic regulation and defence responses, however, the relative importance in these roles varies depending on the plant species (Borges and Martienssen, 2016). sRNA biogenesis occurs via one of two main pathways, miRNA biogenesis (20-22nt), and siRNA (21-24nt) biogenesis.

1.3.1 MicroRNA

Plant miRNAs are a vital for the control of a large number of biological pathways including development, feedback mechanisms, biological functions and response to biotic and abiotic stresses (Várallyay *et al*, 2010). miRNA biogenesis occurs in the nucleus where miRNAs are transcribed from *MIR* genes by RNA polymerase II (PolII). PolII transcribes long single stranded poly-adenylated RNA molecules known as primary microRNAs (pri-miRNAs) which then fold into hairpin-like structures. In canonical miRNA generation the pri-miRNAs are cleaved by DCL1, to form precursor microRNAs (pre-miRNAs), which are then cleaved further to form mature a miRNA duplex (miRNA/miRNA*) (Borges and Martienssen, 2016; Várallyay *et al*, 2010). Different DCL proteins produce miRNAs of various lengths, with DCL1 and DCL4 producing miRNA of 21nt, DCL2 and DCL3 produce miRNAs of 22nt and 24nt respectively (Axtell *et al*, 2011). The miRNA/miRNA duplex* is then methylated at the 3' end by RNA methyl-transferase HUA-Enhancer 1 (HEN1) (Yang *et al*, 2006) and transported from the nucleus where miRNA guide strand is incorporated onto AGO 1 (Katiyar-Agarwal and Jin, 2010), while the miRNA* is degraded (Várallyay *et al*, 2010). The miRNA-AGO1 complex forms the central unit of RISC or RITS, and directs translation inhibition or mRNA degradation (Budak and Akpinar, 2015). miRNAs mediate both transcriptional gene silencing and post-transcriptional gene silencing (PTGS) by mediating transcription and translational repression as well as messenger RNA (mRNA) cleavage and are essential for correct plant development.

Although RNA interference (RNAi) is primarily seen as eukaryotic defence against invasion by harmful genetic elements and viruses (Ramesh *et al*, 2014), it has more recently been shown that some plant viruses also express miRNA (Maghuly *et al*, 2014). Most viruses that have been shown to express miRNA are DNA viruses, with a nuclear component to their replication cycle, however there are exceptions to this

(Kincaid and Sullivan, 2012). Virus-encoded miRNAs are assumed to perform one of three functions in the host: (1) aid in the avoidance of the host anti-viral response, (2) regulate host and viral genes and (3) prolong longevity in infected cells (Kincaid and Sullivan, 2012). Virus-encoded miRNAs are classified into two classes: virus-specific miRNAs and host-analogous miRNAs (Kincaid and Sullivan, 2012). Host-analogous miRNAs are miRNAs which mimic host miRNAs, particularly within the 5' seed region, and allow viruses to regulate host transcripts (Budak and Akpinar, 2015; Kincaid and Sullivan, 2012).

1.3.2 Small interfering RNA

Small interfering RNAs are produced from long dsRNA, which can arise through the folding back of invert-repeat sequences, the hybridisation of sense and anti-sense sequences (which can either be transcripts, or unrelated RNA's) or through synthesis by RNA-dependent RNA polymerases (RDRs) (Borges and Martienssen, 2016; Elvira-matelot *et al*, 2017). There are 6 RDRs (RDR 1-6) involved in the synthesis of siRNAs in plants. Double-stranded RNA is processed by DICER-like proteins (DCL2, DCL3 or DCL4), in conjunction with DOUBLE-STRANDED RNA BINDING (DRB) proteins into siRNAs of between 21-24nt in length (Hamilton and Baulcombe, 1999; Parent *et al*, 2015). The siRNAs are subsequently methylated at the 3' end by HEN1(Yang *et al*, 2006) and transported from the nucleus by HASTY (HST) (Csorba *et al*, 2015; Park *et al*, 2005). The siRNA are then loaded onto AGO proteins to form either a RISC (Kim *et al*, 2007) or RITS, which silence targets through PTGS and TGS respectively (Csorba *et al*, 2015).

Endogenous siRNAs are processed by one of three DCL proteins (DCL2, DCL3 and DCL4) and are categorised as either heterochromatic siRNAs (het-siRNAs) or secondary siRNAs. Heterochromatic siRNAs mediate transcriptional silencing of transposons and pericentromeric repeats via RNA-directed DNA methylation (RdDM) and are the most abundant sRNAs(Borges and Martienssen, 2016). Heterochromatic siRNAs are transcribed by PolIV, after which RDR2 produces dsRNA which is processed by DCL3. The broader class of secondary siRNAs requires Pol II, RDR6 and either DCL2 and DCL4.

1.3.3 Viral sRNA

Many plants use RNA silencing to control viral infection via miRNA-mediated or siRNA-mediated PTGS, or transcriptionally, via sRNA-directed DNA methylation (Raja *et al*, 2014). Many organisms have been shown to produce viral sRNAs (vsRNA) in response to infection by viruses (**Figure 1.3**). Viral sRNA are then recruited by AGOs, incorporated into RISC and degrade viral RNA in a sequence specific manner (Zhang *et al*, 2015).The biogenesis of vsRNAs requires RDR, DCL and AGO proteins, however its exact mechanism varies depending of the species of plant and the infecting virus (Pooggin, 2017; Zhang *et al*, 2015). Viral sRNAs are produced from the processing of dsRNA which can be dsRNA (from RNA viruses); hairpin intermediates or RDR polymerase derived dsRNA (Elvira-matelot *et al*, 2017; Llave, 2010; Molnár *et al*, 2005). All virus infected plants produce siRNAs which target the whole length of the RNA or DNA virus-encoded transcripts in both the sense and anti-sense direction (Pooggin, 2017). These vsRNA are differentially expressed, which results in 'hotspot' which are more highly targeted than others (Sharma *et al*, 2014).

The production of vsRNAs is dependent on the slicing activity of DCL, which cut dsRNA into fragments of 21-24nt. There are 4 DCL involved in the production of virus-derived vsRNAs (DCL1, DCL2, DCL3 and DCL4) and while all four DCL proteins are likely to have access to the dsRNAs, there is a hierarchal order to DCL production of vsRNAs (Blevins *et al*, 2006; Csorba *et al*, 2015; Parent *et al*, 2015; Zhang *et al*, 2015). In the presence of RNA viruses, both DCL4 and DCL2 form part of RNA silencing, DCL4 major role is in the production of primary (21nt) siRNAs, while DCL2 is involved in the production of secondary (22nt) siRNAs. However, their roles are redundant, and they act hierarchically, with DCL4 obscuring DCL2. Only when DCL4 is silenced (*dcl4* mutant), does DCL2 become critically important (Blevins *et al*, 2006; Parent *et al*, 2015). DCL1 is mainly involved in the production of miRNA, however it also regulates the production of siRNA, by increasing the availability of viral dsRNA to other DCL proteins (Blevins *et al*, 2006; Moissiard and Voinnet, 2006) and negatively regulating the production of DCL3 and DCL4 via the miRNA pathway. In the presence of a DNA virus, all four DCL proteins are involved in the production of vsRNAs, however, DCL3 specifically is important for silencing, where in conjunction with DRB4, DRB2 and methyl-transferase 1 (MET1) it induces

the methylation of the viral genome(Raja *et al*, 2014; Vinutha *et al*, 2018). The siRNA profiles of plants can vary depending on the virus present. In *Arabidopsis thaliana* infected with a DNA virus, 21nt and 24nt accumulate at a higher rate than 22nt vsRNA, but in the presence of RNA viruses 21nt and 22nt accumulate in higher amount than the 24nt vsRNA (Aregger *et al*, 2012).

Double stranded RNA binding proteins are involved in a number of pathways, including miRNA processing, the cytoplasmic PTGS siRNA pathway and viral genome methylation (Raja *et al*, 2014). There are a number of different DRB proteins involved in these different pathways, however, DRB2, DRB4 and DRB3 appear to be the most important (Raja *et al*, 2014). DRB2 primary role has been ascribed as the processing of miRNAs involved in plant development, however, DRB2 (and DRB4) have also been shown to inhibit the accumulation of sRNAs associated with the RdRM pathway, and therefore could also be involved in viral genome methylation (Pélissier *et al*, 2011). In addition to this role, DRB4 interacts with DCL4, and is involved in the siRNA-mediated defence against RNA viruses (Qu *et al*, 2008). Finally, DRB3 has specific role in the defence against DNA viruses, where it interacts with AGO4 and DCL3, to methylate their genomes (Raja *et al*, 2014).

Once DCL proteins have processed dsRNA into siRNA, it is loaded onto AGO proteins. The coupling of AGO proteins and sRNA is guided by the 5' terminal nucleotides as well by the thermodynamic properties, length and duplex structure of the siRNA (Baumberger and Baulcombe, 2005; Havecker *et al*, 2010; Morgado *et al*, 2017; Schott *et al*, 2012). sRNA-programmed AGOs are responsible for the silencing of complementary RNA and DNA through TGS and PTGS (Carbonell and Carrington, 2015). There are several AGO proteins involved in the silencing of viral genomes and their importance and activity vary based on the virus and plant host (Carbonell and Carrington, 2015). Generally, it is acknowledged that AGO1 and AGO2 are the most important AGO proteins involved anti-viral defence (Carbonell and Carrington, 2015). In some cases, AGO1, acts as the primary anti-viral AGO protein (Baumberger and Baulcombe, 2005) and as a negative regulator of AGO2 (Harvey *et al*, 2011). AGO1 is targeted by many viruses through translational repression (Várallyay *et al*, 2010), and in the case that AGO1 is repressed, AGO2 acts as a second layer of defence, performing a similar role as AGO1(Harvey *et al*, 2011). There are exception to this however, for example, AGO2 has also been

shown to be the main antiviral AGO against viruses such as *Tobacco rattle virus* (Ma *et al*, 2015), where AGO1 is not targeted by the virus. AGO4 has also been shown to play a role in defence against DNA viruses, specifically geminiviruses, where AGO4 recruits 24nt viral siRNAs and induces viral genome methylation (Mallory and Vaucheret, 2010; Raja *et al*, 2014). In order to achieve a robust defence response, plants require a mechanism to amplify the anti-viral response. This is achieved via the RDR-mediated production of secondary siRNAs from ssRNAs generated by AGO-mediated cleavage of viral RNA (Csorba *et al*, 2015; C. Zhang *et al*, 2015). In RNA viruses this process is carried out by RDR6 and RDR2, facilitated by Suppressor of GENE SILENCINIG 3 (SGS3), SILENCING DEFECTIVE 5 (SDE5) and SDE3, where the RDR combines the ssRNA fragments into a single dsRNA precursor that is then processed by DCL2 and DCL4 to produce siRNAs in an amplification cycle (Csorba *et al*, 2015; Yu *et al*, 2009). This process has been shown to be less efficient against DNA viruses, where the majority of vsRNA were found to be RDR-independent primary siRNAs (Aregger *et al*, 2012). However the suppression of RDR6 has been linked to an increase in vsRNA (Aregger *et al*, 2012). βC1 encoded by the DNA satellite of *Tomato yellow leaf curl China virus* was shown to up-regulate *Nicotiana benthamiana* calmodulin-like protein, which regulates RNA silencing through the regulation of RDR6, showing that secondary siRNA synthesis plays a role in the viral suppression (Li *et al*, 2014).Secondary siRNAs are further classified as trans-acting siRNAs (ta-siRNAs), natural antisense short interfering RNAs (nat-siRNAs), epigenetically-activated siRNAs (ea-siRNAs) and phased siRNAs (pha-siRNAs), based on their size and biogenesis (DCL and RDR interactions).

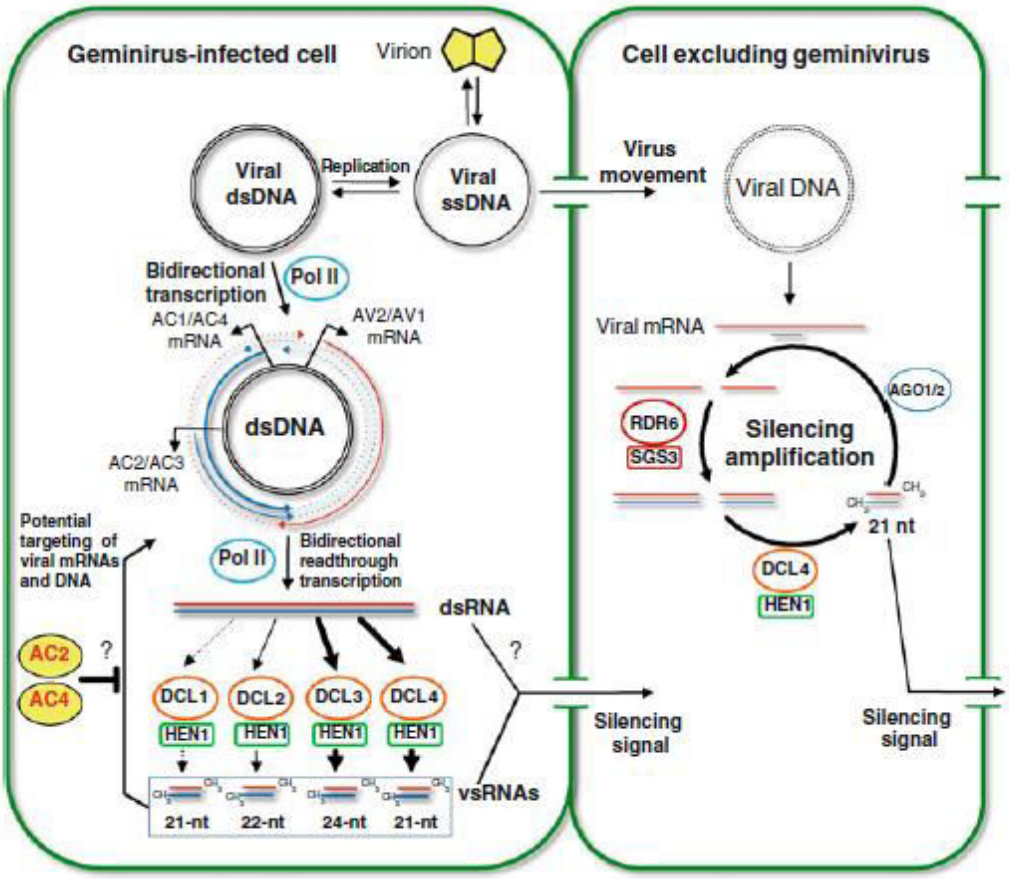


Figure 1.3: Overview of biogenesis of vsRNA for geminiviruses. The two cells show the initial site of infection (left) and secondary siRNA production (right). At the site of first infection dsRNA arises from the bidirectional transcription of the viral genome by Pol II. The dsRNA is then digested by various DCL proteins, creating siRNA 21-, 22 and 24nt in length. Viral DNA as well as the vsRNA is then transferred to neighbouring cells. The vsRNA is processed by RDR6/SG3/DCL4 to produce secondary siRNA, helping to control viral titres (Rajeswaran and Pooggin, 2012)

Viral siRNAs can be recruited by RISC and degrade viral RNA in a sequence specific manner, but are also involved in RNA-directed DNA methylation (RdDM), where 24nt siRNAs guide DNA methylation of viral DNA genomes (Chellappan *et al*, 2004b;Zhang *et al*, 2015). This has been shown to occur in anti-viral defence against DNA viruses. Studies conducted on geminivirus infections, have shown that when the methyl-transferase gene is silenced, *Arabidopsis* plants becomes hypersensitive to infection (Raja *et al*, 2008).

1.4 Suppression of vsRNA biogenesis

In response to the anti-viral defences of plants, most plant viruses have developed silencing suppressors (known as viral suppressors of RNA silencing or VSRs) which counteract the biogenesis and actions of viral si- and mi- RNAs. VSRs are mechanistically diverse and have been shown to block almost all steps of RNA silencing (Csorba *et al*, 2015) including blocking the biogenesis of vsRNA and secondary siRNA and, regulating endogenous factors involved in RNA silencing (Csorba *et al*, 2015; Ye *et al*, 2015). In addition to the production of VSRs, some viruses avoid silencing through the production of mRNA with complex secondary structures which prevents their inclusion into RISC complex, and production of decoy dsRNA which saturate the silencing machinery allowing viral proteins to be expressed (Rajeswaran *et al*, 2014).VSRs have number of targets including DCL inhibition; dsRNA and siRNA sequestration, and AGO proteins. For example, the V2 protein of *Tomato yellow leaf curl virus* interacts directly with SGS3 (the co-factor of RDR6), and competes to bind to dsRNA (Zrachya *et al*, 2007). VSRs can act on a single target, or on multiple targets within siRNA production. The *Potyvirus* HC-Pro plays a number of roles in the suppression of vsRNA production, including binding to siRNA (Torres-Barceló *et al*, 2010), blocking HEN1 methyl-transferase binding (Mlotshwa *et al*, 2016) and RDR6 down-regulation (Kasschau *et al*, 2003). The HC-Pro polyprotein synergizes the co-infection of other viruses as it inhibits anti-viral response of the plant (Torres-Barceló *et al*, 2010). VSRs can also achieve silencing suppression by blocking the assembly of the RISC complex, or targeting the RISC complex RNA component.

1.4.1 Viral suppressors of Geminiviruses

Geminiviruses replicate in the nucleus via a dsDNA intermediate, which acts as a replication and transcription intermediate (Hanley-Bowdoin *et al*, 1999). Viral mRNAs can also be transcribed bi-directionally and can form dsRNA. Therefore, geminiviruses are both the target of TGS (targeting the dsDNA) and PTGS (targeting dsRNA) in plants (Brodersen and Voinnet, 2006; Li *et al*, 2017; Wang *et al*, 2005; Zhang *et al*, 2011). Geminiviruses, as a result, have a number of suppressor proteins that are involved in the suppression of both TGS and PTGS by binding to sRNAs,

and binding and inhibiting enzymes involved in the host methylation cycle (Vinutha *et al*, 2018).

In addition to its role in the viral replication, Rep has been shown to play a role in the inhibition of anti-viral TGS in plants (Raja *et al*, 2008). *Arabidopsis thaliana* plants where the cytosine methyltransferases MET1, CHROMOMETHYLASE 3 (CMT3) were mutated, were shown to be hyper susceptible to geminivirus infection and viral genomes had reduced levels of methylation indicating a reduction of TGS (Rodríguez-Negrete *et al*, 2013). This is specifically achieved by the Rep, in conjunction with AC4 and in concert with AC2 (Rodríguez-Negrete *et al*, 2013).

The ACMV AC2 protein was shown to be able to suppress established RNA silencing by interacting with other proteins within the nucleus of the cell (Trinks *et al*, 2005; Voinnet *et al*, 1999). Subsequent studies have shown that AL2 and L2 of *Tomato golden mosaic virus* and *Cabbage leaf curl virus* were shown to be involved in the suppression of TGS, targeting histone methyl-transferase and Su(var)3-9 homolog 4/Kryptonite (SUVH4/KYP) proteins(Zhang *et al*, 2011). The AC2 protein can also repress and reverse TGS by inhibiting Adenosine kinase (Wang *et al*, 2005) and interfering with the production of S-adenosyl-methionine and attenuation of proteasome-mediated degradation of S-adenosyl-methionine decarboxylase 1 (Zhang *et al*, 2011), which are involved in the establishing and maintenance of methylation in plants.

The AC4 protein generally recognised as a silencing suppressor in geminiviruses, and plays a role in suppression of both TGS and PTGS in plants (however not all geminivirus AC4 are PTGS suppressors) (Chellappan *et al*, 2005). A study conducted on *Tomato leaf curl New Delhi virus showed* that when *AC4* gene of Tomato leaf curl New Delhi virus (ToLCNDV) was constitutively expressed, it showed a reduced intensity of host-induced methylation of the viral genome (Vinutha *et al*, 2018). AC4 directly interacts with AGO4 inhibiting its ability to methylate the viral genome. AC4 has also been shown to reverse PTGS, however, the exact mechanism has not been elucidated as yet, although it is known that it is not through directly interacting with AGO1 (Vinutha *et al*, 2018).

Geminiviruses DNA-B encoded BV1 has also been shown to play a role in the suppression of PTGS. Aberrant viral RNAs serve as a template for RDR6, which

synthesises complementary RNA which is then used to produced siRNAs to mediate viral RNA cleavage (Voinnet, 2001). However, the aberrant viral RNA can also be cleared by host RNA degradation machinery, which depletes the substrate for dsRNA synthesis thereby attenuating PTGS (Ye *et al*, 2015). In plants the stability of the mRNA is dependent on the 5' cap. During mRNA degradation a Decapping protein 2 (DCP2) protein removes the 5' cap and the RNA is subsequently degraded (Xu *et al*, 2006). The BV1 can act as viral suppressor, by binding to the promoter region of ASSYMETRICAL LEAVES 2 (AS2) protein, and exporting it from the nucleus, where it interacts with DCP2 to promote the decapping of viral RNAs, leading to weak accumulation of siRNA and weakened siRNA silencing (Ye *et al*, 2015).

1.5 Application of siRNA technology

An observation that the introduction of a transgene mRNA with high sequence similarity to an endogenous mRNA resulted in the degradation of both mRNAs lead to research into the phenomena observed in transgenic plants (Elvira-Matelot *et al*, 2017). It was found that the mechanism underlying the induced silencing was an RNA-activated sequence specific RNA degradation now referred to as RNA interference (Simón-Mateo and García, 2011).

Traditional breeding of plants for enhanced traits is time consuming and in most cases, inefficient. The advent of whole genome sequencing, in addition to the discovery of RNAi, has created systems whereby specific genetic traits can be controlled and manipulated to create transgenic crops more quickly and efficiently. Further, pathogen-derived resistance (PDR) or RNAi, where partial or whole genes from pathogens are transformed into the host, has proved successful for protecting plants against a number of biotic agents including viruses.

1.5.1 RNAi-mediated virus resistance in plants

The RNAi approach as a means of conferring resistance to a plant has been shown to be effective in numerous plants against many viral species. The underlying mechanism was first elucidated by Lindbo *et al* (1993) who showed that tobacco plants transformed with the coat protein of *Tobacco etch virus* recovered from viral infection (Lindbo *et al*, 1993; Pooggin, 2017). It was shown that these plants,

although initially susceptible to the virus, displayed a recovery phenotype and had a lower level of viral mRNA in the transgenic leaves. This tolerance was also highly specific to Tobacco etch virus (TEV), and plants did not show any broad-spectrum resistance to *potato virus Y*. Subsequent studies have shown that RNAi can induce PTGS, as demonstrated by the Lindbo study (1993), as well TGS via RNA-directed DNA methylation (Jones *et al*, 2001).

The RNAi-mediated anti-viral approach relies on activating the RNA silencing mechanisms of the plant, through the introduction of dsRNA associated with the target virus. This has been achieved through the introduction of a transgene expressing RNA homologous to the viral genome, or transiently with the application of dsRNA that is sufficiently similar to target virus (Khalid *et al*, 2017; Pooggin, 2017). Three different transgene orientations have been use to induce PTGS in plants; sense (S-PTGS) or anti-sense (AS-PTGS) orientation (Brodersen and Voinnet, 2006), and an inverted repeat of the transgene, separated by spacer or intron, producing double-stranded or hairpin RNA (hp-RNA) (Béclin *et al*, 2002; Pooggin, 2017). In the S-PTGS system, the dsRNA substrate needed to produce the siRNAs is produced by the host-encoded RDR, which recognises the highly abundant transgene RNAs (Ho *et al*, 2007). Transgenes with the hairpin/ inverted repeats produce dsRNA and therefore do not need host RDRs for efficient RNA silencing; however, they do require RDR for amplification of secondary siRNAs which re-enforces silencing (Simón-Mateo and García, 2011). Small RNA sequencing has shown that the production of secondary hp-RNA derived siRNA can be inefficient (Chen *et al*, 2010) and therefore the majority of siRNA produced by hp-RNA is in the form of primary siRNA, which are produced at a higher rate than viral-induced siRNA in plants (Blevins *et al*, 2006; Fuentes *et al*, 2016).Transgene encoded intron-spliced hp-RNA has been shown to be a highly efficient silencing trigger and produces high levels of viral resistance (Wesley *et al*, 2001), therefore, this approach has been used in many plant species (Simón-Mateo and García, 2011). Fuentes *et al* (2006) developed a transgenic tomato, transformed with the C1 protein of *Tomato Yellow Leaf Curl virus* which was found to be resistant to the virus, even in extreme conditions. Further testing of this line, under field conditions elucidated some of the features of RNAi-induced resistance (Fuentes *et al*, 2006). Blot hybridisation showed the presence of 21nt, 22nt and (less abundantly) 24nt

siRNAs which corresponded to the hp-RNA. This shows that the hp-RNA is efficiently processed by the cell, where it is spliced and transported from the nucleus by DCL2 and DCL4 to RISC-AGO complex for PTGS, as well as held in the nucleus, where it interacts with DCL3 to produce 24nt siRNAs which act in TGS (due to the levels of 24nt, it is assumed that proportionally less is held in the nucleus).

In addition to S/AS-PTGS and hp-PTGS, artificial miRNA induced resistance (AMIR) has also been introduced as a means of controlling viruses in plants. In this system natural miRNAs are replaced with RNA sequences which are specific to viral transcript within a naturally occurring MIR gene. The mature miRNA is then processed by the cellular machinery, and can be used to silence the viral genome (Khalid *et al*, 2017; Niu *et al*, 2006). AMIR has been used to induce resistance a large number of plant viruses including *Turnip mosaic virus* , *Turnip yellow mosaic virus* and *Cucumber mosaic virus* in *Arabidopsis thaliana* and *N. benthamiana* (Duan *et al*, 2008; Lin *et al*, 2009; Niu *et al*, 2006).The use of AMIR has a number of advantages, including the high specificity and accuracy and a high level of stability (Liu *et al*, 2017). However use of this technique has been limited, most likely because it has been shown that where viruses are allowed to replicate freely, they can quickly overcome the amiRNA transcriptional repression through mutation of the short target RNA (Lin *et al*, 2009; Pooggin, 2017). MicroRNA-mediated virus resistance is also non-transmissible, which is assumed to be because it does not induce the production of 24nt siRNAs (Pooggin, 2017; Schott *et al*, 2012; Tiwari *et al*, 2014).

One of the greatest advantages of hp-RNA-based strategies is that it allows the stacking of multiple cassettes into a single vector (Khalid *et al*, 2017). Although several different strategies have been tried including co-transformation with multiple constructs, constructs with multiple transcription units and constructs with multiple hairpins, the most widely used strategy is to create a single chimeric hairpin, stacking fragments from multiple viruses together under a single promoter which is then transformed into the plant (Khalid *et al*, 2017; Lin *et al*, 2012a). This minimises complex integration patterns and has been shown to have applications in developing resistance to multiple viruses with a single construct (Ma *et al*, 2011), as well targeting multiple areas of the same virus, which could increase durability and efficiency of resistance (Duan *et al*, 2012).

The accumulation of high levels of virus-targeted siRNAs from hp-RNA is usually associated with the triggering of the anti-viral machinery in plants (Fuentes *et al*, 2016; Leibman *et al*, 2011; Ntui *et al*, 2014). Analysis of deep-sequencing data from transgenic plants has shown that although the viral-specific siRNA profiles from transgenic and non-transgenic lines are quite similar, siRNA is present in higher quantities in the uninfected resistant transgenic lines (Fuentes *et al*, 2016; Leibman *et al*, 2011; Pooggin, 2017; Vanderschuren *et al*, 2009), which is assumed to confirm that virus-targeted hp-RNA prepares the natural anti-viral mechanism in the plant. The higher levels of transgene-derived virus-targeted siRNA correlates with the production of RDR1-dependent mobile 24 nt siRNAs, which can lead to long range silencing and increased resistance (Leibman *et al*, 2011). Further, 24nt siRNAs are also known to induce methylation, and may contribute to the TGS of the viral genome (Dalakouras *et al*, 2011). In addition to this, increased levels of transgene derived siRNAs have been linked to a rise in RDR1 levels in the plant, which may help prepare the plant against other similar viruses (Leibman *et al*, 2011). It is important to note however that both 21-22 nt siRNA and 24 nt mobile siRNAs contribute to viral resistance (Pooggin, 2017) and the ability of hp-RNA to trigger silencing in the plant is dependent on a number of factors.

1.5.2 Design of an hp-RNA construct

Virus-targeted hp-RNA does not always lead to effective triggering of the anti-viral RNAi (Dalakouras *et al*, 2011; Vanderschuren *et al*, 2009). The efficiency of the RNAi resistance is dependent on number of factors including the RNAi target region, structural characteristics of the hp-RNA, as well as the site of integration of the transgene (Dalakouras *et al*, 2011; Duan *et al*, 2012).

The efficiency of RNAi based silencing can also be influenced by the choice of viral target. There are a number of targets which have been shown to induce PTGS and TGS of viruses including viral replicase or replication associated proteins (Vanderschuren *et al*, 2007a); viral coat proteins or nuclear capsids proteins; and VSR proteins. In addition to these, some studies have shown that targeting the promoters or untranslated regions (UTR) can also lead to silencing of the viral genome (Khalid *et al*, 2017). Not all targets successfully induce silencing and this is believed to be due to a number of factors. Firstly, in order for a siRNA to successfully

silence a viral gene, there needs to be a high level of homology between the virus and the target. It has been shown that where the sequence and target differ by between 10% to 20%, the virus will break to resistance mechanism (de Haan *et al*, 1992; Khalid *et al*, 2017). It is also possible that if the target region has a high GC content, it could form secondary structures which would prevent its inclusion in a RISC complex, or if the target forms part of a hotspot already targeted by the plant, it may have developed a mechanism of avoiding silencing (Pooggin, 2017; Rajeswaran *et al*, 2014).In the initial studies conducted on RNAi-mediated resistance, RNA silencing was induced using transgenes in either the sense or anti-sense orientation (Brodersen and Voinnet, 2006). However, subsequent studies hp-RNA is a more efficient means of producing siRNAs (Béclin *et al*, 2002; Pooggin, 2017; Smith *et al*, 2000; Wesley *et al*, 2001). The presence of the spacer region is important for the stability and correct processing of the ds-RNA; however, it has been shown that where the spacer is replaced with a functional intron, the RNA processing and subsequent anti-viral resistance is substantially increased (from 50 to 100% in some cases) (Wesley *et al*, 2001). The position of the intron in relation to the promoter and the length of the intron have also been shown to play a role in gene expression (Rose, 2004).

The target region and the length of arms of the hp-RNA can affect the efficiency of the hp-RNA RNAi-mediated silencing. Generally, target hp-RNA arms are between 100- 800nt in length, but shorter (50 bp) and larger constructs (2.5kb) have been used (Liu *et al*, 2007). Larger constructs are assumed to be more effective, as they target longer sequences and could proportionally produce more siRNAs than shorter sequences; however, the risk of possible off-target effects is increased in these constructs. As mentioned previously, shorter constructs (such as those found in AMIR constructs) can also induce resistance, but this resistance is usually broken by the virus (Duan *et al*, 2012). There can also be a minimum required number of base pairs for efficient processing of the hairpin (Duan *et al*, 2012). In a study conducted on *Tomato spotted wilt virus,* constructs (387–453 bp) which targeted the capsid protein were shown to efficiently induce PTGS of the virus. However, when the size of the construct was reduced (92- 235 bp), the construct no longer efficiently conferred resistance. Interestingly, when a *GFP* gene was attached to the shorter sequences, the constructs ability to induce PTGS was restored, indicating that the

limiting factor in the smaller constructs was size and not the loss of specific siRNAs (Duan *et al*, 2008, 2012; Pang *et al*, 1997).

Target selection is an important component of designing RNAi constructs that will elicit resistance in a plant. Plants naturally produce siRNAs that target the viral genome (Patil *et al*, 2016; Vanderschuren *et al*, 2007a), however; the levels of siRNAs that target specific genes vary, for example in a study conducted on *Tomato chlorotic mottle virus* in tomato and *N. benthamiana* the siRNA populations targeting AC1 were significantly larger than those targeting AC5/AV1 (Ribeiro *et al*, 2007). In addition, it has been shown that hp-RNA which target certain ORFs elicit a much greater anti-viral response than others (Lin *et al*, 2012b; Patil *et al*, 2016). For example, in a comparison of cassava plants transformed with hp-RNA constructs targeting each of the ORFs of ACMV, constructs targeting the full ORF of AC1 and AC2 were shown to elicit much higher siRNA expression levels than other viral proteins, with AC2 giving the highest level of resistance (Patil *et al*, 2016). The level of siRNA production can also be affected by the region of the gene that is targeted by the hp-RNA construct. Patil *et al* (2016) showed that 5'-region of an ORF is a more efficient elicitor of RNAi than the 3'-region. This information is important when deciding on the target of a hp-RNA construct. Ribeiro *et al* (2007) targeted the AC2 region of *Tomato chlorotic mottle virus*, which was under represented in the siRNA population of the wild-type *N. benthamiana* plant, and this increased resistance, possibly because it increases the avenues through which the plant could induce PTGS against the virus (Patil *et al*, 2016). The alternate strategy is to target a region which is already associated with high population levels of siRNAs, that with the addition of the hp-RNA construct, will be sufficient to induce PTGS (Patil *et al*, 2016; Vanderschuren *et al*, 2007a). This information is important when creating chimeric/stacked constructs which target multiple ORFs of a single virus, or multiple viruses. The target region of each ORF must be kept to a minimum size which still confers resistance but still allows for stacking of multiple sequences within the same hp-RNA construct.

Resistance is usually linked to the amount of siRNA produced by a plant and lines that produce high amounts of siRNA are more resistant than lines that produce low number of siRNA (Ribeiro *et al*, 2007; Vanderschuren *et al*, 2009). However, it has been noted that not all integrated copies of the hp-RNA induce the expression of

viral siRNAs, and certain integration sites appear to produce siRNAs more efficiently than others. Kalantidis *et al* (2002) reported that some transgenic tobacco lines containing an hp-RNA targeting the CP of Cucumber mosaic virus (CMV) displayed resistance to the virus. However, not all the lines which expressed the transgene displayed resistance to CMV. Further analysis showed that resistance was linked to the locus of the hp-RNA, with all resistant lines containing a copy of the hp-RNA transgene at the same locus (Dalakouras *et al*, 2011). Only these lines produced cytoplasmic siRNA and were able to induce RdRM-dependent methylation in CMV. In this case, the locus was linked to more efficient production of siRNA, thereby inducing better resistance in the plant.

This information is important when designing an effective RNAi hp-RNA construct, as a strong immediate anti-viral response is imperative to stopping viral replication in the early stages of infection. Resistance is dependent on viral load (Vanderschuren *et al*, 2009), and therefore, a virus must be controlled before it is able overwhelm the anti-viral siRNAs response.

1.5.3 Off-target effects

One of the main challenges and concerns when designing an efficient RNAi construct is the avoidance of 'off-targets', which could affect the host plants morphology and fitness. The off-target effect could occur through many mechanisms including the production of siRNAs which target host proteins/ mRNAs, or through the interaction of the viral protein with host siRNAs.

As mentioned previously, a number of viral proteins act as silencing suppressors; interacting with plant mRNA, inhibiting methylation, or binding plant sRNA in order to achieve systemic infection of the plant. These proteins have been shown to be effective RNAi targets, however due to the mode of action of some VSRs it is possible that when constitutively expressed these VSRs can interfere with the development of transgenic plants (Chellappan *et al*, 2005). In *Arabidopsis,* the constitutive expression of ACMV-Cameroon AC4 resulted in stunted growth and lack of developmental tissue in transgenic lines (Chellappan *et al*, 2005). Praveen *et al* (2010) showed that a hp-RNA which did not express the AC4 viral protein still produced siRNAs which targeted host proteins involved in the development of the plant. RNAi constructs targeting the AC4 of ACMV-Cameroon and *Tomato leaf curl*

virus New Delhi have also been shown to cause developmental defects in transgenic plant lines (Praveen *et al*, 2010; Vinutha *et al*, 2018), which was linked to decreased levels host miRNAs important to plant development including miR159, miR167/166 and miR171.

Hairpin-RNAi constructs can produce a number of unintended siRNAs which can target the host genome and perfect sequence complementarity is not always required for interference to occur. For example, one predicted siRNA showed 12nt out of 21nt targeted miRNA3 from tomato, which could lead to abnormalities in plant development (Praveen *et al*, 2010). This is especially true of miRNA targets, where complete complementarity is only required in the 'seed region' of the miRNA (Budak and Akpinar, 2015). Transcriptomic analysis of RNAi-transgenic tomatoes by Fuentes *et al* (2016) showed a common set of differentially expressed genes, associated with the presence of the transgene. While it could not be determined whether these effects were direct or indirect, it is important to note that it is possible for a RNAi construct to have an unintended effect on the host plant (Pooggin, 2017).

1.5.4 RNAi-mediated resistance against geminiviruses

Geminiviruses are an important group of viruses in global crop production and cause huge yield and economic losses annually (Patil *et al*, 2016). These losses have highlighted the importance of developing an efficient means of controlling viruses and as such, several geminivirus resistant crops have been developed using the hp-RNA RNAi approach (Pooggin, 2017). The complementary strands of bi-partite geminiviruses DNA-A and DNA-B components contain the genes for viral proteins AC1, AC2, AC3, AC4 and BC1 respectively. Expression of the genes is driven by strong promoters and they are expressed in the early stages of virus replication (Hanley-Bowdoin *et al*, 2013) and as such they make good targets for RNAi. Geminivirus resistant crops have been developed targeting several regions of the viral genome including: the bidirectional promoter located in the IR (Vanderschuren *et al*, 2007a); Rep (Vanderschuren *et al*, 2009); TrAP(Patil *et al*, 2016); AC4(Patil *et al*, 2016), the CP and BC1 (Patil *et al*, 2016).

1.5.4.1 Coat Protein mediated resistance:

The first report of virus resistance mediated by the expression of a viral protein was in tobacco plants transformed with the CP of *Tobacco mosaic virus* (Abel *et al*,

1986). Later it was shown that the construct interfered with uncoating of virus, thus interfering with virus replication, producing plants that were either resistant or showed recovery (Beachy, 1999). In geminiviruses the use of CP-derived resistance seems to be more effective in monopartite begomoviruses, which require the CP for systemic infection (Hanley-Bowdoin *et al*, 1999; Kunik *et al*, 1994; Rojas *et al*, 2001). In bipartite begomoviruses where the CP is not always required for movement between cells, construct which target multiple areas of the virus including the CP have shown to confer resistance, Ntui *et al*(2015) developed transgenic *Sri Lankan cassava mosaic virus* resistant cassava using a hp-RNA construct which targets the AV1/AV2 overlap. These results were similar to those found by Vu *et al* (2013), who found that a AMIR construct targeting the AV1/AV2 overlap conferred resistance to ToLCNDV in tomatoes, but a AMIR construct that only targeted the CP only conferred mild tolerance. The degree to which the CP is dispensable in bipartite geminiviruses seems to be dependent of the specific virus-host interaction and the genetic background of the virus (Bull *et al*, 2007; Pooma *et al*, 1996).

Although the CP is not required for systemic spread of geminiviruses in the plant, the CP is required for correct transmission of the virus from the vector to the host (Liu *et al*, 1997). In a study conducted by Liu *et al* (1997) it was noticed that two sub-clones ACMV and TGMV were not transmissible by whitefly, and further investigation showed that these clones were CP-deficient. This study suggests that although the CP may not be useful as a trigger of PTGS, it could still be used in the control of CMGs in the field.

1.5.4.2 AV2- mediated resistance

The AV2 ORF plays a number of important roles in Old World begomoviruses. AV2 is responsible for viral accumulation and symptom development and also plays a role as a silencing suppressor (Chowda-Reddy *et al*, 2008; Padidam *et al*, 1996). AV2 was first shown to play a role in virus infection in *N. benthamiana*, transformed with antisense-RNA constructs targeting the AV2/AV1 overlap of TYLCV were resistance to the virus (Padidam *et al*, 1996).These results have been replicated in by number of other groups who have shown that silencing of AV2 results in resistance or attenuated symptoms in the host and lower viral loads (Basu *et al*, 2018; Bull *et al*, 2007; Ntui *et al*, 2015; Vu *et al*, 2013).

1.5.4.3 AC1-mediated resistance

The replication-associated protein plays a vital role in the replication of geminiviruses, regulating transcription from the bidirectional promoter and controlling rolling-circle amplification (Hanley-Bowdoin *et al*, 2004). Rep also interacts with other viral proteins such as REn (which is involved in viral DNA accumulation) as well as cellular proteins and as such has been the target of many RNAi-based control strategies (Patil *et al*, 2016; Vanderschuren *et al*, 2009).

Resistance to Rep has been induced through the expression of truncated portions of the protein and through induction of PTGS using anti-sense as well as hp-RNA constructs (Brunetti *et al*, 1997; Fuentes *et al*, 2006; Noris *et al*, 1997; Vanderschuren *et al*, 2009). A N-terminal truncated Rep protein (from *Tomato Yellow Leaf Curl Sardinia virus*) was shown to inhibit Rep transcription, but it was also shown that the virus could eventually overcome this and infect the plant (Lucioli *et al*, 2003). This would indicate that the expression of a protein as a means of resistance is not the most robust form of protection. In addition, as mentioned previously, Rep interacts with a number of proteins in plants, and the recovery of phenotypically normal transformants could be hindered by the constitutive expression of the Rep (Kong and Hanley-Bowdoin, 2002).

Rep is required for the replication of geminiviruses and as such is a promising target for RNAi mediated resistance. It was shown that *N. benthamiana* and cassava varieties infected with ACMV, only varieties that accumulated high levels of siRNA targeting Rep showed recovery (Chellappan *et al*, 2004b). Anti-sense and hp-RNA constructs targeting Rep have also exhibited resistance which was either only broken at high levels of virus (Vanderschuren *et al*, 2009), or showed durable resistance even in the field (Fuentes *et al*, 2016).

1.5.4.4 AC2-mediated resistance

The AC2 protein has also been shown to be an effective target for RNAi based resistance. AC2 is an important target for the control of geminivirus, as it functions as both transcriptional activator(Hanley-Bowdoin *et al*, 2004) and as a silencing suppressor (Trinks *et al*, 2005). A study conducted on siRNA levels in cassava and *N. benthamiana* infected with ACMV showed plants that produced high levels of siRNA targeting the AC1/AC2 overlap showed recovery (Chellappan *et al*, 2004b).

Hairpin-RNA construct which target the full length AC2 ORF of ACMV-CM (Patil *et al*, 2016) and C2 of *Tomato leaf curl Taiwan virus*(Lin *et al*, 2012b) produce high levels of siRNA, and corresponding levels of resistance. These results are similar to those by Moralo, (2015) and Ribeiro *et al*, (2007) who found that targeting the AC2 *Tomato chlorotic mottle virus* and ACMV respectively, could induce tolerance in their host plants. 1.5.4.5 AC4-mediated resistance

AC4 acts as a silencing suppressor in bipartite geminiviruses. It is encoded by an ORF found within or overlapping the AC1 ORF of the DNA-A molecule (Chellappan *et al*, 2005), therefore most RNAi constructs which target AC4 target some portion of the AC1 too. RNAi constructs which target AC4 have been shown to induce resistance *Mungbean yellow mosaic virus* (Sunitha *et al*, 2013) and transient resistance against ACMV (Patil *et al*, 2016). In a study conducted by Patil *et al* (2016), transient *Agrobacterium*-infiltration assays of hp-RNAi constructs targeting ACMV showed that a hp-RNAi targeting AC4 was an efficient inducer of silencing and a potential target for resistance in cassava. The only caveat to using hp-RNA targeting AC4 is that some constructs targeting AC4 have been shown to induce developmental abnormalities in transformed plants, specifically AC4 from ACMV and MYMV AC4 produced abnormalities in *Arabidopsis* and tomatoes, respectively. For this reason, all transgenic lines should be screened thoroughly for off-target effects from the construct (Chellappan *et al*, 2005; Praveen *et al*, 2010).

1.5.4.6 IR-mediated resistance

Double-stranded RNA has been shown to induce both PTGS and TGS in plants (Mette *et al*, 2000). In addition to targeting the ORFs of viral proteins, RNAi has also been use to target the viral protein promoters, inducing TGS of viral proteins. Targeting the promoter regions of protein induces the production of 24nt siRNAs which results in the methylation of viral promoters. Hairpin-RNA constructs which target the bidirectional promoter and common regions of geminivirus have been shown to successfully induce resistance to number of virus including MYMV and ACMV-Nigeria (ACMV-NOg) (Pooggin *et al*., 2003; Vanderschuren *et al*., 2007a).

1.6 Improvement of cassava

Cassava is grown for its roots, have one of the highest starch content of all crops (Westerbergh *et al*, 2009). When grown under optimum conditions cassava can produce 50Mt of fresh root weight per hectare however, biotic and abiotic stresses can significantly lower this yield (Legg and Thresh, 2000). In order to improve yields, cassava has been genetically modified using a number of methods including traditional breeding, RNA-based functional genomics approaches as well as reverse genetic and gene-editing approaches.

1.6.1 Breeding for improved traits in cassava

Cassava is mainly propagated through stem cuttings or seeds, although due the low fertility and irregular flowering, stem cutting are preferred by farmers (Ceballos *et al*., 2004). However, in order to introduce useful genetic diversity, breeding programs rely on propagation from true seeds, with cross-breeding (Alves, 2002). Cassava is monocious with male and female flowers opening at different time points (females 10-14 days before males) which allows for cross-breeding (Alves, 2002; Jennings and Iglesias, 2009). In order to improve cassava crops, usually the germplasm base is evaluated for desired traits, after which parental lines are selected and cross-breeding occurs, either through controlled pollination or in poly-cross nurseries with open pollination (Ceballos *et al*., 2004). A number of traits have been improved through traditional breeding including starch content, cyanogenic potential (Dixon *et al*., 1994) as well as improved resistance to biotic and abiotic stresses (Ceballos *et al*., 2004).

Although breeding is the most commonly used method of introducing new traits into cassava (Fondong and Rey, 2018), it is made challenging by a number of different factors including; the highly heterozygous nature of cassava genomes, irregular flowering (which affects both cross-pollination and seed production), low fertility and germination rates and the dearth of information on genes in the core germplasm collections (Fondong and Rey, 2018).

1.6.2 RNA-based functional genomics

Hairpin-RNA, co-suppression and anti-sense RNA silencing approach has been widely adopted as a means of improving crops (Cao *et al*., 2013; Wesley *et al*.,

2001). In cassava it has been used to improve viral resistance to a number of viruses including ACMV (Vanderschuren *et al*., 2007a), Sri-Lankan CMV (Ntui *et al*., 2015) and CBSUV (Yadav *et al*., 2011). These technologies have also been used to increase starch yield (Ihemere *et al*., 2006) and quality as well as nutritional content (Chavarriaga-aguirre *et al*., 2016).

Although hp-RNA (and anti-sense) technology has been widely adopted, there are several disadvantages associated with the technology including the stability and expression of the inserted transgene (Fondong and Rey, 2018) as well as off target effects. To over-come these limitations, small -RNA mediated silencing where the construct produces a single sRNA (Wagaba *et al*., 2016) have also been developed and used to improve cassava. This includes miRNA and tasiRNA which have been used to improve resistance to CBSV (Wagaba *et al*., 2016) and cassava bacterial blight caused by *Xanthamonas axonopodis* pv. *manihotis* (*Xam*) (Shybut, 2015).

1.6.3 Reverse genetic and genome editing techniques

A number of targeted gene-editing technologies have been used to edit the cassava genome, to improve disease resistance (Fondong, 2017; Gomez *et al*., 2018; Mehta *et al*., 2019; Shybut, 2015) as well as other physiological traits such a improved starch synthesis and quality (Bull *et al*., 2018). These technologies, including transcription activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeats (CRISPR), can be used to edit the genome using engineered sequence specific nucleases which induce double stranded breaks in the DNA strands of the genome, which are the repaired using either non-homologous end joining (NHEJ) and homology-directed repair (HDR) (Belhaj *et al*., 2015; Odipio *et al*., 2017; Ran *et al*., 2013).

Xam is one of the most important bacterial diseases in cassava (Chavarriaga-aguirre *et al*., 2016). *Xam* infection is reliant on transcriptional activator-like (TAL) effectors, which bind to host genes and regulate host proteins to allow infection (Bogdanove *et al*., 2010). In order to efficiently bind to a host promoter, the promoter must have a effector binding element (ELE) which contains a central region of 34-35 almost identical residue repeats (Boch, 2009). The ELE are present in a number of susceptible and resistance genes and can either improve infection or activate resistance. Using TAL effector nucleases (TALEN), the genome of cassava can be

modified to activate R genes or to remove ELE from susceptibility genes (Shybut, 2015).

Prokaryotes possess a nucleic acid-based defence, known as CRISPR which acts as an adaptive immune response to plasmids, phages and transposons (Kincaid and Sullivan, 2012). Cas9 nuclease (CRISPR-associated protein 9) introduces double stranded cleavage in the DNA, using the CRISPR sequence as a guide. This break is then repaired using NHEJ generating either an insertion or deletion mutation disrupting the normal function of the target gene (Kincaid and Sullivan, 2012; Pennisi, 2013). The CRIPSR/Cas9 system can be recruited to a specific site using a guide RNA (gRNA) strand of ~20 nt complementary to the target sequence in conjunction with a Cas9 nuclease that cleaves 4nt and can be used to control the gene expression (Schaeffer and Nakata, 2015). The CRISPR/Cas9 has been used to target a number of specific traits in both monocots and dicots including: biotic and abiotic stresses and yield improvement (Jaganathan *et al*, 2018).

As mentioned previously, cassava has a very large number of genotypes which have been bred with the specific aim of improvement of various aspects such as improved virus resistance, improved drought tolerance and improved nutrition. One of the limitations of traditional breeding for the improvement of cassava is the lack of knowledge of the genetic profile of these unique germplasm. Using this gene editing technique in combination with reverse genetic system TILLING (Targeting Induced Local Lesions IN Genomes), genes of interest can be indentified in cassava genotypes, which can then be exploited for further crop improvement (Henikoff *et al*., 2004). This technique may help identify unique alleles in cassava lines which confer tolerance and resistance to abiotic and biotic stresses for crop improvement.

1.7 Cassava mosaic disease resistance

Cassava mosaic disease is the most economically important disease of cassava in sub-Saharan Africa (Rabbi *et al*, 2014) and which, despite the significant efforts to curb its influence on crop yields, still causes crop losses of up to 30% annual (FAO 2009). In order to control CMD effectively, three important control measures must be implemented; removing symptomatic (infected) plants through systematic rogueing, deploying resistant varieties and controlled use of virus-free plant material (Rabbi *et al*, 2014). Most of the cassava in sub-Saharan is grown by subsistence farmers in

poor areas, which make rogueing and distribution of virus-free plant material difficult to achieve, so the focus of many cassava improvement projects has been the development of virus-resistant cultivars through genetic breeding (Rabbi *et al*, 2014).

1.7.1 Breeding for resistance against CMGs

Traditionally hybridisation has been used to breed high yielding cassava varieties which are either tolerant or resistant to CMGs (Okogbenin *et al*, 2013; Rabbi *et al, 2014)*. Currently CMD resistance is associated with three genetically distinct mechanisms referred to as CMD1, CMD2 and CMD3 (Fondong, 2017). CMD1 is recessive, polygenic trait which was introgressed into cassava from *Manihot glaziovii* Muell-Arg (Akano *et al*, 2002; Fondong, 2017). The presence of CMD1-associated resistance in a cassava line (referred to as TMS lines) is characterised by lower rates of infection and less severe symptoms in resistant lines, with the symptoms being restricted to isolated shoots and branches in infected plants. In addition to this, the viral levels in the infected lines are lower than those in susceptible cultivars (Fondong, 2017; Patil and Fauquet, 2015). As CMD1 is a polygenetic trait, its usefulness in breeding for resistance in cassava lines is limited by the biology and heterozygous nature of cassava.

Following the discovery of CMD1, a monogenic resistance locus was discovered in some West African landraces (referred to as tropical *M. esculenta* or TME), which is now referred to as CMD2 (Akano *et al*, 2002). CMD2 is easily heritable and confers stable resistance to number of CMGs and as such these landraces were used in a number of breeding programs especially in Latin America and Africa (Fondong, 2017; Kuria *et al*, 2017; Rabbi *et al*, 2014). TME3, TME7 and TME14 are completely resistant to ACMV and show recovery to EACMV after 45 d.p.i (Okogbenin *et al*, 2013). TME3 displays a tolerant-recovery (at 67 dpi) phenotype (Allie *et al*, 2014) when infected with SACMV. However, CMD2 resistance has more recently been found to have a number of limitations, firstly it was recently discovered that CMD2 resistance can be overcome by a number of CMGs in the presence of SEGS-1 and SEGS-2 (sequences enhancing geminivirus symptoms) (Maredza *et al*, 2016; Ndunguru *et al*, 2016). Secondly, the CMD2 locus is lost during somatic embryogenesis of cassava lines, which limits the use of these lines in cassava improvement projects (Beyene *et al*, 2015).

Finally, CMD3 was developed through cross-breeding of TMS30572 carrying CMD1 and TME6 carrying CMD2 (Kuria *et al*, 2017). This highly resistant cultivar (TMS 97/2205) was screened for polymorphic markers which showed the presence of the CMD2 marker, in conjunction with a quantitative trait locus (QTL) (which is now designated CMD3). This line is highly resistant to CMD, with field trials in Nigeria recording incidence levels of below 1% (Okogbenin *et al*, 2010). A genome-wide association study by Wolfe *et al*, (2016) showed that the CMD2 locus is on chromosome 8 (now chromosome 12 cassava v6.1, phytozome), but that two other possible epistatic loci or multiple resistance alleles may modify the resistance response of the major CMD2 QTL.

Although breeding has yielded a number of resistant (at best) or at least CMD tolerant lines, the biological characteristics of cassava including its heterozygous nature, its long growth cycle and a poor understanding of the mechanisms of resistance in the crop, have made breeding for resistance a slow and laborious task (Okogbenin *et al*, 2013). Plant genetic transformation technology, where resistance genes can be transferred to traditional cultivars without the possible appearance of undesirable traits has provided a new avenue for creating resistant cassava cultivars (Patil and Fauquet, 2009).

1.7.2 Resistance, tolerance and recovery

Over the last three decades, plant research groups have worked towards engineering plants which better survive infection by viruses. Within these groups, three terms; resistance, tolerance and recovery are used, but have not been completely defined. For the purpose of clarity, the terms are defined here as those found in Lapidot and Friedmann (2002). In order to establish systemic infection of a plant, a virus must counteract the defence response and take control of host factors which are required for its replication cycle. Where a virus fails to establish infection, symptoms are usually restricted to the infection site (Chellappan *et al*, 2004b), and no systemic symptoms or virus replication can be observed; these plants are referred to as resistant. When a virus establishes systemic infection, the symptoms can vary; some plants show reduced symptoms (tolerant) while others recover (different levels) from symptoms but continue to have low virus loads (Ghoshal and

Sanfaçon, 2015). Symptom variation is dependent on the proteins the virus interferes with and/or accumulation of viral nucleic acids and proteins.

The phenomenon referred to as 'recovery' was first reported in tobacco plants in 1928, in tobacco plants infected with *Tobacco ringspot virus* (Wingard, 1928). During the initial stages of infection, the entire plant was symptomatic however after several growth cycles, the new upper leaves of the infected plants began to show milder or no symptoms of infection (Ghoshal and Sanfaçon, 2015). Recovery can be partial, where the plant displayed much milder symptoms than in the early stages of infection, or the plant can become completely asymptomatic. Recovery has been observed in a number of plant species, in response to both RNA and DNA viruses (Chellappan *et al*, 2004b; Ghoshal and Sanfaçon, 2015). Recovery has been shown to be linked to the reduction of viral RNA and DNA, and the accumulation of siRNA (Chellappan *et al*, 2004b; Ghoshal and Sanfaçon, 2015). While this phenomenon is usually linked with the induction of RNA silencing, Bengyella and Rey (2015) showed that resistant gene analogs (R genes) also play a role in the recovery phenotype.

1.7.3 Transgenic technologies for developing disease resistance

Cassava is an economically important plant and has wide variety of uses. Traditional breeding has been used to improve disease resistance and increase levels of starch and nutrients in the roots but breeding for new traits has proven difficult due to the heterozygous nature of the plant (Bull *et al*, 2009; El-Sharkawy, 2004; Hahn *et al*, 1980). For this reason, the advent of transformation technologies which can be used to integrate desired traits into cassava has been highly advantageous for production. Several strategies have been used to improve specific traits in cassava, including the over-expression of native or heterologous genes under the control of constitutive promoters like *Cauliflower mosaic virus* 35S or *Nopaline synthase (*NOS). Tissue specific promoters, such as potato patatin class I promoter (B33), have also been used to promote expression of genes in specific organs (Zhang *et al*, 2003, 2017). Genes have also been down regulated, through the use of anti-sense and RNAi technologies, in order to modify nutrient and starch levels in the plant, as well as improve virus resistance (Ihemere *et al*, 2006; Vanderschuren *et al*, 2009, 2007a).

1.7.3.1 Cassava mosaic virus resistance

Cassava mosaic disease is one of the most important diseases affecting cassava production in Africa (Legg and Fauquet, 2004). RNAi is a proven technology which can be used to improve cassava's resistance to many of the CMGs responsible for CMD (Vanderschuren *et al*, 2007a), and both anti-sense and hp-RNA technology have been used to improve resistance in model crop TMS. 60444 (Vanderschuren *et al*, 2009, 2007a; Zhang *et al*, 2005).

The AC1, AC2 and AC3 proteins are all involved in the early stages of viral replication and have been shown to be effective targets for anti-sense construct mediated PTGS of ACMV (Zhang *et al*, 2005). Transgenic plants challenged with ACMV-NOg were found to be either completely resistant, or to display recovery when inoculated with 100ng of viral DNA. The resistance broke when viral loads exceeded this value, however, there was still a reduction in the amount of virus detected and symptom severity in the plant and most lines displayed recovery(Zhang *et al*, 2005).

1.7.3.2 Challenges associated with development of a CMD resistant cassava

The field conditions associated with CMD are complex and include both environmental factors and natural variations in the viral population. Disease incidence in the field is affected a number of key factors: rainfall and climatic conditions in the field, the age of the plants at infection (with young plants being more susceptible to infection) and vector populations (Legg and Thresh, 2000; Vanderschuren *et al*, 2007b). In addition to the challenges of the field, CMD is usually caused by a complex of different CMGs, and any resistance strategy must take this into account (Vanderschuren *et al*, 2007b). Geminiviruses are prone to recombination, pseudo-recombination and as well as synergism, which allows for their rapid evolution in response to pressure in the field (Duffy and Holmes, 2009; Pita *et al*, 2001; Vanderschuren *et al*, 2007b). Any resistant crop must be able to withstand these pressures and provide durable resistance in the field. A recent study using PACBIO CIDER-Seq enriched individual full-length genome sequencing, has shown that cassava geminiviruses mutate rapidly in field-grown conditions (Mehta *et*

al, 2017) and that hot spots for mutation were regions of the viral genome where the siRNAs targeted.

1.7.3.3 CRISPR-base technology

The proof of concept study was conducted on cassava TMS.60444 and TME204, where the *phytoene desuturase* gene was silenced using the CRISPR/Cas9 system (Odipio *et al*, 2017) and transgenic lines produced albino plants. Since this study the CRISPR/Cas9 system has also been used to engineer improved starch production and quality, and flowering in cassava lines (Bull *et al*, 2018b). A number of research groups have produced cassava, with CRISPR/Cas9 systems designed to confer resistance to cassava viral pathogen, including CBSV and ACMV*,* however to date no fully resistant lines have been produced (Gomez et al., 2018; Mehta et al., 2019). Mehta *et al* (2019) also noted that ACMV quickly developed resistance to the CRISPR/CAS9 system through a single point mutation, which highlighted the importance of target design in this system (Mehta et al., 2019).

1.8 Biofortification of cassava

Cassava is the most important source of calories for over 250 million people in sub-Saharan Africa, and this number is likely to grow as it is one of the fastest growing food source crops globally (Chetty *et al*, 2013; FAO, 2016; Leyva-Guerrero *et al*, 2012). Although it is high in starch, it is lacking in other key micronutrients including zinc (Gaitán-Solís et al., 2015), iron (Narayanan et al., 2015) and vitamins A, B6 and E (Leyva-Guerrero *et al*, 2012; Zhang *et al*, 2017). Several research groups, including BioCassava Plus, and the Root, Tuber and Banana consortium have invested in the bio-fortification of cassava, in order to improve its potential as food crop (Zhang *et al*, 2017).

1.9 Starch

Starch is an insoluble glucan made up two polymers of glucose, amylopectin and amylose, and is a primary storage molecule of all higher plants (Zeeman *et al*, 2010). Starch is synthesised in both the photosynthetic and non-photosynthetic parts of the plant. Starch is the most abundant carbohydrate in plants (Zeeman *et al*, 2010), however, starches from different plant species have different polymer compositions and structures and therefore have different properties and applications.

Starch synthesis in plants has been extensively studied and is relatively well understood (Smith, 2008). Sucrose enters the cell and is converted to glucose-6-phosphate (this occurs in all studied plants except for cereal endosperms). It is then transported into the plastid using a glucose-6-phophate transporter (Kammerer *et al*, 1998), where it is converted to ADP-glucose. This process requires ATP, which is transported into the plastid by adenylate transporter (Tjaden *et al*, 1998), where it is converted into starch. In graminaceous species, starch synthesis differs in that the synthesis of ADP-glucose occurs in the cytosol using cytosolic form of AGPase (**Figure 1.5**). The ADP- glucose is then imported into the cytosol using a specific sugar nucleotide transporter (Tomlinson and Denyer, 2003).

Plants with high starch content form most of the world's staple food crops with cereal seeds such as rice, maize and wheat, being our most important crops, followed by storage roots (e.g. potato, yams) and storage roots (e.g. cassava). A large part of the world agricultural land is devoted to the production of these crops (Zeeman *et al*, 2010). The majority of these crops are consumed directly or are used as animal feed, however there is also a growing need for starch in industries such as biofuels. These secondary uses of starch are placing an increased demand on food stocks. In order to deal with the increased demand on this resource, there is an increased interest in producing high yields of quality starch. Cassava starch is recognised as a valuable commodity and is used for the production of bioethanol, feedstock and industrial and food applications.

Attempts to increase the amount of starch have focussed on three main areas of starch synthesis: ADP-glucose pyrophosphorylase, starch synthase and the supply of ATP to the plastid (Smith, 2008). ATP levels have been manipulated using two different mechanisms; increasing the activity of adenylate transporter of the plastid envelope (Geigenberger *et al*, 2001) and altering of the adenylate pools by down-regulating plastidial adenylate kinase (Regierer *et al*, 2002). It was demonstrated that increasing the supply of ATP to the plastid stimulates the production of ADP-glucose (thereby increasing the rate of starch synthesis) and could lead to increase in starch yields of up between 16–36% in transgenic potatoes, compared with control storage roots (Geigenberger *et al*, 2001). Also, when plastidial adenylate kinase, which is responsible for the inter-conversion of two molecules of ADP into ATP and AMP, was down regulated it resulted in a tenfold increase in ADP-glucose levels which

translated into a 200% increase in the starch content in potato tuber (Regierer *et al*, 2002).

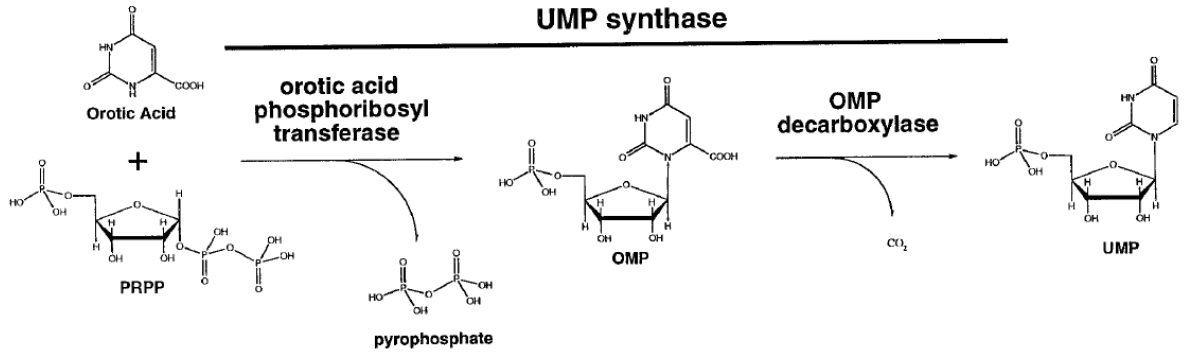


Figure 1.4: Action of UMP synthase during the De Novo synthesis of pyrimidine nucleotides in plants (Santoso and Thornburg, 2002)

Another approach which has shown promise in altering the starch levels in potatoes and may have applications in other crops, is the manipulation of de novo pyrimidine synthesis pathway (Geigenberger *et al*, 2005). Transgenic potatoes, expressing an antisense construct targeting UMP synthase, a key enzyme involved in the synthesis of pyrimidines (**Figure 1.4)**, under the control of a root-specific promoter, showed increased starch levels in potatoes. These approaches could have a large advantage for cassava where an increase in starch in the roots could be hugely beneficial.

1.9.1 Modification of cassava starch

Cassava is one of the most important sources of calories in sub-Saharan Africa and is also a vital source of industrial starch globally (Karlström *et al*, 2016). Cassava starch has a wide range of industrial applications, as it has a bland taste and produces a clear paste. While the potential starch yield of cassava is higher than either maize or rice, sub-optimal condition means that production levels are usually lower. Native cassava starch also has limitations, including low solubility and retrogradation (Zhang *et al*, 2017). The demand for high yields of high-quality cassava starch is increasing and there have been many studies which have modified cassava starch physically, chemically or through the use of biotechnology (Taylor *et al*, 2004).

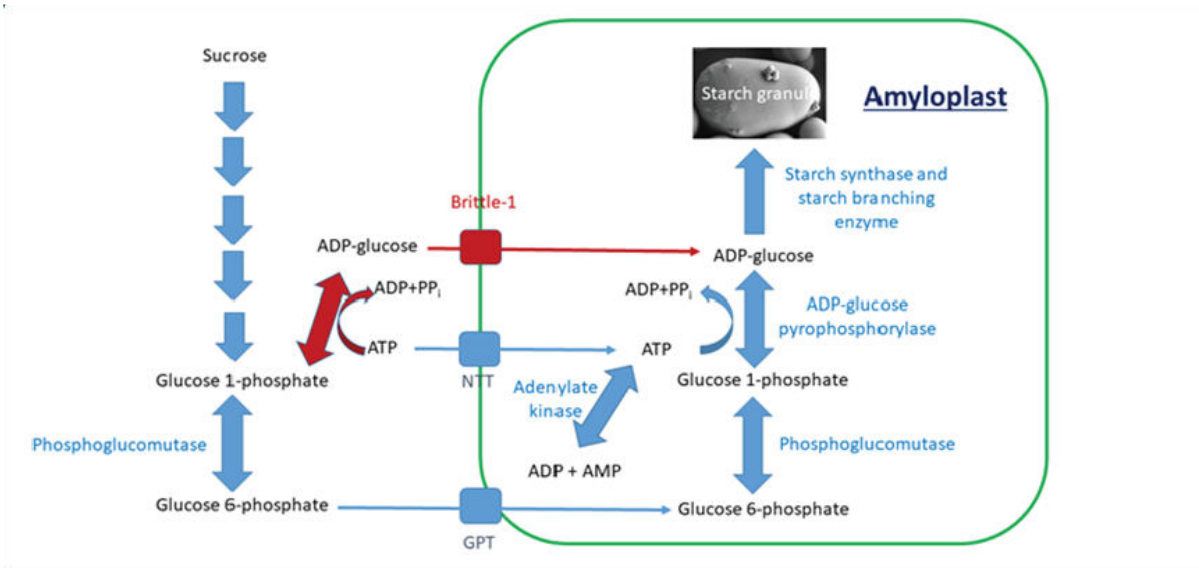


Figure 1.5: Schematic diagram of starch synthesis in plastids. Sucrose is imported into the cell where it is broken down into UDP-glucose and Fructose by Sucrose synthase. The UDP-glucose is then converted to glucose-6-phosphate by sequential manipulation by UDP-glucose pyrophosphorylase and cytosolic phosphoglucomutase. Glucose-6-phosphate is then transported into the amyloplast using a hexose phosphate translocator. The Glucose-6-phosphate is then converted into ADP-glucose via plastidic phosphoglucomutase (G-6-P to G-1-P) and ADP-glucose pyrophosphorylase. The ADP-glucose is then converted to starch by granule-bound starch synthase; soluble starch synthase and branching enzymes. The conversion of G-1-P into ADP-glucose requires ATP, which is transported into the amyloplast by a adenylate translocator (Modified from Lloyd and Kossmann, (2019).

One of the most important characteristics of cassava starch is the ratio of amylose to amylopectin, which determines its use in different industrial applications (Liu *et al*, 2011). In cassava (depending on the genotype and growth environment) the amylopectin content ranges from 70%-80% while its amylose content ranges from 20%-30%. Amylose is a low molecular weight polysaccharide, which produces a gel when boiled, amylopectin is a higher molecular weight polysaccharide with a high viscosity (Liu *et al*, 2011). Amylose is synthesized by the glycosyltransferase GRANULE BOUND STARCH SYNTHASE (GBSS), which tightly associates with the starch granules (Shure *et al*, 1983).Zhao *et al*(2011) developed an amylose free cassava which through the down regulation of granule bound starch synthase (GBSSI) under the control of the CaMV35S promoter. In addition to GBSS, amylose synthesis is mediated by PROTEIN TARGETING TO STARCH (PTST1) which

localises the GBSS to the starch granules (Seung *et al*, 2015). Using the CRISPR/Cas9, Bull *et al* (2018) developed a transgenic amylose-free cassava starch. This starch has a higher viscosity than normal starch as is in high demand for industrial applications (Zhang *et al*, 2017).

Cassava's roots are its main commercial product and the development of high quality, high yielding roots has been the focus of much of the research on cassava. Proteomic and transcriptomic data from cassava has shown that during storage root development, a large proportion of the up-regulated genes are related to starch and sucrose metabolism, and the majority of the differential expressed genes were binding-related enzymes (Wang *et al*, 2016). Ihemere *et al* (2006) showed that the expression of a modified *Escherichia coli* ADP-glucose pyrophosphorylase gene *(glgC*) in cassava caused a 2.6-fold increase in root storage biomass. Ligaba-Osena *et al*(2018) developed transgenic cassava which expressed the *E. coli glgC* gene as well as *Sulfolobus solfataricus* glucoamylase and hyperthermophilic archaeal starch-hydrolyzing enzymes, α-amylase and amylopullulanase from *Pyrococcus furiosus,* producing cassava roots with a 60% higher yield than the wild-type control. The starch is more easily hydrolysed due to the presence of the enzymes.

1.10 Rationale

Cassava is mainly grown for its roots, which are rich in starch (El-Sharkawy, 2004). Cassava's carbohydrate yield is the third highest per cultivated area, of all crop plants and can be grown in poor soils with low rainfall. Cassava therefore it has great value as a food security crop, but it also has value in industrial starch applications such as bioethanol production (Legg and Fauquet, 2004). In South Africa cassava is currently grown by small-hold farmers, as a secondary crop (ARC, 2010) however an international increase in demand for cassava (ARC, 2010) represent a opportunity for increased cassava cultivation. Small-hold farmers usually have limited budgets for pest control and soil improvement, so research in cassava should focus on improving current yields and reducing the threat of diseases such as CMD. The most efficient means of developing improved cassava varieties is through the use of transgenic technologies such as RNAi, which has been used to improve resistance to a number of CMGs (Vanderschuren et al., 2007b) and increase the starch

production (Ihemere et al., 2006). This would increase the economic value cassava and make it a more viable economic crop for subsistence farmers.

1.10.1 Objectives

In South Africa, cassava is usually only grown on a small scale by subsistence farmers, but in several other African countries, its industrial potential is realized. For example in Mozambique it is grown commercially for starch used for cooking-grade ethanol and beer fermentation. Cassava is mainly grown as a famine crop in conjunction with maize, however cassava crops produce a high yield starch and have potential use in a number of areas including bioethanol production (Dai *et al*, 2006). Cassava is usually grown under sub-optimal conditions; however a bigger threat to cassava crops is CMD, which can devastate crops, sometimes leading to losses of 100%. In order for cassava to be an economically viable crop, it is important that CMD is brought under control. The most viable method of controlling CMD is through the production of resistant crops using transgenic technologies. In order to ensure that these technologies are safe, the 'off target effect' must be fully understood. Further, in order to make cassava an economically valuable crop outside of a famine crop, the starch production needs to be optimal which can also be engineered through the use transgenic technologies.

The objective of this study is aligned with the priority to improve the yield of cassava by reducing the impact of CMD and increasing starch production:

1. Evaluate the ability of SACMV AC1/AC4 hp-RNA constructs to confer resistance via RNA silencing in cassava to SACMV. 2. Investigation of the 'off-target effect' observed in FECs transformed with a triple-stacked construct targeting the AC1/ AC4 and IR/AC1 region of three geminiviruses associated with CMD in South Africa, namely SACMV, ACMV and EACMV

3. The third objective was to increase the starch production of cassava through the down-regulation of UMP synthase levels and plastidial adenylate kinase orthologs. Hp- RNA silencing of expression of each gene was evaluated separately.

1.10.2 Specific Aims

Aim 1: Screening of cv.60444 SACMV AC1/AC4 transgenic cassava lines for SACMV resistance

Aim 2: Transform cassava cv.60444 with a triple stacked hp-RNA construct targeting SACMV+ACMV+EACMV, double hp-RNA construct targeting EACMV/ACMV and a single hp-RNA construct EACMV to elucidate mechanisms of 'off-target effects.

Aim 3: Transform cassava cv.60444 with hp-RNA construct targeting ADK and UMP, to improve starch production in transgenic line.

Chapter 2: RNA silencing of South African cassava mosaic virus in transgenic cassava expressing AC1/AC4 hp-RNA induces tolerance

2.1 Introduction

Cassava (*Manihot esculenta* Crantz) is a perennial shrub grown for its tuberous roots, which are a major source of food for almost 700 million people worldwide (El-Sharkawy, 2004; Patil and Fauquet, 2009). One of the greatest threats to cassava crop security is cassava mosaic disease (CMD) which causes deformation and chlorotic mosaic in the leaves and is responsible for huge yield losses (Hahn *et al*., 1980; Legg, 1999; Obambi *et al*., 2011; Rey and Vanderschuren, 2017). Although yield losses due to CMD are dependent on the geographical region of cultivation and cultivar susceptibility, in Africa the average annual yield loss losses due to CMD range between 30 and 40% (Thresh and Cooter, 2005). However losses as great as 90% have been recorded in some areas (Owor and Legg, 2004). In sub-Saharan Africa, CMD is caused by at least 7 cassava mosaic geminivirus (CMG) species including South African cassava mosaic virus (SACMV) (Berrie *et al*., 2001), and many genetic strains/variants (Bart and Taylor, 2017; Fondong, 2017; Legg *et al*., 2015). *South African cassava mosaic virus* (genus: *Begomovirus)* is a bipartite circular single stranded DNA (ssDNA) virus transmitted by the whitefly *Bemisia tabaci* (Genn.) in a persistent manner. Its genome consists of two ssDNA molecules, DNA-A and DNA-B, that are separately encapsidated in twin icosahedral particles (Briddon and Markham, 1995; Gutierrez, 1999; Hanley-Bowdoin *et al*., 1999; Rojas *et al*., 2001). DNA-A is required for transcription and replication of the virus, while DNA-B is required for cell-to-cell and long distance movement (Castillo *et al*., 2004; Chellappan *et al*., 2005; Gutierrez, 1999). DNA-A contains 6 open reading frames (ORF) which encode for 6 proteins namely: AC1 (Replication associated protein/Rep), AC2 (Transcriptional activator protein/TrAP), AC3 (Replication enhancer protein/ REn), AC4, AV1 (coat protein/CP) and AV2 (pathogenicity

determinant). Rep is a highly conserved multifunctional protein which is not only essential for viral replication, but also regulates transcription of viral proteins (Fontes *et al*., 1992; Hanley-Bowdoin *et al*., 2013). Rep also interacts with several host proteins, and acts as a viral suppressor of host response to geminiviral pathogens by lowering the transcription levels of plant methyltransferases which can methylate viral genomes (Rodríguez-Negrete *et al*., 2013). The AC4 ORF overlaps the AC1 ORF and encodes a silencing suppressor that plays a role in the suppression of both transcriptional and post-transcriptional gene silencing in plants (Vanitharani *et al*., 2004; Vinutha *et al*., 2018). Additionally, SACMV encodes a putative AC5 ORF, which functions as a virus suppressor of RNA silencing (VSR) in *Mungbean yellow mosaic India virus* (F. Li *et al*., 2015).

RNA interference (RNAi) is a highly conserved mechanism of gene regulation and plays a role in a variety of biological processes including defence against invading nucleic acids (Jones *et al*., 2001; Parent *et al*., 2015). RNAi is triggered by the presence of double stranded RNA (dsRNA), which is processed by several RNase III-like enzymes known as Dicer-like proteins (DCLs) to produce so-called small interfering RNAs (siRNAs) (Qu *et al*., 2008). The siRNAs associate with a number of effector proteins, including ARGONAUTES (AGOs), to form a RNA-induced Silencing complex (RISC) (Hammond, 2005; Snead and Rossi, 2010). AGOs are responsible for the cleavage of the passenger RNA from the siRNA duplex, which triggers the unwinding of the guide siRNA strand and activates the siRNA-RISC complex (Vaucheret and Fagard, 2001). The activated complex then uses the guide siRNA strand to target and mediate cleavage of homologous viral mRNA in a process known as Post-transcriptional gene silencing (PTGS) (Ketting *et al*., 2001). siRNA can also mediate transcription gene silencing (TGS), where the siRNA is incorporated into a RITS (RNA-induced initiation of transcriptional silencing) complex (Ekwall, 2004) and guides target gene and histone methylation (Borges and Martienssen, 2016). Small interfering RNA has also recently been shown to mediate translational repression in plants (Machado *et al*., 2017) and was also demonstrated in *mad6* and *ago1-27 Arabidopsis* mutants (Brodersen *et al*., 2008; Jain *et al*., 2016). AGO1, AGO2 and AGO19 have been implicated in translational repression in plants (Brodersen *et al*., 2008; Fátyol *et al*., 2016; Ghoshal and Sanfaçon, 2014). Translational repression of viral mRNA was first observed in association with the

defence response activated by a viral elicitor interaction with a resistance gene in *Nicotiana benthamiana* (Bhattacharjee *et al*., 2009), and also in *Tomato ringspot virus* infection of *N. benthamiana* (Ghoshal and Sanfaçon, 2014).

Geminivirus infection has been shown to induce the production of virus-derived siRNAs, which trigger both PTGS and 24 nt-mediated TGS silencing of the virus that has been linked to resistance and recovery (Patil and Fauquet, 2009; Vanitharani *et al*., 2005). Recovery is phenomenon found in plants, where a plant that is initially symptomatic shows milder symptom in the new growth cycle (Ghoshal and Sanfaçon, 2015). Plants which show recovery are referred to as tolerant. West African landrace TME3 displays a tolerant phenotype against *South African cassava mosaic virus* (Allie *et al*., 2014). Symptom recovery in virus-infected plants has been linked to the induction of RNA silencing and generally a concomitant reduction in virus levels (Baulcombe, 2004; Ghoshal and Sanfaçon, 2014). However, in contrast, recovery from tomato ringspot virus (ToRSV) in *N. benthamiana* is not associated with viral clearance in spite of active RNA silencing being triggered (Jovel *et al*., 2007). Recovered leaf tissues, infected with geminiviruses, have been shown to have lower small RNA (21-23nt) levels (Hagen *et al*., 2008), and this was also shown in SACMV-tolerant cassava TME3 (Rogans *et al*., 2016).

Cassava is heterozygous and has a strong inbreeding depression (Bredeson *et al*., 2016; de Freitas *et al*., 2016; Zhang *et al*., 2005), which has made production of CMG resistance cultivars through traditional breeding difficult. However, the genetic modification of cassava using a RNAi-mediated approach has shown great potential for increasing cassava resistance to a number of viruses including African cassava mosaic virus (ACMV) (Taylor *et al*., 2004; Vanderschuren *et al*., 2009; Zhang *et al*., 2005) and Sri-Lankan cassava mosaic virus (SLCMV) (Ntui *et al*., 2015). In this approach, transgene constructs are transformed into the plant, which then induce TGS or PTGS of the virus through the production of virus specific siRNAs derived from a hairpin or inverted repeat transgene. While siRNAs are produced from the entire begomovirus bipartite genome (Aregger *et al*., 2012), several studies have shown hot spots that can be exploited for developing anti-geminiviral resistance (Aregger *et al*., 2012; Lin *et al*., 2012b; Sharma *et al*., 2014). Further, AC1 and AC4 on DNA-A of geminiviruses have been identified as hotspots, with higher level of siRNAs associated with these regions, specifically targeting the AC1/AC4

overlapping region (Aregger *et al*., 2012; Miozzi *et al*., 2013). In most studies, the Rep ORF has been used as a preferred target due to its critical multifunctional activities (Chellappan et al., 2004b; Vanderschuren et al., 2009). However the overlapping C2/C3 (monocots) and AC2/AC3 (dicots) region has been reported to be effective for induction of resistance (Lin *et al*., 2012a). Viral suppressors of gene silencing, such as AC2 and AC4 (Vanitharani *et al*., 2004; Voinnet and Baulcombe, 1997), have also been shown to be hotspots for PTGS (Sharma *et al*., 2014). In cassava, AC1 and the bidirectional promoter region in the Intergenic region (IR) of ACMV have been targeted (Vanderschuren *et al*., 2009, 2007b).

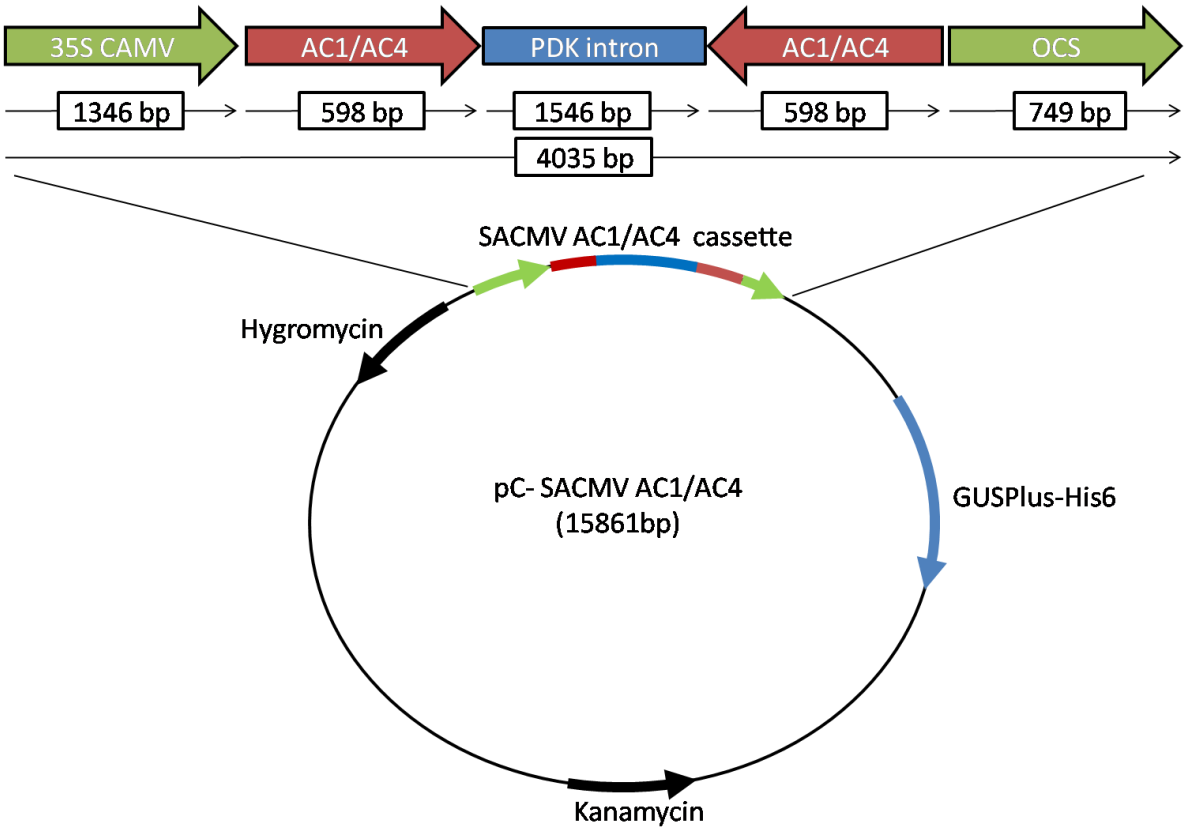
While engineering resistance to ACMV and SLCMV in cassava, and to Indian cassava mosaic virus (ICMV) in *Jatropha curcus* (Ye *et al*., 2014) has been accomplished, resistance against SACMV, has not yet been achieved by genetic engineering. High throughput sequencing of small RNAs associated with SACMV-infected cassava identified siRNAs targeting ORFs and IR on both sense and antisense DNA-A and B strands (Rogans *et al*., 2016). This study reports increased tolerance in genetically modified cassava cv.60444 expressing a hp-RNA from a transgene homologous to the AC1/AC4 overlap sequence from SACMV. Three transformed lines showed a decrease in viral load and symptom severity when compared to infected wild-type (untransformed cv.60444), and similar to the tolerance response of TME3 to SACMV. This increased tolerance was shown to be the result of enhanced PTGS, as confirmed by the expression of transgene-specific siRNAs.

2.2 Material and Methods

2.2.1 Plasmid construction and cassava transformation

A double stranded hairpin RNA (hp-RNA) construct targeting a 598bp region over-lapping the AC1/AC4 region of SACMV DNA-A, separated by a PDK intron was constructed previously by Taylor, (2009) using the pHannibal system (Wesley *et al*, 2001) (**Figure 2.1**). The AC1/AC4 hp-RNA construct, including the CaMV 35S promoter and OCS terminator was then inserted into the pCambia 1305.1 binary vector (AF354045.1) using directional cloning. The AC1/AC4 binary vector, transformed into *Agrobacterium tumefaciens* LBA4404 was used to transform

cassava cultivar 60444 friable embryonic callus (FECs) using the method described in Bull *et al* (2009).



.

Figure 2.1 Schematic representation of the SACMV hp-RNA AC1/AC4 construct in pCambia 1305.1 vector (pC-SACMV AC1/AC4). The construct targets the overlap between the ORFs of the replication-associated protein (AC1) and silencing suppressor AC4. The forward and reverse arms (AC1/AC4) of the hairpin are separated by the PDK intron and are under the control of the CaMV35S promoter (35S promoter) and OCS terminator.

2.2.2 Molecular characterisation of transgenic lines

2.2.2.1 Nucleic acid extraction

Total DNA was extracted from approximately 50mg of leaf material from 14 transgenic line using the CTAB method described in (Doyle, 1991). Total RNA was extracted from approximately 50mg of leaf tissue, collected from infected and mock inoculated plants and flash frozen in liquid nitrogen. The leaf tissue was ground into a fine powder using the QiagenTissueLyser II system (Qiagen) and total RNA was

extracted using Qiazol lysis reagent (Thermo Fisher Scientific), according to the manufacturer's instructions. Total RNA was treated with RiboLock (Thermo Fisher Scientific) and stored at -80°C.

2.2.2.2 PCR

Transgenic cv.60444 plants were screened by PCR for the presence of reporter genes *Hyg* and *Gus* as well as the forward arm of the AC1/AC4 construct with the following primers; HygF and HygR, GusPlusF and GusPlusR, and NBSACMVF and NBSACMVR respectively (**Table 2.1**). Purified empty vector control pCambia1305.1 and hp-RNA pC-AC1/AC4 vector DNA and wild-type untransformed cv.60444 DNA were included as controls.

Table 2.1: Primers used for screening of transgenic cassava cv.60444 transformed with SACMV hp-RNA construct AC1/AC4.

Target	Primer	Primer sequence (5'-3')	Tm
Gus reporter gene	GusPlusF	CAACATCCTCGACGACGATAGCA	54°C
	GusPlusR	GGTCACAACCGAGATCTCCT	
Hyg reporter gene	HygF	TCTCGATGAGCTCATGCTTTGG	56°C
	HygR	AGTACTTCTACACAGCCATGGG	
AC1/AC4 forward arm	pHANREPXhoI	CCTCGAGGTACTCGGTCTCCATGGCC	56°C
	pHANREPEcoRI	GGAATTCACTCTCTCGAAAGAAGCGG	
AC1/AC4 Northern probe	NBSACMVF	TAATACGACTCACTATAGGGTACTCGGTCTCCATGGCC	68°C
	NBSACMVR	AATTAACCCTCACTAAAGGGTCGAAAGAAGCGG	

2.2.2.3 Southern Blot Hybridisation

In order to determine transgene integration events in the 14 transgenic lines, Southern Blot hybridisation was carried out using DIG-High Prime DNA Labelling and Detection Starter Kit II (ThermoFisher Scientific) according to the manufacturer's instructions. Twenty micrograms of DNA from each line was digested overnight with HindIII (15053-15059) which cuts once in the cassette, and EcoRI, restriction enzymes (ThermoFisher Scientific) and separated using gel electrophoresis, in a 1

% agarose gel in 1X TAE. The DNA was then transferred to positively charged nylon Hybond-N+ membrane (Amersham), pre-hybridised at 38°C, and hybridised at 64°C overnight with a DIG-labelled *Hyg* gene probe, which was labelled using the DIG-High Prime DNA Labelling and Detection Starter Kit II (Roche). Pre-hybridization (3 h) and hybridization (overnight) were carried out using DIG EasyHyb buffer (ThermoFisher Scientific). Any unbound probe was washed from the membrane by sequential washing with buffers 2xSDS, 0.1% SDS twice, and 0.1xSDS, 0.1% SDS buffer twice (incubated at 40°C). Signal detection (CDP star) was performed following DIG-High Prime DNA Labelling, and the Detection Starter Kit II (Roche) protocol. Results were visualised using GelDoc XR+ (Biorad) after 10 min.

2.2.2.4 Expression of *Gus*, *Hyg* and SACMV AC1/AC4 transgene

In order to determine transgene expression, cDNA was synthesised from 1µg of total RNA (treated with DNase I) with random hexamer primers, using the RevertAid First strand cDNA synthesis kit (Thermo Fisher Scientific). Two microliters of cDNA product was used as the template for PCR amplification of *Gus, Hyg* and the AC1/AC4 transgene insert using primers GusPlusF and GusPlusR, HygF and HygR and pHANREPXhoI and pHANREPEcoRI, using DreamTaq (ThermoFisher Scientific) according to protocol, with annealing temperatures specified in Table 2.1. The relative expression of the transgene, compared to 1µg of purified PCR AC1/AC4 product amplified from pC-AC1/AC4 (reference band), was calculated using LabImage 4.0 (Bio-Rad).

2.2.3 Northern Blot for expression of AC1/AC4 hairpin-derived siRNA

2.2.3.1 Probe construction

RNA probes were produced from a SACMV AC1/AC4 PCR fragment using the DIG Northern Started kit (Roche) according to manufacturer's instructions. The PCR fragment was produced using Phusion Master mix (ThermoFisher Scientific) with primers NBSACMVF (including T7 promoter)and NBSACMVR (including T3 promoter) (Table 2.1) using 50 ng of pBS-SACMV DNA A plasmid as a template. Two hundred nanograms of purified PCR product quantified using NanoDropOne

(ThermoFisher Scientific), was used to produce the RNA DIG-labelled SACMV AC1/AC4 probe in both the sense and antisense orientation

2.2.3.2 Northern Hybridisation

Thirty micrograms of total RNA from transgenic lines was separated on 15% polyacrylamide gel (8M urea, 30% APS, 1x TBE, 0.4% TEMED) and transferred to a positively charged nylon membrane (Amersham, Hybond N+). Hybridisation was performed using DIG Northern Starter Kit (Thermo Fisher Scientific), according to the manufacturer's instructions. Pre-hybridization (60 min) and hybridization (over-night) were carried out at 60°C , using DIG EasyHyb hybridisation buffer (Thermo Fisher Scientific). The post-hybridization removal of excess probe and signal detection (CDP star) were performed according to Northern Starter Kit instructions (Thermo Fisher Scientific).Excess probe was removed by sequential washing of membrane in (2x SSC, 0.1 SDS) and (0.1 × SSC, 0.1% SDS) buffers incubated at 25° and 60°C,respectively.Northern Blots were exposed for 10 min and imaged using BioRadTransluminator. Labelled 21 nt miRNA (mes-miR169) was used as a positive control.

2.2.4 Evaluation of transgenic AC1/AC4 transgenic lines for resistance to SACMV

2.2.4.1 Agro-inoculation of transgenic lines

From fourteen lines initially screened in a first trial, six transgenic lines (O14-1, O4-1, O12-2, O4-4 O13-5 and O13-8) which displayed lower viral loads and reduced symptom severity, in comparison to the untransformed cassava cv.60444 control as well as untransformed cv.60444 and tolerant landrace TME3 were selected for further virus resistance trials. Twenty to thirty plantlets from each line were micro-propagated in a controlled environment at 28 °C, with 16 h light (8000-10000 lux) and 8 h dark cycles and 50 % humidity. After 6 weeks, 12 plants of similar size were selected from each transgenic line. Twelve plants from each line were agro-inoculated with SACMV pBIN-DNA-A and pBIN-DNA-B infectious clones in *A. tumefaciens* Agl1 (Agl1) (Berrie *et al*., 2001). The infectious clones were prepared by inoculating 50 ml of YEP broth containing appropriate antibiotics (50 µg/ml carbenicillin, 50 µg/ml kanamycin) with DNA-A and DNA-B components (**Figure 2.3)**,

which was then incubated at 28°C under constant agitation (200 r.p.m) until the OD_{600} = 2.0. The bacterial cultures were then pelleted by centrifugation (12,000xg) and re-suspended in fresh YEP supplemented with Acetosyrigone (200 mM) to a final OD_{600} = 2.0. DNA-A and DNA- B components were combined in equal amounts and 200 µl of the culture was used to inoculate the petioles and stems of each plant using a fine needle (5 mm). Untransformed *A. tumefaciens* Agl1 was cultured in the same manner and was used to agro-infect 3 plants from each line as a negative control. The trial was repeated twice more on the three lines that showed highest SACMV tolerance in the first transgenic trial.

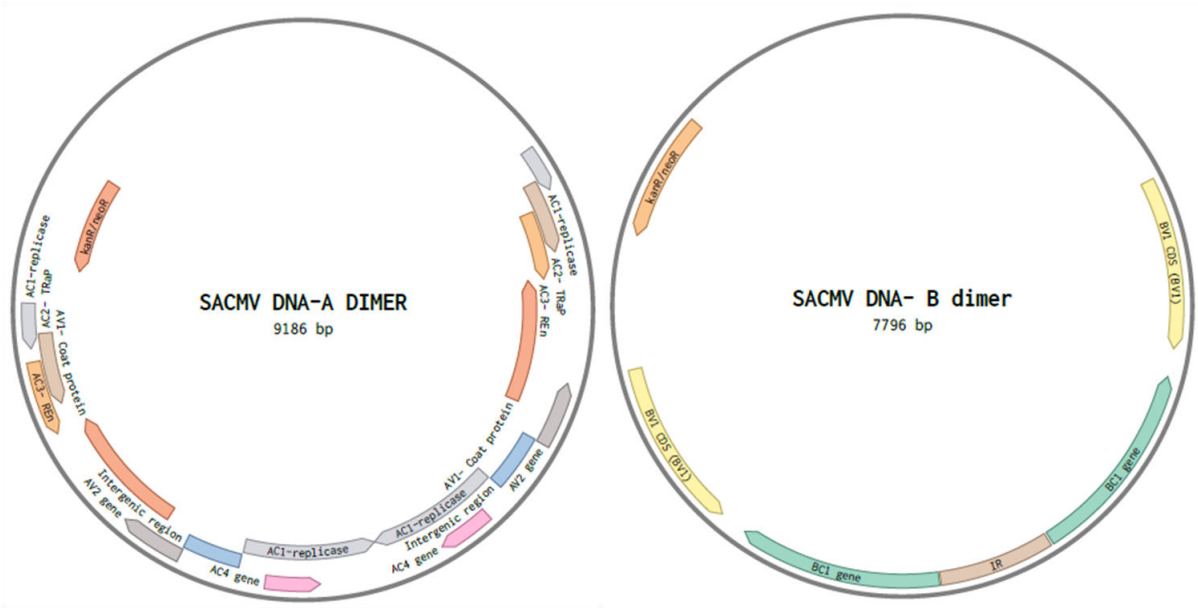


Figure 2.2: Diagram of South African cassava mosaic virus infectious DNA-A (left) and DNA-B (right) dimers inserted into pBIN-10 transformation vector

2.2.4.2 Monitoring of disease progression

After 14 days, six infected plants were selected from each line and were monitored at three time points post infection (14, 32 and 65 DPI). The symptom severity of each plant was evaluated by examining the two upper most leaves, below the apical meristem and scoring the symptom severity based on the following system modified from Legg and Fauquet (Legg and Fauquet, 2004) (0= no symptoms, 1= faint mosaic, 2= mosaic with mild curling, 3= severe mosaic with severe curling) (**Figure 2.4 A**) then an average was calculated to score each plant. The leaves were

harvested and frozen in liquid nitrogen and stored at -70°C for further evaluation. The transgenic and control plants were phenotyped by measuring height at the beginning (0 DPI) and the end of the trial (65 DPI).

2.2.4.3 Quantifying viral load in infected lines

Relative real-time quantitative PCR was used to determine the amounts of viral DNA-A in the collected leaf samples in relation to internal control *Ubiquitin 10* (*UBQ10*) at 32 and 65 DPI. For each trial (3 trials) each transgenic line comprised 6 plants. For each biological replicate (trial) 2 sub-apical leaves from 2 plants were pooled resulting in 3 replicate samples. Total DNA was extracted from 50 mg of each pooled leaf replicate sample [41]. Real-time PCR was conducted according to the Maxima SYBR green (ThermoFisher Scientific) protocol. Two technical replicates per group were performed. Each reaction contained 1x Maxima SYBR green Master Mix, primer SACMV-CCP-F (5'-GCACAAACAAGCGTCGA-3') and SACMV-CCP-R (5'-CTGCCAGTATGCTTAACGTCA-3') (0.5 mM) and 50 ng of pooled DNA. The PCR conditions were 5 min at 95 °C followed by 30 cycles of 95 °C for 15 s, 60 °C for 30 s and 72 °C for 1 min. To normalize the virus level, an internal control of *UBQ10* was quantified using primers UBQ10F (5′ TGCATCTCGTTCTCCGATTG 3′) and UBQ10R (5′ GCGAAGATCAGTCGTTGTTGG 3′).

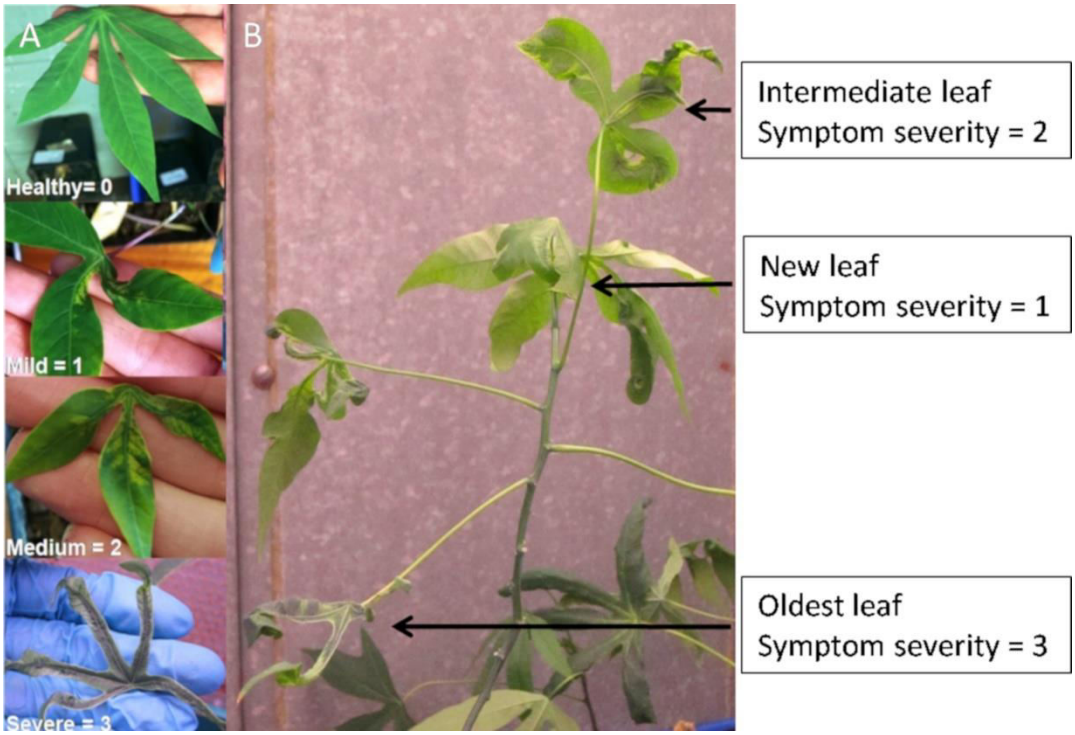


Figure 2.3: Symptom severity (A) and recovery (B) transgenic cassava cv.60444 infected with South African cassava mosaic virus. A) modified symptom severity scale used to determine symptom severity score of plants infected with SACMV. On the scale 0: Healthy (no symptoms), 1: light mosaic or curling, 2: severe mosaic, or curling, 3: severe mosaic and curling with reduction in leaf size. B) recovery phenotype in transgenic cv.60444 line O13-8 in infected with SACMV, where oldest leaves show severe symptoms (symptom severity: 3) while youngest leaves show milder symptoms (symptom severity: 1)

2.2.5 Statistical analysis

The data on symptom severity, plant height and viral load was subjected to a student t-test to determine if there was a statistically significant difference between the infected transgenic lines and wild-type cv.60444 and TME3. Pearson correlation coefficient (r) test was used to establish if there was any correlation between symptom severity and viral load.

2.3 Results

2.3.1 Molecular characterisation of transgenic lines

2.3.1.1 Presence of *Gus, Hyg* and transgene insert

In order to confirm successful stable transformation of cassava cv.60444 with the SACMV AC1/AC4hp-RNAconstruct (pC-AC1/AC4), the transgenic lines were screened for the presence of reporter genes *Gus* and the antibiotic resistance gene *Hyg*, and the forward arm of the AC1/AC4 hairpin construct using PCR. Fourteen lines were confirmed to contain the copies of the *Gus* (**Figure 2.4A**), *Hyg* (**Figure 2.4B**) and pC-AC1/AC4 constructs (**Figure 2.4C**), and were then analysed using Southern Blot hybridisation to determine the number of transgene integration events

2.3.1.2 Southern Blot hybridisation

The total DNA from the transgenic lines was digested with restriction enzymes EcoRI and, HindIII which cuts once within the T-DNA. Southern Blot hybridisation was then carried out using a DNA probe specific to the *Hyg* gene found within the T-border cassette. Purified pCambia 1305.1containing pC-AC1/AC4 was used as a positive control and wild-type untransformed cv.60444 was included as a negative control (**Figure 2.5**). While the majority of the transformed lines were shown to have transgene integration at one site of pC-AC/AC4 (10 lines), some lines (6 lines) had multiple integration sites. Six lines with a single copy of the construct were selected for further testing.

.

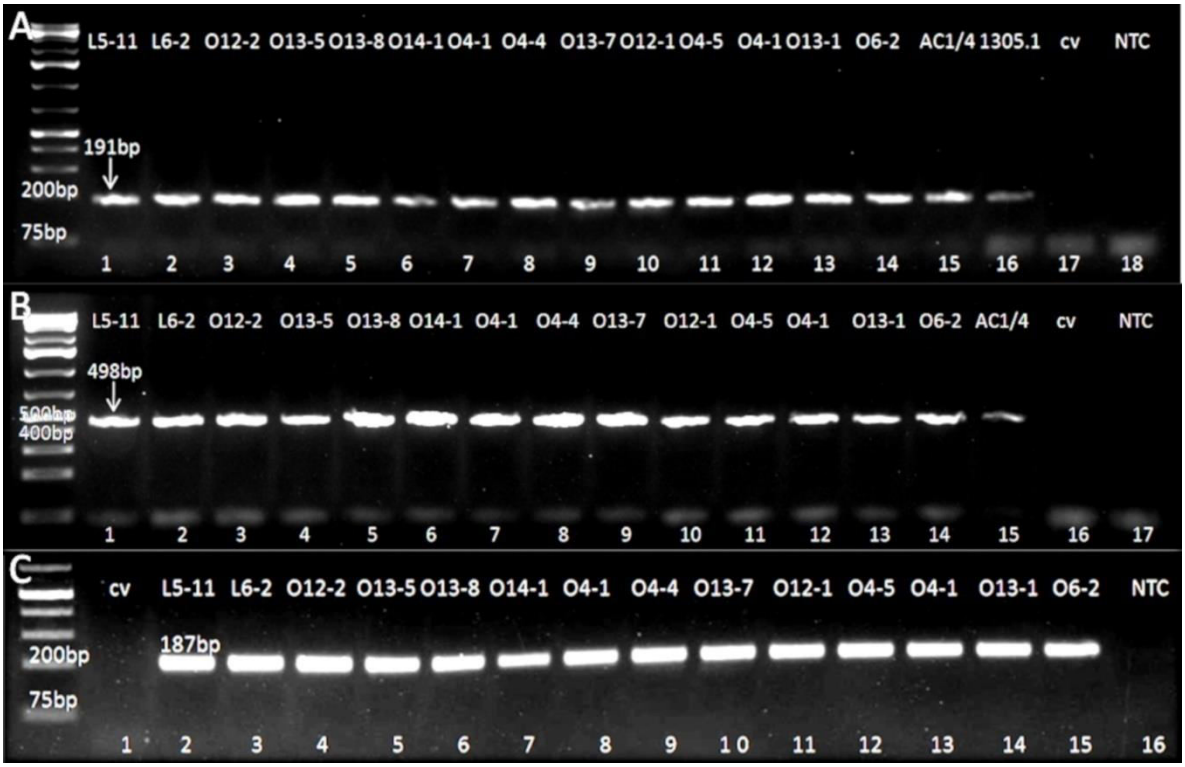


Figure 2.4: PCR detection of *Gus* (A), *Hyg* (B) and AC1/AC4 forward arm(C) in genomic DNA extracted from transgenic cv.60444 transformed with pC-AC1/AC4. The GeneRuler 1kb Plus (Thermo Fisher Scientific) molecular marker was loaded in the first lane of each gel. A: lane 1-14: transgenic cassava lines, lane 15: pCambia-AC1/AC4, lane 16: pCambia 1305.1, lane 17: wild-type untransformed cv.60444 and lane 18: No template control (water). B: lane 1-14: independent transgenic cassava lines, lane 15: pCambia-AC1/AC4, lane 16: wild-type untransformed cv.60444 and lane 17: No template control (water). In C; lane 1: wild-type untransformed cv.60444, lane 2-15: independent transgenic cassava lines, lane 16: No template control (NTC).

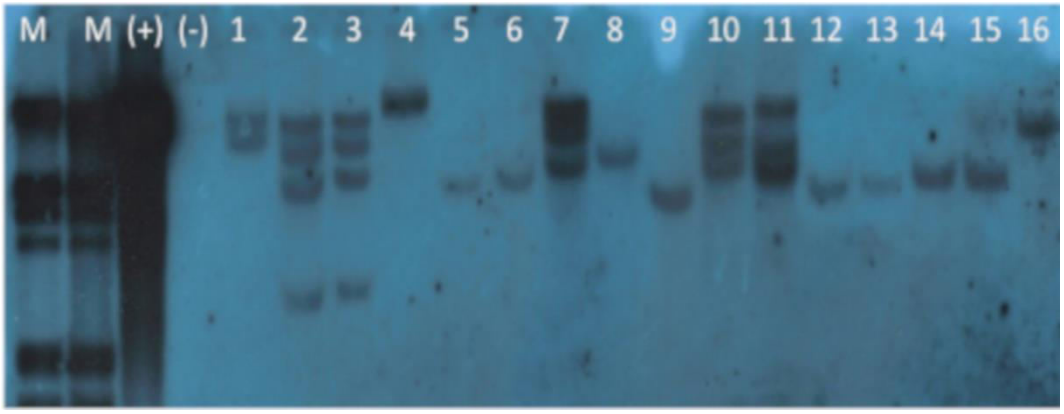


Figure 2.5: Southern hybridisation of total DNA from transgenic cassava cv.60444 transformed with SACMV AC1AC4 construct, digested with restriction enzymes EcoRI

and HindIII, and hybridized with DNA probe specific to *Hyg* gene. Lane M: DIG-labelled molecular weight parker (Thermo Fisher Scientific), lane 1-16: independent transgenic cassava lines. Purified binary vector pC1305.1-AC1/AC4 and wild-type untransformed cv.60444 were included as positive (+) and negative (-) controls respectively.

2.3.1.3 Expression of transgene in transformed lines

Transgene expression was measured in 6 independent transgenic lines (L5-11, O12-2, O4-1, O4-4, O13-5 and O13-8) by RT-PCR. All transgenic lines showed expression of *Gus*, *Hyg* and the pC-AC1/AC4 insert (**Figure 2.3).** In order to infer the relative expression of the transgenes, the relative quantification of pC-AC1/AC4 PCR product in each of the lines compared to the AC1/AC4 control (**Figure 2.3C lane 8**) was calculated. The ratio ranged from 0.37-0.81, with line L5-11 (**Figure 2.3C lane 6**) having the lowest concentration of pC-AC1/AC4 (relative quantity ratio = 0.37) which indicates it has the lowest expression of the transgene. Line O13-8 (**Figure 2.3C lane 5)** had the highest concentration (relative quantity ratio =0.81), which would demonstrate it had the highest expression.

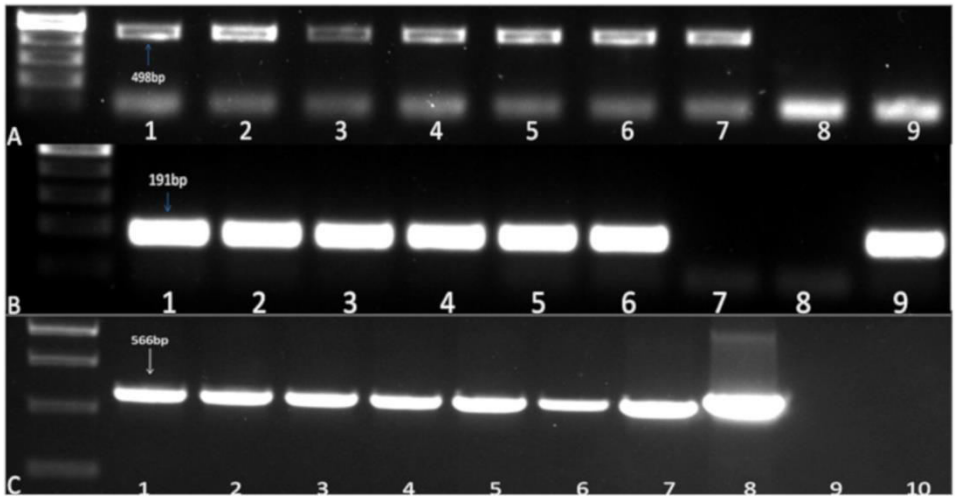


Figure 2.6: Expression of *Hyg* (A), *Gus* (B) and AC1/AC4 transgene (C) in transformed cassava cv.60444 lines (O4-1, O4-4, O12-2, O13-5, O13-8 and L5-11). Molecular weight marker was loaded into the first lane of each gel; 1kb Plus molecular weight marker (Thermoscientific) was used for (A) and (B) and 1kb ladder molecular weight marker (Promega) was used for (C). A: Expression of *Hyg* in transgenic lines. Lane 1: O4-1, lane 2: O4-4, lane 3: O12-2, lane 4: O13-5, lane 5: O13-8, lane 6: L5-11, lane 7: pC-AC1/AC4 (positive control), lane 8: cv.60444, and lane 9: NTC. B: Expression of *Gus* in transgenic lines. Lane 1: O4-1, lane 2: O4-4, lane 3: O12-2, lane 4: O13-5, lane 5: O13-8, lane 6: L5-11, lane 7: cv.60444, lane 8: NTC and lane 9: pC-AC1/AC4. C: Expression of

AC1/AC4 transgene in transgenic lines. Lane 1: O4-1, lane 2: O4-4, lane 3: O12-2, lane 4: O13-5, lane 5: O13-8, lane 6: L5-11, lane 7: pC-AC1/AC4, lane 8: purified AC1/AC4 PCR fragments (1µg), lane 9: cv.60444 and lane 10: NTC.

2.3.2 Evaluation of transgenic AC1/AC4 transgenic lines for resistance to SACMV

2.3.2.1 Symptom severity and plant height

In order to evaluate whether the pC-AC1/AC4 construct confers resistance to the transgenic cassava lines, 6 transgenic lines with a single construct copy, as well as wild-type untransformed cv.60444 and CMD-tolerant TME3, were agro-inoculated with SACMV infectious clones and evaluated for resistance over a period of 65 days for resistance. Three transgenic lines (O12-2, O13-8 and O13-5) which showed decreased viral symptoms and load in the initial screening (Supplementary Data A, B, C) were subjected to further testing. Twelve six week old plants from transgenic lines O12-2, O13-5 and O13-8 were infected with SACMV. After 14 days, leaves were collected from each of the plants, and were screened using coat protein (CP) PCR to establish the success of the infection (Supplementary Data D). Six infected plants were selected for further evaluation and at 32 and 65 days post infection (DPI) the symptom severity and viral load was evaluated. The difference in plant height between 0 DPI and 65 DPI was also measured.

After infection, all plants remained symptomless until 20 DPI and by 32 DPI all infected plants displayed leaf curling and mosaic (**Figure 2.7A**). The average symptom score for wild-type untransformed cv.60444 (1.59 ±0.12) and tolerant TME3 (1.7 ±0.08) was similar at 32 DPI, however at 65 DPI, TME3 displayed recovery (1.36±0.24) which did not occur in wild-type cv.60444 (1.57±0.18) (**Figure 2.7A)**. Lines O12-2, O13-5, O13-8 had lower symptom severity at 32 DPI (1.18±0.40, 1.36 ±0.40 and 1±0.20) when compared to infected wild-type cv.60444. These levels were maintained at 65DPI (1.44±0.28, 1.33±0.23 and 0.85±0.22), with most plants only showing mild curling in the upper leaves. Line O13-8 average symptom score remained low throughout the course of the trials, with very mild curling of the leaves. O13-8 displayed recovery at 65 DPI with new leaves having no visible symptoms (**Figure 8)**. At 65 DPI, line O12-2 and O13-5, had a similar symptom severity score as tolerant TME3, but did not display any signs of recovery. The difference in

symptom severity between the transgenic lines and wild-type cv.60444 was shown to be significant at both 32 and 67 DPI ($p<0.05$).

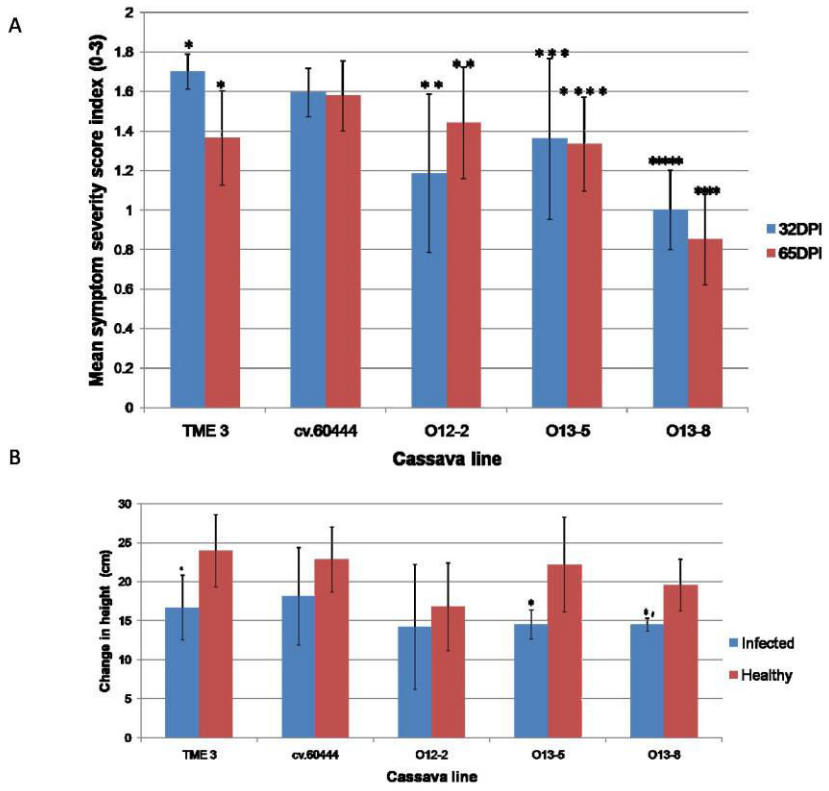


Figure 2.7: Average symptom severity scores (A) and change in plant height (B) in SACMV AC1/AC4 transgenic (O12-2, O13-5, and O13-8), wild-type untransformed cv.60444 and tolerant TME3 cassava plants *Agro*-inoculated with South African cassava mosaic virus. Values represent the mean of three independent biological replicates each with 6 plants per treatment and bars indicate standard error (SE). Asterisks indicate statistical significant results from student t-test.

There was no significant difference between the plant height of the cv.60444 infected and transgenic lines (**Figure 2.7B**).Transgenic lines O13-5 and O13-8 had a reduction in height between infected (14.48±1.87 and 14.45±0.81) and mock inoculated (22.19±6.09 and 19.56±3.29) of 34.7% and 26% respectively. The average height difference between transgenic line O12-2 infected (14.19±7.99) and mock-inoculated (16.78±5.60) lines varied by 15.4%, however this difference was not statistically significant and height does not appear to be affected by the presence of the virus in this transgenic line.

The symptoms displayed by wild-type cv.60444 and O13-5 at 65 DPI remained similar to those displayed at full systemic infection, with O13-5 symptoms being less severe than that in wild-type cv.60444.

Figure 2.8: Symptoms in apical leaf of transgenic lines O12-2, O13-5 and O13-8 as well as TME3 and untransformed cv.60444 infected with SACMV at 68 DPI. A) TME3, B) cv.60444, C) O13-8, D) O13-5 and O12-2. Line O13-8 and TME3 had low symptom severity and showed recovery at 65 DPI. Lines O12-2 and O13-5 had similar symptom severity to tolerant TME3 but did not show any recovery during the course of the trial.

2.3.2.2 Viral load

The relative viral load of each sample was determined by real time qPCR on DNA samples extracted from the plants at 32 and 67 DPI. The data was normalised using internal control gene *UBQ10* (Allie *et al*., 2014) (**Figure 2.9**). The viral loads were analysed for statistically significant differences between transgenic lines O12-2, O 13-8 and O13-5 as well as TME 3 and wild-type infected cv.60444 and which showed a statistical significant difference between viral loads in leaf samples at 32 DPI and 67 DPI ($P<0.05$).

In comparison to non-transgenic cv.60444, the viral loads of transgenic lines O13-5 and O13-8 were significantly lower at 67 DPI ($p<0.05$). Overall, O13-5 had the lowest viral accumulation (average of 47-fold ± 32) at 32 DPI and the relative viral load remained low at 65DPI. Both O12-2 (180-fold ±90) and O13-8 (219±109) had lower relative viral loads than cv.60444 (480-fold±48) at 32DPI and the viral load in both decreased at 67 DPI (which correlates to the recovery observed in line O13-8). The viral loads in TME3 (264-fold± 200), O12-2 and O13-8 were similar at 32 DPI however this was shown not to be significant. The viral loads were significantly lower in TME3 compared to cv.60444 in the first two trials but in the later third trial the viral loads were unexpectedly higher in TME3. This led to a large standard deviation

between the technical replicates and affected the statistical results. Variation and an increase in disease severity in TME3 has recently been observed in our and other labs, where while TME3 disease severity is generally still lower than susceptible cv.60444 and T200, symptoms and viral loads have increased. Although there is no current explanation for this, one hypothesis is the TME3 plants may have undergone genome wide methylation, which has been shown to occur in plants maintained in tissue culture (Kitima et al. 2015) .Methylation has been linked to the loss of the CMD2 resistance in some TME landraces during somatic embryogenesis (Beyene, 2016).

Pearson's correlation test was used to test the correlation between viral load and symptom severity and showed a low positive correlation between the symptom score and the viral load for all samples ($R<0.3$ in all sample comparisons, $p<0.05$) except TME3 which showed a negative correlation ($R= -0.0135$). The relative viral load of O13-5 was also lower than infected wild-type cv.60444, correlating with less severe symptoms. Transgenic line O12-2 symptoms increased over the course of the study (similar to wild-type cv.60444), however its height was not affected by presence of SACMV and its relative viral load decreased at 67 DPI which indicates viral attenuation.

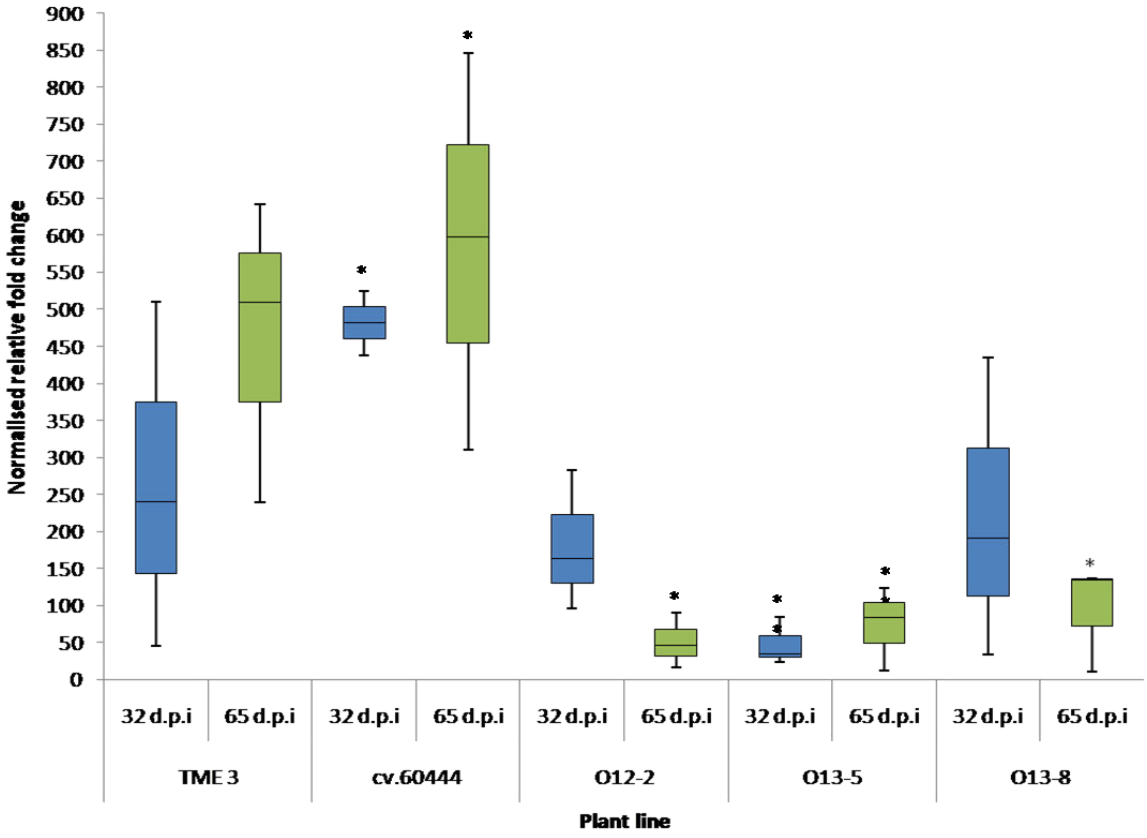


Figure 2.9: The viral load of SACMV, relative to reference gene UBQ10 in transgenic (O12-2, O13-5 and O13-8), wild-type cv.60444 and TME3 infected with SACMV, at 32 and 67 DPI. Values represent the mean of three independent replicates each with 3 plants per treatment and bars indicate SEM. Asterisks (*) indicate statistically significant results ($p<0.05$).

2.3.3 Northern Blot for expression of hairpin-derived siRNA

In order to establish whether the tolerance observed in transgenic lines was linked to the expression of siRNAs, Northern blot hybridisations were performed on the three tolerant lines (O12-2, O13-5 and O13-8) as well as the susceptible line L5-11. Cultivar 60444 and TME 3 were also included as susceptible and tolerant controls, respectively (**Figure 2.10)**. The uninfected wild-type cv.60444 controls did not produce SACMV-targeted siRNAs, while SACMV infected cv.60444produced low levels of siRNA (relative to 200ng of the 21nt marker) targeting SACMV DNA-A AC1/AC4 region (relative quantity ratio: 2.0). Small interfering RNAs specific to AC1/AC4 were detected in mock-inoculated (relative quantity ratio: 1.68) TME3, and

SACMV-infected tolerant TME3 landrace (relative quantity ratio: 2.67), but significantly higher relative levels of siRNAs were detected in infected TME3 plants.

Small interfering RNAs targeting the AC1/AC4 region of SACMV were detected in the uninfected transgenic lines O13-5, O13-8 and O12-2 (Figure 9), indicating their origin from the transgene-derived hp-RNA. Notably, there was an observed increase in the amount of siRNA produced by the infected tolerant transgenic lines. Uninfected line O13-8 had the highest siRNA level (relative quantity ratio: 2.0), which increased (2.77) in the presence of SACMV. A lower quantity of siRNA was detected in uninfected O13-5 and O12-2 (1.17 and 1.14 respectively) which increased after SACMV challenge, although O12-2 (2.04) siRNA levels increased comparatively less than O13-5 (2.62). In contrast to the tolerant lines, in susceptible line L5-11, although a relatively high level of siRNA was present in uninfected plants (1.81), the siRNA levels decreased (1.57) following infection with SACMV.

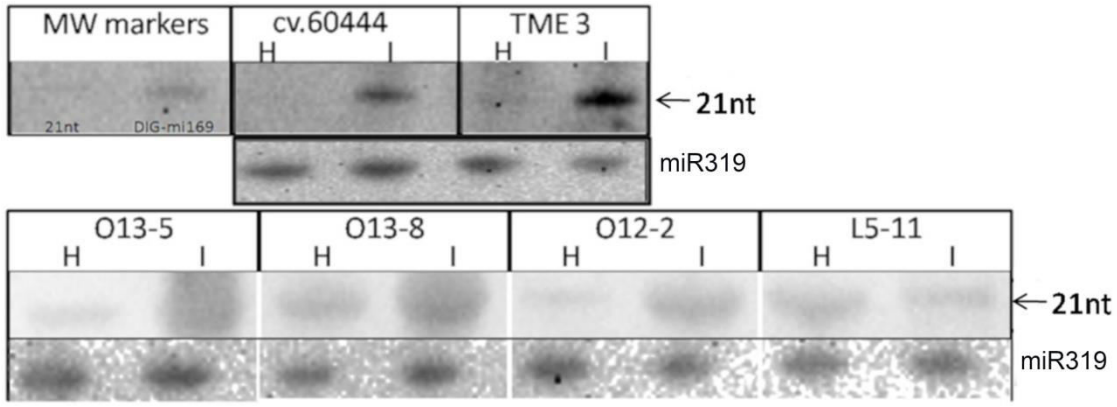


Figure 2.10: Northern blot hybridisation of RNA extracted from infected (I) and healthy (H) (non-infected) transgenic lines (L5-11, O12-2, O13-5 and O13-8), wild-type cv.60444 and TME3. A DIG-labelled miRNA marker (MW) (Thermo Fisher Scientific) and DIG-labelled miR169 were included as size controls.

2.4 Discussion

Globally, the threat presented by CMD is one of the greatest hurdles to cassava production. In southern Africa, SACMV is widely spread in South Africa, Zimbabwe, Mozambique and Madagascar [68], and to date CMD resistance has primarily been engineered against ACMV using antisense (Chellappan et al., 2004a; Zhang et al.,

2005) or hp-RNA PTGS (Ntui *et al*., 2014; Vanderschuren *et al*., 2007a; Zainuddin *et al*., 2012). This study reports, for the first time, SACMV tolerant transgenic cassava. Significantly, enhanced PTGS in SACMV-transgenic cv.60444 resulted in a tolerant host response similar to the wild-type West African landrace TME3 which has the monogenic *CMD2* locus associated with CMD resistance (Wolfe *et al*., 2016). Tolerance was displayed by three transgenic cv.60444 lines (O13-5, O13-8 and O12-2) transformed with the hp-RNA construct targeting the AC1/AC4 region on DNA-A of SACMV. When challenged with SACMV, these three lines displayed less severe symptoms compared with wild-type cv.60444 and tolerant landrace TME3, and had reduced viral loads. Further, one of the lines (O13-8) showed consistent lower viral loads and recovery at 65 DPI, demonstrating that the AC1/AC4 transgene was effective at reducing virus infection.AC1 and AC4 were chosen as effective targets for RNAi-induced PTGS in this study as AC1 is a critically important protein for viral replication, and interactions with host proteins, and AC4, a viral suppressor, has been shown to inhibit both TGS and PTGS in host plants (Ramesh *et al*., 2007; Vinutha *et al*., 2018). Hairpin-RNA constructs targeting the AC1/AC4 overlap has been shown to confer resistance to other geminiviruses, including *Tomato leaf curl virus* in transgenic tomatoes (Ramesh *et al*., 2007). Cassava lines transformed with antisense and hp-RNA constructs targeting the AC1 of ACMV-NOg and ACMV-KE have exhibited increased resistance to the virus (Vanderschuren *et al*., 2009, 2007a; Zhang *et al*., 2005). Transgenic resistance and tolerance achieved in cassava to ACMV and SACMV, respectively, may reflect different host-virus co-adaptability since SACMV and ACMV evolved in different geographical regions, namely East and West Africa, respectively (Berrie *et al*., 2001; Ndunguru *et al*., 2005; Patil and Fauquet, 2009). Adaptability or fine tuning of PTGS efficacy between cassava host and geminivirus may be evolutionary specific.

Tolerance and recovery have been shown to be linked to the production of siRNAs which target the virus genome (Kuria *et al*., 2017). A RNAseq study (Rogans *et al*., 2016) demonstrated that while siRNA targeting SACMV was detected in both infected T200 and TME3, siRNA populations were lower in infected TME3 during recovery, compared to susceptible T200. In particular SACMV-infected TME3 produced high levels of siRNAs targeting the AC4 region of SACMV, (Rogans *et al*., 2016), however these results were not confirmed experimentally. AC1 and AC4

genes being have also been reported as siRNA "hot spots" (Patil *et al*., 2016; Rogans and Rey, 2016; Sharma *et al*., 2014) in other geminivirus studies. In this study the SACMV AC1/AC4 transgenic lines O13-5, O13-8 and O12-2 only displayed tolerance in response to infection with SACMV. Hairpin-derived siRNAs corresponding to the AC1/AC4 region of SACMV were detected in the uninfected three tolerant transgenic lines, as well as in the uninfected susceptible L5-11 transgenic line (**Figure 2.10**) demonstrating that siRNAs are being produced from the transgene in the absence of the virus. Infected susceptible non-transgenic cassava T200 landrace has also been shown to generate virus-derived siRNAs that target SACMV, but these do not confer resistance (Rogans *et al*., 2016). Similarly, the highly susceptible transgenic L5-11 line did generate siRNAs, but levels of siRNAs declined when L5-11plants were inoculated with SACMV. This suggests some interference or suppression of PTGS by SACMV, leading to lower siRNA levels and increase in symptoms and virus load. While induction of transgene-derived siRNAs has been linked to decreased viral loads (Guo *et al*., 2015), in this transgenic line the presence of AC1/AC4-derived siRNAs did not lead to lower virus loads or resistance. It is suggested that SACMV may be partially suppressing both natural (Shamandi *et al*., 2015) and transgene-mediated PTGS in transgenic lines, through viral suppressors such as AC4 and AC2 (Vanitharani *et al*., 2004). Tolerance and symptom recovery in virus challenged O13-5, O13-8 and O12-2 transgenic lines may also represent a battle between virus interference and antiviral silencing, but in this case transgene-mediated PTGS appears to be more effective against SACMV than in L5-11. Viral suppressor proteins inhibit RNA-mediated defence using several mechanisms including inhibition of siRNA generation, inhibition of siRNA incorporation into RISC complex and direct interference with the RISC complex (Rajeswaran and Pooggin, 2012; Rogans *et al*., 2016). Rogans *et al* (2016) (Rogans *et al*., 2016) hypothesised that SACMV avoided RNA silencing in the highly CMD-susceptible T200 landrace through inhibition of incorporation of siRNA into RISC complex. It is possible that this inhibition may also occur in the tolerant transgenic lines in this study..Cultivar cv.60444 is also highly susceptible to cassava geminiviruses, but the plants from the three transgenic cv.60444 lines (O13-5, O13-8 and O12-2) displayed lower symptom severity scores than wild-type cv.60444 and some plants showed recovery, a typical response found in SACMV infected TME3 (Allie *et al*., 2014).

Tolerance in the three transgenic lines correlated with lower viral loads compared to both cv.60444 and TME3, and northern blots of siRNAs confirmed that this was associated with siRNA-mediated PTGS silencing of the virus. Interestingly, at 65 DPI while both wild type TME3 plants and O13-8 displayed signs of recovery, the viral load in line O13-8 was lower than CMD-tolerant TME3. Although decreased symptom severity correlated with a decrease in viral load in transgenic lines O12-2, O13-5, and O13-8, as shown in other studies (Ammara *et al*., 2017), there are several reports where no correlation between viral load and symptom severity was found (Kaweesi *et al*., 2014). The most promising SACMV-transgenic cv.60444 line O13-8 warrants further investigation to unravel contributing factors to its tolerance to SACMV. Some of these factors may include the influence of different host genetic backgrounds (Kuria *et al*., 2017). Interestingly, a study in cowpeas (*Vigna unguiculata L. Walp*) transformed with a hp-RNA construct targeting the AC1/AC4 of Mungbean yellow mosaic virus, also showed milder symptoms rather than systemic resistance (Kumar *et al*., 2017). Different geminiviral species may also respond differently in similar transgenic plant systems. Hairpin or antisense-AC1 RNA which targeted ACMV AC1 (Vanderschuren *et al*., 2007a; Zhang *et al*., 2005) conferred resistance in cv.60444 through PTGS, whereas in this study, only tolerance was achieved with the SACMV-AC1/AC4 hairpin construct. These above-mentioned studies, amongst others, demonstrate that several host and virus factors can determine the outcome in PTGS-transgenic crops (Kuria *et al*., 2017).

Transgenic-induced RNA silencing enhances the natural plant host antiviral defences, but highly pathogenic viruses are able to suppress or evade the production or action of siRNAs (Pooggin, 2017). Virus-induced PTGS is highly dependent on plant-virus interactions and well as the intrinsic features of the virus (Chellappan et al., 2004b). SACMV is a recombinant virus, showing high virulence in cassava, with its AC2 from an unknown virus source, and its IR is closely related to TYLCV-Israel (Berrie *et al*., 2001). The predicted origin of SACMV is the East Africa/south-west Indian Ocean islands region (Lefeuvre *et al*., 2007), and SACMV is able to induce severe symptoms and as well as high numbers of virus derived siRNAs targeting DNA-A and B components in non-transgenic T200 landrace (Rogans and Rey, 2016). Similarly, cabbage leaf curl virus (CaLCuV) also induces high numbers of 21 nt virus-derived siRNAs in *Arabidopsis* (Aregger *et al*., 2012). A

number of studies have shown that in order for a host to display either resistance or recovery, it must produce a high level of siRNAs against the virus [50,56]. Chellappan *et al.*[53] showed that recovery in cassava infected with either ACMV-CM or Sri-Lankan cassava mosaic virus was linked to an increase in siRNA targeting the virus. This has been shown in PTGS-induced resistance studies, where resistance was induced by the constitutive expression of the transgene-derived siRNA. Fuentes *et al.* (Fuentes *et al*., 2016) and Vanderschuren *et al.* (Vanderschuren *et al*., 2009) also demonstrated that the constitutive expression of transgene derived small RNAs, prior to ACMV challenge, was important for engineered virus resistance. Three tolerant transgenic lines in this study were shown to produce transgene-derived siRNAs, however they produced varying levels of tolerance to SACMV, and there did not appear to be a correlation between pre-inoculation siRNAs levels and tolerance. Transgene expression can be affected by several factors including the position of insertion in the genome, which can affect the levels of siRNA produced (Pooggin, 2017; Vanderschuren *et al*., 2009). This may be a factor in this study, but requires further investigation.

A number of other factors can also play a role in the efficiency of siRNA-mediated PTGS silencing in plants. The genetic background of plant cultivar and type of geminivirus can affect effectivity of siRNA-mediated PTGS. For example, CMD resistance was shown to be linked to the ability of different cassava genotypes to induce RNA silencing, evidenced by varying levels of virus-derived siRNA (vsiRNA) (Kuria *et al*., 2017). Additionally, the same plant genotypes can behave differently with various viruses. For example, Kuria *et al.* [65] reported that symptomatic ACMV and EACMV infected cassava cultivars TMS30001 and TMS30572 both showed recovery in the upper leaves after 65 DPI. After cut back, resistance (no symptoms or virus replication)was observed in new shoot growth in the ACMV-infected plants, however, in contrast, in EACMV-infected plants, new shoot tissue was symptomatic and virus was detected. Resistance also depends on the ability of plants to overcome the viral suppressors (Simón-Mateo and García, 2011). The response of *Ty* resistant tomato to beet curly top Iran virus (BCTIV) is variable and depends on the type of begomovirus (Ghanbari *et al*., 2016). The number of integration events or integration site of the hp-RNA may also affect the efficacy of siRNA-mediated PTGS (Butaye *et al*., 2005). In this study the highest relative level of siRNA was detected in

transgenic line O13-8, which showed recovery. The variation in tolerance shown by the three lines may also be due to location and expression levels of the transgene within the three lines (Vanderschuren *et al*., 2009), where hairpin transcription may not necessarily lead to efficient siRNA production and the RNA silencing pathway. A study in Cucumber mosaic virus-infected tobacco showed that the location of transgene insertion played a role in whether 21nt RNAs were processed (Dalakouras *et al*., 2011). Northern blots using an AC1/AC4 specific probe showed that small interfering RNA targeting the AC1/AC4 region of SACMV was also expressed uninfected non-transgenic tolerant TME3 (**Figure 2.10)**. This finding was also observed in TME3 in a RNAseq study (Rogans and Rey, 2016) and in TME3 that showed recovery from infection with ACMV-NOg that had high levels of siRNA targeting the virus (Kuria *et al*., 2017). We conclude that the AC1/AC4 transgene may enhance tolerance to SACMV in transgenic cv.60444 lines, similar to natural CMD tolerance in TME3 which showed similar siRNA levels.

The role of diverse RNA degradation pathways and DNA methylation (Andika *et al*., 2007) in variable transgene expression and hpRNA-derived siRNA production could also play a role in efficiency of RNA silencing. The lines may additionally require RNA-dependent RNA polymerase 6 (RDR6) mediated secondary siRNA production to accumulate enough siRNA to mount an effective resistance response, and RDR2-mediated DNA methylation to silence geminiviruses (Aregger *et al*., 2012). RDR6 contributes to anti-viral immunity by converting viral transcripts to dsRNA precursors of secondary vsRNAs (Zvereva and Pooggin, 2012). Efficient resistance to geminiviruses is proposed to rely on both primary and secondary siRNA amplification (Aregger *et al*., 2012), however cassava RDR2 (cassava v4.1_013755m and 021122m) and RDR6 (Manes 16G121400 v6.1) homologs were not found to be up-regulated in TME3 (Allie *et al*., 2014) and bisulphite sequencing did not reveal a role for DNA (cytosine) methylation of SACMV at recovery in TME3 (Rogans and Rey, 2016).Tomato yellow leaf curl virus was also shown to evade host RNAi defence through a population of *de novo* synthesised unmethylated viral DNA (Butterbach *et al*., 2014). This does not rule out the possibility of histone methylation which needs to be investigated in future. Our results support the suggestion that transgene-induced PTGS was not sufficient to suppress SACMV replication completely in cv.60444.

Additionally, with regards to the most efficient strategy of inducing antiviral RNA silencing in cassava, are considerations with respect to the environment in which RNA silencing operates. Temperature has been shown to influence geminivirus-induced RNA silencing in plants (Chellappan *et al*., 2005; Mohamad Hamed Ghodoum *et al*., 2018). For example, African cassava mosaic virus-induced silencing increased by elevating the temperature from 25^{o}C to 30^{o}C (Chellappan *et al*., 2005). While SACMV displays non-recovery in wild type cv.60444, and transgenic cv.60444 plants were grown at 28^{o}C and it is possible that this may have influenced resistance efficiency. It is also possible the lower resistance levels displayed by the transgenic lines may be due to high agro-infectious SACMV inoculum pressure. In an ACMV bombardment assay, resistance was shown to be broken when the virus load was increased (Zhang *et al*., 2005). Some dsAC1 ACMV-transgenic cassava lines showed an increase in infection rate when the virus pressure was increased from 350 to 700ng (Vanderschuren *et al*., 2009). The needle-mediated agro-inoculation method performed in this study, used because cassava is recalcitrant to mechanical inoculation and leaf-infiltration, delivers high virus titres directly into the plant vascular system. It is that infecting the plantlets with high viral load in this direct and rapid method could over-come the RNAi engineered resistance. Additionally, the effectiveness of RNAi-mediated response is reliant on a there being a high level of sequence homology between the transgene and the viral target (>90% homology). Fuentes *et al*. (Fuentes *et al*., 2016) showed that RNAi-mediated resistance in crops can be stably maintained, when the viral population remains stable. However, recently Mehta *et al.* (Mehta *et al*., 2018a) showed that field cassava geminivirus populations changed in response to RNAi-mediated resistance pressure which could have massive consequences for the development for resistance plants in the field.

In conclusion, this study has shown that a hp-RNA construct targeting the AC1/AC4 region of SACMV DNA-A can confer tolerance to SACMV. This tolerance is ascribed to the induction of enhanced PTGS and primary hpRNA-derived siRNAs in lines expressing AC1/AC4 specific siRNA, leading to reduced symptoms and viral load, and recovery in the latter stages of infection. It is suggested that total resistance is not achieved as SACMV is able to maintain some level of counter defence via PTGS suppression. This finding, and other studies, show variability in PTGS efficacy related to many factors, which will prove challenging for implementing this control

strategy under variable environmental and field conditions. While direct comparisons cannot be made between SACMV and existing resistance studies on two other cassava viruses, namely ACMV and SLCMV, as the exact virus genome target sequences (constructs) and host genome integrations in cv.60444 are not identical, it is notable that distinct cassava viruses behave differently in the same cassava cultivar (cv.60444). Other complementary natural resistance mechanisms need to be explored and manipulated through techniques such as gene editing if robust and enduring resistance to cassava mosaic disease is to be successful. Tolerance may offer some advantages over total resistance, as it is more durable and less likely to break down under high virus pressure in field conditions. It has been suggested that symptom recovery can be regarded as an inducible form of tolerance (Paudel and Sanfaçon, 2018), and is associated with a diversity of mechanisms which could be exploited in developing approaches to control plant virus disease. A balance between virus and host defence mechanisms which reduces the fitness cost of the plants (Bengyella and Rey, 2015; Paudel and Sanfaçon, 2018) more closely resembles durable resistance in wild undomesticated crops. However one disadvantage is that if there are still virions present in the leaves, they could be picked up by the whitefly vector. It has been proposed that developing dual resistance to whitefly and geminiviruses could provide a more stable long term solution to reducing the impact of CMD (Zaidi *et al*., 2017) or combining natural and genetically engineered resistance (Vanderschuren *et al*., 2012).

Chapter 3: Genetic engineering of cassava with a triple construct targeting ACMV, SACMV and EACMV is toxic to friable embryogenic callus and causes developmental abnormalities in regenerated plants

3.1 Introduction

RNA silencing refers to a group of molecular mechanisms which are involved in the regulation of gene expression, chromatin state and defence against invading nucleic acid such as viral nucleic acids (Vaucheret *et al*, 2001). This includes RNA-mediated post-transcriptional gene silencing (PTGS), transcriptional gene silencing (TGS) and RNA interference (RNAi) (Baulcombe, 2004; Pooggin, 2017). RNA silencing is mediated by small RNA (sRNA) (19-24nts) which can be categorised based on difference in their biogenesis and function (Axtell *et al*, 2011). Broadly, sRNA are classified into microRNAs (miRNAs) produced from hairpin-like *MIR* genes and small interfering RNAs (siRNAs) which are produced from double stranded RNA (dsRNA) precursors. Both siRNA and miRNA are processed by Dicer-like proteins (DCL) after which they are sorted by ARGONAUT (AGO) proteins based on size and 5' nt sequence, to form RNA-induced silencing complexes (RISC), which mediate PTGS via translational repression or sRNA-directed mRNA degradation or TGS by sRNA-directed DNA methylation (Aregger *et al*, 2012; Borges and Martienssen, 2016; Pooggin, 2017).

Antiviral RNA silencing occurs naturally in plants, where dsRNA, from both RNA and DNA viruses can activate RNA silencing which then produces virus-derived interfering RNA (vsiRNA) (Akbergenov *et al*, 2006). This signal can then be amplified through secondary vsiRNA produced from long perfect dsRNAs that are synthesised by RNA Dependent RNA polymerase 6 (RDR6) (Zhang *et al*, 2013). Virus-derived siRNAs are associated with AGO proteins in a similar manner to endogenous siRNA

and target viral genomes resulting in both PTGS and TGS of viruses (Baumberger and Baulcombe, 2005; Llave, 2010; Parent *et al*, 2015). Virus-derived siRNA are assumed to be highly specific to their homologous miRNA or protein targets, however when there is a high level of homology between the vsiRNA and the host, down-regulation of host genes can occur (Cao *et al*, 2014).

RNA silencing has been used to enhance virus resistance in a number of crops (Aregger *et al*, 2012) by introducing virus-derived sequences (under the control of a promoter) into the plant genome which then triggers the production of transgene-derived siRNA. The hp-RNA targets either a partial or full length sequence of a viral gene and produces a population of siRNAs of between 21-24nt which have homology to the viral genome. One of the commonly cited advantages of PTGS-induced virus resistance is that it is highly specific, with hp-RNA-derived siRNA only targeting the viral genome, however several studies have noted off-target effects characterised by phenotypic variation in transgenic organisms (Brunetti *et al*, 1997; Fuentes *et al*, 2016; Xu *et al*, 2006). Transcriptomic data from transgenic plants transformed with hp-RNA constructs targeting viral proteins have been shown to have altered transcriptomic profiles compared to the wild type untransformed host (Fuentes *et al*, 2016). The off-target effects are characterised by changes (both up and down regulation) in expression levels of mRNAs as a result of homology between transgene-derived siRNA produced from hp-RNA and host genes (Praveen *et al*, 2010). siRNA-mediated silencing of unintended mRNAs can be triggered by as little as 8nt complementarity between the mRNA 3'-UTR region and the 5' of siRNA (Jackson *et al*, 2006, 2003; Lin *et al*, 2005). This effect has been noted in several studies where plants transformed with the hp-RNA constructs targeting viral proteins have displayed low transformation efficiencies, poor growth and development abnormalities (Fuentes *et al*, 2016, 2006; Praveen *et al*, 2010; Vanderschuren *et al*, 2007a). While no study has shown a single target gene target responsible for these changes, Fuentes *et al*(2016) showed that tomato lines transformed with the C1 protein of *Tomato yellow leaf curl virus* displayed phenotypic and developmental abnormalities (Fuentes *et al*, 2006) which were shown to be associated with a common set of differentially expressed genes in the transgenic lines.

Cassava mosaic disease (CMD) is one of the most detrimental diseases of cassava (*Manihot esculenta*) (Legg *et al*, 2015). In sub-Saharan Africa it is caused by a

number of geminiviruses known collectively as Cassava Mosaic Geminiviruses including *African cassava mosaic virus*- Nigeria, *East African cassava mosaic virus* and *South African cassava mosaic virus*. Cassava mosaic geminiviruses (CMGs) are bipartite, single stranded DNA viruses (genus: *Begomovirus*), with two DNA strands (DNA-A and DNA-B) which encode for 8 open reading frames (ORFs). DNA-A contains 6 ORFs which code for the following proteins AC1 (Replication associated protein), AC2 (Transcription activation protein), AC3 (Replication Enhancer protein) and AC4 in the complementary sense and AV1 (Coat protein) and AV2 in the virion sense (Fondong, 2013), and a putative 7th ORF coding for a protein AC5 (Rey and Vanderschuren, 2017). DNA-B encodes for two ORFs which code for BC1 (movement protein) and BV1 (Nuclear shuttle protein). Both DNA-A and B contain an Intergenic region (IR), which contains the nanonucleotide region required for the binding of Rep (Yadav and Chattopadhyay, 2011).

It has been widely shown that several viral proteins including AC1, AC2 and AV2 act as pathogenicity determinants in hosts and their roles in infection and host interactions have been widely studied (Castillo *et al*, 2003; Hanley-Bowdoin *et al*, 2013; Yang *et al*, 2019). It has been well documented that viral proteins, specifically virus suppressor proteins (VSRs) interfere with RNAi pathways in the plant, preventing PTGS and TGS of the virus (Buchmann *et al*, 2009). VSRs have also been shown to disrupt normal cellular activity (Chapman *et al*, 2004; Trinks *et al*, 2005).

Virus-derived siRNA can play a direct role in antiviral immunity in plants where vsiRNA are incorporated into RISC to down-regulate the expression and translation of the virus and the production vsiRNA are associated with resistance and recovery, (Kuria *et al*, 2017).Virus-derived siRNA have been studied in a number of host-virus models which have shown that vsiRNAs are not produced against the entire length of the virus, but rather are produced in "hotspots" which vary depending on the host and the virus (Akbergenov *et al*, 2006; Kutnjak *et al*, 2015; Yang *et al*, 2018).

In the field, these viruses can often infect the same plant simultaneously (Fondong *et al*, 2000; Pita *et al*, 2001) and an effective anti-viral strategy should take this into account. Vanderschuren *et al*, (2007) showed that a dsRNA construct targeting the Intergenic region (IR) including the bidirectional promoter of *African cassava mosaic*

virus improved recovery of transgenic cassava lines infected with ACMV. Further studies conducted in our lab have shown that cassava transgenic lines, transformed with hp-RNA construct targeting the AC1/AC4 overlap of DNA-A of ACMV or SACMV had improved resistance to ACMV (Moralo, 2015) and SACMV (unpublished data). In order to improve the resistance of cassava to EACMV, ACMV and SACMV, a stacked construct targeting the IR/AC1 region of SACMV and EACMV and AC1/AC4 of ACMV (pC-AES) was used to transform cassava cv.60444. The transformation efficiency of friable embryogenic callus (FEC) in these experiments was low, and transformants displayed developmental abnormalities including stunted growth, leaf deformation and yellowing, leading to a wider study of possible off-target effects associated with these viral targets. Eleven genes were identified as possible off-targets by pC-AES derived siRNAs. Further studies, which included two constructs, which targeted either the EACMV or EACMV-ACMV regions of AES, confirmed that six genes, which had over-lapping functions in development and disease resistance, were differentially regulated in the presence of the hp-RNA constructs. The identification of these siRNA host targets elucidates the importance of understanding of specific virus-host interactions, and provides evidence that fastidious selection of viral sequence targets is critical for future virus resistance engineering. This study also highlights the possible long- term limitations of PTGS for virus control in the field where mixed infections often occur.

3.2 Materials and Methods:

3.2.1 Construction of hp-RNA vectors pC-AES, pC-E-ACMV and pC-EACMV

3.2.1.1 Target selection and primer design

Viral targets which would potentially confer resistance to ACMV, EACMV and SACMV were selected based on previous studies (Moralo, 2015; Vanderschuren *et al*, 2007a) and siRNA hotspot deep RNA sequencing data (Rogans, 2016). Areas targeting the AC1/AC4 overlap regions on DNA-A of ACMV, and IR/ AC1 region of EACMV and SACMV were selected, and primers complementary to EACMV, ACMV and SACMV were designed (**Table 3.2**). The primers included restriction enzyme cut sites to enable directional cloning.

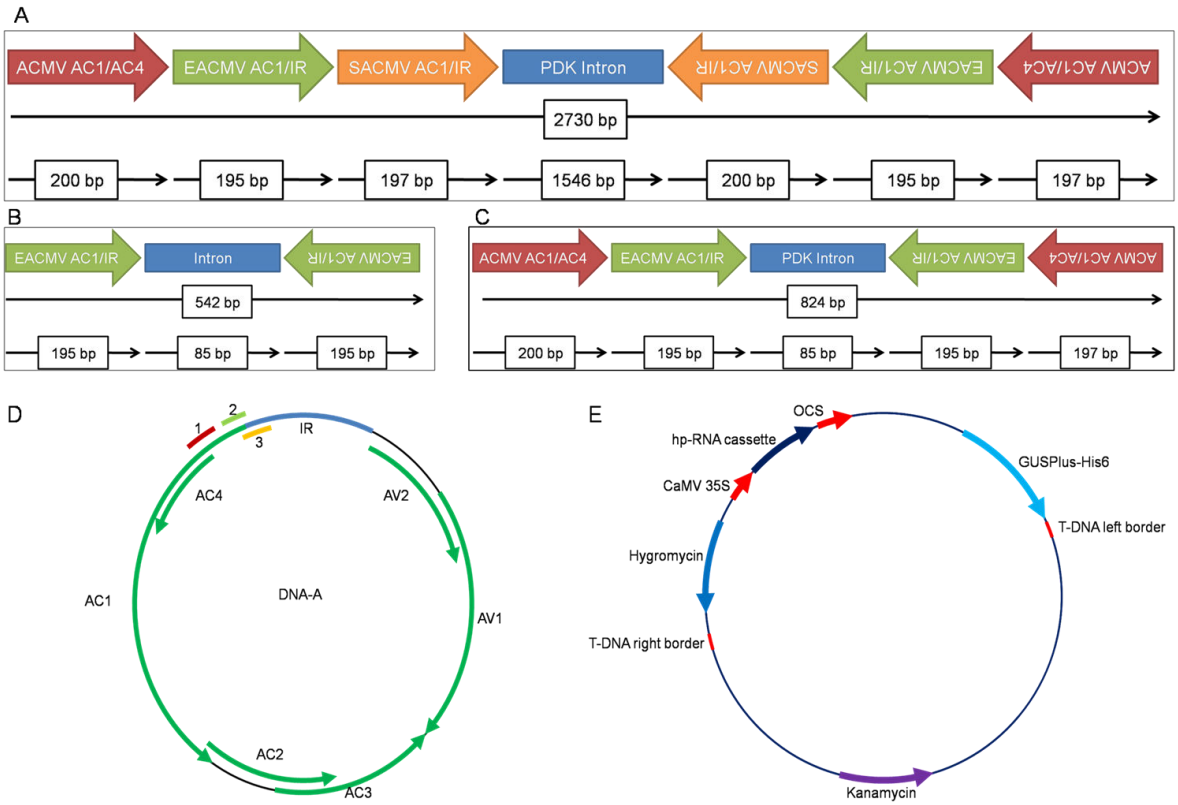


Figure 3.1: Schematic representation of the AES (A), EACMV (B) and E-ACMV (C)hp-RNA constructs, targeting the AC1/AC4 region (D1) of ACMV, and/or the AC1/IR region of EACMV (D2) and SACMV (D3). Each hp-RNA construct was inserted into pCambia 1305.1 (E) under the control of the CaMV 35S promoter and OCS terminator. The sense and anti-sense arms of pC-AES are separated by a PDK intron while pC-EACMV and pC-E-ACMV contain a 85bp intron (Vanderschuren *et al*, 2009).

3.2.1.2. pC-AES

In order to construct the pC-AES vector, three fragments corresponding to the target sequence of AC1/IR SACMV (197 bp) and EACMV (195bp) and AC1/4 (200bp) region of ACMV were amplified via PCR from the DNA-A of an infectious clone of the three viruses: SACMV (**Figure 3.1D)**, ACMV-NOg and EACMV-Cameroon, using primers ACMV Xba F and R, EACMV F and R and SACMV XbaF and R, respectively (**Table 3.2**).The PCR was carried out using Phusion Hot Start II High-Fidelity Taq DNA Polymerase according to the manufacturer's instructions (Thermoscientific). The amplified fragments were ligated together using T4 ligation enzyme (Thermoscientific) according to the manufacturer's instructions and cloned into

pJET1.2/blunt cloning vector (pJET-AES). The forward and reverse arms of the hp-RNA construct were amplified using pJET-AES as a template with both the ACMV XbaI F and SACMV XbaI R or ACMV XhoI F and SACMV XhoI R primers. The PCR fragments were digested with either FastDigest XhoI or XbaI RE (Thermoscientific) and ligated into pHellsgate-8 using directional cloning in an inverted repeat form separated by the PDK intron (**Figure 3.1A**). The hp-AES triple construct was then isolated from pHellsgate-AES using restriction digestion with FastDigest NotI (Thermoscientific) and sub-cloned into binary vector pCambia 13051.1 (referred to a pC-AES).

3.2.1.3 pC-EACMV and pC-E-ACMV

In order to construct pC-EACMV, a DNA fragment corresponding to IR/AC1 of EACMV-Cameroon and an 85bp intron including EcoRI and KpnI RE cut sites was synthesised (GeneArt synthesis, Thermoscientific). The reverse EACMV arm was synthesised via PCR, using primers EACMV KpnI and EACMV HindIII (**Table 3.2.1**) including the HindIII and KpnI RE cut sites. The DNA fragments were inserted into pCambia 1305.1 via directional cloning to form hp-EACMV construct (**Figure 3.1B**). In order to construct pC-E-ACMV (**Figure 3.1C)**, the AC1/4 ACMV DNA fragment and AC1/IR including an 85bp Intron from EACMV DNA fragment were isolated from the pC-AES and pC-EACMV vectors using primers: ACMV EcoRI and ACMV SacI and, EACMV SacI and EACMV KpnI respectively. The fragments were ligated together using the T4 ligase (Thermoscientific) protocol and inserted into pCambia 1305.1 via directional cloning (pC-A-E_fwd). The inverted arm was synthesised via PCR, using the forward arm as a template with primers E-ACMV HindIII and A-EACMV KpnI inserted into pC-E-A Fwd via directional cloning. All vectors and clones were sequenced by Inqaba Biotec© using Sanger sequencing to confirm the sequence integrity.

3.2.2 Agrobacterium-mediated transformation of cv.60444 FECs

Friable embryonic callus was transformed with pC-AES, pC-E-ACMV and pC-EACMV according to Bull *et al* (2009). Briefly, chemically competent *Agrobacterium tumefaciens* LBA4404 was transformed with pC-AES, pC-E-ACMV or pC-EACMV binary vectors using the freeze-thaw method (Hofgen and Willmitzer, 1988).

Transformed *Agrobacterium tumefaciens* LBA4404 was used to inoculate 7 plates of cv.60444 FEC clusters (10 clusters per plate; average size of 1 cm per cluster) according to method described in Bull *et al* (2009) (GD stages of each transformation experiment found in **Table 3.1**) .

Table 3.1: GD stage of FECs transformed with pc-AES, pC-E-ACMV or pC-EACMV

Vector	Transformation	GD stage
pC-AES	1	13
	2	15
	3	11
pC-E/ACMV	1	11
pC-EACMV	1	11

Empty pCambia 1305.1, untransformed *A. tumefaciens* LBA4404 and Greshoff and Doy (GD) buffer were included as controls. The inoculated FECs were co-cultivated at 28°C for 4 days on GD media (2.7g/L Gresshof and Doy medium including vitamins (Duchefa), 20g/L sucrose, picloram (12 mg/ml) (SigmaAldrich) in 16h light/8h dark photoperiod after which the *Agrobacterium* was removed by repeated washing with GD liquid media (2.7 g/L GD medium including vitamins, 20 g/L sucrose, picloram (12 mg/ml) and carbenicillin (SigmaAldrich) (500 mg/ml)). The transformed FECs were then subjected to selection through a series of transfers to GD media supplemented with carbenicillin (250µM/ml) and increasing concentrations (5 µg/ml, 8 µg/ml and 15 µg/ml) of hygromycin. At each stage the FECs were maintained at 28°C for 7 days in 16h light/8h dark photoperiod. The control untransformed cv.60444 FECs were not subjected to hygromycin selection. In order to regenerate transgenic plantlets, the FECs were transferred to regeneration media containing MS2 medium supplemented with 1 µg/ml 1-Naphthaleneacetic acid (NAA) and 250 µg/ml carbenicillin, 15 µg/ml hygromycin and solidified with 8 g/L Noble agar. The FECs were maintained on media in 10 day cycles at 28 °C with 16 h light/8 h dark photoperiod. This cycle was continued until the FECs stopped producing cotyledons. While the FECs were maintained on regeneration media, emerging green cotyledons were moved to cassava elongation medium (CEM) (4.8g/L MS supplemented with 0.4 µg/ml BAP, 2 µM CuSO4, 100 µg/ml carbenicillin and 8 mg/ml Nobel agar). The cotyledons were maintained on CEM on a 14 day

cycle until juvenile leaves and shoots appeared. The shoots were then transferred to sterile bottles on solid cassava basic medium (CBM) (4.8g/L MS with 2 µM CuSO4, 50 µg/ml carbenicillin and 8g/L plant agar). Plates were incubated for 28°C with 16h light/8h dark photoperiod. Regenerated lines were preliminarily screened for presence of the transgene cassette by GUS assay and a rooting test according to method described in Bull *et al* (2009).

Table 3.3: Primers used for the construction of the pC-AES, pC-E-ACMV and pC-EACMV hp-RNA constructs

Construct	**Virus**	**Target**	**Primer Name**	**Primer (5'-3')**
pC-AES	ACMV	AC1/AC4	**ACMV Xbal F**	GCTCTCTAGACTTGATTTGGCACCTTGAATGTTG-
			ACMV XholF	ATATCTCGAGCTTGATTTGGCACCTTGAATGTTG
			ACMV R	GCCCTTTCCGCTGTCATTCATTCAAACACTATC
	EACMV	AC1/IR	**EACMV F**	TGAATGAAGGAAAGGGCCTCTTCTTTTG
			EACMV R	GCATAAAAGCCCTACCTATTTACACATATGCCATTG
	SACMV	IR/AC1	**SACMV F**	GGTAGACTTCGGCTTTTATGCTAAAACGACCA
			SACMV XhoI R	ATGTCTCGAGGGCCATCCGGTAATATTAGAC
			SACMV XbaI R	TCCGTCTAGAGGCCATCCGGTAATATTAGAC
pC-E-ACMV	EACMV	AC1/IR+ Intron	**EACMV SacI**	CGTGTGAATTCGGAAAGGGCCTCT

			EACMV KpnI	3'-GGAAGAGCTCGGCCTCTTCTTTGGTTAATGA
	ACMV	AC1/AC4	**ACMV EcoRI**	TGATGAATTCGGCACCTTGAATGTTGGGGT
			ACMV SacI	TTCCTTGAGCTCTATCTCTCCCATCAA
	EACMV	AC1/IR (Reverse arm)	**E-ACMV HindIII**	AATTCGAAGCTTGCACCTTGAATGTTGGGGTCT
			E-ACMV KpnI	GCAGGGTACCATGGCATATGTGTAAATAGGTAG
pC-EACMV	EACMV	IR/AC1	**EACMV KpnI**	CTCGAAAGCTTCGGAAAGGGCCT
			EACMV HindIII	AATTTGGTACGCTACCTATTTACA

3.2.3 Molecular characterisation of AES lines

Putative transgenic lines and transformed FECs were screened for the presence and expression of the transgene using RT-PCR. Total RNA was extracted from 0.1g of leaf material of transgenic AES plant lines and FECs transformed with pC-AES, pC-EACMV, pC-E-ACMV and pCambia 1305.1 as well as untransformed cv.60444 FECs and leaves, using Trizol lysis reagent (Qiagen) according to the manufacturer's instructions. One microgram of DNAse I-treated (Thermoscientific) RNA from each sample was used to for cDNA synthesis following the Maxima First strand cDNA synthesis kit and protocol using the random primers. AES lines and FECs were then screened for expression of the transgenes using 2 microliters of synthesised cDNA as a substrate, with either the ACMV XbaI F and SACMV XbaI R (Tm=58°C) or EACMV F and R primers respectively (**Table 3.2**) (Tm=54°C). Untransformed cassava cv.60444 was included as a negative control and purified pC-AES vector was included as a PCR control. PCR were conducted with DreamTaq (Thermoscientific), with an initial denaturation step of 95°C for 3 minutes, followed by 35 cycles of 95°C for 30s, annealing at specified Tm, elongation at 72°C for 30 seconds. The final elongation was carried out at 72°C for 10 minutes.

3.2.4 Northern Blots

3.2.4.1 Probe synthesis

An EACMV fragment was amplified from the pC-AES clone with primers EACMV F (**Table 3.2**) and EACMV-T3 (5'-AATTAACCCTCACTAAAGGGCTACCTATTTAC-3') using PhusionTaq polymerase (Thermoscientific), according to protocol (Tm of 56°C). The 200ng of purified PCR fragment (PCR clean-up Gel extraction kit (Machery-Nagel)) was then labelled using DIG Northern Starter kit II protocol (Thermoscientific) according to the manufacturer's instructions.

3.2.4.2 Hybridizations

Total RNA (20µg) was denatured in 2X RNA loading dye (Thermoscientific) at 65°C for 10 min and separated on a 15% polyacrylamide gel. Total RNA was then transferred to Hybond-N+ membrane (Amersham GE Healthcare) using a semi-dry electro blotter (Biorad) and immobilised via UV cross-linking (CL-508, Uvitec Cambridge) on the RNA side at 0.120J/cm2. Pre-hybridisation (1h) and hybridisation (O/N) was performed at 45°C using EasyHyb hybridisation buffer (Thermoscientific). The membrane was then washed twice in 2X SSC + 0.1 w/v SDS solution at room temperature for 5 min and then washed in 0.1X SSC (with 0.1w/v SDS) at 45°C for 15 min twice. Antibody detection was performed following DIG Northern Starter Kit II (Thermoscientific) protocol. Results were visualised using GelDoc XR+ (Biorad).

3.2.5 Acclimatisation of selected transgenic plants

AES lines 1,2, 4 and 5, 6 cv.60444 transgenic lines, as well as untransformed cv.60444 were micro-propagated on CBM (MS medium including vitamins (4.4 g/L), sucrose (20 g/L) and CuSO4 (2 mM)), for viral challenge (20-30 nodal cuttings were made for each line). Rooting plantlets (after 2 weeks) were removed from growth media and transferred to peat jiffies (Jiffy© International) where they were acclimatised in plastic trays sealed with cellophane for a further 3 weeks. The trays were maintained at 28 °C, with 16 h light (8000-10000 lux) and 8 h dark cycles in 60 % humidity. After 3 weeks the cellophane was gradually removed over a period of days. After cellophane was removed from the plants, the jiffy coverings were slit to allow the spread of the roots and were transferred soil (3:1 mixture of potting soil: vermiculite).

3.2.6 siRNA off-target prediction

The forward arm sequence of the triple construct AES, E-ACMV and EACMV was used as input query to predict the most probable siRNAs using prediction software: Jack Lin's siRNA Sequence Finder (Lin 2002), and Ambion siRNA Target Finder (URL:www.ambion.com/techlib/misc/siRNA finder. html). Predicted siRNAs were mapped to cassava genome (*Manihotesculenta*v6) using BLAST software adjusted for short nucleotide sequences (https://blast.ncbi.nlm.nih.gov/Blast.cgi) to identify possible gene targets. Targets were selected based on the following criteria: 1) No mismatches in positions 2-12 of siRNA/target duplex 2) No more than two adjacent mismatches in the siRNA/target complex and 3) length of complementarity with a cut-off length of <14nt.

3.2.7 RT-qPCR for expression of siRNA targets

To determine expression of identified off-target genes, total RNA was isolated from 0.1g of plant material using Trizol reagent according to the manufacturer's instructions. The total RNA was treated with DNase I (Thermoscientific). One microgram of total RNA from each sample was then used for cDNA synthesis using random hexamers following the RevertAid First Strand cDNA Synthesis Kit (Thermoscientific) protocol. Two microlitres of the synthesised cDNA products were then used to for Real-time analysis using Maxima SYBR protocol (Thermoscientific). Each reaction contained: 1x Maxima SYBR green Master Mix, 0.5mM of the target specific primer (**Table 3.2**). The PCR conditions were 5 min at 95 °C followed by 30 cycles of 95 °C for 15 s, 60 °C for 30 s and 72 °C for 1 min. Normfinder (https://moma.dk/normfinder-software) was used to compared expression of reference genes including *β-actin*, *glyceraldehyde-3-phosphate dehydrogenase* (GAPDH), *β-tubulin* and *UBQ 10* (Hu *et al*, 2016) and *UBQ10* was selected as the internal control. Untransformed cv.60444 FECs, as well as FECs transformed with empty pCambia 1305.1 were also included as controls.

Table 3.4: Primers used for real-time analysis of gene expression

Gene Target	Forward (5'-3')	Reverse (5'-3')
Ubiquitin 10	TGCATCTCGTTCTCCGATTG	TGCATCTCGTTCTCCGATTG
WRKY transcription factor 14	CCAATCAGATGCTCCTGTATCC	TATCCCTGGTATCCTCCATAGC
WRKY transcription factor 76	CCGCATCCAAGAGGCTACTA	CAGAGCTCTCTCCACATGCT
Phenylalanine ammonia-lyase 2 (PAL2) gene	AAGATGGTGGCTGAGTTTAGG	AGACCAGATTCACGAGCAATAG
UVB sensitive 2	AGACCAGATTCACGAGCAATAG	AAGCCAGTTCCCAAGTCATAAA
Peroxidase 15-like	GGGCCAACATGGACTAATCTAC	AGCTCCTCGAGGGTCATAAA
mes Pre-miRNA 319b	GCCGACTCATTCATTCAAACAC	GTCCAAGCACAGAGAAGAGAAG
Late embryogenesis abundant protein 14	TCCCAATCCTGTTCCAATCC	AGTTCCAGCGTCTGGTATTAAG
AMP deaminase-like	GGGCTAAGCTAGGGATTGATTT	CCTTCTCCTGAACTCCACAATC
GEM-like protein 4	CTCCCATCTCCAAGCAAAGT	CATGCTCTCGAACTCCTTGT
Farnesyl pyrophosphate synthase 1	GTGCTCTTGGTTGGTTGTAAAG	CAGCTTTCCCATAGTGCTCATA
Farnesyl pyrophosphate synthase 2	GACCAAGATGTCTCCCAGAATAG	GCTCCCTTTGGATACTCAGAAA
GEM-like protein 4	CTCCCATCTCCAAGCAAAGT	CATGCTCTCGAACTCCTTGT

3.2.8 Statistical analysis

In order to determine whether the transformation efficiency results are statistically significant, an student t-test was performed according to **section 2.2.5**

3.3 Results

3.3.1 Transformation of cv.60444 FECs with pC-AES construct

Three separate transformations of cassava cv.60444 FECs were conducted using pC-AES, with pCambia 1305.1 and untransformed *A. tumefaciens* LBA4404 was used as controls. In all three experiments, the transformation efficiency of the FECs transformed with pC-AES did not exceed 7% which was significantly lower than the pCambia 1305.1 control (67%) and did not correlate with the reported levels (Bull *et al*, 2009). Further FECs transformed with pC-AES had lower multiplication/proliferation in comparison to FECs transformed with pCambia 1305.1 at the same time point (**Figure 3.2 B and C**). The 6 regenerated lines were screened for expression of molecular markers gene *Gus* (**Figure 3.3A)** and *Hyg* (**Figure 3.3B**) as well as the forward arm of the insert (**Figure 3.3C)**. The transformed lines showed a number of developmental defects including stunted growth (in comparison to untransformed cv.60444) and leaf deformation and discolouration (**Figure 3.2D**).

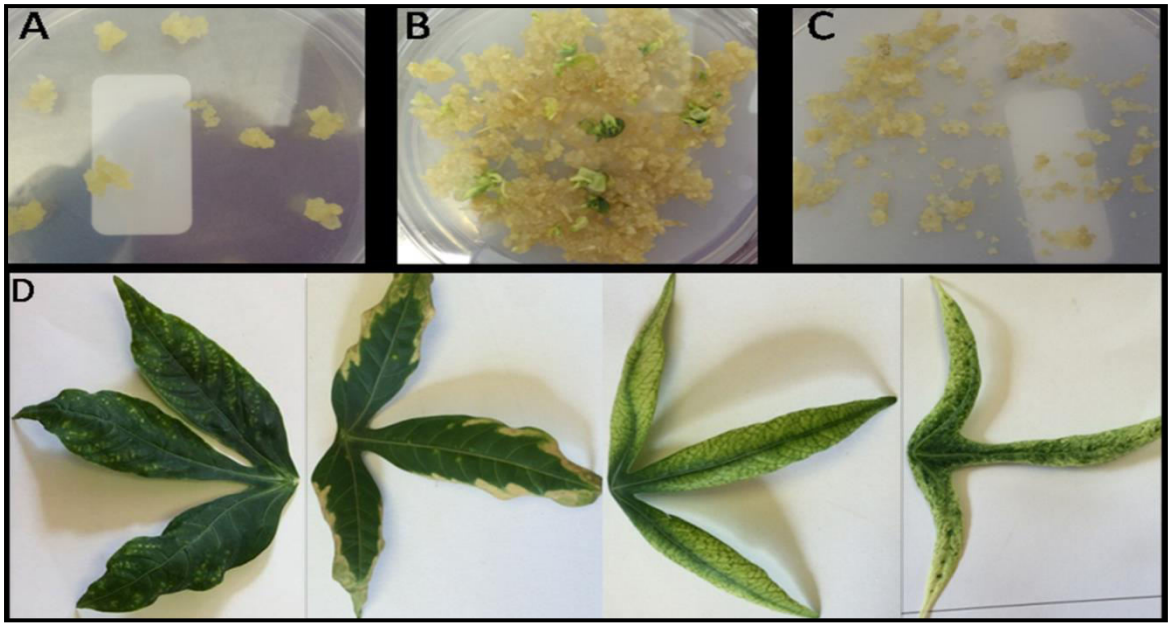


Figure 3.2: Effect of transformation of cv.60444 FECs with pC-AES and cv.60444 transformed with pC-AES A) Untransformed FECs clusters. B) Cassava cv.60444 FECs transformed with pCambia 1305.1 (control). C) Necrotic FECs transformed with triple construct. D) Aberrant foliar symptoms (from left to right): laminar clearing, margin clearing, filiform and misshapen leaves.

Based on these results, it was hypothesised that the pC-AES construct is toxic to cassava cv.60444. It was hypothesised that the toxic effect of the construct could be due to the EACMV sequence, as previously a construct targeting two overlapping regions (AC2/AC3 + AC1/AC4) of EACMV had been shown to have similar effects (unpublished). Alternatively, the toxic effects could be due to the production of off-target siRNAs which target genes important to development in the plant. Vanderschuren *et al*(2007) noted that a cassava cv.60444 transformed with a construct targeting the IR region of ACMV also had low transformation rates, although their transgenic plants did not display the aberrant symptoms displayed in the AES lines.

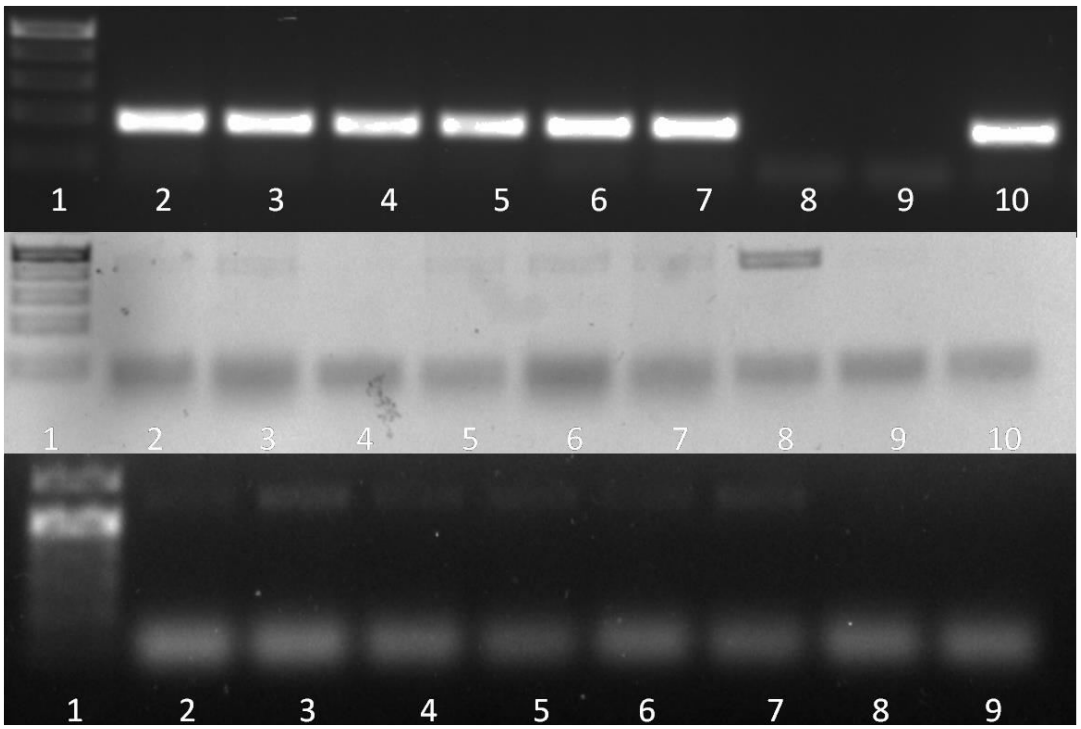


Figure 3.3 Expression of *Gus* (A), *Hyg* (B) and AES transgene (C) in transformed cassava cv.60444 lines (Line 1, 2, 3, 4, 5, 6). Molecular weight marker was loaded into the first lane of each gel; 1kb Plus molecular weight marker (Thermoscientific) A: Expression of *Gus* in transgenic lines. Lane 2: Line 1, lane 3: Line 2, lane 4: Line 3, lane 5: Line 4, lane 6: Line 5, lane 7: Line 6, lane 8: untransformed cv.60444, lane 9: NTC, and lane 10: pC-AES (positive control). B: Expression of *Hyg* in transgenic lines. Lane 2: Line 1, lane 3: Line 2, lane 4: Line 3, lane 5: Line 4, lane 6: Line 5, lane 7: Line 6, lane 8: pC-AES (positive control), lane 9: untransformed cv.60444, and lane 10: NTC. C: Expression of AES transgene in transformed cv.60444. Lane 2: Line 1, lane 3:

Line 2, lane 4: Line 3, lane 5: Line 4, lane 6: Line 5, lane 7: Line 6, lane 8: untransformed cv.60444, lane 9: NTC.

3.3.2 Comparison of transformation efficiency of FECs transformed with pC-EACMV, pC-E-ACMV and pC-AES

In order to test whether the presence of the EACMV AC1/IR region was toxic to the transformed FECs, or whether the toxic phenotype noted in the transformed FECs is due to the stacked constructs, pC-EACMV targeting EACMV (AC1/IR) and pC-E-ACMV targeting ACMV (AC1/AC4) and EACMV (AC1/IR), were constructed and the transformation efficiency of the three constructs (pCambia 1305.1 and untransformed LBA4404 were included as controls) was compared. The transformation efficiencies of pC-AES, pC-E/ACMV and pC-EACMV were 7.05%, 16.7% and 21,9% respectively which was significantly lower than the control pC-1305.1 (73.3%)($F_{3,8}$=41.54, $p<0.05$). While FECs transformed with each of the hp-RNA constructs exhibited lower transformation rates than the controls, the rates were comparatively lower in pC-AES and pC-E-ACMV compared to pC-EACMV alone, which indicated that the toxic effects of the constructs are more likely due to the effects of stacking of more than one virus sequence (pC-AES and pC-E-ACMV), which could cause off-target effects, rather than the presence of EACMV sequence. In order to confirm expression of the transgene, RT-PCR was conducted on RNA extracted from FECs transformed with the three constructs, as well as untransformed cv.60444 (**Figure 3.4).** Transgene expression was confirmed in FECs transformed with all three constructs.

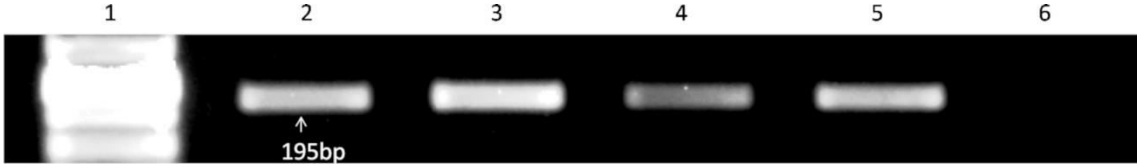


Figure 3.4: RT-PCR to confirm expression of transgenes in FECs transformed with pC-AES (Lane 2), pC-E-ACMV (Lane 3) and pC-EACMV (Lane 4). Lane1: 1kb plus Molecular weight marker (Thermoscientific), Vector pC-AES (positive control for PCR) (Lane 5) and untransformed cv.60444 (negative control) (Lane 6) were also included.

3.3.3 Northern Blots

In order to confirm that siRNAs corresponding to the three pC-EACMV, pC-E-ACMV and pC-AES constructs is being expressed in FECs, a Northern blot hybridisation was performed (**Figure 3.5)**. RNA extracted from untransformed cv.60444 FECs was used as a control. The relative levels of siRNAs were compared to the density of the 5 ng of 21nt miRNA169 marker (**Figure 3.5,** lane 1) using the ImageLab v6.01 (Biorad), and results showed that pC-EACMV produced the highest relative amount of siRNA (2.65), pC-AES had a similar amount(2.32) and pC-E-ACMV produced the least amount of siRNAs, compared to the control (1.98). However, this analysis of the relative siRNA expression levels is only an estimation.

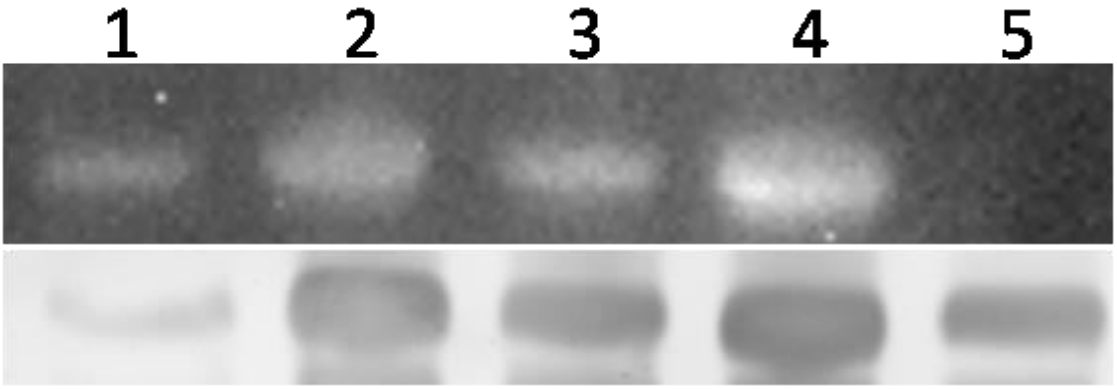


Figure 3.5: Northern blot hybridisation of FECs transformed with pC-AES, pC-E-ACMV and pC-EACMV constructs. Lane 1: DIG-labelled 21nt marker, Lane 2: pC-EACMV, Lane 3: pC-E-ACMV, Lane 4: pC-AES, Lane 5: cv.60444

3.3.4 siRNA off-target prediction

Analyses of the full length triple construct AES, E-ACMV construct and EACMV construct with siRNA prediction software predicted 55, 34 and 19 possible siRNAs respectively. The 19 predicted siRNAs produced by EACMV were common to all three constructs as expected. Twenty of the predicted siRNAs from AES were produced by the SACMV AC1/IR region and were therefore not predicted for either E-ACMV or EACMV. The E-ACMV and triple construct AES shared 15 siRNAs not produced by EACMV. The predicted siRNAs were then compared to cassava genome, to identify possible gene targets. Small interfering RNAs were required to have completely homogeneity between bases 3-8 at 3' end and have a minimum of 14 complementary bases. Ten possible siRNAs with corresponding which had

homology to 9 cassava gene targets were identified (**Table 3.3**). Three of the predicted siRNAs are produced by ACMV AC1/4, which forms part of the pC-AES and pC-E-ACMV constructs and bind to a variety of targets including Peroxidase 15-like, Late-embryogenesis protein and mes pre-miRNA 319b (**Figure 3.6**). Four siRNAs with homology to the cassava genome are predicted to be produced by the EACMV AC1/IR region and have homology to Farnesyl pyrophosphate synthase -1 and Farnesyl pyrophosphate-2, GEM-like protein and Adenosine monophosphate deaminase (AMP) deaminase-like protein. Three siRNAs with homology to UVB sensitive protein, Phenylalanine ammonia lyase 2 and Transcription factor WRKY 14 are predicted to be produced by the SACMV AC1/IR region.

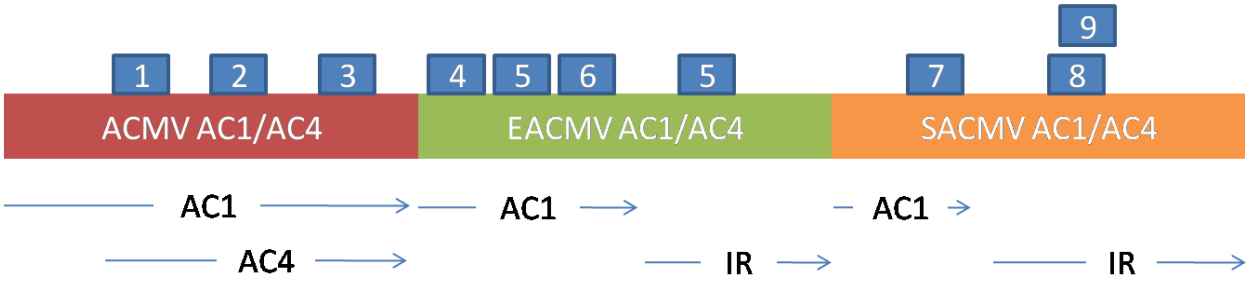


Figure 3.6: Positions of predicted siRNAs produced in relation to the forward arm of pC-AES. The position the siRNAs indicate the origin of the siRNA on the AES triple construct and is labelled numerically to the indicate the homologous cassava gene 1) Peroxidase 15-like (Prxs-15) 2) Late embryogenesis abundant protein 14 (Late-Embryo) 3) mes pre-miR319, 4) Farnesyl pyrophosphate synthase -1 and Farnesyl pyrophosphate-2, 5) GEM-like protein 4 6) AMP deaminase-like protein, 7) UVB sensitive protein 8) Phenylalanine ammonia lyase 2 and 9) Transcription factor WRKY 14.

Table 3.5: Putative siRNA sequences produced by pC-AES, pC-E-ACMV and pC-EACMV.

No.	siRNA (5'-3')	Region	Possible targets	Function of Target	Binding Site
1	AAGUGAGGUUCCCCAUUCUGAUG	ACMV	Peroxidase 15-like	Removal of H_2O_2Oxidation of toxic reductants Biosynthesis and degradation of lignin Auxin catabolism Response to environmental stresses (Gómez-Vásquez *et al*, 2004)	mRNA coding region
2	GAUGCAGCUCUCGACAGAUUUUA	ACMV	Late embryogenesis abundant protein 14	Role in dehydration tolerance (Karami *et al*, 2009)	3' UTR
3	GAGAGAUAGUGUUUGAAUGAAUG	ACMV	Pre-microRNA 319c	Jasmonic acid signaling Response to auxin Leaf morphogenesis Cell differentiation Regulation of development (Schommer *et al*, 2014)	Precursor miRNA

4	GAAAUAAUUUUUGGCAUUUAUUU	EACMV	Farnesyl pyrophosphate synthase 1	Isoprenoid synthesis (Zhang *et al*, 2015)	3'UTR
			Farnesyl pyrophosphate synthase 2		3'UTR
5	AAAUAAUUUUUGGCAUUUAUUUG	EACMV	GEM-like protein 4	Embryo development in *Arabidopsis* (Mauri *et al*, 2016)	mRNA coding region
	AAGGGCCUCUUCUUUGGUUAAUG				
6	AAACGACCGGCUCUUGGCAUAUU	EACMV	AMP deaminase-like	Energy metabolism. Essential for the transition from zygote to embryo. Expressed during somatic embryogenesis (Xu *et al*, 2005)	3' UTR
7	CAGCCCUCGGCAUUUUCGCUGUC	SACMV	UVB sensitive 2	Root UV-B sensing pathway Protection against the hypersensitivity to very low-fluence-rate(Peer *et al*, 2010)	5'UTR
8	CAAAUGGCAUAUGUGUAAUUUUG	SACMV	phenylalanine ammonia-lyase 2 gene	Catalyzes first reaction in the biosynthesis from L-phenylalanine(Huang *et al*, 2010)	5' UTR

9	AAUGGCAUAUGUGU AAUUUUGUG	SACMV	WRKY transcription factor 14	Involved in drought tolerance in cassava (Agarwal *et al*, 2011)	3' UTR

3.3.5 Expression of siRNA targets

Computational analysis of the pC-AES, pC-E-ACMV and pC-EACMV constructs showed that there are 10 possible siRNAs with predicted protein targets which could bind to targets in the cassava genome. In order to assess whether the expression of any of these gene targets have been regulated in response to transformation with the constructs, expression analysis using RT-qPCR was conducted. Cassava cv.60444 FEC tissues transformed with pC-AES, pC-EACMV and pC-E-ACMV, were collected after the transformed FECs had been incubated on regeneration media for 4 weeks (when pCambia 1305.1 FECs started to produce cotyledons). *Ubiquitin 10* was used as an endogenous gene control to normalize the relative expression of the genes and gene expression was compared to FECs transformed with pCambia 1305.1. The relative RT-qPCR showed that 6 of the predicted targets; transcription factor WRKY14, GEM-like 4 protein, AMP deaminase-like, peroxidase 15 (Prxs 15)), Farnesyl pyrophosphatase-1 and Farnesyl pyrophosphatase-2 were significantly down-regulated in the response to transformation in comparison to FECs transformed with pC-1305.1 vector only (**Figure 3.7**).

Two cassava genes targeted by siRNAs predicted from the AC4 region of ACMV, namely Peroxidase 15 and mes pre-microRNA 319, showed a variation in regulation. Peroxidase-like 15 was shown to be significantly down-regulated ($p<0.01$) in response to pC-AES (4.46±0.38) and pC-E-ACMV (1.63 ±1.03) compared to pC-1305.1 (5.53±1.63). The mes pre-miRNA319 transcript was shown to be significantly up-regulated in response to transformation with pC-AES (2.41±1.31) and pC-E-ACMV (2.35±1.39) in comparison to pC-1305.1 (0.60±0.23).

Four of the cassava proteins, Farnesyl pyrophosphate 1 and 2, GEM-like 4 and AMP deaminase, targeted by siRNA predicted from the EACMV construct were down-regulated in FECs transformed with the three constructs. AMP deaminase-like

protein was significantly down-regulated in FECs transformed with pC-E-ACMV in comparison with pC-1305.1 vector control (2.11±0.44). GEM-like protein 4 was significantly down-regulated by pC-AES (-3.8±0.05), pC-E-ACMV (-5.8±0.1), pC-EACMV (-4.6±0.06) in comparison to pC-1305.1 (1.13±1.5) Farnesyl pyrophosphate 1 was significantly down-regulated in pC-AES (-0.96±0.28) and pC-E-ACMV (-0.28±0.33) compared to pC-1305.1 (1.55±1.93) while Farnesyl pyrophosphate 2 was significantly down-regulated in all three FECs lines (-0.58±0.53, -5.32±0.18, 0.47±1.07) in comparison to pC-1305.1 (2.56±0.91). WRKY transcription factor 14 targeted by siRNA produced by IR region of SACMV was shown to be significantly down-regulated in FECs transformed with pC-AES (-2.58±0.11) compared to pCambia 1305.1 (0.30±0.28).

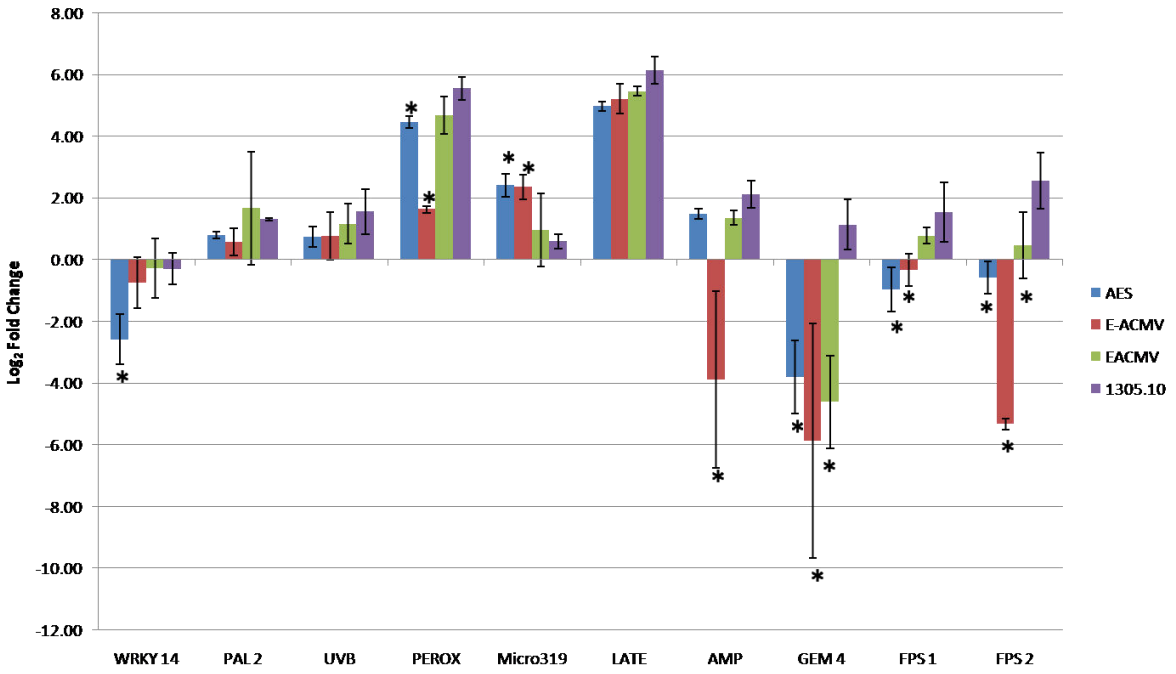


Figure 3.7: Relative gene expression fold changes (Log_2Fold) of gene targets associated with siRNA produced by constructs pC-AES, pC-E-ACMV and pC-EACMV in FECs transformed with the corresponding constructs. The error bars show standard variation and asterisks indicate statistical significant results ($p<0.01$) (Appendix I).

3.4 Discussion

RNA-induced silencing where hp-RNA homologous to target viral mRNA are transformed into plants, is a widely used technique designed to improve resistance. Cassava mosaic disease is the largest threat to cassava cultivation and is caused by a number of cassava mosaic viruses including ACMV, EACMV and SACMV. Cassava mosaic disease is usually associated with mixed infections in the plant (Bull *et al*, 2007) and an effective resistance strategy should enhance resistance to all the CMG species. Stacked constructs offer durable and broad spectrum resistance for crops exposed to mixed virus infections and are an effective way of producing crops that have multiple resistance simultaneously (Fuchs, 2017). In order to engineer a cassava variety that shows durable resistance to EACMV, ACMV and SACMV, a triple stacked construct (pC-AES) targeting the overlapping region of Rep (AC1) and AC4 region of ACMV, and the AC1 and Intergenic regions of EACMV and SACMV was developed and used to transform cassava cv.60444 lines. However transformation efficiency of FECs was very low in comparison to the control lines and confirmed AES lines transgenic showed an abnormal phenotype not present in the control lines. Previous studies targeting either ACMV (Moralo, 2015) or SACMV (Taylor, 2009) did not lead to severe abnormalities, which lead to the hypothesis that either the EACMV region, or the stacked construct could be responsible for the abnormal phenotype, producing off-target siRNAs with high level of homology to cassava genes. Two further constructs targeting the ACMV-EACMV and EACMV portions of the triple construct respectively, were included in the study to test this hypothesis. Regeneration rates in all three constructs were significantly lower than in the control lines, and pC-AES and pC-E-ACMV constructs were significantly lower than in the pC-EACMV constructs. This confirmed the hypothesis that the low regeneration rates could be due to a combined effect of stacked construct, producing siRNA-associated gene off targets, which has been reported in many other studies (Fuentes *et al*, 2006; Praveen *et al*, 2010; Xu *et al*, 2006). Predicted siRNAs from the pC-AES triple construct had homology with a number of genes associated with development, disease resistance and stress-tolerance the cassava transformed with the three constructs (**Table 3.3)**. Analysis of these genes expression in FECs transformed with the three constructs was shown to be differential expressed. The results from this research are significant as triple constructs targeting multiple CMGs

have not been attempted to date. Furthermore, although phenotypic variation in transgenic lines has been reported before in lines expressing viral proteins (Chellappan *et al*, 2004a; Zhang *et al*, 2005), this is the first report of siRNA specific gene 'off-targets' identified in cassava in response to cassava mosaic geminiviruses.

Seven genes, associated with transgene-derived siRNAs produced by pC-AES were shown to be differentially-regulated in FECs transformed with the stacked constructs. Two siRNAs corresponding to the AC4 region of ACMV (1 and 3) have homology to the mRNA region of Prxs-15 and mes pre-miRNA 319 respectively (**Figure 3.6**) and were shown to be regulated in FECs transformed with pC-AES and pC-E-ACMV, but not by pC-EACMV. siRNAs (4, 5 and 6) corresponding to the AC1/IR EACMV region had homology to 4 cassava genes, FPS1, FPS 2, GEM-like protein 4, and AMP-deaminase. siRNA 4 and 6 corresponded to the AC1 region of EACMV and have homology to the 3' UTR region of FPS 1 and FPS2 and, AMP deaminase respectively. Two siRNAs corresponding to the AC1 and IR regions of EACMV have homology to the mRNA and 3' UTR regions of GEM-4 protein. These genes were down regulated in FECs transformed with all three constructs. One cassava gene, WRKY transcription factor 14, was shown to be down-regulated by a siRNA (9), which corresponds to the IR of the SACMV portion of pC-AES and was only found to be down regulated in FECs transformed with pC-AES construct.

RNA silencing technology has been widely adopted as means of improving crops (Senthil-kumar and Mysore, 2011). Plants are genetically modified using binary vectors which contain partial or full gene fragments, cloned in an inverted repeat separated by an intron (Waterhouse and Helliwell, 2003). The gene target, which usually corresponds to the mRNA sequence, is between usually between 200-350bp in length but can be increased (Xu *et al*, 2006). Transgene length has been correlated with efficiency of gene silencing, where longer gene targets result in more efficient silencing due to increased populations of siRNAs (Wesley *et al*, 2001). However, the dsRNA length can also impact the likelihood of producing off-target siRNAs where there is less control over individual siRNAs produced by longer constructs (Ossowski *et al*, 2008). Gene stacking, as the case in this study, where multiple target genes are combined into a single hp-RNA transgene, to silence multiple genes simultaneously, has the potential disadvantage that the hp-RNA construct are larger, producing larger more diverse, siRNA populations. The arm of

AES construct is 632 bp long and could generate 55 different possible siRNAs (according the Jack Lin's siRNA Sequence Finder). Although the majority of these siRNA were shown to have no homology with host genes, ten were shown in this study to have partial homology with identified genes in the cassava genome and could potentially regulate their expression.

siRNA have been shown to mediate both transcriptional gene silencing (TGS) and PTGS (Angaji *et al*, 2010). PTGS can also be used to target the 3' and 5' untranslated regions (UTRs) with great effect (Brummell *et al*, 2003; Kusaba, 2004). One of the commonly cited advantages of RNAi is that it is highly specific, due to the high level of homology required between the target and the siRNA. However, it has been shown that there does not have to be 100% homology between the two strands of RNA in order for the siRNA to bind (Praveen *et al*, 2010; P. Xu *et al*, 2006). In addition to this, siRNA which target the 3'UTR region of a gene requires very low levels of over-all homology, with complete homology only required at bases 2-8 at 3' end (Birmingham *et al*, 2006; Doench *et al*, 2003). This does not occur when similar homology is observed in the 5 'UTR region or the mRNA. Notably, five of the siRNA targeted genes (**Table 3.3**) identified in cv.60444 which were shown to be down-regulated by siRNAs produced by the hp-RNA constructs had high levels of homology to the 3'UTR region with 100% homology within the seed region, which could indicate why down-regulation was shown for these and not for those where homology was shown in 5'UTR region. Two of the other predicted siRNA (1 and 5) had high levels of homology to the mRNA coding region of host genes Prxs-15 and GEM-4 respectively which were shown to be down-regulated in response to transformation with pC-AES, and pC-E-ACMV and, all three constructs respectively. As little as 14 nt similarity between a siRNA and mRNA has been shown to be sufficient to cause silencing, with this being more efficient if homology is higher at the 3' end of the mRNA sequence (Wang *et al*, 2006). Both GEM-like 4 and Peroxidase-like 15 identified here in have high levels of homology (15 and 16 bp respectively) at the 3' end with the predicted siRNA, which could cause down-regulation of these genes and was demonstrated by RT-qPCR (**Figure 3.7**).

Geminiviruses have been shown to induce both PTGS and TGS in plants, where the dsRNA intermediates produced during viral replication are targeted by the cell silencing machinery (Blevins *et al*, 2006). Virus-derived siRNAs (vsiRNAs) are a

class of exogenous siRNAs produced from the viral genome using cell machinery including DCLs, RDRs and AGOs (Zhang *et al*, 2015). Previous studies have shown that the introduction of hp-RNA homologous to viral protein AC1, the AC1/AC4 overlap and the bidirectional promoter (Moralo, 2015; Vanderschuren *et al*, 2009, 2007a) can induce resistance through the production of siRNA, similar to vsiRNA produced in response to viral infection. The AES triple construct in this study targets the AC1/AC4 overlap as well as the intergenic region including the bidirectional promoter ACMV, EACMV and SACMV respectively and was predicted to induce high levels of resistance to all three viruses. However, the triple stacked construct exhibited a high level of toxicity in the cassava FECs, and caused developmental defects in the transgenic lines, which had not been previously reported. Regeneration of cassava, using the via the FECs method has been shown to produce a low level (<5%) of plants which show a variable phenotype, including weaken stems and lanceolate leaves (Ligaba-Osena *et al*, 2018; Moralo, 2015; Taylor *et al*, 2012) but the severity of physiological abnormalities observed with the triple construct pC-AES were severe and presented in all the transgenic lines, however these symptoms were not present in the either the FECs transformed with pCambia 1305.1 or untransformed cv.60444 FEC controls. Geminivirus viral suppressors AC2, AC4 and AV2 have all been linked to disease severity and when over-expressed have been linked to developmental abnormalities (Chellappan *et al*, 2005; Vanitharani *et al*, 2004). While VSR are commonly linked to developmental abnormalities, the regulation of host genes by vsiRNA is less commonly discussed. Although vsiRNA play a role in resistance, several studies have recently shown that some vsiRNA also interact with host proteins and transcription factors (Yang *et al*, 2018). A vsiRNA derived from CMV-Y satellite Y RNA was shown to target a magnesium chelatase subunit (*CHLI*) mRNA which resulted in chlorosis in the leaves, due the impairment of chlorophyll biosynthesis (Shimura *et al*, 2011). A single vsiRNA derived from *Potato spindle tuber virus* d was also shown to target several callose synthase genes in tomatoes, resulting in stunted plants with curled leaves (Adkar-Purushothama *et al*, 2015).

The triple construct pC-AES targets the AC4, AC1 and the IR of CMGs ACMV, EACMV and ACMV, all of which have been shown to be vsiRNA hotspots in other geminiviruses (Kuria *et al*, 2017; Miozzi *et al*, 2013). Interestingly, seven genes

related to abiotic and biotic stress tolerance and were shown to be differentially expressed in cv.60444 FECs transformed with the three constructs. Evidence strongly suggests that these off-target effects are most likely due to homology between the hp-RNA derived siRNA, and the host genome. A possible reason for the developmental defects observed in the cassava FECs, is that the constructs may be producing siRNA which mimic vsiRNAs, and which could interact with the identified host proteins and transcription factors. In order for vsiRNA to be produced, a viral dsRNA intermediate must be produced during viral replication. The three constructs used in this study produce hp-RNA which can be processed in order to produce a population of siRNAs and therefore could produce siRNA homologous to vsiRNA which have been previously identified (Kuria *et al*, 2017; Miozzi *et al*, 2013; Rogans and Rey, 2016) in high quantities.

A number of studies have been conducted on virus-derived RNA generated from geminiviruses, including a study conducted by Akbergenov *et al* (2006) which showed that a number of hosts including cassava, tobacco and *N. benthamiana* produce vsiRNA via TGS or PTGS in response to infection with ACMV as well as *Cabbage leaf curl virus*. While most identified vsiRNA have been shown to target specific regions of the viral genome, recently vsiRNAs that interact with host factors associated with biotic and abiotic stress including WRKY and MYB factors have also been identified (Wang *et al*, 2016). WRKY 14, targeted by a siRNA associated with the IR of SACMV (siRNA 9) was shown to be down regulated in FECs transformed with AES construct (**Figure 3.6**). WRKY transcription factors have a diverse range of functions and have been shown to play key roles in a broad range of biological functions in plants, ranging from pathogen-induced defence (Chen, 2002) to developmental processes such as embryogenesis (Agarwal *et al*, 2011; Chen, 2002; Jiang and Yu, 2009; Zhou *et al*, 2008). In cassava plants put under drought stress conditions WRKY14 was highly up-regulated, indicating that it is involved in the regulation of a number of genes involved in stress response (Wei *et al*, 2016). Allie *et al* (2014) showed that a number of WRKY transcription factors, including WRKY 14 were down-regulated in response to infection with SACMV in susceptible cultivar T200 and was up-regulated in resistance cultivar TME 3, exposed to CMGs in the (Fregene *et al*, 2004). WRKY 14 has over-lapping function in defence and

development and down-regulation could result in stunted growth, similar observed in the AES transgenic lines, as well as in the AES transformed FECs.

The ability of vsiRNA to regulate host proteins, is indicative of the constant balance found within virus-host interactions, and is employed by viruses in order to shift host biology in the viruses favour (Cao *et al*, 2013). The IR region of geminiviruses has been identified as "hot spot" for vsiRNA (Miozzi *et al*, 2013; Yang *et al*, 2019) and vsiRNA from this region have been shown to interact with host proteins and long non-coding regions responsible for resistance (Yang *et al*, 2019).The IR can be bi-directionally transcribed (Yang *et al*, 2019) and the expression of IR has been shown to be associated with abnormal phenotypes in TYLCV-susceptible tomato cultivars (Yang *et al*, 2019). Interestingly, a single vsiRNA from this region was found to be responsible for an abnormal phenotype, which mimics the curled phenotype associated with the disease and down-regulate the *SlLNR1* long non-coding (lnc) region (Yang *et al*, 2019). In Vanderschuren *et al* (2007) cassava cv.60444 transformed with a hp-RNA targeting the bidirectional promoter of ACMV was also shown to have low transformation efficiency, similar to those shown in AES lines, which they hypothesises could either be due to plant material quality or be due to the fact the construct could sequester host replicase and replication factors which are involved in plant development. The two siRNAs (5 and 9) associated derived from the IR of EACMV and SACMV were found to have homology with genes (GEM-4 and WRKY14) that were down regulated in the transformed FECs. GL2-EXPRESSION MODULATOR (GEM) is part of the GRAM (Glycosyltransferases, Rab-like GTPase Activators, Myotubularins) family, which is induced by Abscisic acid (ABA) (Caro *et al*, 2007; Liu *et al*, 1999). GEM-1 is important for early root development and cell differentiation, where down-regulation of GEM results in restricted cell division potential (Caro *et al*, 2007). GL2 (GLABRA2), which is negatively regulated by GEM proteins is a homeo-box leucine zipper rich box protein which has been shown to be important in the development of trichomes as well other developmental tissue in *Arabidopsis* (Rerie *et al*, 1994; Tominaga-wada *et al*, 2009). In *Arabidopsis,* cell differentiation was effected by GL2 activity, where cells with increased levels of GL2 showed increased differentiation. GL2 levels fluctuate as the cell cycle moves through the mitotic states, which allows the cells to differentiate and remodel. Interestingly, GL2 has been shown to regulate xyloglucan endotransglucosylase

(XET), an enzyme responsible for cross-linking of glucans in cell-walls (Tominaga-wada *et al*, 2009), with gl2 mutants showing increase levels of XET. Ma *et al* (2015) showed that during the transition from 'fresh FECs, to 'old FECs' there was a significant down-regulation in XET expression. Although no studies have currently been done on the link between GEM and cell development in cassava, it is possible that the down-regulation of GEM results the up-regulation of genes involved in cell differentiation, inhibiting regeneration in the FECs. In addition to the siRNA produced by IR region of EACMV, another siRNA produced by the AC1 region of EACMV was also shown to have a high level of homology to the GEM-like 4 protein. Wang *et al* (2006) showed that the efficiency of gene silencing can be increased through the production of multiple siRNA targeting a single gene. It is possible that the significant down-regulation seen here is due to production of multiple siRNA which target the same gene.

In addition to vsiRNA which have been shown to regulate host proteins and transcription factors (Miozzi *et al*, 2013), viral suppression of RNA silencing in host has also been linked to adverse effects on host biology (Praveen *et al*, 2010). In fact, the expression of the AC4 repressor can interfere with the production and action of miRNAs and results in abnormalities in plant development (Naqvi *et al*, 2010; Voinnet, 2005). The triple construct targets the full length AC4 region of ACMV and could be producing siRNA that could mimic vsiRNA known to be produced in this area. Two of the siRNAs predicted in this region (1 and 3) (**Figure 3.6**) have homology to Prxs-15 and mes pre-miRNA319 respectively. Peroxidases are Hydrogen Peroxide (H2O2) decomposing enzymes that play a vital role in a wide range of biological functions (Pandey *et al*, 2017). Peroxidase 15 (Prxs 15) belongs to the Class III peroxidases, which are involved specifically in growth and development including cell wall metabolism, auxin metabolism and well as in the defence against pathogens (Bindschedler *et al*, 2006; Cosio and Dunand, 2009; Pandey *et al*, 2017). Class III peroxidases play a role in defence through (1) the increased production of reactive oxygen species (ROS) which play a role in signal mediation and (2) the re-enforcement of cell wall polysaccharides (Bindschedler *et al*, 2006). A number of studies on the interaction between geminiviruses and their host have shown that resistance in linked to an increase in peroxidase levels in the plant (Li *et al*, 2017; Siddique *et al*, 2014). Lockhart and Semancik (1970) also noted

a decrease in Indole-3-acetic acid (IAA) in cowpea seedlings infected with Cowpea mosaic virus. Peroxidases are responsible for the oxidative carboxylation of IAA, which is found predominantly in the apical and meristematic regions of plants (Kramer and Ackelsberg, 2015). In Ma *et al* (2015) a number of IAA proteins were found to be down regulated during the transition from FFEC to OFEC. Changes in the expression Peroxidase levels could result in the IAA expression modification of FECs, which could be affecting the FECs ability to regenerate.

MicroRNAs originate from non-coding *MIR* genes, where the dsRNA originates from the single stranded transcript which forms an imperfect fold- back known as a primary miRNA (pri-miRNA), which is then processed further to form a shorter precursors miRNA (pre-miRNA) which is then processed by DICER 1 (DCL1) to release the miRNA/miRNA* duplex (Brodersen and Voinnet, 2006). The miRNA is released from the duplex and is then used by RISC to guide cleavage of mRNA (Narry, 2014). One of the more interesting findings in this study is the correlation between the increased levels of mes pre-miRNA 319 and transformation of FECs with AES and E-ACMV constructs. MicroRNA319 negatively regulates a number of transcription factors within the TCP(*TEOSINTE BRANCHED, CYCLOIDEA*, and *PROLIFERATING CELL NUCLEAR ANTIGEN BINDING FACTOR)* family (Carla *et al*, 2012; Rodriguez *et al*, 2016) which are responsible for the regulation of several biological pathways, including hormone biosynthesis, and cell proliferation and differentiation (Carla *et al*, 2012; Efroni *et al*, 2013; Schommer *et al*, 2008). The over expression of miRNA 319 can lead to leaf deformation including either crinkly (Efroni *et al*, 2013) or serrated leaves (Koyama *et al*, 2017) or result in delayed seed germination (Smith *et al*, 2017).The up-regulation of miR319 in pC-AES and pC-E-ACMV transformed cv.60444 FECs could account for the serration of the leaves (**Figure 3.2D**) observed in the AES lines, as well as contribute to the toxic effect in the FECs. MicroRNAs play a vital role in a number of biological processes, and as such their expression is tightly regulated. MicroRNAs are defined by their biogenesis, which usually involves the excision of the mature miRNA from the stem of the pre-miRNA by DCL-1. The microRNA 319 family is encoded by 6 genes and is widely conserved in plants (Palatnik *et al*, 2007). Unusually, the miRNA precursors have large fold-back sequences which contain three different regions which could potentially produce small RNAs (Talmor-Neiman *et al*, 2006), and are processed

through three sequential steps, beginning with cleavage of loop, instead of cleavage at the base (Bologna *et al*, 2009). This cleavage is vital for the proper processing of miRNA 319, and the removal of the terminal loop inhibits the processing of miRNA 319 (Bologna *et al*, 2009). Thereafter three phased miRNA duplexes are released, however only the mature miRNA 319 is released, while the alternate miRNA are expressed at low levels (Bologna and Voinnet, 2014). Interestingly, in a study conducted by Bologna *et al* (2009) the deep sequencing data associated with miRNA319 processing isolated a number of small RNAs derived from miR319a precursor region, in addition to miRNA319, which mapped to the cleavage regions of the precursor (Talmor-Neiman *et al*, 2006). Further investigation of these small RNA showed that miRNA 319 precursor has a number of internal bulges, close to the terminal loop, to prevent the accumulation of these small RNAs. Over expression of one of these sRNA (referred to as small RNA#1) in *Arabidopsis* mutants resulted in a crinkly leaf mutant. The siRNA produced by the AC4 region of ACMV was found to bind to a similar region of the miRNA 319c (105-122), before the terminal loop. If the siRNA is similar to one of the alternative splicing of mes miRNA 319, it is possible that its over-expression may be responsible for the phenotypic effects in cassava.

The mechanism by which geminivirus AC4-associated siRNA could regulate miRNA is not clear as the siRNA regulation of miRNA has never been reported in plants, although some studies have shown the expression of certain siRNAs in animal cells has resulted in the up-regulation of miRNAs (Khan *et al*, 2009; Liang *et al*, 2013). Expression is regulated at both a transcriptional and post-transcriptional level (Winter *et al*, 2009), where both the pri-miRNA and pre-miRNA are subject to regulation by proteins and long non-coding RNA (lncRNA) sequences (Lu and Cullen, 2004). A study on the variation in expression miRNA319 in algae and moss showed that the expression of miRNA319 was increased in siRNA deficient mutants (Li *et al*, 2011) which may point the evolution of a miRNA-specific post-transcriptional regulatory mechanism in these eukaryotic organisms. Furthermore, previous studies conducted on geminivirus infection have shown that AC4 inhibits PTGS, in order to establish infection and its over expression results in an increase in miRNA 319 levels (Naqvi *et al*, 2010). While this is all speculative, it does suggest that it might be possible that a siRNA could regulate miRNA expression. This would be very interesting to study further in future studies.

In addition to a siRNA which targets the GEM-like 4 protein (target 5), the AC1 region of EACMV also produces two siRNAs which have a high level of homology to FPS (target 4) and AMP deaminase-like protein (target 6). Farnesyl Diphosphate synathase is a key enzyme in isoprenoid synthesis (Zhang *et al*, 2015) where it catalyses the synthesis of Farnesyl diphosphate (FDP) (Closa *et al*, 2010). Isoprenoids form a large class of compounds which are important to a number of biological functions including photosynthesis and the regulation of normal growth and development (Manzano *et al*, 2016). *Arabidopsis thaliana* contains two genes which code for FPS namely *fps1* and *fps2,* which have been shown to have overlapping functions. Farnesyl Diphosphate synathase 1 is localised in the cytosol and mitochondria while FPS2 is only localised to the mitochondria. Knock-out studies in *Arabidopsis* showed that *fps1*/*fps2* double knockouts mutants it result in a lethal phenotype and arrest of early embryo development (Closa *et al*, 2010). Single knockouts did not produce this phenotype, but it was noted that FPS2 specifically plays an important role in embryo development (Closa *et al*, 2010). Further studies showed that in *Arabidopsis* plants where FPS was down-regulated after seed development, cotyledons and leaves showed chlorosis and altered chloroplasts (Manzano *et al*, 2016). Farnesyl Diphosphate synathase down-regulation also resulted in the mis-regulation of genes involved in the Jasmonic acid pathway and stress response (Manzano *et al*, 2016). A study conducted by Ma *et al* (2015) showed that a number of phytohormones including cytokinines and brassinosteroids are important for the formation of friable embryonic callus where they are involved in the regulation of cell elongation and division. As FECs transition from fresh FEC (FFEC) to old FEC (OFECs) the synthesis pathways are down-regulated and FECs lose their regenerative ability. It is possible that the down-regulation *fps*1 (in FECs transformed with pC-E-ACMV) and *fps*2 (in FECs transformed with all three constructs) could result in lower levels of isoprenoid synthesis, which could result in the transition of FFEC to OFEC reducing the number of FECs capable of regenerating, or down-regulation of *fps1* and/or *fps 2* could cause a lethal phenotype similar to that shown in *Arabadopsisfps1/fps* double mutants.

AMP deaminase-like protein is responsible for the irreversible hydrolytic deamination of AMP which produces IMP and NH_3. This reaction drives an increase in the production of ATP by altering the adenylate pool and increasing adenylate kinase

activity. In a study conducted in *Arabidopsis,* EMBRYONIC FACTOR 1 (*FAC1)* was shown to be important for the embryo development, with null mutants having a lethal phenotype. *FAC*1 was shown to code for a putative AMP DEAMINASE-LIKE PROTEIN , and was expressed in high levels in 7-day-old seedlings as well as during somatic embryo development (Xu *et al*, 2005). It is possible that the down-regulation of AMP DEAMINASE-LIKE PROTEIN could have a similar lethal effect on development of cassava. The predicted siRNA is predicted to be produced by the IR/AC1 of EACMV therefore all FECs should show down-regulation of AMP DEAMINASE-LIKE PROTEIN. While there was a decrease in expression in all three FEC experiments, only pC-E-ACMV showed significant down regulation. Praveen *et al* (2010) showed that small constructs produce more significant siRNA effects, due a variation in number of specific siRNAs produced by each construct (**Figure 3.6**), which are influenced by the size of the construct and the position in relation to the intron (Praveen *et al*, 2010), but this would have to studied further to confirm this result. The threat of CMD to crop security has been well documented and a large number of studies, developing CMD resistant varieties through breeding and transgenics have been conducted (Chellappan *et al*, 2004a; Lapidot and Friedmann, 2002; Ntui *et al*, 2015; Vanderschuren *et al*, 2007a) with various levels of success. More recently the expansion of genomics has helped to elucidating some of the key factors involved in virus-host interactions (Allie *et al*, 2014; Kuria *et al*, 2017; Pierce and Rey, 2013), which is being used to design more efficient targets to improve virus resistance in cassava. This study highlights the importance of the understanding the interaction between cassava and CMG vsiRNA, specifically as it impacts on the feasibility of stacking multiple viral targets constructs for virus resistance. Each of each of the regions targeted by the AES construct have been tested individually in hp-RNA constructs, and none were reported to cause developmental defects. Therefore it is likely that the constructs' toxicity is due to the amalgamated consequences of each of the 'off-targets' as opposed to the effect of an individual siRNA. This is similar to the synergism effect which occurs when multiple geminiviruses infect a single plant, resulting in increased symptom severity (Masuta and Shimura, 2013).We hypothesis that the expression of the triple construct transgene produces siRNAs from the IR region in combination with the AC1/AC4, which mimic vsiRNAs which are able to interact with multiple cassava transcripts which results in developmental abnormalities in the plants. A lack of understanding

of how host and virus derived-siRNA interact hampers the ability to effectively combat viral disease employing RNA silencing, and select efficient viral targets that minimally affect the plant.

Chapter 4: Increase in Starch Production through down regulation of starch-related genes in cassava

4.1 Introduction:

Cassava (*Manihot esculenta*) is perennial shrub, grown in tropical and sub-tropical regions around the world for its tuberous roots which are high in starch (El-Sharkawy, 2004). Cassava is of particular importance in sub-Saharan Africa where it is a source of calories for over 250 million people (Leyva-Guerrero *et al*, 2012). In addition to its role as a calorie source, starch is used in a number of industrial processes and as cassava has one of the highest yields of starch (Nassar and Ortiz, 2006), it has been the focus of a number of studies to increase the quality and the quantity of the starch (Lloyd and Kossmann, 2019) (Zhang *et al*, 2017).

Starch is stored as osmotically inactive, water insoluble granules in chloroplasts (transitory) and storage roots known as storage roots. It is stored in specialised plastids known as amyloplasts in both the roots and leaves (Slattery *et al*, 2000). The starch content of storage roots is variable depending on the plants species but is usually between 73-84% (Hoover, 2001). Starch is fractioned into two types of polymers; amylose and amylopectin which vary based on their linear chains. Amylopectin is the predominate polymer found in starch (usually 75% of starch granule) (Zeeman *et al*, 2010) and is made up of chains of glucosyl residues (of between 6-100 residues) linked by α-1-4-bonds which are linked together by α-1-6 bonds (Smith, 2008). Amylose makes up about 30-35% of the starch granule, is primarily linear, and is made up of α-1-4 bond linked chains (Slattery *et al*, 2000; Zeeman *et al*, 2010).

The biosynthesis of starch takes place exclusively in the plastids of plants (**Figure 1.5**). Sucrose enters the cytosol where it is converted to UDP-glucose by sucrose synthase, which is then converted to glucose-6-phosphate (G-6-P) which is transported into the plastid using a glucose-6-phophate transporter (Kammerer *et al*, 1998). Once in the plastid, it is reversibly transformed into Glucose-1-phosphate (using enzyme phosphoglucomutase) which is then converted to ADP-glucose using

ADP-glucose pyrophosphorylase (AGPase) (Tjaden *et al*, 1998). This process requires ATP, which is transported across the plastid membrane using a Adenylate kinase transporter (Tjaden *et al*, 1998). The ADP-glucose is then converted to starch by starch synthase (SS) enzymes, mainly granule bound starch synthase (GBSS),which converts ADP-glucose into amylose and Branching Enzyme, and Soluble starch synthase (SSS), which are mainly responsible for the synthesis of amylopectin (Smith, 2008).

Stark *et al* (1992) showed that the over expression of *Escherichia coli Glg*C gene which codes for a AGPase, in potatoes produced up to 30% more starch. After a study conducted by Munyikwa *et al* (1998) showed that the down-regulation of AGPase in cassava resulted in decrease in starch in cassava storage roots, AGPase was focussed on as means of regulating starch production in cassava. In a study conducted by Ihemere *et al* (2006) cassava was transformed with a modified *glgC* gene from *E. coli,* which coded for AGPase resistant to feedback inhibition from fructose-1,6-biphosphate. These plants showed a 70% increase AGPase activity and a 2.6 increase in over-all biomass. Building on this study, Ligaba-Osena *et al* (2018) transformed cassava with a modified AGPase, in conjunction with α-amylase and amylopullulanase from *Pyrococcus furiosus* and glucoamylase from *Sulfolobus solfataricus* which improved both the yield (>60%) and digestibility of the starch.

AGPase is a vital enzyme for starch synthesis (Lloyd and Kossmann, 2019) and increasing its activity has been shown to lead to an increase in starch synthesis (Tiessen et al., 2002). AGPase is highly regulated through both allosteric effectors and by redox, and as such over-expression of the native enzyme does not always lead to an increase in starch. Several studies where modified AGPase was over-expressed have shown an increase in starch (Ligaba-Osena et al., 2018; T. R. I. Munyikwa et al., 1998; Stark et al., 1992). One of the other mechanisms shown to increase AGPase activity is to increase the supply of either Adenosine triphosphate (ATP) or/and glucose-6-phosphate to the plastidial AGPase (Geigenberger et al., 2001; Trethewey et al., 1998). This can be achieved either through over-expression of adenylate translocator (Tjaden et al., 1998) or through the regulation of the plastidial isoform of adenylate kinase (Geigenberger *et al*, 2001; Loef *et al*, 2001; Regierer *et al*, 2002; Stark *et al*, 1992; Tjaden *et al*, 1998). Tjaden *et al* (1992) showed he over-expression of the ADP/ATP translocator (Tjaden *et al*, 1998) results

in an alteration in the adenylate pools, and resulted in increased starch in potatoes (Loef *et al*, 2001). Adenylate kinase is responsible for the inter-conversion of ATP and AMP into ADP and is responsible for the maintenance of equilibrium of adenylate pools (Noda, 1973). Regeirer *et al* (2002) showed when ADK was down-regulated in potatoes, there was a increase in starch of up to 60% in some lines, as well as increase in tuber yields. This was shown to be due to the post-translational redox-activation of ADP-glucose pyrophosphorylase (Oliver *et al*, 2008)..

Although the majority of studies aimed at increasing the starch content in storage roots have focussed on the expression and manipulation of enzymes directly involved in starch synthesis, some studies have shown that carbon fluxes are regulated by nucleotide cofactors. Geigenberger *et al* (1993) showed that sucrose synthase mediated degradation of sucrose is dependent on the available uridine pool and that the starch levels in potatoes could be increased through the manipulation of the de novo pyrimidine synthesis pathway (Geigenberger *et al*, 2005). Specifically, the down-regulation of *Uridine-5'-monophosphate synthase* (*UMP synthase*) expression was shown to force potatoes into the pyrimidine salvage pathway, which increased the levels or uridine available in the storage roots and increased starch production (Geigenberger *et al*, 2005).

Cassava cultivation has a number of advantages over other starch crops; it can be grown under sub-optimal conditions with minimum input, and is seen as climate-change resistant crop (De Souza *et al*, 2016). Traditional breeding has been used to create new cultivars with improved yield, nutrition and starch content (Ceballos *et al*, 2004), however this method is laborious and difficult. More recently it has been shown that cassava can be modified using a biotechnological approach, using *Agrobacterium*-mediated transformation of cassava FECs in conjunction with RNA silencing. The manipulation of cassava starch has focussed on two main areas, improvement of starch quality, where the amylose and amylopectin ratios have been manipulated, and on starch quantity (Bull *et al*, 2018b; Ihemere *et al*, 2006; Ayalew Ligaba-Osena *et al*, 2018; Munyikwa *et al*, 1998). A number of studies have also focussed on the improvement of waxy starch content, controlled by GBSS enzymes (Bull *et al*, 2018b; Raemakers *et al*, 2005; Zhao *et al*, 2011) where the down-regulation of GBSS resulted in lower levels or amylose free starch.

RNA silencing is a process found in all eukaryotes, where dsRNA triggers the production of small RNA which bind to gene targets, leading to suppression of gene expression (Pérez-quintero and López, 2010). RNA silencing can be triggered by RNA by both sense and anti-sense RNA with high level of homology to target gene, however the most efficient response is triggered by hp--RNA molecule under the control of a promoter (Waterhouse and Helliwell, 2003). During this process, the dsRNA is processed by a Dicer-like protein, which cleaves the dsRNA into small interfering RNA (21-24nt) which is then bound to Argonaut protein and other cellular proteins to form a RNA-induced silencing complex. The RISC uses the siRNA as a guide, binding to and cleaving mRNA with a high level of homology to the siRNA guide strand. This process, known a post-transcriptional gene silencing has been used both to study and enhance biological traits in plants (Angaji *et al*, 2010).

This study reports on the transformation of cassava cv.60444 with hp--RNA constructs targeting the cassava plastidial *adenylate kinase* (ADK) and *uridine-5'-monophosphate synthase* (UMP synthase), in order to increase the starch content of the storage roots. Using *Agrobacterium*-mediated transformation of cassava cv.60444 FECs. Ten UMPS lines and eight ADK lines were produced and micro-propagated in order to further characterise the effect of down-regulation of plastidial ADK and UMPS in cassava storage root yield. Three UMPS lines showed down-regulation of *UMP synthase,* which correlated to increase in storage root yields. Two ADK lines also showed differential expression of *ADK* in comparison to the control cv.60444 lines, but this was not shown to be linked to an increase in storage root yields.

4.2 Materials and Methods

4.2.1 Construction of pCB33

4.2.1.1 Isolation of Patatin promoter (B33) from pCambia2200

The B33 promoter cassette (1987bp) was isolated from pCambia 2200-B33 (J. Lloyd, unpublished data), using Fast Digest EcoRI and HindIII (Thermoscientific) restriction enzymes according to manufacturer's instructions. Five hundred nanograms of pCambia2200-B33 vector was added to 1X FastDigest Green restriction buffer, 1U of EcoRI and 1U of HindIII in total volume of 20µl. The mixture

was incubated for at 37°C for 1h after which the reaction was deactivated by heating to 80°C for 5min. The digested products were then separated using gel electrophoresis on a 1% agarose gel, and the B33 promoter (1987bp) was isolated using GeneJet gel purification kit (Thermoscientific) according to the manufacturer's instructions.

4.2.1.2 Ligation of B33 promoter into pCambia 1305.1

Plant transformation vector pCambia 1305.1 was digested with FastDigest EcoRI and HindIII restriction enzymes (Thermoscientific) according to the manufacturer's instructions (**see 3.2.1.2**)**.** The B33 promoter was ligated into linearised pCambia 1305.1 using T4 DNA ligase kit (Thermoscientific) according to the manufactures' instructions to form pCB33. Twenty nanograms of linearised pCambia 1305.1 and 100ng of B33 linearised DNA was added 1x T4 ligase buffer, 1U of T4 ligase in a total volume of 20µl. The mixture was then incubated at 22°C for 10min.

The ligated pCB33 (**Figure 4.1C)** constructs were transformed into chemically competent *E. coli* DH5α. Five microlitres of ligation mixture was added to 50µl of competent cells and incubated on ice for 10min. The cells were then heat shocked at 37°C for 45s and then placed on ice for 2min. Transformation control of untreated pCambia 1305.1 was also included. The cells were then spread onto LB agar plates containing 50µg/ml of kanamycin (Sigma) and incubated overnight at 37°C. Ten single colonies from the overnight plates were isolated and incubated overnight in LB broth containing 50µg/ml kanamycin (Sigma).

4.2.1.3 Screening of pCB33 clones

The pCB33 plasmid was isolated using the GeneJet plasmid extraction kit (Thermoscientific) according to the manufacturer's instructions. Clones were then screened for the presence of insert using restriction digestion with FastDigest EcoRI and HindIII according to the manufacturer's instructions (**see 4.2.1.2**). Results were then analysed using gel electrophoresis on a 1% agarose gel.

4.2.2 Construction of pCB-ADK and pCB-UMPS transformation vectors

4.2.2.1 Identification of ADK and UMP synthase:

The amino acid sequence of Plastidial *ADK* (XP_006367264) and *UMP synthase* (XP_006362855) for *Solanum tuberosum* (potato) were obtained from the NCBI database. Using BLASTP (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins), sequences were then compared to cassava genome to identify possible *UMP synthase* and *ADK* homolog targets.

4.2.2.2 Construction of pCB-ADK and pCB-UMPS

DNA hairpins complementary to the conserved sequence of the two *ADK* and one *UMP synthase* target sequences in the sense and anti-sense orientation separated by an 81bp intron (Vanderschuren et al., 2009) were synthesised (Life technology) (**Figure 4.1A and B**). Each sequence included KpnI restriction cut site at the 5' and 3' end for easy insertion into the pCB33 vector. The synthesised *ADK* (831bp) and *UMP synthase* (1087bp) hp-DNA was inserted into pCB33 through sticky end cloning with restriction enzyme KpnI. Both hp-DNA constructs, as well as pCB33 were digested with FastDigest KpnI restriction enzyme (Thermoscientific) according to the manufacturer's instructions (**see 3.2.1.2**). The ADK and UMPS hp-RNA cassettes were ligated into linearised pCB33 using T4 DNA ligase kit (Thermoscientific) according to the manufactures' instructions. The pCB-ADK and pCB-UMPS constructs were transformed using heat-shock into chemically competent *E. coli* DH5α (**see 3.2.1.2**). The cells were then spread onto LB agar plates containing 50µg/ml of kanamycin (SigmaAldrich) and incubated overnight at 37°C. Colonies from the overnight plates were isolated screened by colony PCR.

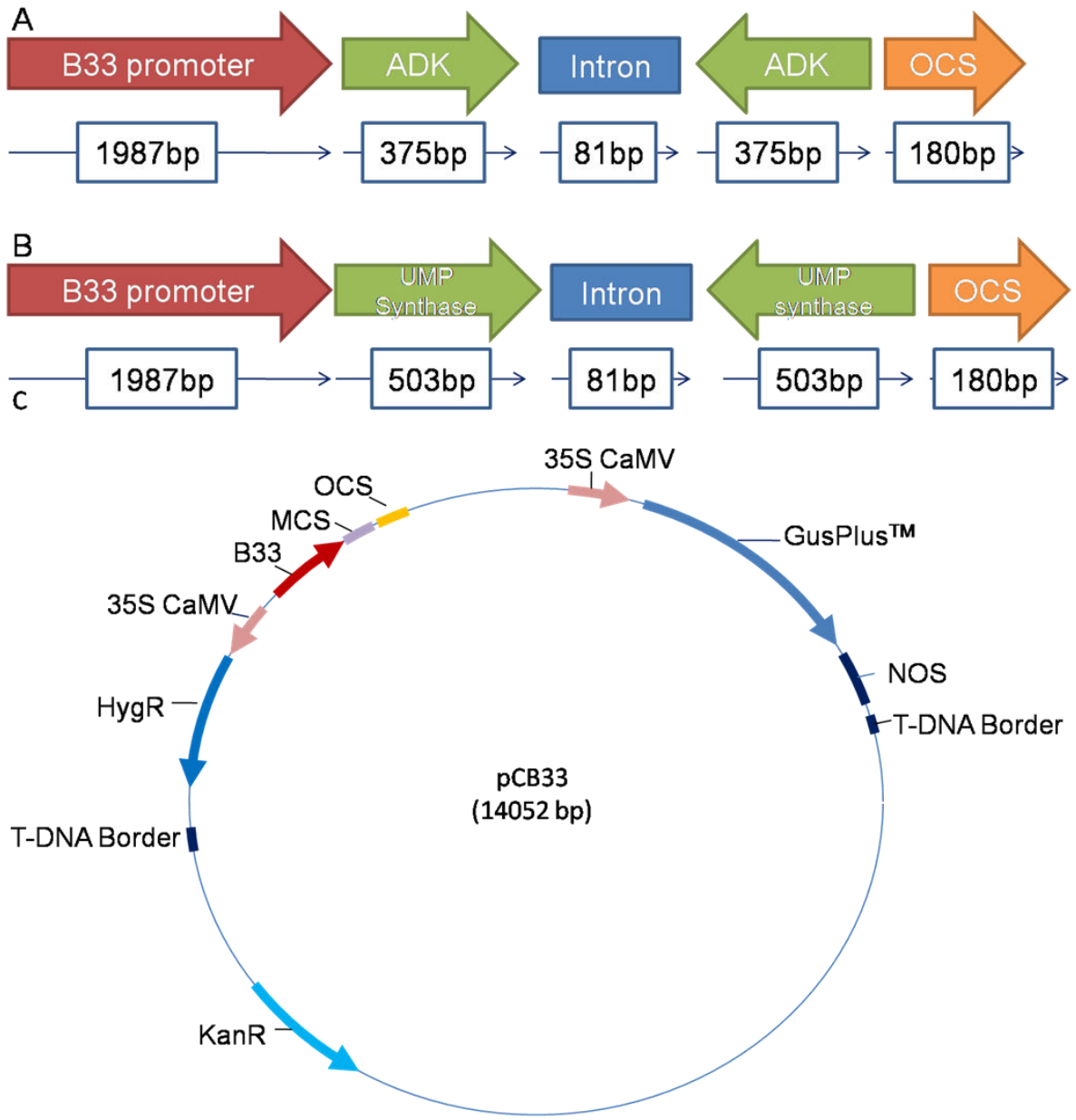


Figure 4.1: Diagrammatic representation of the synthesised Adenylate kinase (ADK) (831bp) (A) and (B) Uridine monophosphate synthase (UMPS) (1087bp) hairpins which were inserted into pCB33 at the MCS site using sticky end cloning . Each hairpin is flanked by the patatin root promoter (B33) and a octopine synthase transcription terminator (OCS). The forward and reverse arms of the hairpins are separated by a 81bp intron (I).

4.2.2.3 Screening of pCB-ADK and pCB-UMPS clones

Single colonies were selected and transferred to 10µl of LB liquid media (containing 50µg/ml Kanamycin) and incubated for 1 hour at 37°C. One microlitre of culture was used as the substrate in colony PCR reaction which contained the following reagents (per reaction); DreamTaq Green buffer (1X) (Thermoscientific), 1U DreamTaq DNA polymerase (Thermoscientific), 0.4mM dNTPS (Thermoscientific), 1mM forward and reverse primer, 4mM MgCl2. ADK qPCR forward (5' CCATGGGCAATTAGTTTCTGATG'3) and reverse (5' TTCCGCTTGTCTGATAGTTCG'3) were used to screen ADK colonies and UMPS qPCR forward (5'TTGCCACTGGAGATTACACAG'3) and reverse (5' GGGTTGCCTGAATAAATGCTG'3) primers were used to screen the pCB-UMPS colonies. The cycling conditions were: 95°C for 5min; followed by 35 cycles of 95°C for 30s, 60°C for 30s; and 72°C for 45s followed by a final extension of 72°C for 5 minutes. Five µl of each reaction was subjected to electrophoresis on a 1% agarose gel. The direction of the insert was confirmed using PCR with SC Forward (5' TCTCAACTTGTTTACGTGCCT'3) and SC Reverse (5' CGCTCGTACCTGCATGATATC'3) (Tm=57°C) with DreamTaq according to the manufacturer's instructions. The integrity of the confirmed pCB-ADK and pCB-UMPS clones was confirmed using by Sanger sequencing by Inqaba Biotechnology.

4.2.3 *Agrobacterium*-mediated transformation of cassava cv.60444 FECs with pCB-ADK and pCB-UMPS

4.2.3.1 Production of cassava cv.60444 FECs

Friable embryogenic callus (FECs) were produced according to Bull *et al* (2009). Cassava cv.60444 was maintained in a controlled growth chamber at 28°C with 16h/8h light dark cycle. Nodal cutting were taken from 6 week old plantlets and cultured on CAM media (4.46 g/L MS salt with vitamins (Sigma-Aldrich), 20g/L sucrose, 2 mM CuSO4, 6-Benzylaminopurine (BAP) (1 mg/ml) (Sigma-Aldrich)). The swollen auxiliary buds were removed from the stems and continuously sub-cultured on CIM (4.46 g/L MS salt with vitamins (Sigma-Aldrich), 20g/L sucrose, 2 mM CuSO4, picloram (12 mg/ml) (Sigma-Aldrich)) until homogenous finger-like structures appeared. The clusters were then transferred to GD media (2.7g/L GD

with vitamins (Duchefa Biochemie), 20g/L sucrose, and 12mg/ml picloram) and maintained until they were transformed.

4.2.3.2 Transformation of *Agrobacterium tumefaciens* LBA4404 with pCB-ADK and pCB-UMPS

Purified pCB-ADK and pCB-UMPS transformation vectors were used to transform *A. tumefaciens* LBA4404 using the freeze-thaw method (Holsters *et al*, 1978). *A. tumefaciens* LBA4404 (chemically competent) cells stored at -70°C were placed on ice to thaw. Once thawed, 250ng of pCB-ADK, pCB-UMPS constructs and pCambia 1305.1 vector construct were added to 100µl of the competent cells. The cells were snap frozen in liquid nitrogen for 5min followed by heat shock at 37°C for 5min after which, the cells were incubated in 1ml Yeast extract-phosphate (YEP) at 28 °C for 2-4 h with gentle shaking (50 rpm). 100µl of transformed cells were then spread on YEP plates containing 50µg/ml rifampicin, 50µg/ml kanamycin and 100µg/ml streptomycin and incubated in the dark at 28 °C for 1-2 days until colonies appeared.

4.2.3.3 Screening of *A. tumefaciens* LBA4404 transformants for presence of pCB-ADK and pCB-UMPS

A colony was selected from each of the pCB-ADK and pCB-UMPS *A. tumefaciens* LBA4404 transformation plates and inoculated into YEP broth with appropriate anti-biotics (50µg/ml rifampicin, 50µg/ml kanamycin and 100µg/ml streptomycin) and incubated overnight at 28°C under constant agitation (200rpm). The transformation vector was extracted from the overnights cultures and screened for the presence of the pCB-ADK and pCB-UMPS by restriction digestion with FastDigest EcoRI and HindIII (Thermoscientific) according to the manufacturer's instructions. Purified pCB-ADK and pCB-UMPS transformation vector were included as controls.

4.2.3.4 Preparation of Agrobacterium inoculum

pCB-ADK and pCB-UMPS constructs, transformed into A. *tumefaciens* LBA4044, and *A. tumefaciens* harbouring empty pCambia 1305.1 plant transformation vector, were prepared according to Bull *et al* (2009). *A. tumefaciens,* harbouring pCB-ADK, pCB-UMPS or pCambia 1305.1 as well as untransformed *A. tumefaciens* LBA4404 was streaked onto YEP plates containing appropriate antibiotics (50 µg/ml kanamycin (except for untransformed *A. tumefaciens* LBA4404), 50 µg/ml rifampicin

and 100 µg/ml streptomycin) and plates were incubated for 48h, in the dark at 28°C. A single colony from each plate was selected and used to inoculate 5ml of YEP liquid media, supplemented with appropriate antibiotics and incubated at 28 °C under constant agitation (200rpm) until the optical density (OD) measured between 0.7-1.0, at λ=600nm. Five hundred microliters of each culture were used to inoculate 25 ml of YEP which was incubated as before until the OD_{600} ≈0.7-1.0. The bacteria was then pelleted by centrifugation (4000g for 10min, at room temperature) and resuspended in liquid GDS medium (supplemented with Acetosyringone 200µM) at final OD_{600} = 0.5. Bacterial suspensions were then incubated at room temperature for 45 min on a horizontal shaker (50rpm).

4.2.3.5 Transformation of cassava cv.60444 FECs

Cassava cv.60444 FECs (GD15) were transformed according to Bull *et al* (2009). Seven plates with approximately 10 FEC clusters (approximately 0.8- 1cm in diameter) per plate were used for the transformation of pCB-ADK and pCB-UMPS, 2 plates were used for the transformation of pCambia 1305.1 and 1 plate was transformed with untransformed *A. tumefaciens* LBA4404. All FECs were co-cultivated with the bacterial suspensions for 4 days at 28 °C in 16 h light/8h dark photoperiod. After 4 days the FECs were washed with GD liquid media containing carbenicillin (500µg/ml). FECs were then transferred to nylon mesh (100µm pore size) and placed on GD media supplemented with 250µg/ml carbenicillin (to allow for the recovery of the FECs) and incubated at 28°C for 5 days in 16 h light/8h dark photoperiod. In order to increase the recovery of transformants, FECs were then transferred to GD containing increasing amounts of Hygromycin. FECs (on mesh) were transferred to GD containing 250µg/ml carbenicillin and 5µg/ml hygromycin (for 1 week), followed by 1 week on GD containing 250µg/ml and 8µg/ml hygromycin, followed by 1 week on GD containing 250ug/ml and 15µg/ml hygromycin. All plates were incubated at 28 °C with 16 h light/8 h dark photoperiod. Untransformed cv.60444 FECs were not cultured with hygromycin.

In order to facilitate the regeneration of the transformed FECs, the FECs (on mesh) were transferred to MSN medium containing 250µg/ml carbenicillin, 1µg/ml 1-Naphthaleneacetic acid (NAA) and 15 µg/ml Hygromycin, and 8 g/L Noble agar and maintained for 10 days at 28 °C with 16 h light/8 h dark photoperiod. The FECs were

transferred to fresh MSN media every 10 days for a maximum for 6 cycles. Emerging green cotyledons were transferred to cassava elongation media (CEM) (MS2 supplemented with 2μM $CuSO_4$, 0.4μg/ml BAP and 100 μg/ml carbenicillin). The cotyledons were maintained on CEM until shoots and leaves appeared. Shoots were removed and placed in sterile small bottles on solid cassava basic medium (CBM) (MS2 with 2 μM CuSO4; 50 μg/ml carbenicillin and solidified 3 mg/ml phytagel). Plates were incubated for 28 °C with 16 h light/8h dark photoperiod

4.2.4 Screening of regenerated plantlets for pCB-ADK and pCB-UMPS constructs

4.2.4.1 Gus assay

Gus assays were conducted according to Bull *et al* (2009). Two leaflets were selected from each regenerated line and submerged in GUS assay solution (100 mM Tris/ NaCl buffer, 1 mg/ml 5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid (X-Gluc) and 1 % Triton X-100). Untransformed cv.60444 was included as a negative control. All samples were incubated at 37 °C for 12hr in the dark. The leaves were then removed from GUS assay solution and destained by soaking in 70% ethanol.

4.2.4.2 Rooting test

In order to screen regenerated plantlets for presence and expression of the *Hyg* gene, a rooting test was conducted using CBM media supplemented with Hygromycin (10μg/ml) and carbenicillin (50μg/ml). An apical shoot from each line was isolated and placed in sterile tissue culture containing the rooting media. Untransformed, regenerated cv.60444 was included as a negative control.

4.2.5 Molecular screening of transgenic plantlets

4.2.5.1 PCR screening

A total DNA extraction was performed on 50 mg of leaf collected from each sample, flash frozen in liquid nitrogen and crushed into a fine powder. DNA was extracted from all plantlets using the CTAB method described in Doyle (1991). The plantlets were screened for the presence of the reporter genes *Gus* and *Hyg*, which are found in T-border region of pC1305.1. To amplify the *Gus*, GUSPlus F and GUSPlus R (**Table 2.1**) primers (producing a 181bp fragment) were used. To amplify the *Hyg*,

Hyg F and Hyg R (**Table 2.1)** primers were used, producing a 444bp amplicon. The reaction mixture contained 1X Green DreamTaq buffer, 2mM MgCl2, 0.2mM dNTPs, 0.2µM of each primer, 1U DreamTaq (Thermoscientific), 50ng of template DNA and nuclease-free water to a final volume of 20µl. Reactions were cycled in a thermal cycler (Bio-Rad) set at 94 °C for 2min and 35 cycles 94°C for 30s, specified annealing temperature for 30s, primer extension at 72°C for 30s and final extension step of 72°C for 10min. Purified pC-1305.1 plasmid (positive control) and untransformed cv.60444 DNA (negative control) was used in each PCR.

4.2.5.2 Southern Blot Hybridisation

The number of transgene integration events was determined using Southern blot hybridisation using DIG-High Prime DNA Labelling and Detection Starter Kit II (Thermoscientific) according to the manufacturer's instructions. Total DNA was extracted from 50mg of leaf material using CTAB method (Doyle, 1991)**.** Twenty micrograms (quantity determined using nanodrop (NanoOne)(Thermoscientific) of RNAse treated DNA from each line was digested overnight with HindIII restriction enzyme (Thermoscientific) and separated using gel electrophoresis, in a 1 % agarose gel in 1X TAE. The DNA was then transferred to positively charged nylon Hybond-N+ membrane (Amersham),pre-hybridised at 38°C (3h), and hybridised at 38°C overnight with a DIG-labelled *Hyg* gene probe, which was labelled using the DIG-High Prime DNA Labelling and Detection Starter Kit II (Roche) using DIG Easy Hyb buffer (Thermoscientific). Unbound probe was removed by sequential washing with Washing buffers 2xSDS, 0.1% SDS twice, and 0.1xSDS, 0.1% SDS buffer twice (incubated at 40°C) and signal detection (CDP star) was performed following DIG-High Prime DNA Labelling, and the Detection Starter Kit II (Roche) protocol. Results were visualised using GelDoc XR+ (Biorad) after 10 min.

4.2.6 Characterisation of storage roots in cassava cv.60444 transformed with pCB-ADK and pCB-UMPS

4.2.6.1 Micro-propagation

In order to test the starch content of the roots, the selected plant lines were bulked up through nodal culture. Thirty nodal cuttings, from each of the starch lines, as well as cassava cv.60444 were taken from 2 month old regenerated plants and cultured

on MS2 medium (4.31 g/L of MS medium with vitamins (Sigma); 20g/L of sucrose; 20mM $CuSO_4$ (pH 5.8), 8g/L Noble Agar) and incubated at 28°C under 16 h light for 4 weeks. Rooted plantlets were then transferred to Jiffies (Jiffy®) and maintained in a sealed environment until plants reached 4 leaf stage. Plants (7 from each line) were then acclimatised over 1 week, after which they were transferred to pots ($8cm^2$) and grown at 28°C, with 50% humidity for 16h light, 8h dark photoperiod. After 3 months plants were transferred to large circular pots (15cm) and allowed to grow for a further 4 months, after which plants were moved back into smaller pots to force storage root growth. The plants were fertilised once a month and were monitored for appearance of storage roots. After 10 months, 6 plants from each line were removed from pots, and storage roots were collected for further characterisation

4.2.6.2 Expression of *ADK* and *UMPS*

4.2.6.2.1 Total RNA extraction

Total RNA was extracted from storage roots using method according to Kumar *et al* (2007). Storage roots were collected from six independent plants and ground (mortar and pestle) to a fine powder in liquid Nitrogen. Total RNA then extracted from the pooled samples (2g per sample). Four millilitres of Lysis buffer (1.75 ml of borate-Tris buffer ((0.2 M, pH 8.0), 10 mM EDTA), 0.5 ml of 10% (w/v) SDS, 1.0 ml of 5 M NaCl, 1.65 ml of 1.95% (w/v) Na2SO3, and 0.1 ml of β-mercaptoethanol) was added to each sample and sample was incubated at 65 °C for 5 min after which it was centrifuged for 5 min (1800g). Equal volumes of Tris-saturated phenol (pH 7.9) was added to the supernatant and mixture was centrifuged (1800g) until phase separation was observed. The upper phase was extracted and treated with an equal volume of chloroform-isoamyl alcohol (24:1) and centrifuged for 15 min at (1800g). One millilitre of the upper phase was transferred to a 2 ml microfuge tube containing 0.9 ml of 2-propanol and mixture was incubated for 1 h at 4°C. The RNA pellet was then extracted by centrifugation (20 000g for 15 min at 4 °C), washed with 70% EtOH and solubilised in DEPC treated water.

4.2.6.2.2 Real-time RT-PCR

Relative real-time quantitative PCR was used to determine the expression of *UMP synthase* and *ADK* in relation to reference gene α-tubulin. cDNA was synthesised from 10µg of total RNA (treated with DNase I) with random hexamer primers, using

the RevertAid First strand cDNA synthesis kit (Thermoscientific). Two microliters of cDNA product was used as the template for RT-qPCR amplification of either *ADK* or *UMP synthase* using qADK or qUMPS forward and reverse primers (**see 3.2.3 above**) or β-tubulin F (3' CAAGTGCGATCCTCGACATG'5) and R (3'GATACCGCACTTGAACCCAG'5) (Ligaba-Osena *et al*, 2018). Real-time PCR was performed according to the Maxima SYBR green (Thermoscientific) protocol with three technical replicates per group. The PCR conditions were 5 min at 95 °C followed by 30 cycles of 95 °C for 15 s, 60 °C for 30 s and 72 °C for 1 min. Untransformed cassava cv.60444 was included as a negative control.

.

4.2.6.3 Wet and Dry mass

The moisture content of the root samples was determined using percentage weight. One storage root from each plant (6 plants in total) was weighed and then heated to 100 °C for 24 h and weighed again (dry weight (dw)). The moisture content (of the 6 storage roots) was determined as a percentage of the original weight.

4.2.7 Statistical significance

In order to determine whether the relative change in expression of *ADK* and *UMP* synthase were statistically significant, an student T-test was performed using Microsoft Excel. A Pearson correlation test was performed to determine if there was a correlation between the down regulation of *ADK* or UMP *synthase* and the dry weight percentage.

4.3 Results:

4.3.1 Construction of pCB33 vector

The B33 promoter cassette (1987bp) was isolated from pCambia2200 (**Figure 4.2A**) and inserted into pCambia 1305.1 using directed cloning. The vector was isolated from single colonies and were screened for the presence of B33 cassette by restriction digestion with EcoRI and HindIII (**Figure 4.2B**). Two of the selected colonies contained the B33 insert (lane 4 and lane 7). These clones (referred to as pCB33) were used in the construction of pCB-ADK and pCB-UMPS.

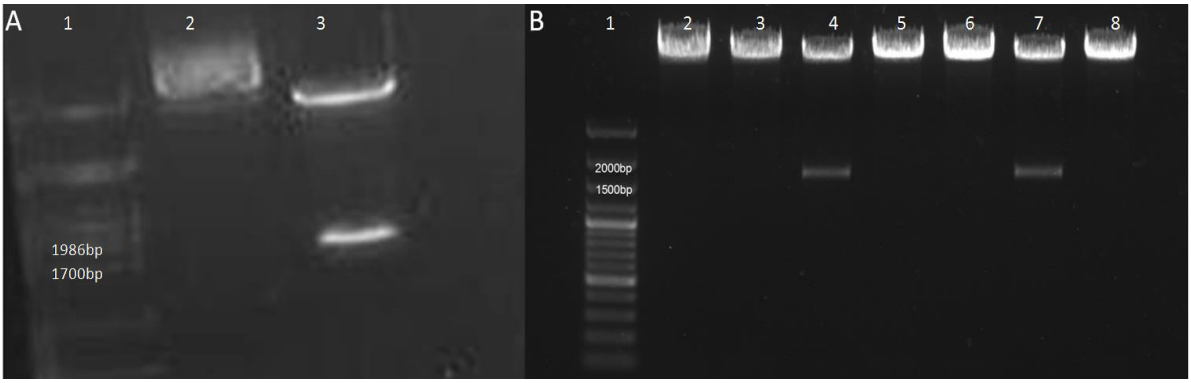


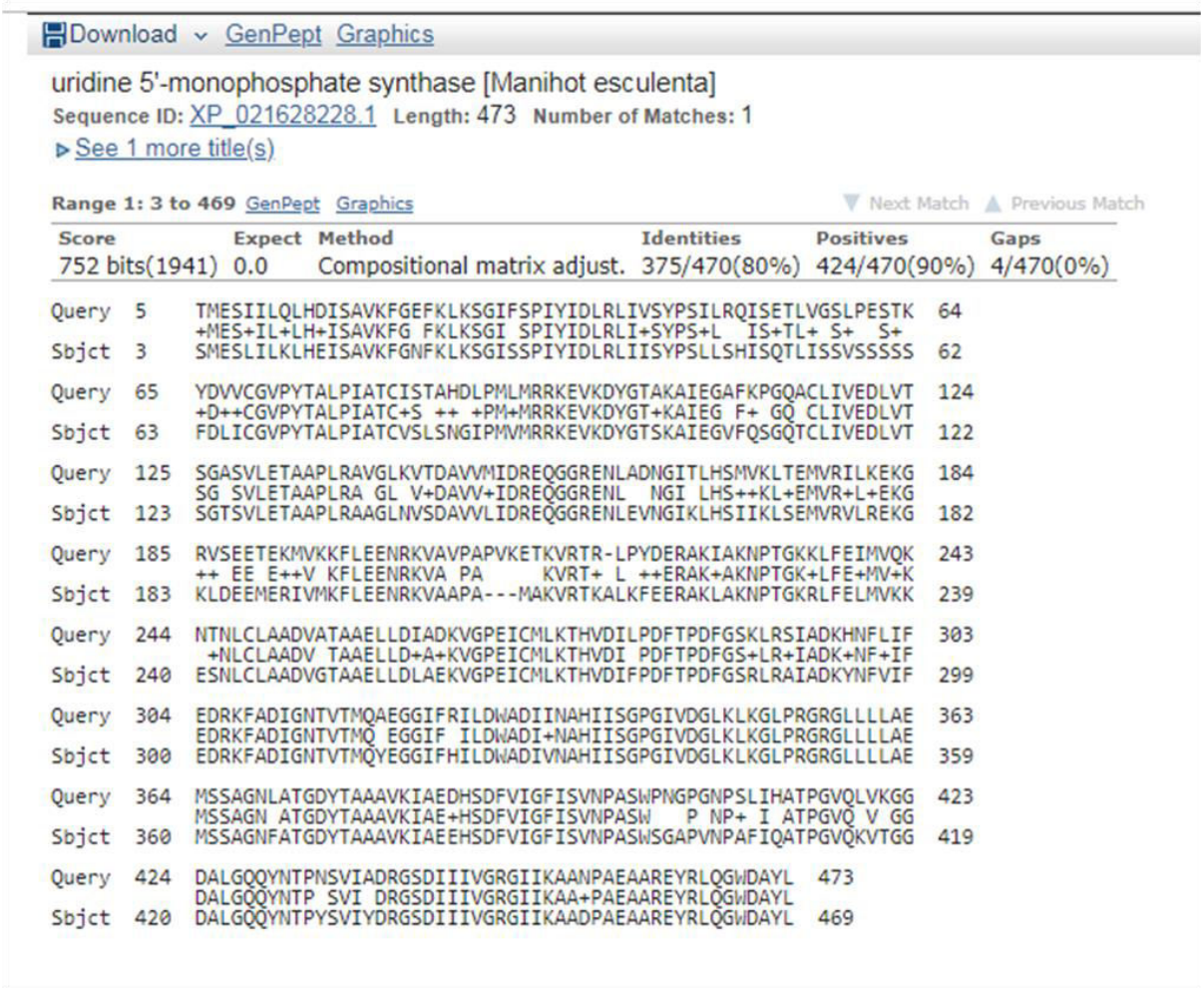
Figure 4.2: Insertion of B33 patatin promoter into pCambia 1305.1. A) Restriction digestion of pC2200-B33 with FastDigest EcoRI and HindIII restriction enzymes. Lane 1: Molecular marker Phage Lambda digest with PstI, lane 2: undigested pC2200-B33, lane 3: pC2200-B33 digested with EcoRI and HindII. B) Restriction digest of possible pCB33 clones with EcoRI and HindIII. Lane 1: Molecuar maker- Gene Ruler 100b plus (Thermoscientific), lane 2: pCambia 1305.1, lane 3, 5, 6, 8: negative clones (pCambia 1305.1 without B33 insert), lanes 4, 7: positive clones (pCambia 1305.1 containing B33 insert).

4.3.2 Construction of pCB-ADK and pCB-UMPS

4.3.2.1 Identification of *Adenylate kinase* and *Uridine-5-monophosphate synthase* in cassava and design of ADK and UMPS hairpins

Plastidial *ADK* and *UMP synthase* amino acid sequences for *S. tuberosum* were obtained from NCBI database and compared to cassava genome to identify possible UMP synthase and ADK homolog targets. One possible UMP synthase homolog was identified on chromosome 11 (cassava4.1_006843m.g) which had 80% similarity to UMP Synthase from *S. tuberosum* (**Figure 4.3**). Functional annotation of this transcript shows a Oritidine 5'-phosphate decarboxylase and Phosphoribosyl transferase and it is most likely UMPS for cassava, however it has not been

confirmed experimentally.


Download GenPept Graphics

uridine 5'-monophosphate synthase [Manihot esculenta]

Sequence ID: XP_021628228.1 Length: 473 Number of Matches: 1

See 1 more title(s)

Range 1: 3 to 469 GenPept Graphics Next Match Previous Match

Score	Expect	Method	Identities	Positives	Gaps
752 bits(1941)	0.0	Compositional matrix adjust.	375/470(80%)	424/470(90%)	4/470(0%)

```
Query  5    TMESIILQLHDISAVKFGEFKLKSGIFSPIYIDLRLIVSYPSILRQISETLVGSLPESTK  64
            +MES+IL+LH+ISAVKFG FKLKSGI SPIYIDLRLI+SYPS+L  IS+TL+ S+  S+
Sbjct  3    SMESLILKLHEISAVKFGNFKLKSGISSPIYIDLRLIISYPSLLSHISQTLISSVSSSSS  62

Query  65   YDVVCGVPYTALPIATCISTAHDLPMLMRRKEVKDYGTAKAIEGAFKPGQACLIVEDLVT  124
            +D++CGVPYTALPIATC+S ++ +PM+MRRKEVKDYGT+KAIEG F+ GQ CLIVEDLVT
Sbjct  63   FDLICGVPYTALPIATCVSLSNGIPMVMRRKEVKDYGTSKAIEGVFQSGQTCLIVEDLVT  122

Query  125  SGASVLETAAPLRAVGLKVTDAVVMIDREQGGRENLADNGITLHSMVKLTEMVRILKEKG  184
            SG SVLETAAPLRA GL V+DAVV+IDREQGGRENL  NGI LHS++KL+EMVR+L+EKG
Sbjct  123  SGTSVLETAAPLRAAGLNVSDAVVLIDREQGGRENLEVNGIKLHSIIKLSEMVRVLREKG  182

Query  185  RVSEETEKMVKKFLEENRKVAVPAPVKETKVRTR-LPYDERAKIAKNPTGKKLFEIMVQK  243
            ++ EE E++V KFLEENRKVA PA     KVRT+ L ++ERAK+AKNPTGK+LFE+MV+K
Sbjct  183  KLDEEMERIVMKFLEENRKVAAPA---MAKVRTKALKFEERAKLAKNPTGKRLFELMVKK  239

Query  244  NTNLCLAADVATAAELLDIADKVGPEICMLKTHVDILPDFTPDFGSKLRSIADKHNFLIF  303
             +NLCLAADV TAAELLD+A+KVGPEICMLKTHVDI PDFTPDFGS+LR+IADK+NF+IF
Sbjct  240  ESNLCLAADVGTAAELLDLAEKVGPEICMLKTHVDIFPDFTPDFGSRLRAIADKYNFVIF  299

Query  304  EDRKFADIGNTVTMQAEGGIFRILDWADIINAHIISGPGIVDGLKLKGLPRGRGLLLLAE  363
            EDRKFADIGNTVTMQ EGGIF ILDWADI+NAHIISGPGIVDGLKLKGLPRGRGLLLLAE
Sbjct  300  EDRKFADIGNTVTMQYEGGIFHILDWADIVNAHIISGPGIVDGLKLKGLPRGRGLLLLAE  359

Query  364  MSSAGNLATGDYTAAAVKIAEDHSDFVIGFISVNPASWPNGPGNPSLIHATPGVQLVKGG  423
            MSSAGN ATGDYTAAAVKIAE+HSDFVIGFISVNPASW   P NP+ I ATPGVQ V GG
Sbjct  360  MSSAGNFATGDYTAAAVKIAEEHSDFVIGFISVNPASWSGAPVNPAFIQATPGVQKVTGG  419

Query  424  DALGQQYNTPNSVIADRGSDIIIVGRGIIKAANPAEAAREYRLQGWDAYL  473
            DALGQQYNTP SVI DRGSDIIIVGRGIIKAA+PAEAAREYRLQGWDAYL
Sbjct  420  DALGQQYNTPYSVIYDRGSDIIIVGRGIIKAADPAEAAREYRLQGWDAYL  469
```

Figure 4.3: Results of amino acid sequence comparison of plastidial UMP synthase from S. tuberosum to *M. esculenta* v6.1 to identify possible UMP synthase. One UMP synthase was identified which has 80% similarity to cassava.

Two possible ADKs; 6 and 7, were identified on chromosome 7 (chloroplastic form X1) (cassava4.1_0247417m.g) (67.87% similarity) (**Figure 4.4A)** and chromosome 10 (chloroplastic form X2) (cassava4.1_013232) (70.85% similarity) (**Figure 4.4B**). In order to design a construct that would silence both ADK transcripts, a consensus sequence was created using Bioedit (version 7.2)(Hall, 1999) and the region with highest homology was identified (313bp–688bp), 98.8% with ADK isoform X1 and 94.6% with ADK isoform X2. Using this information hp-RNA constructs were designed and sent for synthesis. These hairpins were inserted into pCB33.

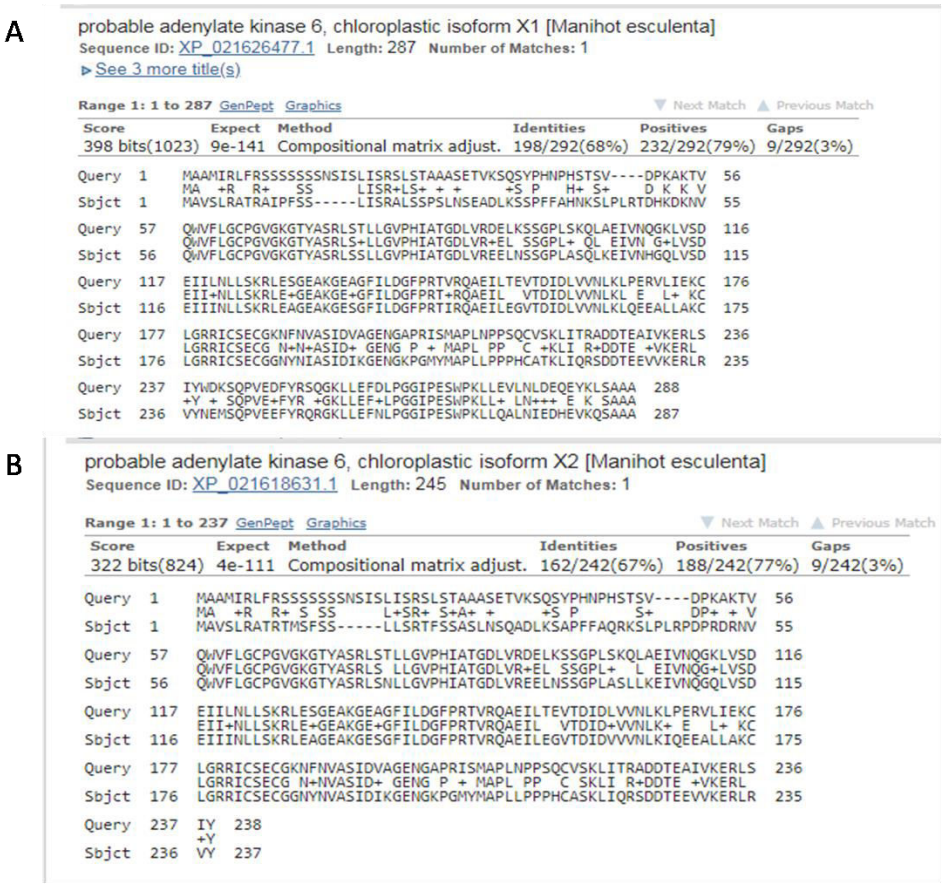


Figure 4.4: Results of amino acid sequence comparison of *S. tuberosum* plastidial ADK to cassava to identify possible plastidial ADK. Two similar sequences corresponding to adenylate kinase isoform X1 and X2 were identified.

4.3.2.2 Insertion of ADK and UMPS constructs into pCB33

The synthesised ADK and UMP synthase DNA fragments were inserted into the pCB33 vector using restriction digestion and sticky end ligation, after which colony PCR was performed to screen transformants for the presence of the insert. Seventeen transformants were screened for the presence of the ADK insert, and 6 transformants were screened for the presence of UMP. The insert was found to be present in 4 of the ADK transformants (**Figure 4.5**) and 1 of the UMPS transformants (**Figure 4.6**).

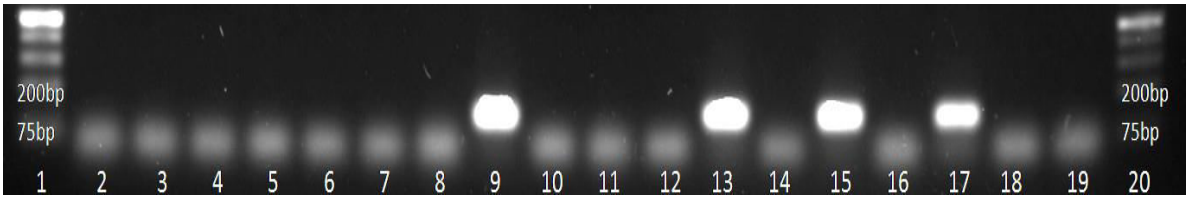


Figure 4.5: Colony PCR amplification of *ADK* gene to screen possible ADK transformants for the presence of the ADK insert. Lane 1, 20: GeneRuler 1kb Plus ladder (Thermoscientific). Lane 9, 13, 15, 17: positive ADK transformants with amplification of portion (145bp) of ADK insert. Lane 2-8, 10-12, 14: negative ADK transformants. Lane 18: Empty pCB33 (negative control. Lane 19: Non-template control (NTC).

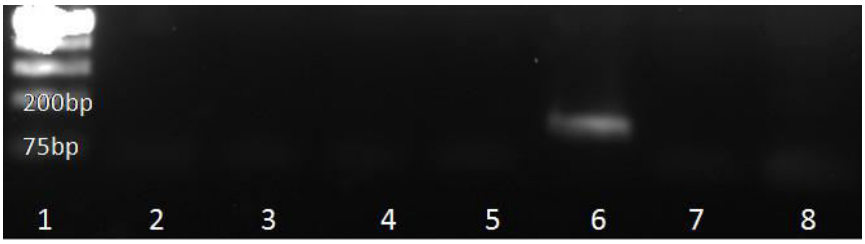


Figure 4.6: Colony PCR amplification of *UMP synthase* gene to screen possible UMPS transformants for the presence of the UMPS insert. Lane 1: GeneRuler 1kb Plus ladder (Thermoscientific). Lane 6: positive UMPS transformants with amplification of portion (123bp) of UMPS insert. Lane 2-5: negative ADK transformants. Lane 7: Empty pCB33 (negative control). Lane 8: Non-template control.

4.3.3 Agrobacterium-mediated transformation of cassava cv.60444 FECs

4.3.3.1 Screening of *A. tumefaciens* LBA4404 transformants for presence of pCB-ADK and pCB-UMPS

In order to confirm that *A. tumefaciens* LB4404 was successfully transformed with either pCB-ADK or pCB-UMPS, purified vector (from 1 clone pCB-ADK and pCB-UMPS plates) was digested using EcoRI and HindIII restriction enzymes. pCB-ADK produced three fragments (11871bp, 1895bp and 471bp) and pCB-UMPS produced two fragments (11 871bp and 2988bp) (**Figure 4.7**) which correlates to the expected sizes.

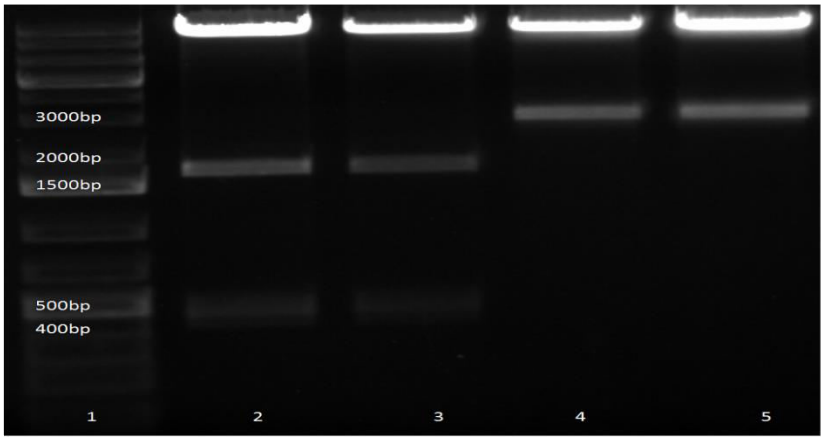


Figure 4.7: Restriction digestion of pCB-ADK and pCB-UMPS transformation vectors, digested with FastDigest EcoRI and Hind III restriction enzymes (Thermoscientific), extracted from *A. tumefaciens*LBA4404. Lane 1: GeneRuler 1kb Plus (Thermoscientific), lane 2: pCB-ADK extracted from transformed *A. tumefaciens*, lane 3: pCB-ADK control, lane 4: pCB-UMPS extracted from *A. tumefaciens*, lane 5: pCB-UMPS control.

4.3.3.2 Screening of regenerated plantlets for pCB-ADK and pCB-UMPS constructs

Cassava cv.60444 FECs were transformed with pCB-ADK and pCB-UMPS constructs using Agrobacterium mediated transformation. Transformed FECs (GD14) were then regenerated through the addition of synthetic auxin NAA and cytokinin BAP. Thirty four UMPS lines, 26 ADK lines, 12 pC1305.2 and 15 LBA 4404 lines were regenerated (**Table 4.1**).

Table 4.1: Results of transformation of cv.60444 FECs with pCB-ADK and pCB-UMPS

Transformation	No. of embryo's	No. of regenerated plantlets	Gus Assay	Rooting test
pCB-ADK	51	26	18	10
pCB-UMPS	43	33	22	9
pC1305.2	21	12	8	5
LBA 4404	27	15	0	0

4.3.3.3 Histological GUS assay and rooting test

In order to establish which of the regenerated lines had been successfully transformed, a number of histological and molecular tests were conducted. Plant material from each line was incubated in GUS assay buffer and then destained to remove the chlorophyll from the leaves (**Figure 4.8**). Leaves were then evaluated for characteristic blue colour associated with expression of the *Gus* gene. Eighteen of the pCB-ADK (69%) and 22 (64%) of the pCB-UMPS lines were positive for GUS expression (**Table 4.1**). The lines also subjected to a rooting test, to test for the expression of *Hyg* gene. Ten pCB-ADK (38%) and 9 (26%) pCB-UMPS lines rooted (**Figure 4.8B**).

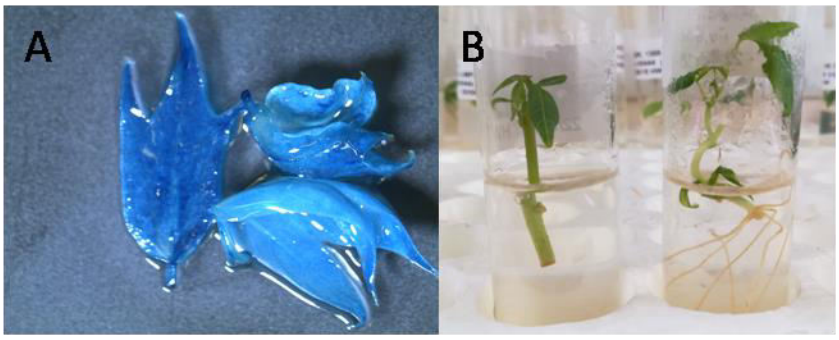


Figure 4.8: Screening of pCB-ADK and pCB-UMPS transformed cv.60444 lines for presence of *Hyg* and *Gus* transgenes. A: GUS assay performed on leaflets of cv.60444 line transformed with pCB-ADK. B: Rooting test performed on cv.60444 transformed with pCB-ADK construct, left: negative control (untransformed cv.60444), right: positive line.

4.3.4 Molecular screening of pCB-ADK and pCB-UMPS lines

4.3.4.1 PCR for *Hyg* and *Gus* genes

In order to confirm the histological tests performed on the transgenic pCB-ADK and pCB-UMPS lines, PCR were conducted to screen for presence of reporter genes *Gus* and selection marker *Hyg*. Of the pCB-ADK lines, 13 lines contained the *Hyg* gene (**Figure 4.9A**) and *Gus* genes (**Figure 4.9B**). Of the 33pCB-UMPS regenerated lines 13 lines contained the *Hyg* gene (**Figure 4.10**), and 18 contained the *Gus* gene (**Figure 4.11**). Based on these results, nine pCB-UMPS lines (1, 2, 10, 11, 12, 13,

17, 24, 26, 28), and eight pCB-ADK (1, 5, 7, 8, 13, 20, 21, 24) were chosen for further studies.

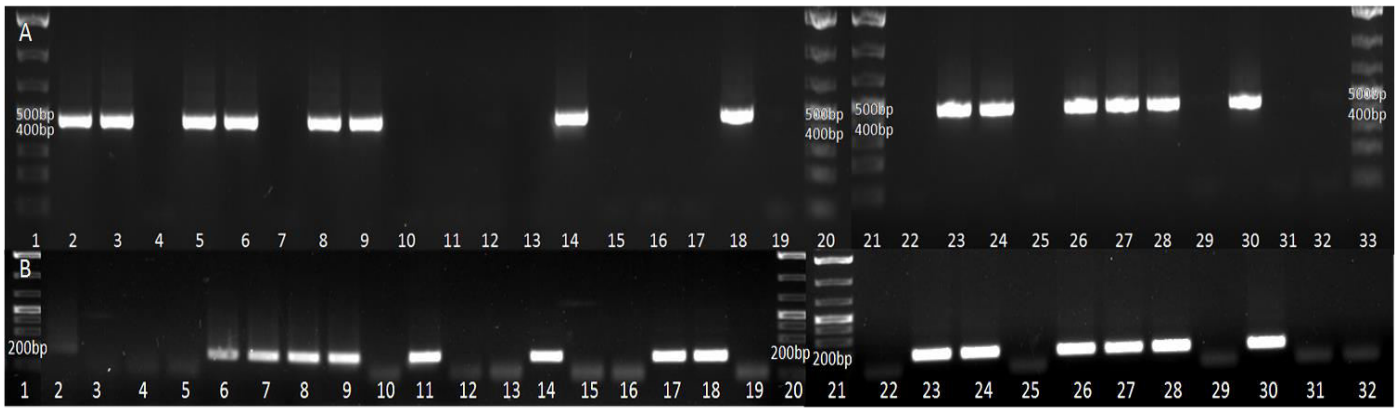


Figure 4.9: PCR amplification of *Hyg (A)* and *Gus (B)* genes in cv.60444 plants transformed with pCB-ADK construct. A) Lane 1, 20, 21, 33: GeneRuler 1kb plus molecular weight marker (Thermoscientific), lane 2-19, 22-29: pCB-ADK lines, lane 30: pC1305.1 (positive control), lane 31: untransformed cv.60444 (negative control), lane 32: NTC. B) Lane 1, 20, 21: GeneRuler 1kb plus molecular weight marker (Thermoscientific), lane 2-19, 22-29: pCB-ADK lines, lane 30: pC1305.1 (positive control), lane 31: untransformed cv.60444 (negative control), lane 32: NTC.

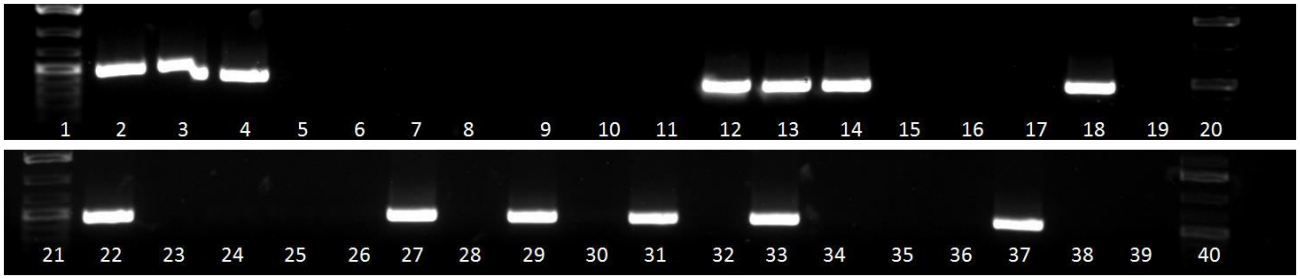


Figure 4.10: PCR amplification of *Hyg* gene in cv.60444 plants transformed with pCB-UMPS construct. Lane 1, 20, 21, 40: GeneRuler 1kb Plus molecular weight marker, lane 2-19: pCB-UMPS lines 1-18, lane 22- 36: pCB-UMPS lines 19- 33, lane 37: Positive control (*pCB-UMPS*) lane 38: untransformed cv.60444 (neg control), lane 39: NTC.

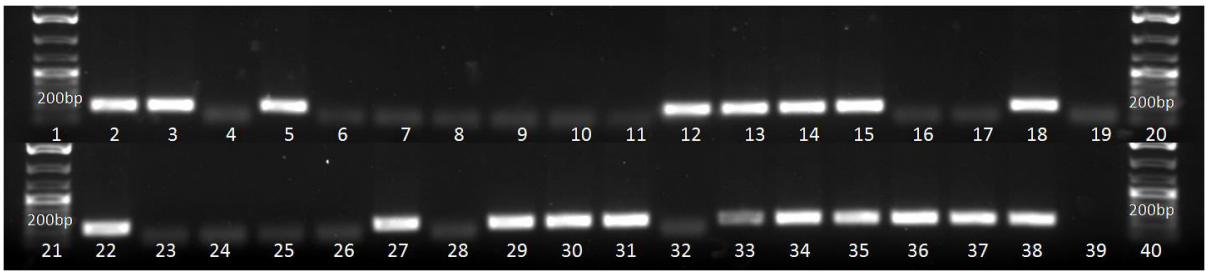


Figure 4.11: PCR amplification of *Gus* gene in cv.60444 plants transformed with pCB-UMPS construct. Lane 1, 20, 21, 40: GeneRuler 1kb Plus molecular weight marker, lane 2-19: pCB-UMPS lines 1-18, lane 22- 37: pCB-UMPS lines 19- 33, lane 38: pC1305.1, lane 39: NTC

4.3.4.2 Southern blot hybridisation

Southern Blot hybridisation was performed in order to determine the number transgene integration events in the transgenic lines which contained both *Hyg* and *Gus* gene markers (**Figure 4.12**).ADK transgenic lines 1, 5, 7, 20 and 21 contained a single copy of the *Gus* gene, line 8 contained 3 copies and line 24 contained 2 copies. Six UMPS transgenic lines (11, 13, 17, 24, 26, and 28) had single copies of *Gus*, 3 lines with double copies (2, 10, and 12) and one line (1) with three copies of *Gus*.

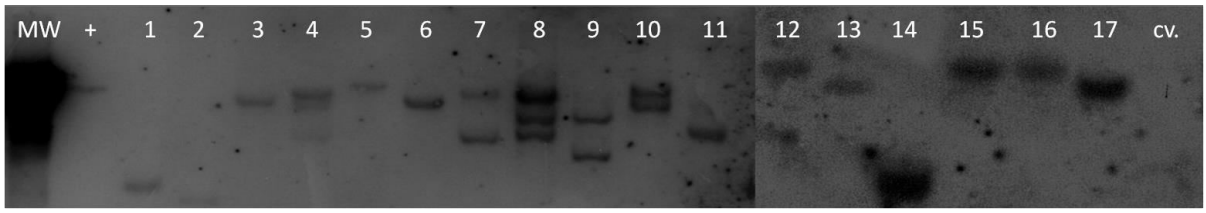


Figure 4.12: Southern blot analysis of cv.60444 transgenic lines transformed with hp-RNA constructs pCB-ADK and pCB-UMPS probed with DIG-labelled DNA Gus probe. MW: DNA molecular weight marker, DIG-labelled (MW) (Thermoscientific), Positive control (+) 1µg of HindIII digested pCambia 1305.1. Lane 1-8: cv.60444 ADK transgenic lines 1, 5, 7, 8, 20, 21, 24 respectively. Lane 9- 18: cv.60444 UMPS transgenic lines 1, 2, 10,11,12,13,17,24,26,28 respectively. Untransformed cv.60444 was included as a control (cv.).

4.3.5 Characterisation of micro-propagated cassava lines

On the MS media, the transgenic UMPS lines took on average 4 days longer to develop primary roots, before transfer into the jiffies than either the regenerated

cv.60444 or transgenic ADK lines. These roots also appeared to be thicker than the control or ADK lines. Once in soil, the ADK lines developed many more leaves than either the control lines or UMPS lines, which grew more slowly. After 3 months in small pots the average size of the UMPS lines (6 plants per line) were 5cm (with a standard deviation of ±2.31cm) shorter than either the cv.60444 or ADK lines. After 10 months the no disenable pattern with regards to height was observed (**Figure 4.13).**

Figure 4.13: Transgenic cv.60444 lines transformed hp-RNA construct targeting either plastidial *ADK* or *UMPS* gene. The heights of each of the transgenic lines; A: ADK-1, B: ADK 5, C: UMPS 1, D: UMPS 2, E: UMP 13 and F: UMPS 17 (right) which produced storage roots were compared to untransformed cv.60444 (left plant).

After a 10 month period only 4 UMPS lines (1, 2, 13 and 17) and two ADK lines (1, 5) as well as control cv.60444 had begun to develop storage roots (**Figure 4.14**). The UMPS lines produced storage roots that were larger than the cv.60444 control lines and weighed on average more than either the untransformed cassava cv.60444 or ADK lines (dry weight) (**Table 4.3**). Only two ADK (1, 5) lines produced storage roots (6 plants per line) which were smaller (2.7±1.3 and 3.06± 2.42 respectively) than the control cv.60444 lines and weighed less (dry weight) (**Table 4.3**).

Figure 4.14: Sample of storage roots produced by cassava cv.60444 transgenic plants transformed with hp-RNA constructs targeting UMPS and ADK. A: ADK 1, B, ADK 2, C: UMPS 1, D: UMPS 2, E: UMPS 13, F: UMPS 17, G: untransformed wild-type cv.60444.

Table 4.2: Characteristics of storage roots produced by cassava cv.60444 transgenic plants (6 plants per line) transformed with hp-RNA constructs targeting UMP synthase and ADK

Construct	Line	Size (cm)	Wet weight	Dry weight	Percentage Dry weight
pCB-ADK	1	2.7±1.3	2.05± 1.66	0.331±0.2	16.14634
	2	2.93±1.78	4.06± 2.42	0.18±0.05	5.882353
pCB-UMPS	1	4.67 ±2.13	2.59± 2.43	1.2±0.3	46.33205
	2	4 ± 3.56	3.9± 1.36	1.654±0.21	16.76923
	13	8.7 ± 2.13	14.09± 0.81	4.254± 0.34	30.19163
	17	7± 3.1	20.3± 3.77	5.33 ±1.76	26.25616
cv.60444		4± 3.7	4.58±1.44	0.806±0.232	17.59825

4.3.7 Expression of *ADK* and *UMP synthase* in transgenic lines

In order to confirm that increased starch levels shown in the developed storage roots was related to down-regulation of either *ADK* or *UMP synthase* expression, their relative expression (relative to β-tubulin) was evaluated using RT-qPCR. cDNA was synthesised from total RNA extracted from each of the storage roots, with cv.60444 used as the control. Of the two ADK lines, ADK 1 and ADK 5 which formed storage roots, ADK 5 had the lowest relative expression levels (0.04±0.015) (**Figure 4.15**) of the three lines, while ADK 1 had slightly raise expression levels (0.072±0.017) compared to cv.60444 (0.05±0.021). While ADK 5 expression levels were lower than cv.60444, these results were found not to be significant.

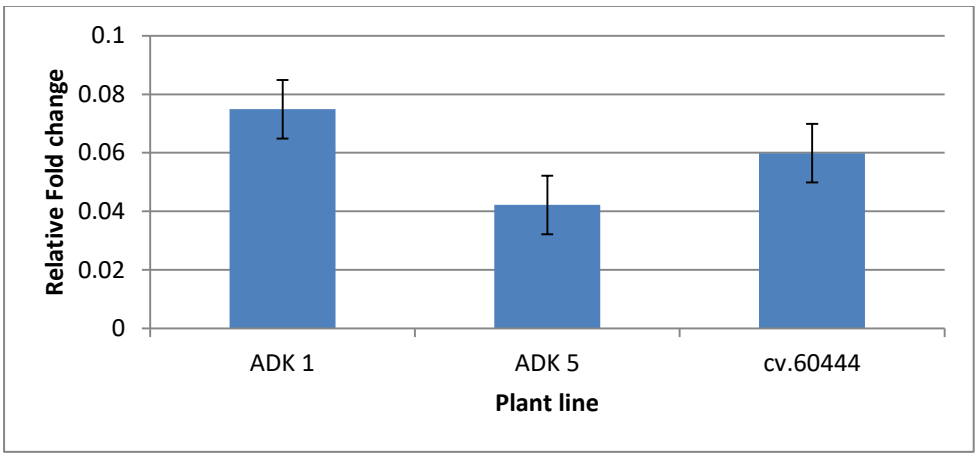


Figure 4.15: Relative expression of Adenylate kinase in transgenic lines (ADK 1, ADK 5) transformed with pCB-ADK construct and untransformed cv.60444. The values represent the mean value of three biological replicates with bars indicating standard error.

After 10 months, 4 UMPS lines produced storage roots. The relative expression of *UMP synthase* in these lines was evaluated by RT-qPCR and expression was compared to untransformed cv.60444 (**Figure 4.16**). Two UMPS lines (13 and 17) had significantly lower levels of expression of *UMP* synthase ($p<0.1$) (Appendix P) (0.03± 0,003, 0.025±0.004, and 0.018±0.016 respectively), compared to untransformed cv.60444 (0.051±0.013). Transgenic line UMPS 2 had the highest relative level of *UMP synthase* expression (0.073±0.016).

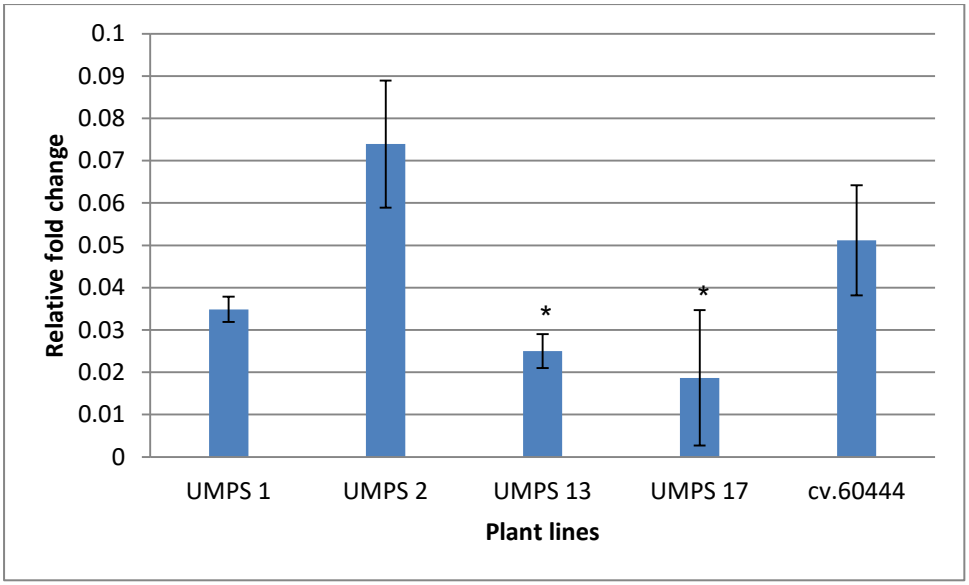


Figure 4.16: Relative expression of *Uridine-5'-monophosphate synthase* in transgenic lines (UMPS 1, 2, 13 AND 17) transformed with pCB-UMPS construct and untransformed cv.60444. The values represent the mean value of three biological replicates with bars indicating standard error. Asterisks indicate statistically significant results.

4.4 Discussion

Cassava's storage roots are its main commercial product and the development of high quality, high yielding storage roots has been the focus of much of the research on cassava. The regulation of nucleotide pools has been shown to be linked to the synthesis of starch in a number of plant species including potatoes and *Arabidopsis* (McKibbin *et al*, 2006; Oliver *et al*, 2008; Tjaden *et al*, 1998). Specifically alterations in the ATP and Uridine pools in amyloplasts have been shown to effect the rate and quantity of starch stored in potato storage roots (Oliver *et al*, 2008; Regierer *et al*, 2002). This study reports, for the first time, the construction and transformation of cassava cv.60444 with hp-RNA constructs which down-regulated two enzymes, ADK and UMPS, which are involved in starch synthesis. Preliminary characterisation of transgenic plants, storage root formation and quantity was carried out and storage roots were formed by 4 UMPS line and two ADK, due to a 10 month period required to generate roots in pots, only preliminary starch assessments were performed, however the increase in storage root yields in UMPS lines was shown to be correlated to down-regulation of *UMP synthase* in the UMPS lines. Further characterisation of these lines was not possible due to the time constraints of the

project. However, in future, measurement of starch production in these lines will be evaluated using analytical methods (Bull *et al*, 2018b; Ihemere *et al*, 2006; Regierer *et al*, 2002).

Cassava is highly heterologous in nature and breeding for improved characteristics is complicated. For this reason, cassava is a good candidate for crop improvement using transgenic technology (Bull *et al*, 2018a). Post-transcriptional gene silencing has been used to enhanced traits in a number of different crops (Nassar and Ortiz, 2006). In cassava it has been used to increase zinc accumulation and iron in roots (Gaitán-Solís *et al*, 2015; Ihemere *et al*, 2012), improve vitamin content (Te Li *et al*, 2015; Welsch *et al*, 2010) as well as alter the physiochemical properties and amount of starch (Ihemere *et al*, 2006; Ligaba-Osena *et al*, 2018; Raemakers *et al*, 2005). In order to increase the amount of starch in cassava, Ihemere *et al* (2006) altered AGPase expression through the over-expression of *E. coli glg*C gene (2-fold increase) which increased starch levels in cassava by 60%. In addition, several studies have been conducted which improve the quality of starch has been manipulated. These include the down-regulation of GBSII (Bull *et al*, 2018b; Raemakers *et al*, 2005) which resulted in lower levels of amylose in order to improve the quality of starch. While the down-regulation of *UMP synthase* and *ADK* have been shown to have potential in increasing the amount of starch in potatoes (Geigenberger *et al*, 2005; Regierer *et al*, 2002), no transgenic cassava varieties have been produced targeting these regions in cassava previously. In this study 3 UMPS lines showed increased storage root yield, compared to untransformed cv.60444. Although the starch content of these storage roots was not tested, it shows that the down-regulation of UMPS may regulate starch syntheses in cassava.

Agrobacterium-mediated transformation of cassava FECs is a widely used technique for the transformation of cassava (Bull *et al*, 2009). However this technique is time consuming (with regeneration times of in excess of 8 months) and production of high quality FECs for good levels of regeneration is difficult. In this study, two rounds of FEC transformations were carried out and both the regeneration and transformation efficiencies were low in both. Using the transformation protocol developed by Bull *et al* (2009) (where FECs were transformed with pCambia 1305.1) should produce in excess of 50 plantlets per 100 FEC clusters however, the highest regeneration efficiency achieved was shown in FECs transformed with pCB-UMPS which

produced 33 plantlets from 140 clusters. The transformation efficiency was also lower than expected with only 10 transgenic lines being produced.. While these transformation rates are lower than reported by Bull *et al* (2009), they are similar to studies conducted by (Ligaba-Osena *et al*, 2018), where only 11 transgenic lines were produced. This procedure was also more efficient than the procedure used by Raemakers *et al* (2005), where somatic embryo's were used, or Ihemere *et al* (2006), where FECs were transformed using particle bombardment and only three transgenic lines were isolated. While these regeneration efficiencies are lower than expected, (Bull *et al* (2009) showed rates of >80%), they are similar to the pC1305.1 transformation rate (lines (GUS- 66% and Hyg- 41%), which indicates that these rates could be due to the quality of FECS or due to stress experienced by the FECs (where cabinets containing FECs in this study went through periods of high heat, which may have affected transformation efficiency.

In order to confirm transformation, all the regenerated plantlets were subjected to both histological and molecular screening for reporter genes. Of the total of 26 ADK and 33 UMPS transformants that were regenerated, ten UMPS lines and eight ADK lines contained both *Hyg* and *Gus* genes confirming transformation. Some interesting results were noted in the histological testing of the transgenic lines. Of the eighteen ADK lines and twenty one UMPS lines that tested positive for Gus activity, the *Gus* gene was only detected in a proportion of the lines (and seventeen respectively), indicating that some of the results were false positive. The Gus assay is highly photosensitive and it is possible that the samples were exposed to light; however the negative control did not show a positive reaction. False positives have been reported in other studies and are known to occur in the leaves, stem and pollen of higher order plants (Thomasset *et al*, 2002) which is why it is important to screen plants using molecular methods. The molecular screening of ADK and UMPS transgenic lines also showed that a proportion of the lines (two ADK lines and four UMPS) contained the *Hyg* gene but did not root when grown on MS media containing hygromycin, which indicates that although the gene is present, it did not confer tolerance to hygromycin in these transgenic lines. This has been observed in other studies, where genetic and epigenetic variations occur due to the random insertion of the transgenic cassette in individual transgenic lines (Ligaba-Osena *et al*,

2018), and it is possible that the locus of transgene insertion affects its expression (Fuentes *et al*, 2006).

In this study, the early stages of development (on MS media) the UMPS lines grew more slowly than the cv.60444 or ADK lines. The primary roots also appeared to be thicker and took more time to develop. This could be due to the down-regulation of de novo synthesis of pyrimidines, caused by the UMPS transgene expression, and has been reported in other plants where de novo synthesis has been down-regulated (Schroder *et al*, 2005). The primary catabolism and uridine salvage pathways play an important part in plant cells metabolism. The ATP created during photosynthesis is used for two main purposes; sucrose biosynthesis (for UDP-glucose) or starch production (for ADP-glucose) (Chen and Thelen, 2011), which implies that these two processes are competitive. The primary synthesis of pyrimidines takes place during period of high abundance of nutrients, such as during hours of light and that this is down-regulated during dark periods, when pyrimidine salvaging then occurs and starch accumulation increases. By down-regulating de novo synthesis in the plant storage roots, starch accumulation occurs at a higher rate (Chen and Thelen, 2011). The down-regulation of the UMP synthase (under the control of the patatin promoter) causes plants to halt de novo synthesis of pyrimidines and increase starch production in the roots. This may cause the plants to grow more slowly than either the ADK or cv.60444 controls.

Once molecular characterisation had confirmed the presence of the reporter genes, transgenic ADK and UMPS lines were then propagated and moved to soil, in order to assess the effect of the down-regulation of ADK and UMPS on starch production in cassava. After 10 months only four UMPS lines and two ADK lines had produced storage roots, with only two of the UMPS lines (13 and 17) showing larger storage roots than the control lines. In order to establish whether the increased storage root production was due down-regulation of either *ADK* or *UMP synthase*, relative expression compared to β-tubulin was established in each of the lines. Of the four UMPS lines, *UMP synthase* was shown to significantly down-regulated in UMPS lines 1, 13 and 17. Interestingly although UMPS 17 had the lowest average relative expression level, it did not produce the largest storage roots, and had a lower dry weight percentage than both UMPS 1 and UMPS 13, indicating lower starch content than either of these lines. Transgenic line UMPS 1 and UMPS 13 also had

significantly lower expression of *UMP synthase*, and larger average roots than the cv.60444 control. These results would indicate that the down-regulation of *UMP synthase* contributes to starch accumulation as UMPS 2 did not show a significant increase in root size and had, on average, increase expression of UMPS compared to the other transgenic lines. These variation could be due to genetic or epigenetic differences between the lines due to the insertion of the transgene which has been shown in other studies (Machczyńska *et al*, 2015). The correlation between transgene expression and starch accumulation was also shown by Ihmere *et al* (2006) who showed that a 2.6 fold increase in the expression of AGPase correlated to an increase of 60% in storage root yields. Stark *et al* (1992) showed that the over expression of AGPase (through the introduction of a sense *E.coli GlgC* gene) increase the production of starch in potatoes by more than 30% in one variety. Although *ADK* was shown to be down-regulated in transgenic line ADK 5 compared to untransformed cv.60444, this was found not to be significant. This could indicate that the transgene is not highly expressed in the root or that in these lines the down-regulation is not significant due to variations between the lines, similar to those seen in UMPS 2. Further investigation of the lines will have to be conducted to establish whether expression of *Gus* and *Hyg* correlate with these findings. The other transgenic lines were not tested for expression of ADK or UMPS as they did not produce storage roots, however rooting trials are currently being repeated in these lines and transgene expression will be tested if they produce storage roots at a later date.

Cassava is mainly grown for its storage roots which are produced at about 6-12 months depending on the cultivar (Hahn and Keyser, 1985). In the field, cassava planted from stakes develop shoots and primary root system over the first two months, after which cassava enters a secondary growth phase where fibrous roots develop and mature into storage roots 2-4 months later (depending on the cultivar) (El-Sharkawy, 2004). In order to induce storage root formation in the transgenic lines in this study, the lines were micropropagated and planted in small pots and grown under controlled conditions for 4 months, fertilised monthly with NPK fertiliser. It was not possible to undertake field trials as GM permits have to be obtained which was not possible in the available timeframe. The plants were then transferred to larger pots and fertilised with root promoting fertiliser (high in phosphorus, low in

Nitrogen). After 6 months, none of the cassava lines (including controls) had produces any secondary fibrous or tuberous roots. This is unusual as most studies carried out on cassava have reported storage roots within 6 month of potting, although Ligaba-Osena *et al* (2018) did report on a variance in the time required for the initiation of storage root development. As the slow development of storage roots occurred in both control (untransformed cv.60444) and transgenic lines, it is most likely that the conditions under which the plants were grown hindered storage root development, and was not due to the transgenes. The level of nitrogen in soil in known to affect root crop development, where high levels of nitrogen in the soil is linked to low yields of root vegetables (De la Morena *et al*, 1994). It is possible that the fertilisation regime followed in this experiment allowed for too much nitrogen in the soil, which lowered the storage root yields. After 6 months, in order to encourage storage root development, the plants were transferred back into smaller pots and were only fertilized once after 2 months after which the small tuberous root (±1cm) started to develop. The plants were also subjected to controlled water restrictions to induce them to produce storage roots under water stress (Bull, pers. comm.). After 10 months, six of the transgenic lines (2 ADK lines and 4 UMPS lines) had produced storage roots (as well as control cv.60444). Two of UMPS lines produced large storage roots (>50% compared with untransformed cv.60444) while all other lines produced storage roots of a similar or smaller size than untransformed cv.60444. This result is promising, and is the first report of UMPS-transgenic cassava to date. Future studies to the optimize conditions for root development should be carried out in order to establish under what condition cassava will produce roots in controlled environments, and larger greenhouse and field trials on UMPS-transgenic cassava need to be performed. The ADK-knockdown lines do not seem to increase starch yield, however a combination of stacked ADK-UMPS hp- RNA constructs may prove interesting in future studies.

The developed storage roots were further characterised based on size and wet and dry weights. Dry weight (as a percentage of total mass) is used as an indicator of the chemical potential of many crops, and is considered a good reflection of the true biological yield of the crop (Teye *et al*, 2011). Cassava dry weight is dependent on the cultivar but is usually between 17-47%, with varieties with greater than 30% dry weight considered high (Barimah *et al*, 1999). Of the 6 transgenic lines which

produced storage roots in this preliminary study only two UMPS lines; 1 and 13, had a dry weight which exceeded 30% (UMPS 1 and UMPS 13). While these results are promising, before any conclusions about the starch content of these roots can be made, more accurate measurements of the starch content, amylose and amylopectin levels will have to be conducted before any further conclusions can be drawn. These will include metabolite analysis including soluble sugars (Fernie, 2001), sucrose and starch (Geigenberger *et al*, 2001) and amino acid analysis (Trethewey *et al*, 1998). In addition, the ADK lines will be evaluated for ADK activity (Regierer *et al*, 2002) and UMPS lines will be evaluated for UMP synthase activity (Geigenberger *et al*, 2005; Regierer *et al*, 2002) Due to time constraints, the transgenic lines were not tested further. Future studies of these lines will have to include a repetition of storage root trials, where conditions for storage root development are optimised. Further, the starch content and composition must be tested, including amylose and amylopectin composition.

This study reports on the successful *Agrobacterium*-mediated transformation of cassava cv.60444 with hp-RNA constructs targeting ADK and UMPS in order to increase starch content. The transgenic lines were characterised and propagated for future studies on the effect of down-regulation of these gene targets on starch accumulation. These approaches could have a large advantage for cassava where an increase in starch in the roots could be hugely beneficial for commercial purposes for the paper, textile, food and bioethanol industries (Blagbrough *et al*, 2010; Nassar and Ortiz, 2006).

Chapter 5: Summary and Conclusions

Cassava is an important food crop as well as an important source of industrial starch (Zhang *et al*, 2017). The increase in global population and therefore food demands, in conjunction with changes in global climate have increased the interest in cassava cultivation because it can be grown under sub-optimal conditions, in nutrient poor soil with little water(De Souza *et al*, 2016; Zhang *et al*, 2017). While cassava has great potential, there are a number of challenges associated with cassava cultivation including viral disease, as well as low yields due to the poor conditions under which cassava is usually grown (Jennings and Iglesias, 2009). Cassava can be improved through traditional breeding (Nassar and Ortiz, 2006), however the heterozygous nature of cassava makes this approach difficult and therefore genetic engineering of cassava has great potential for improved crop production. A number of studies have utilised RNAi technology in order to improve traits in cassava including nutrition in roots and storage roots (Gaitán-Solís *et al*, 2015; Ihemere *et al*, 2012; Li *et al*, 2015; Welsch *et al*, 2010) improved starch (with altered physiochemical properties and well as increased content) (Ihemere *et al*, 2006; Ligaba-Osena *et al*, 2018; Raemakers *et al*, 2005) and improved viral resistance (Vanderschuren *et al*, 2007; Zhang *et al*, 2005). This project focused on two main areas of cassava improvement, increased viral resistance through the introduction of hairpin-RNA (hp-RNA) constructs targeting the AC1/AC4 region of *South African cassava mosaic virus*, and increased starch production through the down-regulation of *ADK* and *UMP synthase* genes, which are involved in starch production. In addition to these two aims, a third study examined the off-target effects associated with a triple stacked hp-RNA construct targeting *East African cassava mosaic virus* , SACMV and *African cassava mosaic virus* , designed to improve resistance to all three viruses but which showed toxic effects when it was used to transform cassava friable embryogenic callus (FECs). While phenotypic variation, and gene regulation as a result of transformation with viral transgenes has been reported previously (Fuentes *et al*, 2006; Praveen *et al*, 2010; Vanitharani *et al*, 2004) to the best of our knowledge this the first study which identifies specific genes which are affected by the transgenes targeting cassava mosaic geminiviruses. Further this is the first study which reports on possible negative effects associated with gene stacking for resistance to multiple viruses. This

has ramifications for future work involving the production of transgenic plants, where transgene target multiple viruses.

Cassava mosaic disease (CMD) is a large hurdle to the cultivation of cassava in sub-Saharan Africa (Legg *et al*, 2015), where it is caused by a at least 9 cassava mosaic geminiviruses (Brown *et al*, 2015). While RNA-based silencing mechanisms have shown great potential in improving cassava resistance to ACMV(Vanderschuren *et al*, 2007; Zhang *et al*, 2005) this study is the first report of a transgene conferring improved SACMV tolerance to susceptible cassava cultivar 60444. In order to improve cassava resistance to SACMV, cassava cv.60444 was transformed with a hp-RNA construct targeting the AC1/AC4 overlapping region of SACMV. Three lines showed increased tolerance to SACMV, where infected plants had decreased symptoms and lower viral titres than the untransformed cv.60444 control lines. One line also showed recovery, similar to CMD resistant line TME3. While these lines show potential, longer more intensive studies need to be carried out on these lines before they could be considered for release. The recovery of these lines may be affected by a number of factors, including the fluctuation of the developmental cycle as the plant ages, light and temperature variation (Patil and Fauquet, 2009) as well as inoculum pressure where repeated exposure over time, or a high inoculum pressure (Vanderschuren *et al*, 2009) may affect their tolerance. These lines also need to be evaluated for the effect of infection on storage root yield, which has not been carried out yet. In the field CMD is usually caused by multiple begomovirus species within a single plant (Hahn *et al*, 1980) which can replicate simultaneously (Morilla *et al*, 2004). For this reason, these lines should also be tested for resistance to other prominent cassava mosaic geminiviruses including ACMV and EACMV.

One of the limitations associated with the use of RNA interference (RNAi) technology to induce resistance to pathogens, including viruses is the limited amount of in-depth information available about the exact mechanism involved in the efficient production of transgene derived siRNA in plants. It has been shown, in our studies, as well as in others, that the expression of the transgene is not always sufficient to induce resistance (Ribeiro *et al*, 2007),and while high levels of expression of transgene-derived siRNA is usually associated with increased resistance (Kalantidis *et al*, 2002), there is very little information available on the reasons why some transgenic lines produce lower levels of siRNA than others. It is assumed that one of the factors

leading to the variation in expression levels is due to unpredictable sites of integration of the transgene(Birch, 1997).This limits the usefulness of this technique as it requires that large numbers of transgenic lines are produced and screened for resistance, which laborious and time consuming. Another limitation associated with RNAi is target selection, while RNAi constructs of different lengths have been shown to successfully induce resistance, longer siRNA hairpins are expected to produce more diverse siRNA species and therefore be more effective. However, these longer hairpins have a higher likelihood of producing off-targets, as was shown with the AES triple construct. While shorter targets can be used, these transgenes produce less diverse siRNA populations and can be more quickly overcome in the field (Simón-Mateo and García, 2011). Another limitation of RNAi technology is its limited acceptance by the public, where there is a distinct gap between demand for disease free planting material by farmers and public perception of the safety of genetically modified crops. Although this perception is changing in some areas (Lucht, 2015), it is still a limitation to the release of genetically modified crops globally. One of the alternatives to RNAi technology is the new plant breeding technology known a CRISPR (clustered regularly interspaced short Palindromic Repeats)- Cas9 (CRISPR-associated protein 9) system (Ran *et al*, 2013). The CRISPR system is an efficient means of altering genomes in a site specific manner. The CRISPR system has been used to modify a number of genes in cassava including GRANULAR BOUND STARCH SYNTHASE, phytoene desaturase (Bull *et al*, 2018; Odipio *et al*, 2017) and has been used to improve resistance to ACMV (Mehta *et al*, 2018b).

Another of the limitations of this technology is its applicability and longevity in the field. Begomoviruses have high levels of intra-and inter-specific diversity, which allows them to adapt to new host and pressures (Monci *et al*, 2002). This diversity is mostly driven by high levels of recombination and pseudo-recombination events but also occurs through mutations which occur at high rates and are subject to positive selection (Duffy and Holmes, 2009; Monci *et al*, 2002). This genetic diversity allows geminiviruses to adapt quickly to environmental pressures and hosts (Berrie *et al*, 2001; Padidam *et al*, 1999; Zhou *et al*, 1997). In field trials conducted on *Tomato yellow leaf curl virus* resistant tomatoes, lines which had been completely resistant in greenhouse trials were shown to be infected by a variant of TYLCV (as well as another previously unidentified virus) with mutations and deletions in the genome

(Fuentes *et al*, 2016). Although the plants did not show symptoms is does point to the fact that viruses will adapt to host pressures. More recently Mehta *et al* (2018) showed that a antiviral dsRNA constructs targeting ACMV in cassava caused a shift in viral populations, towards species with lower levels of homology to the transgene. This adaptive behaviour has major implications for antiviral dsRNA technology, where any transgenic plant would have to be constantly monitored for signs of shifts in viral populations, and it cannot be assumed that the antiviral immunity would be permanent. Possible future studies in this area could include multiple variants of the transgene, or degenerate bases, however this could increase the likelihood of off-targets. Alternatively, as is the case with the SACMV AC1/AC4 construct, if lines are generated to confer tolerance, where the virus is maintained at low levels, it may lower the evolutionary pressure on the virus, and help to maintain field resistance for longer periods of time.

The second aim of our study was to study the 'off-target effect' observed in cultivar 60444 FECs transformed with a triple stacked construct, targeting the AC1/AC4 overlap of ACMV, as well as the AC1/IR of SACMV and EACMV. Friable embryogenic callus transformed with this construct showed high levels of mortality and very low transformation efficiency. Further examination of predicted transgene associated siRNAs showed that some of the siRNA had partial homology with cassava genes, which were shown to be differential regulated in the presence of the transgene. This study has significant implications for RNAi technology, which has traditionally been seen as safe, due to the high levels of specificity associated with small RNA target binding. RNA silencing trigged by dsRNA regulates a number of biological processes including defence against invading nucleic acids, plant development and environmental stresses (Lin *et al*, 2009; Pooggin *et al*, 2003). It involves the suppression of genes by small RNAs including siRNA and miRNA which are produced from dsRNA. RNAi technology has been widely adopted as a mechanism to improve crops, without affecting other agronomic traits (Mamta and Rajam, 2018; Zhang *et al*, 2017). However, a recent transcriptomic study of transgenic tomatoes, transformed with hp-RNA targeting TYLCV, identified a common set of up and down-regulated genes associated with the expression of the transgene (Fuentes *et al*, 2016), indicating that RNAi-induced silencing may affect host genes unintentionally and may limit the use of this technology. In our study, a

hp-RNA construct targeting ACMV, SACMV and EACMV, had a similar effect, where a number of host genes were regulated in response to transformation with the construct. The triple construct produced significant phenotypic variation in transformed cv.60444 plants and had a toxic effect on FECs. All of the regulated host genes were involved in stress or defence responses and it was further hypothesises that the triple construct hp-RNA produced siRNAs could be mimicking virus-derived siRNAs which target host genes (Yang *et al*, 2019). These results highlight the need for further exploration of viral target selection for PTGS-induced anti-viral technology in plants, where the expression of some viral targets could have undesired effects on the host.

One of this limitations of this work is that it relied heavily on manual identification of off-targets following previously identified criteria (Jackson *et al*, 2006; Praveen *et al*, 2010), and only examined proteins with known functions. Although a number of cassava cultivars have been sequenced (Bredeson *et al*, 2016) the entire genome has yet to be annotated and, due the high level of diversity within and between germplasm and the small sequence length of sRNA, this could mean that important targets were misidentified or not identified (Vaucheret *et al*, 2015). These results could be expanded through transcriptomic or degradome analysis(Jani *et al*, 2018). In this study, off-targets were indentified through bioinformatics and confirmed using RT-qPCR to analyses expression the identified gene targets, however this approach is only a pilot study and in order to more fully understand the impact of triple-stacked construct on cassava, a full transcriptomic analysis of all genes and/or a degradome study, which would better show regulation of genes by small RNA, would give better coverage of the effects on the genome.

The final aim of the study was to increase starch in the storage roots of cassava through the down-regulation of two genes, *ADK* and *UMP synthase*, which have shown potential in other crops (Geigenberger *et al*, 2005, 2001). Two other studies have focused on starch accumulation in cassava, by increasing the expression of *glgC*, which codes for AGPase. While both of these studies showed an increase in starch levels, AGPase studies in potatoes has shown that not all potato cultivars respond to an increase expression of AGPase (Sweetlove *et al*, 2015), so this technology may not be transferable to other cassava cultivars. Furthermore, the down-regulation of *ADK* and *UMP synthase* produced larger increases in starch in

potato, than the up-regulation of AGPase, which may also be the case in cassava. Other studies on starch in cassava have focussed on modifying starch characteristics such as amylose and amylopectin levels (Bull *et al*, 2018; Ayalew Ligaba-Osena *et al*, 2018). Future studies of the UMPS and ADK lines may include creating stacked constructs to combine these features. Although transformation efficiencies were lower than expected, eighteen transgenic lines (ten UMPS lines and eight ADK lines) were developed. These lines were micropropagated and allowed to grow for 10 months, in order to analyse storage root formation. After 10 months very few of the lines had produced storage roots (four UMPS lines, two ADK lines and control cv.60444). Although it has been shown previously that the time to produce storage roots can vary significantly between cassava cultivars other studies on cassava cv.60444 showed storage root formation in 6-8 months (Bull *et al*, 2018; Lloyd *et al*, 2018). Future studies of these lines will include optimisation of storage root formation. Two UMPS lines; 1 and 13, showed increased storage root yields which correlated (R-value: -.56) with a down-regulation in the expression of *UMP synthase* in the plants. Two ADK lines also showed down-regulation of *ADK*, but this did not result in a greater number or greater weight of storage roots. Further studies on the two UMPs lines, including starch analysis (quality and quantity) will be conducted but were not possible due to time constraints. Also, enzyme activity (*ADK* and *UMP synthase*) will be examined all transgenic lines, to establish whether down-regulation of gene has a correlation to enzyme level and activity, as some studies have shown that this is not always the case (Geigenberger *et al*, 2005).

Cassava has one of the highest starch contents of any plant (De Souza *et al*, 2016), however in sub-Saharan Africa is mainly grown by subsistent farmers as a survival crop, which impacts its yield. The increase in demand for high starch substrates for bioethanol has increased the demand for cassava (Smith, 2008), and high-starch varieties of cassava may have potential as a cash-crop for these farmers. The UMPS lines, in particular may have potential for use in bioethanol production as they produced larger storage roots with a higher dry weight percentage than the untransformed model cv.60444. However these results are only preliminary, a number of further trials, including repetition of the pot trials as well as large scale greenhouse studies and field trials will have to be conducted to confirm these results.

In summary, the aim of this study was to improve cassava cv.60444 using RNAi technology. It focused on two important aspects, viral tolerance where a hp-RNA construct targeting the AC1/AC4 region of SACMV increase tolerance to SACMV in a controlled environment. Further the study expanded on interactions of hp-RNA targeting CMG's with host genes and highlighted the importance of selection of viral targets for silencing, due to off-target effects. These off-target effects can cause phenotypic variation in cassava as well as induce cell-death in FECs. The final aim of the project was to increase starch production in cassava cv.60444, through the down-regulation of two genes *ADK* and *UMP synthase*, which have been shown to have potential in other crops. Two UMPS lines, 1 and 13, were shown to produce larger storage roots with a higher dry weight percentage , which correlated with the down-regulation of *UMP synthase*. These lines may have potential for commercial crop development, but will be further tested in future to characterise the starch in detail. While both the results of the AC1/AC4 SACMV lines and the UMPS lines are positive and are the first reports of either SACMV tolerance and increased starch in cassava respectively, these studies will have to be expanded in future to include large scale greenhouse testing and field trials. Future studies could also include stacked constructs for SACMV tolerance and *UMP synthase*, to ensure high crop yields in the field.

References

Abel, P.P., Nelson, R.S., De, B., Hoffmann, N., Rogers, S.G., Fraley, R.T., Beachy, R.N., 1986. Delay of disease development in transgenic plants that express the Tobacco Mosaic Virus coat protein gene. Science (80-.). 232, 738–743. doi:10.1126/science.3457472

Adkar-Purushothama, C.R., Perreault, J.-P., Giguère, T., Moffett, P., Sano, T., Brosseau, C., 2015. Small RNA derived from the virulence modulating region of the *Potato spindle tuber viroid* silences *callose synthase* genes of tomato plants. Plant Cell 27, 2178–2194. doi:10.1105/tpc.15.00523

Agarwal, P., Reddy, M.P., Chikara, J., 2011. WRKY: Its structure, evolutionary relationship, DNA-binding selectivity, role in stress tolerance and development of plants. Mol. Biol. Rep. 38, 3883–3896. doi:10.1007/s11033-010-0504-5

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

Akbergenov, R., Si-Ammour, A., Blevins, T., Amin, I., Kutter, C., Vanderschuren, H., Zhang, P., Gruissem, W., Meins, F., Hohn, T., Pooggin, M.M., 2006. Molecular characterization of geminivirus-derived small RNAs in different plant species. Nucleic Acids Res. 34, 462–71. doi:10.1093/nar/gkj447

Allem, A.C., Postal, C., 1994. The origin of Manihot esculenta Crantz (Euphorbiaceae). Genet. Resour. Crop Evol. 41, 133–150.

Allie, F., Pierce, E.J., Okoniewski, M.J., Rey, C., 2014. Transcriptional analysis of *South African cassava mosaic virus*-infected susceptible and tolerant landraces of cassava highlights differences in resistance, basal defense and cell wall associated genes during infection. BMC Genomics 15. doi:10.1186/1471-2164-15-1006

Alves, A.A.C., 2002. Cassava Botany and Physiology, in: Hillocks, R.J., Thresh, J.M., Bellotti, A.C. (Eds.), Cassava: Biology, Production and Utilization. CAB International, pp. 67–89.

Ammara, U., Al-Sadi, A.M., Al-Shihi, A., Amin, I., 2017. Real-time qPCR assay for the TYLCV titer in relation to symptoms-based disease severity scales. Int. J. Agric. Biol. 19, 145–151. doi:10.17957/IJAB/15.0256

Andika, I.B., Kondo, H., Tamada, T., 2007. Evidence that RNA silencing-mediated resistance to *Beet necrotic yellow vein virus* is less effective in roots than in leaves. Mol. Plant-Microbe Interact. 18, 194–204. doi:10.1094/mpmi-18-0194

Angaji, S.A., Hedayati, S.S., Hosein, R., Samad, S., Shiravi, S., Madani, S., 2010. Application of RNA interference in plants. Plant Omi. J. 3, 77–84.

ARC, 2010. Cassava Production guideline, Agriculture, forestry and fisheries.

Aregger, M., Borah, B.K., Seguin, J., Rajeswaran, R., Gubaeva, E.G., Zvereva, A.S., Windels, D., Vazquez, F., Blevins, T., Farinelli, L., Pooggin, M.M., 2012. Primary and Secondary siRNAs in geminivirus-induced gene silencing. PLoS Pathog. 8. doi:10.1371/journal.ppat.1002941

Axtell, M.J., Westholm, J.O., Lai, E.C., 2011. Vive la différence: Biogenesis and evolution of microRNAs in plants and animals. Genome Biol. doi:10.1186/gb-2011-12-4-221

Balat, M., Balat, H., 2009. Recent trends in global production and utilization of bio-ethanol fuel. Appl. Energy 86, 2273–2282. doi:10.1016/j.apenergy.2009.03.015

Barimah, J., Ellis, W.O., Oldham, J.H., Safo-Kantanka, O., Pawar, G.D., 1999. The effects of drying and varietal differences on the physicochemical properties of cassava starch. J. Ghana Sci. Assoc. 1, 55–59.

Bart, R.S., Taylor, N.J., 2017. New opportunities and challenges to engineer disease resistance in cassava, a staple food of African small-holder farmers. PLoS Pathog. 13, 1–7. doi:10.1371/journal.ppat.1006287

Basu, S., Kushwaha, N.K., Singh, A.K., Sahu, P.P., Kumar, R.V., Chakraborty, S., 2018. Dynamics of a geminivirus-encoded pre-coat protein and host RNA-dependent RNA polymerase 1 in regulating symptom recovery in tobacco. J. Exp. Bot. 69, 2085–2102. doi:10.1093/jxb/ery043

Baulcombe, D.C., 2004. RNA silencing in plants. Nature 431, 356–63.

doi:10.1038/nature02874

Baumberger, N., Baulcombe, D.C., 2005. Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. Proc. Natl. Acad. Sci. 102, 11928–11933. doi:10.1073/pnas.0505461102

Beachy, R.N., 1999. Coat-protein-mediated resistance to tobacco mosaic virus: discovery mechanisms and exploitation. Philos. Trans. R. Soc. 354, 659–664. doi:10.1098/rstb.1999.0418

Béclin, C., Boutet, S., Waterhouse, P., Vaucheret, H., 2002. A branched pathway for transgene-induced RNA silencing in plants. Curr. Biol. 12, 684–8.

Belhaj, K., Chaparro-Garcia, A., Kamoun, S., Patron, N.J., Nekrasov, V., 2015. Editing plant genomes with CRISPR/Cas9. Curr. Opin. Biotechnol. 32, 76–84. doi:10.1016/j.copbio.2014.11.007

Bellotti, A., Smith, L., Lapointe, S., 1999. Recent advances in cassava pest management. Annu. Rev. Entomol. 44, 343–370.

Bengyella, L., Rey, C., 2015. Resistance gene analogs involved in tolerant cassava–geminivirus interaction that shows a recovery phenotype. Virus Genes 51, 393–407. doi:10.1007/s11262-015-1246-1

Berrie, L.C., Rybicki, E.P., Rey, M.E.C., 2001. Complete nucleotide sequence and host range of *South African cassava mosaic* virus: further evidence for recombination amongst begomoviruses. J. Gen. Virol. 82, 53–8.

Beyene, G., Chauhan, R.D., Wagaba, H., Moll, T., Alicai, T., Miano, D., Carrington, J.C., Taylor, N.J., 2016. Loss of CMD2-mediated resistance to cassava mosaic disease in plants regenerated through somatic embryogenesis. Mol. Plant Pathol. 17, 1095–1110.

Bhattacharjee, S., Zamora, A., Azhar, M.T., Sacco, M.A., Lambert, L.H., Moffett, P., 2009. Virus resistance induced by NB-LRR proteins involves Argonaute4-dependent translational control. Plant J. 58, 940–951. doi:10.1111/j.1365-313X.2009.03832.x

Bhattacharyya, D., Gnanasekaran, P., Kumar, R.K., Kushwaha, N.K., Sharma, V.K.,

Yusuf, M.A., Chakraborty, S., 2015. A geminivirus betasatellite damages the structural and functional integrity of chloroplasts leading to symptom formation and inhibition of photosynthesis. J. Exp. Bot. 66, 5881–5895. doi:10.1093/jxb/erv299

Birch, R.G., 1997. Plant transformation: Problems and strategies for practical application. Annu. Rev. Plant Physiol. Plant Mol. Biol. 48, 297–326. doi:10.1146/annurev.arplant.48.1.297

Birmingham, A., Anderson, E.M., Reynolds, A., Ilsley-Tyree, D., Leake, D., Fedorov, Y., Baskerville, S., Maksimova, E., Robinson, K., Karpilow, J., Marshall, W.S., Khvorova, A., 2006. 3′UTR seed matches, but not overall identity, are associated with RNAi off-targets. Nat. Methods 3, 199–204. doi:10.1038/nmeth854

Blagbrough, I.S., Bayoumi, S.A.L., Rowan, M.G., Beeching, J.R., 2010. Cassava: An appraisal of its phytochemistry and its biotechnological prospects. Phytochemistry 71, 1940–1951. doi:10.1016/j.phytochem.2010.09.001

Blevins, T., Rajeswaran, R., Shivaprasad, P. V., Beknazariants, D., Si-Ammour, A., Park, H.S., Vazquez, F., Robertson, D., Meins, F., Hohn, T., Pooggin, M.M., 2006. Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. Nucleic Acids Res. 34, 6233–6246. doi:10.1093/nar/gkl886

Boch, J., 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. Science (80-.). 326, 1509–1512. doi:10.1126/science.1178811

Bogdanove, A.J., Schornack, S., Lahaye, T., 2010. TAL effectors : finding plant genes for disease and defense. Curr. Opin. Plant Biol. 13, 394–401. doi:10.1016/j.pbi.2010.04.010

Bologna, N.G., Mateos, J.L., Bresso, E.G., Palatnik, J.F., 2009. A loop-to-base processing mechanism underlies the biogenesis of plant microRNAs miR319 and miR159. EMBO J. 28, 3646–3656. doi:10.1038/emboj.2009.292

Bologna, N.G., Voinnet, O., 2014. The diversity, biogenesis, and activities of endogenous silencing small RNAs in *Arabidopsis*. Annu. Rev. Plant Biol. 65, 473–503. doi:10.1146/annurev-arplant-050213-035728

Borges, F., Martienssen, R.A., 2016. The expanding world of small RNAs in plants. Nat. Rev. Mol. Cell Biol. 16, 727–741. doi:10.1038/nrm4085.The

Boster, J.S., 1985. Selection for perceptual distinctiveness : evidence from Aguaruna cultivars of *Manihot esculenta*. Econ. Bot. 39, 310–325.

Brand, A., Quimbaya, M., Tohme, J., Chavarriaga-aguirre, P., 2019. Arabidopsis *LEC1* and *LEC2* orthologous genes are key regulators of somatic embryogenesis in cassava. Front. Plant Sci. 10, 1–14. doi:10.3389/fpls.2019.00673

Bredeson, J. V., Lyons, J.B., Prochnik, S.E., Wu, G.A., Ha, C.M., Edsinger-Gonzales, E., Grimwood, J., Schmutz, J., Rabbi, I.Y., Egesi, C., Nauluvula, P., Lebot, V., Ndunguru, J., Mkamilo, G., Bart, R.S., Setter, T.L., Gleadow, R.M., Kulakow, P., Ferguson, M.E., Rounsley, S., Rokhsar, D.S., 2016. Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. Nat. Biotechnol. 34, 562–570. doi:10.1038/nbt.3535

Briddon, R.W., Markham, P., 1995. Family Geminiviridae, in: Murphy, F.A., Fauquet, C.M., Bishop, D.H., Ghabrial, S.A., Jarvis, A.W., Martelli, G.P., Mayo, M.A., Summers., M.D. (Eds.), Virus Taxonomy: Classification and Nomenclature of Viruses. Sixth Report of the International Committee on Taxonomy of Viruses. Springer-Verlag, Vienna and New York, pp. 158–165.

Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y.Y., Sieburth, L., Voinnet, O., 2008. Widespread translational inhibition by plant miRNAs and siRNAs. Science (80-.). 320, 1185–1190. doi:10.1126/science.1159151

Brodersen, P., Voinnet, O., 2006. The diversity of RNA silencing pathways in plants. Trends Genet. 22, 268–80. doi:10.1016/j.tig.2006.03.003

Brown, J.K., Zerbini, F.M., Navas-Castillo, J., Moriones, E., Ramos-Sobrinho, R., Silva, J.C.F., Fiallo-Olivé, E., Briddon, R.W., Hernández-Zepeda, C., Idris, A., Malathi, V.G., Martin, D.P., Rivera-Bustamante, R., Ueda, S., Varsani, A., 2015. Revision of Begomovirus taxonomy based on pairwise sequence comparisons.

Arch. Virol. 160, 1593–1619. doi:10.1007/s00705-015-2398-y

Brummell, D.A., Balint-Kurti, P.J., Harpster, M.H., Palys, J.M., Oeller, P.W., Gutterson, N., 2003. Inverted repeat of a heterologous 3′-untranslated region for high-efficiency, high-throughput gene silencing. Plant J. 33, 793–800. doi:10.1046/j.1365-313X.2003.01659.x

Brunetti, A., Tavazza, M., Noris, E., Tavazza, R., Caciagli, P., Ancora, G., Crespi, S., Accotto, G.P., 1997. High expression of truncated viral Rep protein confers resistance to tomato yellow leaf curl virus in transgenic tomato plants. Mol. Plant-Microbe Interact. 10, 571–579. doi:10.1094/MPMI.1997.10.5.571

Buchmann, R.C., Asad, S., Wolf, J.N., Mohannath, G., Bisaro, D.M., 2009. Geminivirus AL2 and L2 proteins suppress transcriptional gene silencing and cause genome-wide reductions in cytosine methylation. J. Virol. 83, 5005–5013. doi:10.1128/JVI.01771-08

Budak, H., Akpinar, B.A., 2015. Plant miRNAs: biogenesis, organization and origins. Funct. Integr. Genomics 15, 523–531. doi:10.1007/s10142-015-0451-2

Bull, S., Briddon, R.W., Sserubombwe, W.S., Ngugi, K., Markham, P.G., Stanley, J., 2007. Infectivity, pseudorecombination and mutagenesis of Kenyan cassava mosaic begomoviruses. J. Gen. Virol. 88, 1624–1633. doi:10.1099/vir.0.82662-0

Bull, S., Owiti, J., Niklaus, M., Beeching, J., Gruissem, W., Vanderschuren, H., 2009. *Agrobacterium*-mediated transformation of friable embryogenic calli and regeneration of transgenic cassava. Nat. Protoc. 4, 1845–54. doi:10.1038/nprot.2009.208

Bull, S., Seung, D., Chanez, C., Mehta, D., Kuon, J.-E.E., Truernit, E., Hochmuth, A., Zurkirchen, I., Zeeman, S.C., Gruissem, W., Vanderschuren, H.H., 2018. Accelerated ex situ breeding of *GBSS*- and *PTST1*-edited cassava for modified starch. Sci. Adv. 4, eaat6086. doi:10.1126/sciadv.aat6086

Butaye, K.M.J., Cammue, B.P.A., Delauré, S.L., De Bolle, M.F.C., 2005. Approaches to minimize variation of transgene expression in plants. Mol. Breed. 16, 79–91. doi:10.1007/s11032-005-4929-9

Butterbach, P., Verlaan, M.G., Dullemans, A., Lohuis, D., Visser, R.G.., Baib, Y., Kormelink, R., 2014. Tomato yellow leaf curl virus resistance by *Ty-1* involves increased cytosine methylation of viral genomes and is compromised by cucumber mosaic virus infection. Proc. Natl. Acad. Sci. 111, 12942–12947. doi:10.1073/pnas.1400894111

Campo, B.V.H., Hyman, G., Bellotti, A., 2011. Threats to cassava production : known and potential geographic distribution of four key biotic constraints. Food Secur. 3, 329–345. doi:10.1007/s12571-011-0141-4

Cao, M., Zhou, C., Wang, X., Yu, Y.-Q., Gal-On, A., Li, W., Qiu, Y.-H., Du, P., Li, Y., Ding, S.-W., 2014. Virus infection triggers widespread silencing of host genes by a distinct class of endogenous siRNAs in *Arabidopsis*. Proc. Natl. Acad. Sci. 111, 14613–14618. doi:10.1073/pnas.1407131111

Cao, X., Lu, Y., Di, D., Zhang, Z., Liu, H., Tian, L., Zhang, A., Zhang, Y., Shi, L., Guo, B., Xu, J., Duan, X., Wang, X., Han, C., Miao, H., Yu, J., Li, D., 2013. Enhanced virus resistance in transgenic maize expressing a dsRNA-specific endoribonuclease gene from *E. coli*. PLoS One 8. doi:10.1371/journal.pone.0060829

Carbonell, A., Carrington, J.C., 2015. Antiviral roles of plant ARGONAUTES. Curr. Opin. Plant Biol. 27, 111–117. doi:10.1016/j.pbi.2015.06.013

Carla, S., Bresso, E.G., Spinelli, S. V., Palatnik, J.F., 2012. Role of MicroRNA miR319 in plant development, in: MicroRNAs in Plant Development and Stress Responses. Springer Berlin Heidelberg, pp. 29–47. doi:10.1007/978-3-642-27384-1

Caro, E., Castellano, M.M., Gutierrez, C., 2007. A chromatin link that couples cell division to root epidermis patterning in *Arabidopsis*. Nature 447, 213–218. doi:10.1038/nature05763

Castillo, A.G., Collinet, D., Deret, S., Kashoggi, A., Bejarano, E.R., 2003. Dual interaction of plant PCNA with geminivirus replication accessory protein (REn) and viral replication protein (Rep). Virology 312, 381–394. doi:10.1016/S0042-6822(03)00234-4

Castillo, A.G., Kong, L.J., Hanley-Bowdoin, L., Bejarano, E.R., 2004. Interaction between a geminivirus replication protein and the plant sumoylation system. J. Virol. 78, 2758–69. doi:10.1128/JVI.78.6.2758-2769.2004

Ceballos, H., Iglesias, C.A., Pérez, J.C., Dixon, A.G.O., 2004. Cassava breeding: opportunities and challenges. Plant Mol. Biol. 56, 503–516. doi:10.1007/s11103-004-5010-5

Chapman, E.J., Prokhnevsky, A.I., Gopinath, K., Dolja, V. V., Carrington, J.C., 2004. Viral RNA silencing suppressors inhibit the microRNA pathway at an intermediate step. Genes Dev. 18, 1510. doi:10.1101/gad.1201204.cessing

Chavarriaga-aguirre, P., Brand, A., Medina, A., Prías, M., Escobar, R., Martinez, J., Díaz, P., López, C., Roca, W.M., Tohme, J., 2016. The potential of using biotechnology to improve cassava : a review. Vitr. Cell. Dev. Biol. - Plant 461–478. doi:10.1007/s11627-016-9776-3

Chellappan, P., Masona, M. V, Vanitharani, R., Taylor, N.J., Fauquet, C.M., 2004a. Broad spectrum resistance to ssDNA viruses associated with transgene-induced gene silencing in cassava. Plant Mol. Biol. 56, 601–611. doi:10.1007/s11103-004-0147-9

Chellappan, P., Vanitharani, R., Fauquet, C.M., 2005. MicroRNA-binding viral protein interferes with Arabidopsis development. Proc. Natl. Acad. Sci. 102, 10381–10386. doi:10.1073/pnas.0504439102

Chellappan, P., Vanitharani, R., Fauquet, C.M., 2004b. Short interfering RNA accumulation correlates with host recovery in DNA virus-infected hosts, and gene silencing targets specific viral sequences. J. Virol. 78, 7465. doi:10.1128/JVI.78.14.7465

Chen, C., 2002. Potentiation of developmentally regulated plant defense response by AtWRKY18, a pathogen-induced *Arabidopsis* transcription factor. Plant Physiol. 129, 706–716. doi:10.1104/pp.001057

Chen, H.-M., Chen, L.-T., Patel, K., Li, Y.-H., Baulcombe, D.C., Wua, S.-H., 2010. 22-nucleotide RNAs trigger secondary siRNA biogenesis in plants. Proc. Natl. Acad. Sci. 107, 15269–15274. doi:10.1073/pnas.1001738107

Chen, M., Thelen, J.J., 2011. Plastid Uridine salvage activity is required for photoassimilate allocation and partitioning in Arabidopsis. Plant Cell 23, 2991–3006. doi:10.1105/tpc.111.085829

Chetty, C.C., Rossin, C.B., Gruissem, W., Vanderschuren, H., Rey, M.E., 2013. Empowering biotechnology in southern Africa: establishment of a robust transformation platform for the production of transgenic industry-preferred cassava. N. Biotechnol. 30, 136–43. doi:10.1016/j.nbt.2012.04.006

Chowda-Reddy, R. V., Achenjang, F., Felton, C., Etarock, M.T., Anangfac, M.T., Nugent, P., Fondong, V.N., 2008. Role of a geminivirus AV2 protein putative protein kinase C motif on subcellular localization and pathogenicity. Virus Res. 135, 115–124. doi:10.1016/j.virusres.2008.02.014

Closa, M., Vranova, E., Bortolotti, C., Bigler, L., Arro, M., Ferrer, A., Gruissem2, W., 2010. The *Arabidopsis thaliana* FPP synthase isozymes have overlapping and specific functions in isoprenoid biosynthesis, and complete loss of FPP synthase activity causes early developmental arrest. Plant J. 63, 512–525. doi:10.1093/mp/sss015

Csorba, T., Kontra, L., Burgyán, J., 2015. Viral silencing suppressors: Tools forged to fine-tune host-pathogen coexistence. Virology 479–480, 85–103. doi:10.1016/j.virol.2015.02.028

Dai, D., Hu, Z., Pu, G., Li, H., Wang, C., 2006. Energy efficiency and potentials of cassava fuel ethanol in Guangxi region of China. Energy Convers. Manag. 47, 1686–1699. doi:10.1016/j.enconman.2005.10.019

Dalakouras, A., Tzanopoulou, M., Tsagris, M., Wassenegger, M., Kalantidis, K., 2011. Hairpin transcription does not necessarily lead to efficient triggering of the RNAi pathway. Transgenic Res. 20, 293–304. doi:10.1007/s11248-010-9416-3

de Freitas, J.P.X., da Silva Santos, V., de Oliveira, E.J., 2016. Inbreeding depression in cassava for productive traits. Euphytica 209, 137–145. doi:10.1007/s10681-016-1649-7

de Haan, P., Gielen, J.J.., Prins, M., Wijkamp, I.G., van Schepen, A., Peters, D., van Grinsven, Ma.Q.J.., Goldbach, R., 1992. Characterization of RNA-mediated

resistance to *Tomato Spotted Wilt Virus* in transgenic tobacco plants. Plant Mol. Biol. 10. doi:10.1023/A:1005729808191

De la Morena, I., Guillén, A., del Moral, L.F.G., 1994. Yield development in potatoes as influenced by cultivar and the timing and level of nitrogen fertilization. Am. Potato J. 71, 165–173. doi:10.1007/BF02849051

De Souza, A.P., Massenburg, L.N., Jaiswal, D., Cheng, S., Shekar, R., Long, S.P., 2016. Rooting for cassava: insights into photosynthesis and associated physiology as a route to improve yield potential. New Phytol. 213, 50–65. doi:10.1111/nph.14250

Dixon, G.., Aseidu, R., Bokanga, M., 1994. Breeding of cassava for low cyanogenic potential: problems, progress and prospects. Acta Biol. Szeged. 375, 153–162.

Doench, J.G., Petersen, C.P., Sharp, P.A., 2003. siRNAs can function as miRNAs. Genes Dev. 17, 438–442. doi:10.1101/gad.1064703

Doyle, J., 1991. DNA protocols for plants–CTAB total DNA isolation., in: Molecular Techniques in Taxonomy. Springer Berlin Heidelberg, pp. 283–293.

Duan, C., Wang, C.H., Guo, H.S., 2012. Application of RNA silencing to plant disease resistance. Silence 3, 1. doi:10.1186/1758-907X-3-5

Duan, C.G., Wang, C.-H., Fang, R.-X., Guo, H.-S., 2008. Artificial MicroRNAs highly accessible to targets confer efficient virus resistance in plants. J. Virol. 82, 11084–11095. doi:10.1128/JVI.01377-08

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

Eagle, P. a, Orozco, B.M., Hanley-Bowdoin, L., 1994. A DNA sequence required for Geminivirus replication also mediates transcriptional regulation. Plant Cell 6, 1157–1170. doi:10.1105/tpc.6.8.1157

Efroni, I., Han, S.K., Kim, H.J., Wu, M.F., Steiner, E., Birnbaum, K.D., Hong, J.C., Eshed, Y., Wagner, D., 2013. Regulation of leaf maturation by chromatin-mediated modulation of cytokinin responses. Dev. Cell 24, 438–445.

doi:10.1016/j.devcel.2013.01.019

Ekwall, K., 2004. The RITS Complex - A direct link between Small RNA and Heterochromatin. Mol. Cell 13, 304–305. doi:10.1016/S1097-2765(04)00057-7

El-Sharkawy, M., 2004. Cassava biology and physiology. Plant Mol. Biol. 56, 481–501.

Elvira-matelot, E., Martínez de Alba, Á.E., Vaucheret, H., 2017. Diversity of RNA silencing pathways in plants, in: Plant Gene Silencing: Mechansism and Applications. CABI Publication, Boston, pp. 1-31.

FAO, 2018. Food Outlook.

FAO, 2016. Food and Agriculture Organization of the United Nations, Food Outlook Biannual Report on Global Food Markets, October 2016.

FAO, 2013. Save and grow: A guide to sustainable production intensification., Plant Production.

FAO, 2000a. The world cassava economy. Facts, trends and outlook, in: FAO IFAD. Rome.

FAO, 2000b. The global cassava development strategy and implementation plan, V. 1., in: Proceedings of the Validation Forum on the Global Cassava Development Strategy.

Fátyol, K., Ludman, M., Burgyán, J., 2016. Functional dissection of a plant Argonaute. Nucleic Acids Res. 44, 1384–1397. doi:10.1093/nar/gkv1371

Fauquet, C.M., Fargette, D., 1990. African cassava mosaic virus: etiology, epidemiology and control. Plant Dis. 74, 404–411.

Fernie, A.R., 2001. The contribution of plastidial phosphoglucomutase to the control of starch synthesis within the potato tuber. Planta 213, 418–426. doi:10.1007/s004250100521

Fondong, V.N., 2017. The search for resistance to Cassava Mosaic Geminiviruses: how much we have accomplished, and what lies ahead. Front. Plant Sci. 8, 1–19. doi:10.3389/fpls.2017.00408

Fondong, V.N., 2013. Geminivirus protein structure and function. Mol. Plant Pathol. 14, 635–649. doi:10.1111/mpp.12032

Fondong, V.N., Pita, J.S., Rey, M.E.C., Kochko, A., Beachy, R.N., Fauquet, C.M., de Kochko, A., 2000. Evidence of synergism between *African cassava mosaic virus* and a new double-recombinant geminivirus infecting cassava in Cameroon. J. Gen. Virol. 81, 287–97.

Fondong, V.N., Rey, C., 2018. Recent biotechnological advances in the improvement of cassava. Cassava. InTechOpen. doi:http://dx.doi.org/10.5772/intechopen.70758 Abstract

Fontenelle, M.R., Luz, D.F., Gomes, A.P.S., Florentino, L.H., Zerbini, F.M., Fontes, E.P.B., 2007. Functional analysis of the naturally recombinant DNA-A of the bipartite begomovirus *Tomato chlorotic mottle virus*. Virus Res. 126, 262–267. doi:10.1016/j.virusres.2007.02.009

Fontes, E.P., Eagle, P.A., Sipe, P.S., Luckow, V.A., Hanley-Bowdoin, L., 1994. Interaction between a geminivirus replication protein and origin DNA is essential for viral replication. J. Biol. Chem. 269, 8459–65.

Fontes, E.P., Luckow, V.A., Hanley-Bowdoin, L., 1992. A geminivirus replication protein is a sequence-specific DNA binding protein. Plant Cell 4, 597–608.

Fregene, M., Matsumura, H., Akano, A., Dixon, A., Terauchi, R., 2004. Serial analysis of gene expression (SAGE) of host-plant resistance to the cassava mosaic disease (CMD). Plant Mol. Biol. 56, 563–571. doi:10.1007/s11103-004-3477-8

Fuchs, M., 2017. Pyramiding resistance-conferring gene sequences in crops. Curr. Opin. Virol. 26, 36–42. doi:10.1016/j.coviro.2017.07.004

Fuentes, A., Carlos, N., Ruiz, Y., Callard, D., Sánchez, Y., Ochagavía, M.E., Seguin, J., Malpica-López, N., Hohn, T., Lecca, M.R., Pérez, R., Doreste, V., Rehrauer, H., Farinelli, L., Pujol, M., Pooggin, M.M., 2016. Field trial and molecular characterization of RNAi-transgenic tomato plants that exhibit resistance to *Tomato Yellow Leaf Curl* Geminivirus. Mol. Plant-Microbe Interact. 29, 197–209. doi:10.1094/MPMI-08-15-0181-R

Fuentes, A., Ramos, P.L., Fiallo, E., Callard, D., Sánchez, Y., Peral, R., Rodríguez, R., Pujol, M., 2006. Intron-hairpin RNA derived from replication associated protein *C1* gene confers immunity to Tomato Yellow Leaf Curl Virus infection in transgenic tomato plants. Transgenic Res. 15, 291–304. doi:10.1007/s11248-005-5238-0

Gaitán-Solís, E., Taylor, N.J., Siritunga, D., Stevens, W., Schachtman, D.P., 2015. Overexpression of the transporters *AtZIP1* and *AtMTP1* in cassava changes zinc accumulation and partitioning. Front. Plant Sci. 6, 1–11. doi:10.3389/fpls.2015.00492

Geigenberger, P., Langenberger, S., Wilke, I., Heineke, D., Heldt, H.W., Stitt, M., 1993. Sucrose is metabolised by sucrose synthase and glycolysis within the phloem complex of *Ricinus communis* L. seedlings. Planta 190, 446–453. doi:10.1007/BF00224782

Geigenberger, P., Regierer, B., Nunes-Nesi, A., Leisse, A., Urbanczyk-Wochniak, E., Springer, F., Dongen, J.T. van, Kossmann, J., Fernie, A.R., 2005. Inhibition of de Novo Pyrimidine synthesis in growing potato tubers leads to a compensatory stimulation of the Pyrimidine salvage pathway and a subsequent increase in biosynthetic performance. Plant Cell 17, 2077–2088. doi:10.1105/tpc.105.033548.1

Geigenberger, P., Stamme, C., Tjaden, J., Schulz, a, Quick, P.W., Betsche, T., Kersting, H.J., Neuhaus, H.E., 2001. Tuber physiology and properties of starch from tubers of transgenic potato plants with altered plastidic adenylate transporter activity. Plant Physiol. 125, 1667–78.

Ghanbari, M., Eini, O., Ebrahimi, S., 2016. Differential expression of *MYB33* and *AP2* genes and response of *TY* resistant plants to *Beet Curly Top Iran virus* infection in tomato. J. Plant Pathol. 98, 555–562.

Ghoshal, B., Sanfaçon, H., 2015. Symptom recovery in virus-infected plants: Revisiting the role of RNA silencing mechanisms. Virology 479–480, 167–179. doi:10.1016/j.virol.2015.01.008

Ghoshal, B., Sanfaçon, H., 2014. Temperature-dependent symptom recovery in

Nicotiana benthamiana plants infected with tomato ringspot virus is associated with reduced translation of viral RNA2 and requires ARGONAUTE 1. Virology 456–457, 188–197. doi:10.1016/j.virol.2014.03.026

Gomez, M.., Daniel, L.., Moll, T., Luebbert, C., Chauhan, R.., Vijayaraghavan, A., Kelley Renninger, Beyene, G., Taylor, N., Carrington, J.C., Staskawicz, B., Bart, R., 2018. Simultaneous CRISPR/Cas9-mediated editing of cassava *eIF4E* isoforms *nCBP-1* and *nCBP-2* confers elevated resistance to cassava brown streak disease. Mol. Plant Pathol. July. doi:10.1111/pbi.12987

Gonzalez, A.., Schopke, C., Taylor, N.J., Beachy, R.N., Fauquet, C.M., 1998. Regeneration of transgenic cassava plants (*Manihot esculenta* Crantz) through *Agrobacterium*- mediated transformation of embryogenic suspension cultures. Plant Cell Rep. 17, 827–831.

Guo, C., Li, L., Wang, X., Liang, C., 2015. Alterations in siRNA and miRNA expression profiles detected by deep sequencing of transgenic rice with siRNA-mediated viral resistance. PLoS One 10, 1–13. doi:10.1371/journal.pone.0116175

Gutierrez, C., 1999. Geminivirus DNA replication. Cell. Mol. Life Sci. 56, 313–29.

Hagen, C., Rojas, M.R., Kon, T., Gilbertson, R.L., 2008. Recovery from *Cucurbit leaf crumple virus* (Family *Geminiviridae* , Genus *Begomovirus*) infection is an adaptive antiviral response associated with changes in viral small RNAs. Phytopathology 98, 1029–1037. doi:10.1094/phyto-98-9-1029

Hahn, S., Terry, E., Leuschner, K., 1980. Breeding cassava for resistance to cassava mosaic disease. Euphytica 29, 673–683.

Hahn, S.K., Keyser, J., 1985. Cassava: a basic food of Africa. Outlook Agric. 14, 95–99. doi:10.1177/003072708501400207

Haible, D., Kober, S., Jeske, H., 2006. Rolling circle amplification revolutionizes diagnosis and genomics of geminiviruses. J. Virol. Methods 135, 9–16. doi:10.1016/j.jviromet.2006.01.017

Hall, T., 1999. BioEdit: a user-friendly biological sequence alignment editor and

analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. doi:citeulike-article-id:691774

Hamilton, A., Baulcombe, D.C., 1999. A species of small antisense RNA in posttranscriptional gene silencing in plants. Science (80-.). 286, 950.

Hammond, S.M., 2005. Dicing and slicing: The core machinery of the RNA interference pathway. FEBS Lett. 579, 5822–5829. doi:10.1016/j.febslet.2005.08.079

Hanley-Bowdoin, L., Bejarano, E.R., Robertson, D., Mansoor, S., 2013. Geminiviruses: masters at redirecting and reprogramming plant processes. Nat. Rev. Microbiol. 11, 777–88. doi:10.1038/nrmicro3117

Hanley-Bowdoin, L., Settlage, S.B., Orozco, B.M., Nagar, S., Robertson, D., 1999. Geminiviruses: models for plant DNA replication, transcription, and cell cycle regulation. Crit. Rev. Plant Sci. 18, 71–106.

Hanley-Bowdoin, L., Settlage, S.B., Robertson, D., 2004. Reprogramming plant gene expression: a prerequisite to geminivirus DNA replication. Mol. Plant Pathol. 5, 149–156. doi:10.1111/J.1364-3703.2004.00214.X

Harrison, B., Robinson, D., 1999. Natural genomic and antigenic variation in whitefly-transmitted geminiviruses (Begomoviruses). Annu. Rev. Phytopathol. 37, 369–98.

Harvey, J.J.W., Lewsey, M.G., Patel, K., Westwood, J., Heimstädt, S., Carr, J.P., Baulcombe, D.C., 2011. An antiviral defense role of AGO2 in plants. PLoS One 6, 1–6. doi:10.1371/journal.pone.0014639

Havecker, E.R., Wallbridge, L.M., Hardcastle, T.J., Bush, M.S., Kelly, K.A., Dunn, R.M., Schwach, F., Doonan, J.H., Baulcombe, D.C., 2010. The Arabidopsis RNA-Directed DNA methylation Argonautes functionally diverge based on their expression and interaction with target loci. Plant Cell 22, 321–334. doi:10.1105/tpc.109.072199

Henikoff, S., Till, B.J., Comai, L., Division, B.S., Hutchinson, F., Washington, S.H., 2004. TILLING . Traditional mutagenesis meets functional genomics. Perspect.

Transl. Biol. 135, 630–636. doi:10.1104/pp.104.041061.630

Ho, T., Wang, H., Pallett, D., Dalmay, T., 2007. Evidence for targeting common siRNA hotspots and GC preference by plant Dicer-like proteins. FEBS Lett. 581, 3267–3272. doi:10.1016/j.febslet.2007.06.022

Hofgen, R., Willmitzer, L., 1988. Storage of competent cells for *Agrobacterium* transformation. Nucleic Acids Res. 16, 9877. doi:10.1093/nar/16.20.9877

Holsters, M., de Waele, D., Depicker, A., Messens, E., van Montagu, M., Schell, J., 1978. Transfection and transformation of *Agrobacterium tumefaciens*. MGG Mol. Gen. Genet. 163, 181–187. doi:10.1007/BF00267408

Hoover, R., 2001. Composition, molecular structure, and physicochemical properties of tuber and root starches: A review. Carbohydr. Polym. 45, 253–267. doi:10.1016/S0144-8617(00)00260-5

Hu, M., Hu, W., Zhou, X., Xia, Z., Wang, W., 2016. Validation of reference genes for relative quantitative gene expression studies in cassava (*Manihot esculenta* Crantz) by using quantitative real-time PCR. Front. Plant Sci. 7, 1–12. doi:10.3389/fpls.2016.00680

Ihemere, U., Arias-Garzon, D., Lawrence, S., Sayre, R., 2006. Genetic modification of cassava for enhanced starch production. Plant Biotechnol. J. 4, 453–65. doi:10.1111/j.1467-7652.2006.00195.x

Ihemere, U.E., Narayanan, N.N., Sayr, R.T., 2012. Iron biofortification and homeostasis in transgenic cassava roots expressing the algal iron assimilatory gene, *FEA1*. Front. Plant Sci. 3, 1–22. doi:10.3389/fpls.2012.00171

Ilyas, M., Qazi, J., Mansoor, S., Briddon, R.W., 2010. Genetic diversity and phylogeography of begomoviruses infecting legumes in Pakistan. J. Gen. Virol. 91, 2091–2101. doi:10.1099/vir.0.020404-0

Jackson, A.L., Bartz, S.R., Schelter, J., Kobayashi, S. V, Burchard, J., Mao, M., Li, B., Cavet, G., Linsley, P.S., 2003. Expression profiling reveals off-target gene regulation by RNAi. Nat. Biotechnol. 21, 635–638.

Jackson, A.L., Burchard, J., Schelter, J., Chau, B.N., Cleary, M., Lim, L.E.E., Linsley,

P.S., 2006. Widespread siRNA '"off-target"' transcript silencing mediated by seed region sequence complementarity. RNA 12, 1179–1187. doi:10.1261/rna.25706.gene

Jaganathan, D., Ramasamy, K., Sellamuthu, G., Jayabalan, S., Venkataraman, G., 2018. CRISPR for crop improvement: an update review. Front. Plant Sci. 9, 1–17. doi:10.3389/fpls.2018.00985

Jain, R., Iglesias, N., Moazed, D., 2016. Distinct functions of Argonaute slicer in siRNA maturation and Heterochromatin formation. Mol. Cell 63, 191–205. doi:10.1016/j.molcel.2016.05.039

Jani, J., Samad, A.F.A., Ismail, I., Sajad, M., Murad, A.M.A., 2018. Data on degradome sequencing and analysis from mock-inoculated and *Fusarium oxysporum* treated leaves samples in *Persicaria minor*. Data Br. 20, 555–557. doi:10.1016/j.dib.2018.08.034

Jennings, D.L., Iglesias, C., 2009. Breeding for crop improvement., in: Cassava: Biology, Production and Utilization. pp. 149–166. doi:10.1079/9780851995243.0149

Jeske, H., Lütgemeier, M., Preiß, W., 2001. DNA forms indicate rolling circle and recombination-dependent replication of *Abutilon* mosaic virus. EMBO J. 20, 6158–6167. doi:10.1093/emboj/20.21.6158

Jiang, W., Yu, D., 2009. Arabidopsis WRKY2 transcription factor mediates seed germination and postgermination arrest of development by abscisic acid. BMC Plant Biol. 9, 1–14. doi:10.1186/1471-2229-9-96

Jones, L., Ratcliff, F., Baulcombe, D.C., 2001. RNA-directed transcriptional gene silencing in plants can be inherited independently of the RNA trigger and requires Met1 for maintenance. Curr. Biol. 11, 747–757. doi:10.1016/S0960-9822(01)00226-3

Jovel, J., Walker, M., Sanfacon, H., 2007. Recovery of Nicotiana benthamiana plants from a necrotic response induced by a Nepovirus is associated with RNA silencing but not with reduced virus titer. J. Virol. 81, 12285–12297. doi:10.1128/jvi.01192-07

Kalantidis, K., Psaradakis, S., Tabler, M., Tsagris, M., 2007. The occurrence of CMV-Specific short RNAs in transgenic tobacco expressing virus-derived double-stranded RNA is indicative of resistance to the virus. Mol. Plant-Microbe Interact. 15, 826–833. doi:10.1094/mpmi.2002.15.8.826

Kammerer, B., Fischer, K., Hilpert, B., Schubert, S., Gutensohn, M., Weber, A., Flügge, U.I., 1998. Molecular characterization of a carbon transporter in plastids from heterotrophic tissues: the glucose 6-phosphate/phosphate antiporter. Plant Cell 10, 105–17.

Karlström, A., Calle, F., Salazar, S., Morante, N., Dufour, D., Ceballos, H., 2016. Biological implications in cassava for the production of amylose-free starch: impact on root yield and related traits. Front. Plant Sci. 7, 1–8. doi:10.3389/fpls.2016.00604

Kasschau, K.D., Xie, Z., Allen, E., Llave, C., Chapman, E.J., Krizan, K.A., Carrington, J.C., 2003. P1/HC-Pro, a viral suppressor of RNA silencing, interferes with Arabidopsis development and miRNA function. Dev. Cell 4, 205–217. doi:10.1016/S1534-5807(03)00025-X

Katiyar-Agarwal, S., Jin, H., 2010. Role of Small RNAs in Host-Micobe interactions. Annu. Rev. Phytopathol. 48, 225–246. doi:10.1146/annurev-phyto-073009-114457.

Kaweesi, T., Kawuki, R., Kyaligonza, V., Baguma, Y., Tusiime, G., Ferguson, M.E., 2014. Field evaluation of selected cassava genotypes for cassava brown streak disease based on symptom expression and virus load. Virol. J. 11. doi:10.1186/s12985-014-0216-x

Ketting, R.F., Fischer, S.E.J., Bernstein, E., Sijen, T., Hannon, G.J., Plasterk, R.H.A., 2001. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. Genes Dev. 15, 2654–2659. doi:10.1101/gad.927801.2654

Khalid, A., Zhang, Q., Yasir, M., Li, F., 2017. Small RNA based genetic engineering for plant viral resistance: Application in crop protection. Front. Microbiol. 8, 1–11. doi:10.3389/fmicb.2017.00043

Khan, A. a, Betel, D., Miller, M.L., Sander, C., Leslie, C.S., Debora, S., 2009. Transfection of small RNAs globally perturbs gene regulation by endogenous microRNAs. Nat. Biotechnol. 27, 549–555. doi:10.1038/nbt.1543.Transfection

Kim, K., Lee, Y.S., Carthew, R.W., 2007. Conversion of pre-RISC to holo-RISC by Ago2 during assembly of RNAi complexes. RNA 13, 22–29. doi:10.1261/rna.283207

Kincaid, R.P., Sullivan, C.S., 2012. Virus-Encoded microRNAs: An overview and a look to the future. PLoS Pathog. 8. doi:10.1371/journal.ppat.1003018

Kong, L.-J., Hanley-Bowdoin, L., 2002. A geminivirus replication protein interacts with a protein kinase and a motor protein that display different expression patterns during plant development and infection. Plant Cell 14, 1817–1832. doi:10.1105/tpc.003681.AL1

Kong, L.-L., M.Orozco, B., Roe, J.L., Nagar, S., Ou, S., Feiler, H.S., Durfee, T., Miller, A.B., Gruissem, W., Robertson, D., Hanley-Bowdoin, L., 2000. A geminivirus replication protein interacts with the retinoblastoma protein through a novel domain to determine symptoms and tissue specificity of infection in plants. EMBO J. 19, 3485–3495. doi:10.1093/emboj/19.13.3485

Koyama, T., Sato, F., Ohme-Takagi, M., 2017. Roles of miR319 and TCP transcription factors in leaf development. Plant Physiol. 175, pp.00732.2017. doi:10.1104/pp.17.00732

Kumar, G.N.M., Iyer, S., Knowles, N.R., 2007. Extraction of RNA from fresh, frozen, and lyophilized tuber and root tissues. J. Agric. Food Chem. 55, 1674–1678. doi:10.1021/jf062941m

Kumar, S., Tanti, B., Patil, B.L., Mukherjee, S.K., Sahoo, L., 2017. RNAi-derived transgenic resistance to *Mungbean yellow mosaic India virus* in cowpea. PLoS One 12, 1–20. doi:10.1371/journal.pone.0186786

Kunik, T., Salomon, R., Zamir, D., Navot, N., Zeidan, M., Michelson, I., Gafni, Y., Czosnek, H., 1994. Transgenic tomato plants expressing the tomato yellow leaf curl virus capsid protein are resistant to the virus. Nat. Biotechnol. 12, 500–504. doi:10.1038/ng1294-340

Kuria, P., Ilyas, M., Ateka, E., Miano, D., Onguso, J., Carrington, J.C., Taylor, N.J., 2017. Differential response of cassava genotypes to infection by cassava mosaic geminiviruses. Virus Res. 227, 69–81. doi:10.1016/j.virusres.2016.09.022

Kusaba, M., 2004. RNA interference in crop plants. Curr. Opin. Biotechnol. 15, 139–43. doi:10.1016/j.copbio.2004.02.004

Kutnjak, D., Rupar, M., Gutierrez-Aguirre, I., Curk, T., Kreuze, J.F., Ravnikar, M., 2015. Deep sequencing of virus-derived small interfering RNAs and RNA from viral particles shows highly similar mutational landscapes of a plant virus population. J. Virol. 89, 4760–4769. doi:10.1128/jvi.03685-14

Lapidot, M., Friedmann, M., 2002. Breeding for resistance to whitefly-transmitted geminiviruses. Ann. Appl. Biol. 140, 109–127. doi:10.1111/j.1744-7348.2002.tb00163.x

Laufs, J., Jupin, I., David, C., Schumacher, S., Heyraud-Nitschke, F., Gronenborn, B., 1995. Geminivirus replication: genetic and biochemical characterization of Rep protein function, a review. Biochimie 77, 765–73.

Lefeuvre, P., Becker, N., Hoareau, M., Delatte, H., Naze, F., Martin, D.P., Reynaud, B., Thierry, M., Varsani, A., Lett, J.-M., 2007. Begomovirus “melting pot” in the south-west Indian Ocean islands: molecular diversity and evolution through recombination. J. Gen. Virol. 88, 3458–3468. doi:10.1099/vir.0.83252-0

Legg, J., 1999. Emergence, spread and strategies for controlling the pandemic of cassava mosaic virus disease in east and central Africa. Crop Prot. 18, 627–637.

Legg, J.P., Lava Kumar, P., Makeshkumar, T., Tripathi, L., Ferguson, M., Kanju, E., Ntawuruhunga, P., Cuellar, W., 2015. Cassava virus diseases: Biology, epidemiology, and management. Adv. Virus Res. 91, 85–142. doi:10.1016/bs.aivir.2014.10.001

Legg, J.P., Thresh, J.M., 2000. Cassava mosaic virus disease in East Africa: a dynamic disease in a changing environment. Virus Res. 71, 135–49. doi:10.1016/S0168-1702(00)00194-5

Legg, P., Fauquet, C.M., 2004. Cassava mosaic disease in Africa. Plant Mol. Biol. 56, 1–14.

Leibman, D., Wolf, D., Saharan, V., Zelcer, A., Arazi, T., Yoel, S., Gaba, V., Gal-On, A., 2011. A high level of transgenic viral small RNA is associated with broad Potyvirus resistance in cucurbits. Mol. Plant-Microbe Interact. 24, 1220–1238. doi:10.1094/mpmi-05-11-0128

Leyva-Guerrero, E., Narayanan, N.N., Ihemere, U., Sayre, R.T., 2012. Iron and protein biofortification of cassava: Lessons learned. Curr. Opin. Biotechnol. 23, 257–264. doi:10.1016/j.copbio.2011.12.009

Li, F., Huang, C., Li, Z., Zhou, X., 2014. Suppression of RNA Silencing by a plant DNA virus satellite requires a host Calmodulin-Like protein to repress RDR6 expression. PLoS Pathog. 10, 11–14. doi:10.1371/journal.ppat.1003921

Li, F., Wang, Y., Zhou, X., 2017. SGS3 cooperates with RDR6 in triggering geminivirus-induced gene silencing and in suppressing geminivirus infection in *Nicotiana benthamiana*. Viruses 9. doi:10.3390/v9090247

Li, F., Xu, X., Huang, C., Gu, Z., Cao, L., Hu, T., Ding, M., Li, Z., Zhou, X., 2015. The AC5 protein encoded by *Mungbean yellow mosaic India virus* is a pathogenicity determinant that suppresses RNA silencing-based antiviral defenses. New Phytol. 208, 555–569. doi:10.1111/nph.13473

Li, H.-Q., Sautter, C., Potrykus, I., Puonti-Kaerlas, J., 1996. Genetic transformation of cassava (*Manihot esculenta* Crantz). Nat. Biotechnol. 14, 1246–1251. doi:10.1038/ng0496-417

Li, K. Te, Moulin, M., Mangel, N., Albersen, M., Verhoeven-Duif, N.M., Ma, Q., Zhang, P., Fitzpatrick, T.B., Gruissem, W., Vanderschuren, H., 2015. Increased bioavailable Vitamin B6 in field-grown transgenic cassava for dietary sufficiency. Nat. Biotechnol. 33, 1029. doi:10.1038/nbt.3318

Li, Y., Li, C., Ding, G., Jin, Y., 2011. Evolution of MIR159 / 319 microRNA genes and their post-transcriptional regulatory link to siRNA pathways. BMC Evol. Biol. 11, 122. doi:10.1186/1471-2148-11-122

Liang, X. hai, Hart, C.E., Crooke, S.T., 2013. Transfection of siRNAs can alter miRNA levels and trigger non-specific protein degradation in mammalian cells. Biochim. Biophys. Acta - Gene Regul. Mech. 1829, 455–468. doi:10.1016/j.bbagrm.2013.01.011

Ligaba-Osena, A., Jones, J., Donkor, E., Chandrayan, S., Pole, F., Wu, C.-H., Vieille, C., Adams, M.W.W., Hankoua, B.B., 2018. Novel bioengineered cassava expressing an archaeal starch degradation system and a bacterial ADP-Glucose Pyrophosphorylase for starch self-digestibility and yield increase. Front. Plant Sci. 9, 1–15. doi:10.3389/fpls.2018.00192

Lin, C.Y., Ku, H.M., Chiang, Y.H., Ho, H.Y., Yu, T.A., Jan, F.J., 2012a. Development of transgenic watermelon resistant to *Cucumber mosaic virus* and *Watermelon mosaic virus* by using a single chimeric transgene construct. Transgenic Res. 21, 983–993. doi:10.1007/s11248-011-9585-8

Lin, C.Y., Tsia, W.-S., Ku, H.-M., Jan, F.-J., 2012b. Evaluation of DNA fragments covering the entire genome of a monopartite begomovirus for induction of viral resistance in transgenic plants via gene silencing. Transgenic Res. 21, 231–241. doi:10.1007/s11248-011-9523-9

Lin, H.C., Morcillo, F., Dussert, S., Tranchant-Dubreuil, C., Tregear, J.W., Tranbarger, T.J., 2009. Transcriptome analysis during somatic embryogenesis of the tropical monocot *Elaeis guineensis*: Evidence for conserved gene functions in early development. Plant Mol. Biol. 70, 173–192. doi:10.1007/s11103-009-9464-3

Lin, S.S., Wu, H.W., Elena, S.F., Chen, K.C., Niu, Q.W., Yeh, S.D., Chen, C.C., Chua, N.H., 2009. Molecular evolution of a viral non-coding sequence under the selective pressure of amiRNA-mediated silencing. PLoS Pathog. 5. doi:10.1371/journal.ppat.1000312

Lin, X., Ruan, X., Anderson, M.G., Mcdowell, J.A., Kroeger, P.E., Fesik, S.W., Shen, Y., Laboratories, A., Road, A.P., Park, A., 2005. siRNA-mediated off-target gene silencing triggered by a 7 nt complementation. Nucleic Acids Res. 33, 4527–4535. doi:10.1093/nar/gki762

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. Plant Cell 5, 1749–1759.

Liu, J., Zheng, Q., Ma, Q., Gadidasu, K.K., Zhang, P., 2011. Cassava genetic transformation and its application in breeding. J. Integr. Plant Biol. 53, 552–569. doi:10.1111/j.1744-7909.2011.01048.x

Liu, J.H., Luo, M., Cheng, K.J., Mohapatra, S.S., Hill, R.D., 1999. Identification and characterization of a novel barley gene that is ABA-inducible and expressed specifically in embryo and aleurone. J. Exp. Bot. 50, 727–728. doi:10.1093/jxb/50.334.727

Liu, S., Bedford, I.D., Briddon, R.W., Markham, P.G., 1997. Efficient whitefly transmission of African cassava mosaic geminivirus requires sequences from both genomic components. J. Gen. Virol. 78, 1791–1794. doi:10.1099/0022-1317-78-7-1791

Liu, S.R., Zhou, J.J., Hu, C.G., Wei, C.L., Zhang, J.Z., 2017. MicroRNA-mediated gene silencing in plant defense and viral counter-defense. Front. Microbiol. 8, 1–12. doi:10.3389/fmicb.2017.01801

Liu, Z.R., Scorza, R., Hily, J.M., Scott, S.W., James, D., 2007. Engineering resistance to multiple Prunus fruit viruses through expression of chimeric hairpins. J. Am. Soc. Hortic. Sci. 132, 407–414.

Llave, C., 2010. Virus-derived small interfering RNAs at the core of plant-virus interactions. Trends Plant Sci. 15, 701–707. doi:10.1016/j.tplants.2010.09.001

Lloyd, J.R., Hostettler, C.E., Kossmann, J., Wang, W., Zeeman, S.C., Damberger, F.F., 2018. Modification of cassava root starch phosphorylation enhances starch functional properties. Front. Plant Sci. 9, 1–16. doi:10.3389/fpls.2018.01562

Lloyd, J.R., Kossmann, J., 2019. Starch Trek: The Search for Yield. Front. Plant Sci. 9, 1–8. doi:10.3389/fpls.2018.01930

Loef, I., Stitt, M., Geigenberger, P., Planta, S., April, N., Loef, I., Stitt, M., Geigenberger, P., 2001. Increased levels of adenine nucleotides modify the

interaction between starch synthesis and respiration when adenine is supplied to discs from growing potato tubers. Planta 212, 782–791.

Lu, S., Cullen, B.R., 2004. *Adenovirus* VA1 Noncoding RNA can inhibit small interfering RNA and MicroRNA biogenesis. Jounal Virol. 78, 12868–12876. doi:10.1128/JVI.78.23.12868

Lucht, J.M., 2015. Public acceptance of plant biotechnology and GM crops. Viruses 7, 4254–4281. doi:10.3390/v7082819

Lucioli, A., Noris, E., Brunetti, A., Tavazza, R., Ruzza, V., Castillo, A.G., Bejarano, E.R., Accotto, G.P., Tavazza, M., 2003. *Tomato Yellow Leaf Curl Sardinia Virus* Rep-derived resistance to homologous and heterologous geminiviruses occurs by different mechanisms and is overcome if virus-mediated transgene silencing is activated. Microbiology. doi:10.1128/JVI.77.12.6785

Ma, L., Dong, J., Jin, Y., Chen, M., Shen, X., Wang, T., 2011. RMDAP: a versatile, ready-to-use toolbox for multigene genetic transformation. PLoS One 6, e19883. doi:10.1371/journal.pone.0019883

Ma, Q., Zhou, W., Zhang, P., 2015. Transition from somatic embryo to friable embryogenic callus in cassava: dynamic changes in cellular structure, physiological status, and gene expression profiles. Front. Plant Sci. 6, 1–14. doi:10.3389/fpls.2015.00824

Ma, X., Nicole, M.C., Meteignier, L.V., Hong, N., Wang, G., Moffett, P., 2015. Different roles for RNA silencing and RNA processing components in virus recovery and virus-induced gene silencing in plants. J. Exp. Bot. 66, 919–932. doi:10.1093/jxb/eru447

Machado, J.P.B., Calil, I.P., Santos, A.A., Fontes, E.P.B., 2017. Translational control in plant antiviral immunity. Genet. Mol. Biol. 1 (suppl), 292–304.

Machczyńska, J., Zimny, J., Bednarek, P.T., 2015. Tissue culture-induced genetic and epigenetic variation in triticale (× *Triticosecale* spp. Wittmack ex A. Camus 1927) regenerants. Plant Mol. Biol. 89, 279–292. doi:10.1007/s11103-015-0368-0

Maghuly, F., Ramkat, R.C., Laimer, M., 2014. Virus versus host plant microRNAs: Who determines the outcome of the interaction? PLoS One 9. doi:10.1371/journal.pone.0098263

Mallory, A., Vaucheret, H., 2010. Form, function, and regulation of ARGONAUTE proteins. Plant Cell 22, 3879–3889. doi:10.1105/tpc.110.080671

Mamta, B., Rajam, M.V., 2018. RNA Interference: A promising approach for crop improvement, in: Gosal, S.., Wani, S.H. (Eds.), Biotechnologies of Crop Improvement. Springer, Cham, pp. 41–65. doi:10.1007/978-3-319-78283-6

Mansoor, S., Briddon, R.W., Zafar, Y., Stanley, J., 2003. Geminivirus disease complexes: an emerging threat. Trends Plant Sci. 8, 128–34. doi:10.1016/S1360-1385(03)00007-4

Manzano, D., Andrade, P., Caudepón, D., Altabella, T., Arró, M., Ferrer, A., 2016. Suppressing Farnesyl Diphosphate Synthase alters chloroplast development and triggers sterol-dependent induction of Jasmonate- and Fe-related responses. Plant Physiol. 172, 93–117. doi:10.1104/pp.16.00431

Maredza, A.T., Allie, F., Plata, G., Rey, M.E.C., 2016. Sequences enhancing cassava mosaic disease symptoms occur in the cassava genome and are associated with *South African cassava mosaic virus* infection. Mol. Genet. Genomics 291, 1467–1485. doi:10.1007/s00438-015-1049-z

Masuta, C., Shimura, H., 2013. RNA silencing against viruses: molecular arms race between *Cucumber mosaic virus* and its host. J. Gen. Plant Pathol. 79, 227–232. doi:10.1007/s10327-013-0448-4

Mbanzibwa, D.R., Tian, Y.P., Tugume, A.K., Mukasa, S.B., Tairo, F., Kyamanywa, S., Kullaya, A., Valkonen, J.P.T., 2011. 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 171, 394–400. doi:10.1016/j.jviromet.2010.09.024

McKibbin, R.S., Muttucumaru, N., Paul, M.J., Powers, S.J., Burrell, M.M., Coates, S., Purcell, P.C., Tiessen, A., Geigenberger, P., Halford, N.G., 2006. Production of high-starch, low-glucose potatoes through over-expression of the metabolic

regulator *SnRK1*. Plant Biotechnol. J. 4, 409–18. doi:10.1111/j.1467-7652.2006.00190.x

Mehta, D., Hirsch-hoffmann, M., Were, M., Patrignani, A., Were, H., 2018a. A new full-length virus genome sequencing method reveals that antiviral RNAi changes geminivirus populations in field-grown cassava 1–24.

Mehta, D., Stürchler, A., Anjanappa, R.B., Zaidi, S.S.E.A., Hirsch-Hoffmann, M., Gruissem, W., Vanderschuren, H., 2019. Linking CRISPR-Cas9 interference in cassava to the evolution of editing-resistant geminiviruses. Genome Biol. 20, 80. doi:10.1186/s13059-019-1678-3

Mehta, D., Stürchler, A., Hirsch-hoffmann, M., Gruissem, W., 2018b. CRISPR-Cas9 interference in cassava linked to the evolution of editing- resistant geminiviruses. bioRxiv 1–21. doi:10.1101/314542

Melgarejo, T.A., Kon, T., Rojas, M.R., Paz-Carrasco, L., Zerbini, F.M., Gilbertson, R.L., 2013. Characterization of a New World monopartite Begomovirus causing leaf curl disease of tomato in Ecuador and Peru Reveals a new direction in geminivirus evolution. J. Virol. 87, 5397–5413. doi:10.1128/jvi.00234-13

Mette, M.F., Aufsatz, W., van der Winden, J., Matzke, M. a, Matzke, a J.M., Winden, J. Van Der, Matzke, M. a, Matzke, a J.M., 2000. Transcriptional silencing and promoter methylation triggered by double-stranded RNA. EMBO J. 19, 5194–5201. doi:10.1093/emboj/19.19.5194

Miozzi, L., Pantaleo, V., Burgyán, J., Accotto, G.P., Noris, E., 2013. Analysis of small RNAs derived from tomato yellow leaf curl Sardinia virus reveals a cross reaction between the major viral hotspot and the plant host genome. Virus Res. 178, 287–296. doi:10.1016/j.virusres.2013.09.029

Mlotshwa, S., Pruss, G.J., MacArthur, J.L., Reed, J.W., Vance, V., 2016. Developmental defects mediated by the P1/HC-Pro Potyviral silencing suppressor are not due to misregulation of AUXIN RESPONSE FACTOR 8. Plant Physiol. 172, 1853–1861. doi:10.1104/pp.16.01030

Mohamad Hamed Ghodoum, P., Leila, R., Babak Pakdaman, S., 2018. Temperature affected transmission, symptom development and accumulation of *Wheat dwarf*

virus. Plant Prot. Sci. 54, 222–233. doi:10.17221/147/2017-pps

Moissiard, G., Voinnet, O., 2006. RNA of host transcripts silencing by *Cauliflower mosaic virus* coordinated action of the requires four proteins *Arabidopsis* Dicer-Like Proteins. Proc. Natl. Acad. Sci. United States Am. 103, 19593–19598.

Molnár, A., Csorba, T., Lakatos, L., 2005. Plant virus-derived small interfering RNAs originate predominantly from highly structured single-stranded viral RNAs. J. Virol. 79, 7812. doi:10.1128/JVI.79.12.7812

Monci, F., Sánchez-Campos, S., Navas-Castillo, J., Moriones, E., 2002. A natural recombinant between the geminiviruses *Tomato yellow leaf curl Sardinia virus* and Tomato yellow leaf curl virus exhibits a novel pathogenic phenotype and is becoming prevalent in Spanish populations. Virology 303, 317–326. doi:10.1006/viro.2002.1633

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 Pathol. 50, 768–775.

Moralo, M., 2015. Evaluation of transgenic cassava expressing mismatch and non-mismatch constructs derived from *African cassava mosaic virus* and *South African cassava mosaic virus* open reading frames. Univesity of the Witwatersrand.

Morgado, L., Jansen, R.C., Johannes, F., 2017. Learning sequence patterns of AGO-sRNA affinity from high-throughput sequencing libraries to improve in silico functional small RNA detection and classification in plants. bioRxiv 173575. doi:10.1101/173575

Morilla, G., Castillo, A.G., Preiss, W., 2006. A versatile transreplication-based system to identify cellular proteins involved in geminivirus replication. J. Virol. 80, 3624–3633. doi:10.1128/JVI.80.7.3624

Morilla, G., Jeske, H., Bejarano, E.R., Wege, C., 2004. Teˆte a`Teˆte of *Tomato Yellow Leaf Curl Virus* and *Tomato Yellow Leaf Curl Sardinia Virus* in single nuclei. J. Virol. 78, 10715–10723. doi:10.1128/JVI.78.19.10715–10723.2004

Mukharjee, A., 1995. Embryogenesis and regeneration from cassava calli of anther and leaf, in: International Scientific Meeting (2, 1984, Bogor, Indonesia). The Cassava Biotechnology Network: Proceedings..

Munyikwa, T.R.., Raemakers, K., Schreuder, M., Kok, R., Schippers, M., Jacobsen, E., Visser, R.G.., 1998. Pinpointing towards improved transformation and regeneration of cassava (*Manihot esculenta* Crantz). Plant Sci. 135, 87–101. doi:10.1016/S0168-9452(98)00074-0

Munyikwa, T.R.I., Raemakers, C.C.J.M., Schreuder, M., Kreuze, J., Suurs, L., Kok, R., Rozeboom, M., Jacobsen, E., Visser, R.G.F., 1998. Introduction and expression of antisense ADP-glucose pyrophosphorylase of cassava leads to decreased levels of starch and increased levels of sugars, in: Proceedings of the IVth International Science Meeting of the Cassava Biotechnology Network Re. Brazil (Pires de Matos, A. and Vilarinhos, A., Eds.). EMBRAPA, Brasilia DF, Brazi, p. 62.

Naqvi, A.R., Haq, Q.M., Mukherjee, S.K., 2010. MicroRNA profiling of *Tomato leaf curl new delhi virus* (ToLCNDv) infected tomato leaves indicates that deregulation of mir159/319 and mir172 might be linked with leaf curl disease. Virol. J. 7, 1–16. doi:10.1186/1743-422X-7-281

Narayanan, N., Beyene, G., Chauhan, R.D., Gaitán-Solis, E., Grusak, M.A., Taylor, N., Anderson, P., 2015. Overexpression of *Arabidopsis* VIT1 increases accumulation of iron in cassava roots and stems. Plant Sci. 240, 170–181. doi:10.1016/j.plantsci.2015.09.007

Narry, V., 2014. MicroRNA biogenesis: Co-ordinated cropping and dicing. Plant Physiol., Methods in Molecular Biology 6, 1–2. doi:10.1017/CBO9781107415324.004

Nassar, N.M.A., Ortiz, R., 2006. Cassava improvement: challenges and impacts. J. Agric. Sci. 145, 163. doi:10.1017/S0021859606006575

Nawaz-ul-Rehman, M.S., Fauquet, C.M., 2009. Evolution of geminiviruses and their satellites. FEBS Lett. 583, 1825–1832. doi:10.1016/j.febslet.2009.05.045

Ndunguru, J., De León, L., Doyle, C.D., Sseruwagi, P., Plata, G., Legg, J.P.,

Thompson, G., Tohme, J., Aveling, T., Ascencio-Ibáñez, J.T., Hanley-Bowdoin, L., 2016. Two Novel DNAs that enhance symptoms and overcome CMD2 resistance to cassava mosaic disease. J. Virol. 90, 4160–4173. doi:10.1128/JVI.02834-15

Ndunguru, J., Legg, J.P., Aveling, T.A.S., Thompson, G., Fauquet, C.M., 2005. Molecular biodiversity of cassava begomoviruses in Tanzania: Evolution of cassava geminiviruses in Africa and evidence for East Africa being a center of diversity of cassava geminiviruses. Virol. J. 2, 1–23. doi:10.1186/1743-422X-2-21

Niu, Q.W., Lin, S.S., Reyes, J.L., Chen, K.C., Wu, H.W., Yeh, S.D., Chua, N.H., 2006. Expression of artificial microRNAs in transgenic *Arabidopsis thaliana* confers virus resistance. Nat. Biotechnol. 24, 1420–1428. doi:10.1038/nbt1255

Noda, L., 1973. Adenylate kinase. Enzym. 8, 279–305.

Noris, E., Accotto, G.P., Tavazza, R., Brunetti, A., Crespi, S., Tavazza, M., 1997. Resistance to Tomato Yellow Leaf Curl Geminivirus in *Nicotiana benthamiana* plants transformed with a truncated viral *C1* gene. Virology 227, 519-. doi:10.1006/viro.1996.8350

Ntui, V.O., Kong, K., Azadi, P., Khan, R.S., Chin, D.P., Igawa, T., Mii, M., Nakamura, I., 2014. RNAi-mediated resistance to Cucumber Mosaic Virus (CMV) in genetically engineered Tomato. Am. J. Plant Sci. 5, 554–572.

Ntui, V.O., Kong, K., Khan, R.S., Igawa, T., Janavi, G.J., Rabindran, R., Nakamura, I., Mii, M., 2015. Resistance to Sri Lankan cassava mosaic virus (SLCMV) in genetically engineered cassava cv. KU50 through RNA silencing. PLoS One 10, 1–23. doi:10.1371/journal.pone.0120551

Nyaboga, E.N., Njiru, J.M., Tripathi, L., 2015. Factors influencing somatic embryogenesis , regeneration , and *Agrobacterium* -mediated transformation of cassava (*Manihot esculenta* Crantz) cultivar TME14. Front. Plant Sci. 6, 1–13. doi:10.3389/fpls.2015.00411

Obambi, M., Bembe, A., Ntawuruhunga, P., Mvila, J., Okao-Okuja, G., Legg, J., 2011. Incidence and severity of cassava mosaic disease in the Republic of

Congo. African Crop Sci. J. 15, 1–9. doi:10.4314/acsj.v15i1.54405

Odipio, J., Alicai, T., Ingelbrecht, I., Nusinow, D.A., Bart, R., Taylor, N.J., 2017. Efficient CRISPR/Cas9 Genome editing of *Phytoene desaturase* in cassava. Front. Plant Sci. 8, 1–11. doi:10.3389/fpls.2017.01780

Okogbenin, E., Moreno, I., Tomkins, J., Fauquet, C.M., Mkamilo, G., Fregene, M., 2013. Marker-Assisted breeding for cassava mosaic disease resistance. Transl. Genomics Crop Breeding, Vol. I Biot. Stress I, 291–325. doi:10.1002/9781118728475.ch15

Oliver, S.N., Tiessen, A., Fernie, A.R., Geigenberger, P., 2008. Decreased expression of plastidial adenylate kinase in potato tubers results in an enhanced rate of respiration and a stimulation of starch synthesis that is attributable to post-translational redox-activation of ADP-glucose pyrophosphorylase. J. Exp. Bot. 59, 315–25. doi:10.1093/jxb/erm312

Olsen, K., Schaal, B.., 1999. Evidence on the origin of cassava : Phylogeography of Manihot esculenta. Proc Natl Acad Sci USA 96, 5586–5591.

Ossowski, S., Schwab, R., Weigel, D., 2008. Gene silencing in plants using artificial microRNAs and other small RNAs. Plant J. 53, 674–690. doi:10.1111/j.1365-313X.2007.03328.x

Owor, B.E., Legg, J.P., 2004. The effect of cassava mosaic geminiviruses on symptom severity , growth and root yield of a cassava mosaic virus disease-susceptible cultivar in Uganda. Ann. Appl. Biol. 145, 331–337. doi:10.1111/j.1744-7348.2004.tb00390.x

Padidam, M., Beachy, R.N., Fauquet, C.M., 1996. The role of AV2 ('precoat') and coat protein in viral replication and movement in tomato leaf curl geminivirus. Virology 224, 390–404. doi:10.1006/viro.1996.0546

Padidam, M., Sawyer, S., Fauquet, C.M., 1999. Possible emergence of new geminiviruses by frequent recombination. Virology 265, 218–225. doi:10.1006/viro.1999.0056

Palatnik, J.F., Wollmann, H., Schommer, C., Schwab, R., Boisbouvier, J., Rodriguez,

R., Warthmann, N., Allen, E., Dezulian, T., Huson, D., Carrington, J.C., Weigel, D., 2007. Sequence and expression differences underlie functional specialization of *Arabidopsis* microRNAs miR159 and miR319. Dev. Cell 13, 115–125. doi:10.1016/j.devcel.2007.04.012

Pang, S.-Z., Jan, F.-J., Gonsalves, D., 1997. Nontarget DNA sequences reduce the transgene length necessary for RNA-mediated tospovirus resistance in transgenic plants. Plant Biol. 94, 8261–8266. doi:10.1073/pnas.94.15.8261

Parent, J.S., Bouteiller, N., Elmayan, T., Vaucheret, H., 2015. Respective contributions of *Arabidopsis* DCL2 and DCL4 to RNA silencing. Plant J. 81, 223–232. doi:10.1111/tpj.12720

Park, M.Y., Wu, G., Gonzalez-Sulser, A., Vaucheret, H., Poethig, R.S., 2005. Nuclear processing and export of microRNAs in *Arabidopsis*. Proc. Natl. Acad. Sci. 102, 3691–3696. doi:10.1073/pnas.0405570102

Patil, B.L., Bagewadi, B., Yadav, J.S., Fauquet, C.M., 2016. Mapping and identification of cassava mosaic geminivirus DNA-A and DNA-B genome sequences for efficient siRNA expression and RNAi based virus resistance by transient agro-infiltration studies. Virus Res. 213, 109–115. doi:10.1016/j.virusres.2015.11.011

Patil, B.L., Fauquet, C.M., 2015. Studies on differential behavior of cassava mosaic geminivirus DNA components, symptom recovery patterns, and their siRNA profiles. Virus Genes 50, 474–486. doi:10.1007/s11262-015-1184-y

Patil, B.L., Fauquet, C.M., 2009. Cassava mosaic geminiviruses: Actual knowledge and perspectives. Mol. Plant Pathol. 10, 685–701. doi:10.1111/j.1364-3703.2009.00559.x

Paudel, D.B., Sanfaçon, H., 2018. Exploring the diversity of mechanisms associated with plant tolerance to virus infection. Front. Plant Sci. 9, 1–20. doi:10.3389/fpls.2018.01575

Pélissier, T., Clavel, M., Chaparro, C., Pouch-PéLissier, M.N., Vaucheret, H., Deragon, J.M., 2011. Double-stranded RNA binding proteins DRB2 and DRB4 have an antagonistic impact on polymerase IV-dependent siRNA levels in

Arabidopsis. Rna 17, 1502–1510. doi:10.1261/rna.2680711

Pennisi, E., 2013. The CRISPR Craze. Science (80-.). 341, 833 LP – 836.

Pérez-quintero, Á.L., López, C., 2010. Artificial microRNAs and their applications in plant molecular biology. Agron. Colomb. 28, 373–381.

Pierce, E.J., Rey, M.E.C., 2013. Assessing global transcriptome changes in response to *South African Cassava Mosaic Virus* [ZA-99] infection in susceptible *Arabidopsis thaliana*. PLoS One 8. doi:10.1371/journal.pone.0067534

Pita, J.S., Fondong, V.N., Sangaré, a, Otim-Nape, G.W., Ogwal, S., Fauquet, C.M., 2001. Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. J. Gen. Virol. 82, 655–65.

Pooggin, M., Shivaprasad, P. V., Veluthambi, K., Hohn, T., 2003. RNAi targeting of DNA virus in plants. Nat. Biotechnol. 21, 131–132. doi:10.1038/nbt0203-131b

Pooggin, M.M., 2017. RNAi-mediated resistance to viruses: a critical assessment of methodologies. Curr. Opin. Virol. 26, 28–35. doi:10.1016/j.coviro.2017.07.010

Pooma, W., Gillette, W.K., Jeffrey, J.L., Petty, I.T.D., 1996. Host and viral factors determine the dispensability of coat protein for bipartite geminivirus systemic movement. Virology 218, 264–268. doi:10.1006/viro.1996.0189

Pradhan, B., Vu Van, T., Dey, N., Mukherjee, S.K., 2017. Molecular biology of geminivirus DNA replication, in: Viral Replication. pp. 2–35.

Praveen, S., Ramesh, S. V., Mishra, A.K., Koundal, V., Palukaitis, P., 2010. Silencing potential of viral derived RNAi constructs in *Tomato leaf curl virus*-AC4 gene suppression in tomato. Transgenic Res. 19, 45–55. doi:10.1007/s11248-009-9291-y

Preiss, W., Jeske, H., 2003. Multitasking in replication is common among Geminiviruses. J. Virol. 77, 2972–2980. doi:10.1128/JVI.77.5.2972-2980.2003

Qu, F., Ye, X., Morris, T.J., 2008. *Arabidopsis* DRB4, AGO1, AGO7, and RDR6

participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1. Proc. Natl. Acad. Sci. 105, 14732–14737. doi:10.1073/pnas.0805760105

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

Raemakers, C., Sofiari, E., Jacobsen, E., Visser, R., 1997. Regeneration and transformation of cassava. Euphytica 96, 153–161.

Raemakers, K., Schreuder, M., Suurs, L., Furrer-Verhorst, H., Vincken, J.P., De Vetten, N., Jacobsen, E., Visser, R.G.F., 2005. Improved cassava starch by antisense inhibition of granule-bound starch synthase I. Mol. Breed. 16, 163–172. doi:10.1007/s11032-005-7874-8

Raja, P., Jackel, J.N., Li, S., Heard, I.M., Bisaro, D.M., 2014. *Arabidopsis* double-stranded RNA binding protein DRB3 participates in methylation-mediated defense against geminiviruses. J. Virol. 88, 2611–2622. doi:10.1128/JVI.02305-13

Raja, P., Sanville, B.C., Buchmann, R.C., Bisaro, D.M., 2008. Viral genome methylation as an epigenetic defense against Geminiviruses. J. Virol. 82, 8997–9007. doi:10.1128/JVI.00719-08

Rajeswaran, R., Golyaev, V., Seguin, J., Zvereva, A.S., Farinelli, L., Pooggin, M.M., 2014. Interactions of Rice tungro bacilliform Pararetrovirus and its protein P4 with plant RNA-silencing machinery. Mol. Plant-Microbe Interact. 27, 1370–1378. doi:10.1094/MPMI-07-14-0201-R

Rajeswaran, R., Pooggin, M.M., 2012. Role of virus-derived small RNAs in plant antiviral defense: Insights from DNA viruses, in: Sunkar, R. (Ed.), MicroRNAs in Plant Development and Stress Responses. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 261–289. doi:10.1007/978-3-642-27384-1_13

Ramesh, S.V., Mishra, A.K., Praveen, S., 2007. Hairpin RNA-mediated strategies for Silencing of *Tomato Leaf Curl Virus AC1* and *AC4* genes for effective resistance

in plants. Oligonucleotides 17, 251–257.

Ramesh, S.V., Ratnaparkhe, M.B., Kumawat, G., Gupta, G.K., Husain, S.M., 2014. Plant miRNAome and antiviral resistance: A retrospective view and prospective challenges. Virus Genes 48, 1–14. doi:10.1007/s11262-014-1038-z

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., Zhang, F., 2013. Genome engineering using crispr-cas9 system. Nat. Protoc. 8, 2281–2308. doi:10.1007/978-1-4939-1862-1_10

Regierer, B., Fernie, A.R., Springer, F., Perez-Melis, A., Leisse, A., Koehl, K., Willmitzer, L., Geigenberger, P., Kossmann, J., 2002. Starch content and yield increase as a result of altering adenylate pools in transgenic plants. Nat. Biotechnol. 20, 1256–60. doi:10.1038/nbt760

Rerie, G.W., Feldmann, K., Marks, M.D., 1994. The *GLABRA2* gene encodes a homeodomain protein required for normal ovule development in *Arabadopsis*. Genes Dev. 8, 1388–1399.

Rey, C., Vanderschuren, H.V., 2017. Cassava mosaic and brown streak diseases: Current perspectives and beyond. Annu. Rev. Virol. 4, 429–452. doi:10.1146/annurev-virology-101416-041913

Rey, M.E.C., Ndunguru, J., Berrie, L.C., Paximadis, M., Berry, S., Cossa, N., Nuaila, V.N., Mabasa, K.G., Abraham, N., Rybicki, E.P., Martin, D., Pietersen, G., Esterhuizen, L.L., 2012. Diversity of dicotyledenous-infecting geminiviruses and their associated DNA molecules in Southern Africa, including the South-west Indian Ocean Islands. Viruses 4, 1753–1791. doi:10.3390/v4091753

Ribeiro, S.G., Lohuis, H., Goldbach, R., Prins, M., 2007. *Tomato Chlorotic Mottle Virus* is a target of RNA silencing but the presence of specific short interfering RNAs does not guarantee resistance in transgenic plants. J. Virol. 81, 1563–1573. doi:10.1128/JVI.01238-06

Rizvi, I., Choudhury, N.R., Tuteja, N., 2014. Insights into the functional characteristics of geminivirus rolling-circle replication initiator protein and its interaction with host factors affecting viral DNA replication. Arch. Virol. 160, 375–387. doi:10.1007/s00705-014-2297-7

Rodríguez-Negrete, E., Lozano-Durán, R., Piedra-Aguilera, A., Cruzado, L., Bejarano, E.R., Castillo, A.G., 2013. Geminivirus Rep protein interferes with the plant DNA methylation machinery and suppresses transcriptional gene silencing. New Phytol. 199, 464–475. doi:10.1111/nph.12286

Rodriguez, R.E., Schommer, C., Palatnik, J.F., 2016. Control of cell proliferation by microRNAs in plants. Curr. Opin. Plant Biol. 34, 68–76. doi:10.1016/j.pbi.2016.10.003

Rogans, S.J., 2016. The role of small RNAs in susceptibility and tolerance to cassava mosaic disease. University of the Witwatersrand.

Rogans, S.J., Allie, F., Tirant, J.E., Rey, M.E.C., 2016. Small RNA and methylation responses in susceptible and tolerant landraces of cassava infected with *South African cassava mosaic virus*. Virus Res. 225, 10–22. doi:10.1016/j.virusres.2016.08.011

Rogans, S.J., Rey, C., 2016. Unveiling the micronome of cassava (*Manihot esculenta* crantz). PLoS One 11, 1–30. doi:10.1371/journal.pone.0147251

Rogers, D.J., 1963. Studies of *Manihot esculenta* Crantz and related species. Bull. Torrey Bot. Club 90, 43–54.

Rojas, M.R., Hagen, C., Lucas, W.J., Gilbertson, R.L., 2005. Exploiting chinks in the plant's armor: Evolution and emergence of Geminiviruses. Annu. Rev. Phytopathol. 43, 361–394. doi:10.1146/annurev.phyto.43.040204.135939

Rojas, M.R., Jiang, H., Salati, R., Xoconostle-Cázares, B., Sudarshana, M.R., Lucas, W.J., Gilbertson, R.L., 2001. Functional analysis of proteins involved in movement of the monopartite begomovirus, *Tomato yellow leaf curl virus*. Virology 291, 110–125. doi:10.1006/viro.2001.1194

Rose, A.B., 2004. The effect of intron location on intron-mediated enhancement of gene expression in Arabidopsis. Plant J. 40, 744–751. doi:10.1111/j.1365-313X.2004.02247.x

Roth, B.M., Pruss, G.J., Vance, V.B., 2004. Plant viral suppressors of RNA silencing. Virus Res. 102, 97–108. doi:10.1016/j.virusres.2004.01.020

Ruhel, R., Chakraborty, S., 2018. Multifunctional roles of geminivirus encoded replication initiator protein. VirusDisease 1–8. doi:10.1007/s13337-018-0458-0

Santoso, D., Thornburg, R., 2002. Uridine 5′-Monophosphate Synthase is transcriptionally regulated by pyrimidine levels in *Nicotiana plumbaginifolia*. Plant Physiol. 116, 815–821. doi:10.1104/pp.116.2.815

Saunders, K., Lucy, A., Stanley, J., 1992. RNA-primed complementary-sense DNA synthesis of the geminivirus *African cassava mosaic virus*. Nucleic Acids Res. 20, 6311–6315. doi:10.1093/nar/20.23.6311

Schaeffer, S.M., Nakata, P.A., 2015. CRISPR/Cas9-mediated genome editing and gene replacement in plants: Transitioning from lab to field. Plant Sci. 240, 130–142. doi:10.1016/j.plantsci.2015.09.011

Schommer, C., Palatnik, J.F., Aggarwal, P., Chételat, A., Cubas, P., Farmer, E.E., Nath, U., Weigel, D., 2008. Control of jasmonate biosynthesis and senescence by miR319 targets. PLoS Biol. 6, 1991–2001. doi:10.1371/journal.pbio.0060230

Schopke, C., Taylor, N., Carcamo, R., Konan, N.K., Marmey, P., Henshaw, G.G., Beachy, R.N., Fauquet, C.M., 1996. Regeneration of transgenic cassava plants (*Manihot esculenta* Crantz) from microbombarded embryogenic suspension cultures. Nat. Biotechnol. 14, 731–735.

Schott, G., Mari-Ordonez, A., Himber, C., Alioua, A., Voinnet, O., Dunoyer, P., 2012. Differential effects of viral silencing suppressors on siRNA and miRNA loading support the existence of two distinct cellular pools of ARGONAUTE1. EMBO J. 31, 2553–2565. doi:10.1038/emboj.2012.92

Schroder, M., Giermann, N., Zrenner, R., 2005. Functional analysis of the Pyrimidine de Novo synthesis pathway in solanaceous species. Plant Physiol. 138, 1926–1938. doi:10.1104/pp.105.063693

Senthil-kumar, M., Mysore, K.S., 2011. Caveat of RNAi in plants: The off-target effect, in: RNAi and Plant Gene Function Analysis, Methods in Molecular Biology. pp. 13–25. doi:10.1007/978-1-61779-123-9

Seung, D., Soyk, S., Coiro, M., Maier, B.A., Eicke, S., Zeeman, S.C., 2015.

PROTEIN TARGETING TO STARCH Is required for localising GRANULE-BOUND STARCH SYNTHASE to starch granules and for normal amylose synthesis in *Arabidopsis*. PLoS Biol. 13, 1–29. doi:10.1371/journal.pbio.1002080

Shamandi, N., Zytnicki, M., Comella, P., Elvira-Matelot, E., Lepère, G., Shamandi, N., Mallory, A.C., Sáez-Vásquez, J., Charbonnel, C., Bochnakian, A., Vaucheret, H., 2015. Plants encode a general siRNA suppressor that is induced and suppressed by viruses. PLOS Biol. 13, e1002326. doi:10.1371/journal.pbio.1002326

Sharma, V.K., Kushwaha, N., Basu, S., Singh, A.K., Chakraborty, S., 2014. Identification of siRNA generating hot spots in multiple viral suppressors to generate broad-spectrum antiviral resistance in plants. Physiol. Mol. Biol. Plants 21, 9–18. doi:10.1007/s12298-014-0264-0

Shimura, H., Pantaleo, V., Ishihara, T., Myojo, N., Inaba, J. ichi, Sueda, K., Burgyán, J., Masuta, C., 2011. A Viral Satellite RNA induces yellow symptoms on tobacco by targeting a gene involved in Chlorophyll biosynthesis using the RNA silencing machinery. PLoS Pathog. 7, 1–12. doi:10.1371/journal.ppat.1002021

Shure, M., Wessler, S., Fedoroff, N., 1983. Molecular identification and isolation of the Waxy locus in maize. Cell 35, 225–233. doi:10.1016/0092-8674(83)90225-8

Shybut, M.E., 2015. Transcription Activator-Like (TAL) Effectors of the Cassava Bacterial Blight Pathogen. Berkeley.

Simón-Mateo, C., García, J.A., 2011. Antiviral strategies in plants based on RNA silencing. Biochim. Biophys. Acta 1809, 722–31. doi:10.1016/j.bbagrm.2011.05.011

Slattery, C.J., Kavakli, I.H., Okita, T.W., 2000. Engineering starch for increased quantity and quality. Trends Plant Sci. 5, 291–298. doi:10.1016/S1360-1385(00)01657-5

Smith, A.M., 2008. Prospects for increasing starch and sucrose yields for bioethanol production. Plant J. 54, 546–58. doi:10.1111/j.1365-313X.2008.03468.x

Smith, N.A., Singh, S.P., Wang, M.-B., Stoutjesdijk, P.A., Green, A.G., Waterhouse,

P.M., 2000. Total silencing by intron- spliced hairpin RNAs. Nature 407, 319–320. doi:10.1002/(SICI)1097-0177(199909)216:1<1::AID-DVDY1>3.0.CO;2-T

Smith, O., Palmer, S.A., Clapham, A.J., Rose, P., Liu, Y., Wang, J., Allaby, R.G., 2017. Small RNA activity in archeological Barley shows novel germination inhibition in response to environment. Mol. Biol. Evol. 34, 2555–2562. doi:10.1093/molbev/msx175

Snead, N.M., Rossi, J.J., 2010. Biogenesis and function of endogenous and exogenous siRNAs. Wiley Interdiscip. Rev. RNA 1, 117–131. doi:10.1002/wrna.14

Stamp, J.., Henshaw, G.., 1987. Secondary somatic embryogenesis and plant regeneration in cassava. Plant Cell. Tissue Organ Cult. 10, 227–233.

Stamp, J.A., Henshaw, G.G., 1987. Somatic embryogenesis from clonal leaf tissues of cassava. Ann. Bot. 59, 445–450. doi:10.1093/oxfordjournals.aob.a087333

Stamp, J.A., Henshaw, G.G., 1982. Somatic embryogenesis in cassava. Zeitschrift für Pflanzenphysiologie 105, 183–187. doi:10.1016/s0044-328x(82)80011-1

Stark, D., Timmerman, K., Barry, G., 1992. Regulation of the amount of starch in plant tissues by ADP glucose pyrophosphorylase. Science (80-.). 258.

Sunitha, S., Shanmugapriya, G., Balamani, V., Veluthambi, K., 2013. Mungbean yellow mosaic virus (MYMV) AC4 suppresses post-transcriptional gene silencing and an AC4 hairpin RNA gene reduces MYMV DNA accumulation in transgenic tobacco. Virus Genes 46, 496–504. doi:10.1007/s11262-013-0889-z

Sweetlove, L.J., Burrell, M.M., Rees, T. ap, 2015. Starch metabolism in tubers of transgenic potato (*Solanum tuberosum*) with increased ADPglucose pyrophosphorylase. Biochem. J. 320, 493–498. doi:10.1042/bj3200493

Talmor-Neiman, M., Stav, R., Frank, W., Voss, B., Arazi, T., 2006. Novel micro-RNAs and intermediates of micro-RNA biogenesis from moss. Plant J. 47, 25–37. doi:10.1111/j.1365-313X.2006.02768.x

Taylor, N., Chavarriaga, P., Raemakers, K., Siritunga, D., Zhang, P., 2004. Development and application of transgenic technologies in cassava. Plant Mol.

Biol. 56, 671–688. doi:10.1007/s11103-004-4872-x

Taylor, N., Gaitán-solís, E., Moll, T., Trembley, C., Abernathy, V., Corbin, D., Fauquet, C.M., 2012. A High-throughput Ppatform for the production and analysis of transgenic cassava (*Manihot esculenta*) plants. Trop. Plant Biol. 5, 127–139. doi:10.1007/s12042-012-9099-4

Taylor, N.J., Edwards, M., Kiernan, R.J., Davey, C.D.M., Blakesley, D., Henshaw, G.G., 1996. Development of friable embryogenic callus and embryogenic suspension culture systems in Cassava (Manihot Esculenta Crantz). Nat. Biotechnol. 14, 726–730. doi:10.1038/nbt0696-726

Taylor, S.H., 2009. The development of SACMV - Resistant cassava using pathogen - derived resistance strategies, targeting the DNA - A Component. Univ. Witwatersrand, Johannesbg.

Teye, E., Asare, A.P., Amoah, R.S., Tetteh, J.P., 2011. Determination of the dry matter content of cassava (*Manihot esculenta*, Crantz) tubers using specific gravity method. ARPN J. Agric. Biol. Sci. 6, 23–28.

Thomasset, B., Inzé, D., Thomas, D., Boetti, H., Denmat, L.A., Ménard, M., 2002. β-Glucuronidase activity in transgenic and non-transgenic tobacco cells: specific elimination of plant inhibitors and minimization of endogenous GUS background. Plant Sci. 113, 209–219. doi:10.1016/0168-9452(95)04286-5

Thresh, J.M., Cooter, R.J., 2005. Strategies for controlling cassava mosaic virus disease in Africa. Plant Pathol. 54, 587–614. doi:10.1111/j.1365-3059.2005.01282.x

Tiessen, A., Hendriks, J.H.M., Stitt, M., Branscheid, A., Gibon, Y., Farré, E.M., Geigenberger, P., 2002. Starch synthesis in potato tubers is regulated by post-translational redox modification of ADP-glucose pyrophosphorylase. Plant Cell 14, 2191–2213. doi:10.1105/tpc.003640.2192

Tiwari, M., Sharma, D., Trivedi, P.K., 2014. Artificial microRNA mediated gene silencing in plants: Progress and perspectives. Plant Mol. Biol. 86. doi:10.1007/s11103-014-0224-7

Tjaden, J., Möhlmann, T., Kampfenkel, K., Hendrichs, G., Neuhaus, H., 1998. Altered plastidic ATP/ADP- transporter activity influences potato (*Solanum tuberosum* L.) tuber morphology, yield and composition of tuber starch. Plant J. 16, 531–540.

Tominaga-Wada, R., Iwata, M., Sugiyama, J., Kotake, T., Ishida, T., Yokoyama, R., Nishitani, K., Okada, K., Wada, T., 2009. The GLABRA2 homeodomain protein directly regulates *CESA5* and *XTH17* gene expression in *Arabidopsis* roots. Plant J. 60, 564–574. doi:10.1111/j.1365-313X.2009.03976.x

Torres-Barceló, C., Daròs, 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. Arch. Virol. 155, 251–254. doi:10.1007/s00705-009-0563-x

Trethewey, R.N., Geigenberger, P., Riedel, K., Hajirezaei, M.R., Sonnewald, U., Stitt, M., Riesmeier, J.W., Willmitzer, L., 1998. Combined expression of glucokinase and invertase in potato tubers leads to a dramatic reduction in starch accumulation and a stimulation of glycolysis. Plant J. 15, 109–118. doi:10.1046/j.1365-313X.1998.00190.x

Trinks, D., Rajeswaran, R., Shivaprasad, P. V., Oakeley, E.J., Veluthambi, K., Pooggin, M.M., Akbergenov, R., Hohn, T., Oakeley, E.J., Trinks, D., Pooggin, M.M., Veluthambi, K., 2005. Suppression of RNA Silencing by a Geminivirus Nuclear Protein, AC2, Correlates with Transactivation of Host Genes. J. Virol. 79, 2517–2527. doi:10.1128/jvi.79.4.2517-2527.2005

Vanderschuren, H., Akbergenov, R., Pooggin, M.M., Hohn, T., Gruissem, W., Zhang, P., 2007a. Transgenic cassava resistance to *African cassava mosaic virus* is enhanced by viral DNA-A bidirectional promoter-derived siRNAs. Plant Mol. Biol. 64, 549–557. doi:10.1007/s11103-007-9175-6

Vanderschuren, H., Alder, A., Zhang, P., Gruissem, W., 2009. Dose-dependent RNAi-mediated geminivirus resistance in the tropical root crop cassava. Plant Mol. Biol. 70, 265–272. doi:10.1007/s11103-009-9472-3

Vanderschuren, H., Moreno, I., Anjanappa, R.B., Zainuddin, I.M., Gruissem, 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. doi:10.1371/journal.pone.0045277

Vanderschuren, H., Stupak, M., Fütterer, J., Gruissem, W., Zhang, P., 2007b. Engineering resistance to geminiviruses - Review and perspectives. Plant Biotechnol. J. 5, 207–220. doi:10.1111/j.1467-7652.2006.00217.x

Vanitharani, R., Chellappan, P., Fauquet, C.M., 2005. Geminiviruses and RNA silencing. Trends Plant Sci. 10, 144–51. doi:10.1016/j.tplants.2005.01.005

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

Várallyay, É., Válóczi, A., Ágyi, Á., Burgyán, J., Havelda, Z., 2010. Plant virus-mediated induction of miR168 is associated with repression of ARGONAUTE1 accumulation. EMBO J. 29, 3507–3519. doi:10.1038/emboj.2010.215

Vaucheret, H., Beclin, C., Fagard, M., 2001. Post-transcriptional gene silencing in plants by RNA. J. Cell Sci. 114, 3083–3091. doi:10.1007/s00299-003-0699-1

Vaucheret, H., du Jardin, P., Ramon, M., Nogué, F., Casacuberta, J.M., Devos, Y., 2015. Biotechnological uses of RNAi in plants: risk assessment considerations. Trends Biotechnol. 33, 145–147. doi:10.1016/j.tibtech.2014.12.003

Vaucheret, H., Fagard, M., 2001. Transcriptional gene silencing in plants: Targets, inducers and regulators. Trends Genet. doi:10.1016/S0168-9525(00)02166-1

Vinutha, T., Kumar, G., Garg, V., Canto, T., Palukaitis, P., Ramesh, S., Praveen, S., 2018. Tomato geminivirus encoded RNAi suppressor protein, AC4 interacts with host AGO4 and precludes viral DNA methylation. Gene 678, 184–195. doi:10.1016/j.gene.2018.08.009

Voinnet, O., 2005. Induction and suppression of RNA silencing: Insights from viral infections. Nat. Rev. Genet. 6, 206–220. doi:10.1038/nrg1555

Voinnet, O., 2001. RNA silencing as a plant immune system against viruses.

TRENDS Genet. 17, 449–459.

Voinnet, O., Baulcombe, D., 1997. Systemic signalling in gene silencing. Nature 389, 553.

Voinnet, O., Pinto, Y.M., Baulcombe, D.C., 1999. Suppression of gene silencing: A general strategy used by diverse DNA and RNA viruses of plants. Proc. Natl. Acad. Sci. 96, 14147–14152. doi:10.1073/pnas.1513950112

Vu, T. Van, Roy Choudhury, N., Mukherjee, S.K., 2013. Transgenic tomato plants expressing artificial microRNAs for silencing the pre-coat and coat proteins of a begomovirus, *Tomato leaf curl New Delhi virus*, show tolerance to virus infection. Virus Res. 172, 35–45. doi:10.1016/j.virusres.2012.12.008

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. J. Virol. Methods 231, 38–43. doi:10.1016/j.jviromet.2016.02.004

Wang, H., Buckley, K.J., Yang, X., Cody, R., Bisaro, D.M., Buchmann, R.C., 2005. Adenosine Kinase inhibition and suppression of RNA silencing by geminivirus AL2 and L2 proteins. J. Virol. 79, 7410–7418. doi:10.1128/JVI.79.12.7410

Wang, J., Ma, N., Tang, Y., Ling, X., Kan, J., He, Z., Yang, Y., Zhang, B., 2016. *Cotton Leaf Curl Multan Virus*-derived Viral Small RNAs can target cotton genes to promote viral infection. Front. Plant Sci. 7, 1–13. doi:10.3389/fpls.2016.01162

Wang, S., Shi, Z., Liu, W., Jules, J., Feng, X., 2006. Development and validation of vectors containing multiple siRNA expression cassettes for maximizing the efficiency of gene silencing. BMC Biotechnol. 6, 1–7. doi:10.1186/1472-6750-6-50

Wang, X., Chang, L., Tong, Z., Wang, Dongyang, Yin, Q., Wang, Dan, Jin, X., Yang, Q., Wang, L., Sun, Y., Huang, Q., Guo, A., Peng, M., 2016. Proteomics profiling reveals carbohydrate metabolic enzymes and 14-3-3 proteins play important roles for starch accumulation during cassava root tuberization. Sci. Rep. 6, 1–15. doi:10.1038/srep19643

Warburg, O., 1894. Die Kulturpflanzen Usambaras. Mitt. Dtsch. Schutzgeb. 7.

Waterhouse, P.M., Helliwell, C. a, 2003. Exploring plant genomes by RNA-induced gene silencing. Nat. Rev. Genet. 4, 29–38. doi:10.1038/nrg982

Wei, Y., Shi, H., Xia, Z., Tie, W., Ding, Z., Yan, Y., Wang, W., Hu, W., Li, K., 2016. Genome-Wide Identification and Expression Analysis of the WRKY Gene Family in Cassava. Front. Plant Sci. 7, 1–18. doi:10.3389/fpls.2016.00025

Welsch, R., Arango, J., Bär, C., Salazar, B., Al-Babili, S., Beltrán, J., Chavarriaga, P., Ceballos, H., Tohme, J., Beyer, P., 2010. Provitamin A accumulation in cassava (*Manihot esculenta*) roots driven by a Single Nucleotide Polymorphism in a *Phytoene Synthase* Gene. Plant Cell 22, 3348–3356. doi:10.1105/tpc.110.077560

Wesley, S. V, Helliwell, C. a, Smith, N. a, Wang, M.B., Rouse, D.T., Liu, Q., Gooding, P.S., Singh, S.P., Abbott, D., Stoutjesdijk, P. a, Robinson, S.P., Gleave, a P., Green, a G., Waterhouse, P.M., 2001. Construct design for efficient, effective and high-throughput gene silencing in plants. Plant J. 27, 581–90.

Westerbergh, A., Hu, X., Sun, C., Jansson, C., Zhang, J., 2009. Cassava, a potential biofuel crop in (the) People's Republic of China. Appl. Energy 86, S95–S99. doi:10.1016/j.apenergy.2009.05.011

Wingard, S.., 1928. Hosts and symptoms of ring spot, a virus disease of plants. J. Agric. Res. 37, 127–153.

Winter, J., Jung, S., Keller, S., Gregory, R.I., Diederichs, S., 2009. Many roads to maturity: microRNA biogenesis pathways and their regulation Julia. Nat. Cell Biol. 11. doi:10.1109/NSSMIC.2014.7430807

Wolfe, M.D., Rabbi, I.Y., Egesi, C., Hamblin, M., Kawuki, R., Kulakow, P., Lozano, R., Carpio, D.P. Del, Ramu, P., Jannink, J.-L., 2016. Genome-Wide association and prediction reveals genetic architecture of cassava mosaic disease resistance and prospects for rapid genetic improvement. Plant Genome 9, 0. doi:10.3835/plantgenome2015.11.0118

Woodward, B., Puonti-Kaerlas, J., 2001. Somatic embryogenesis from floral tissue of cassava (*Manihot esculenta* Crantz). Euphytica 1–6.

Xu, J., Yang, J.-Y., Niu, Q.-W., Chua, N.-H., 2006. Arabidopsis DCP2, DCP1, and VARICOSE form a Decapping complex required for postembryonic development. Plant Cell 18, 3386–3398. doi:10.1105/tpc.106.047605

Xu, J., Zhang, H.Y., Xie, C.H., Xue, H.W., Dijkhuis, P., Liu, C.M., 2005. EMBRYONIC FACTOR 1 encodes an AMP deaminase and is essential for the zygote to embryo transition in Arabidopsis. Plant J. 42, 743–756. doi:10.1111/j.1365-313X.2005.02411.x

Xu, P., Zhang, Y., Kang, L.L., Roossinck, M.J., Mysore, K.S., 2006. Computational estimation and experimental verification of Off-Target silencing during Posttranscriptional gene silencing in plants. Plant Physiol. 142, 429–440. doi:10.1007/s10728-011-0185-9

Yadav, J.S., Ogwok, E., Wagaba, H., Patil, B.L., Bagewadi, B., Alicai, T., Gaitan-solis, E., Taylor, N.J., Fauquet, C.M., 2011. RNAi-mediated resistance to *Cassava brown streak Uganda* virus in transgenic cassava. Mol. Plant 12, 677–687. doi:10.1111/J.1364-3703.2010.00700.X

Yadav, R.K., Chattopadhyay, D., 2011. Enhanced viral intergenic region-specific short interfering RNA accumulation and DNA methylation correlates with resistance against a geminivirus. Mol. Plant. Microbe. Interact. 24, 1189–97. doi:10.1094/MPMI-03-11-0075

Yang, M., Xu, Z., Zhao, W., Liu, Q., Li, Q., Lu, L., Liu, R., Zhang, X., Cui, F., 2018. *Rice stripe virus*-derived siRNAs play different regulatory roles in rice and in the insect vector Laodelphax striatellus. BMC Plant Biol. 18, 1–13. doi:10.1186/s12870-018-1438-7

Yang, Y., Chen, T., Ling, X., Dou, D., Shen, D., Liu, T., Wang, M.-B., Yu, W., Wang, J., Huang, J., Hu, Z., Hu, J., Zhang, B., 2019. Tomato yellow leaf curl virus intergenic siRNAs target a host long noncoding RNA to modulate disease symptoms. PLOS Pathog. 15, e1007534. doi:10.1371/journal.ppat.1007534

Yang, Z., Ebright, Y.W., Yu, B., Chen, X., 2006. HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2′ OH of the 3′ terminal nucleotide. Nucleic Acids Res. 34, 667–675. doi:10.1093/nar/gkj474

Ye, J., Qu, J., Mao, H.Z., Ma, Z.G., Rahman, N.E.B., Bai, C., Chen, W., Jiang, S.Y., Ramachandran, S., Chua, N.H., 2014. Engineering geminivirus resistance in *Jatropha curcus*. Biotechnol. Biofuels 7, 1–11. doi:10.1186/s13068-014-0149-z

Ye, J., Yang, J., Sun, Y., Zhao, P., Gao, S., Jung, C., Qu, J., Fang, R., Chua, N.H., 2015. Geminivirus activates *ASYMMETRIC LEAVES 2* to accelerate cytoplasmic DCP2-mediated mRNA turnover and weakens RNA silencing in *Arabidopsis*. PLoS Pathog. 11, 1–21. doi:10.1371/journal.ppat.1005196

Yu, J.-L., Li, W.-X., Ding, S.-W.S.-W., Cillo, F., Wang, X.-B.X.-B., Ito, T., Wu, Q., Chen, X., Wua, Q., Itoa, T., Cilloc, F., Lia, W.-X., Chend, X., Yub, J.-L., Ding, S.-W.S.-W., 2009. RNAi-mediated viral immunity requires amplification of virus-derived siRNAs in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. 107, 484–489. doi:10.1073/pnas.0904086107

Zaidi, S.S.-A., Briddon, R.W., Mansoor, S., 2017. Engineering dual begomovirus-*Bemisia tabaci* resistance in plants. Trends Plant Sci. 22, 6–8. doi:10.1016/j.tplants.2016.11.005

Zainuddin, I.M., Schlegel, K., Gruissem, W., Vanderschuren, H., 2012. Robust transformation procedure for the production of transgenic farmer-preferred cassava landraces. Plant Methods 8, 24. doi:10.1186/1746-4811-8-24

Zeeman, S.C., Kossmann, J., Smith, A.M., 2010. Starch: its metabolism, evolution, and biotechnological modification in plants. Annu. Rev. Plant Biol. 61, 209–34. doi:10.1146/annurev-arplant-042809-112301

Zerbini, F.M., Briddon, R.W., Idris, A., Martin, D.P., Enrique, M., Navas-Castillo, J., Rivera-Bustamante, R., Roumagnac, P., Varsani, A., Consortium, Report, I., 2017. ICTV virus taxonomy profile: Geminiviridae. J. Gen. Virol. 98, 131–133. doi:10.1099/jgv.0.000973

Zhang, C., Wu, Z., Li, Y., Wu, J., 2015. Biogenesis, function, and applications of virus-derived small RNAs in plants. Front. Microbiol. 6, 1–12. doi:10.3389/fmicb.2015.01237

Zhang, L., Xie, X., Song, Y., Jiang, F., Zhu, C., Wen, F., 2013. Viral resistance mediated by shRNA depends on the sequence similarity and mismatched sites

between the target sequence and siRNA. Biol. Plant. 57, 547–554. doi:10.1007/s10535-013-0314-4

Zhang, P., Bohl-Zenger, S., Puonti-Kaerlas, J., Potrykus, I., Gruissem, W., 2003. Two cassava promoters related to vascular expression and storage root formation. Planta 218, 192–203. doi:10.1007/s00425-003-1098-0

Zhang, P., Ma, Q., Naconsie, M., Wu, X., Zhou, W., Yang, J., 2017. Advances in genetic modification of cassava. Achiev. Sustain. Cultiv. cassava, 2. doi:10.19103/AS.2016.0014.17

Zhang, P., Puonti-Kaerlas, J., 2000. PIG-mediated cassava transformation using positive and negative selection. Plant Cell Rep. 19, 1041–1048.

Zhang, P., Vanderschuren, H., Fütterer, J., Gruissem, W., 2005. Resistance to cassava mosaic disease in transgenic cassava expressing antisense RNAs targeting virus replication genes. Plant Biotechnol. J. 3, 385–397. doi:10.1111/j.1467-7652.2005.00132.x

Zhang, Y., Li, Z.X., Yu, X.D., Fan, J., Pickett, J.A., Jones, H.D., Zhou, J.J., Birkett, M.A., Caulfield, J., Napier, J.A., Zhao, G.Y., Cheng, X.G., Shi, Y., Bruce, T.J.A., Xia, L.Q., 2015. Molecular characterization of two isoforms of a *farnesyl pyrophosphate synthase* gene in wheat and their roles in sesquiterpene synthesis and inducible defence against aphid infestation. New Phytol. 206, 1101–1115. doi:10.1111/nph.13302

Zhang, Z., Chen, H., Huang, X., Xia, R., Zhao, Q., Lai, J., Teng, K., Li, Y., Liang, L., Du, Q., Zhou, X., Guo, H., Xie, Q., 2011. BSCTV C2 attenuates the degradation of SAMDC1 to suppress DNA methylation-mediated gene silencing in *Arabidopsis*. Plant Cell 23, 273–288. doi:10.1105/tpc.110.081695

Zhao, S.S., Dufour, D., Sánchez, T., Ceballos, H., Zhang, P., 2011. Development of waxy cassava with different Biological and physico-chemical characteristics of starches for industrial applications. Biotechnol. Bioeng. 108, 1925–1935. doi:10.1002/bit.23120

Zhou, Q.Y., Tian, A.G., Zou, H.F., Xie, Z.M., Lei, G., Huang, J., Wang, C.M., Wang, H.W., Zhang, J.S., Chen, S.Y., 2008. Soybean WRKY-type transcription factor

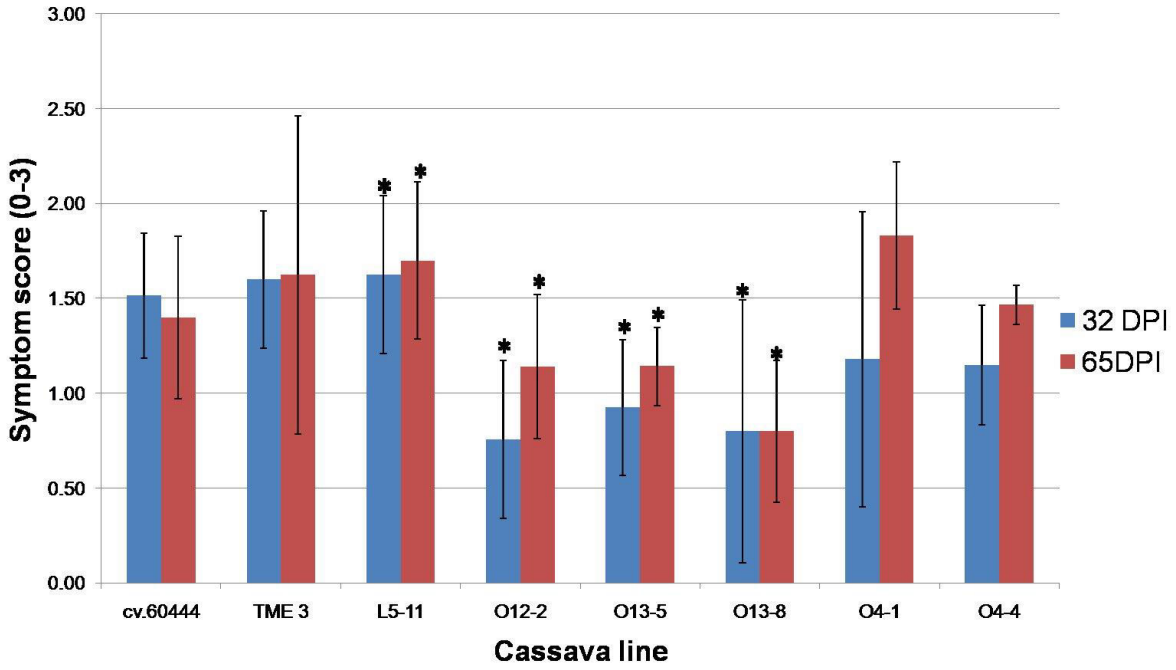
genes, *GmWRKY13, GmWRKY21*, and *GmWRKY54*, confer differential tolerance to abiotic stresses in transgenic *Arabidopsis* plants. Plant Biotechnol. J. 6, 486–503. doi:10.1111/j.1467-7652.2008.00336.x

Zhou, X., Liu, Y., Calvert, L., Munoz, C., Otim-Nape, G.W., Robinson, D.J., Harrison, B.D., Zhou, X., Robinson, D.J., Calvert, L., Munoz, C., Harrison, B.D., 1997. Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. J. Gen. Virol. 78 (Pt 8), 2101–11. doi:10.1099/0022-1317-78-8-2101

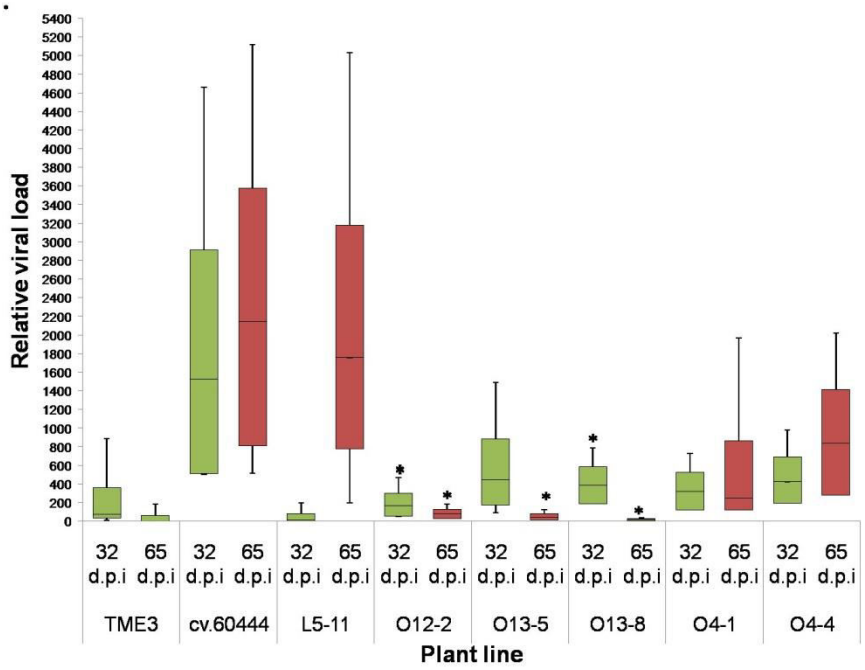
Zrachya, A., Glick, E., Levy, Y., Arazi, T., Citovsky, V., Gafni, Y., 2007. Suppressor of RNA silencing encoded by Tomato yellow leaf curl virus-Israel. Virology 358, 159–165. doi:10.1016/j.virol.2006.08.016

Zvereva, A.S., Pooggin, M., 2012. Silencing and innate immunity in plant defense against viral and non-viral pathogens. Viruses. doi:10.3390/v4112578

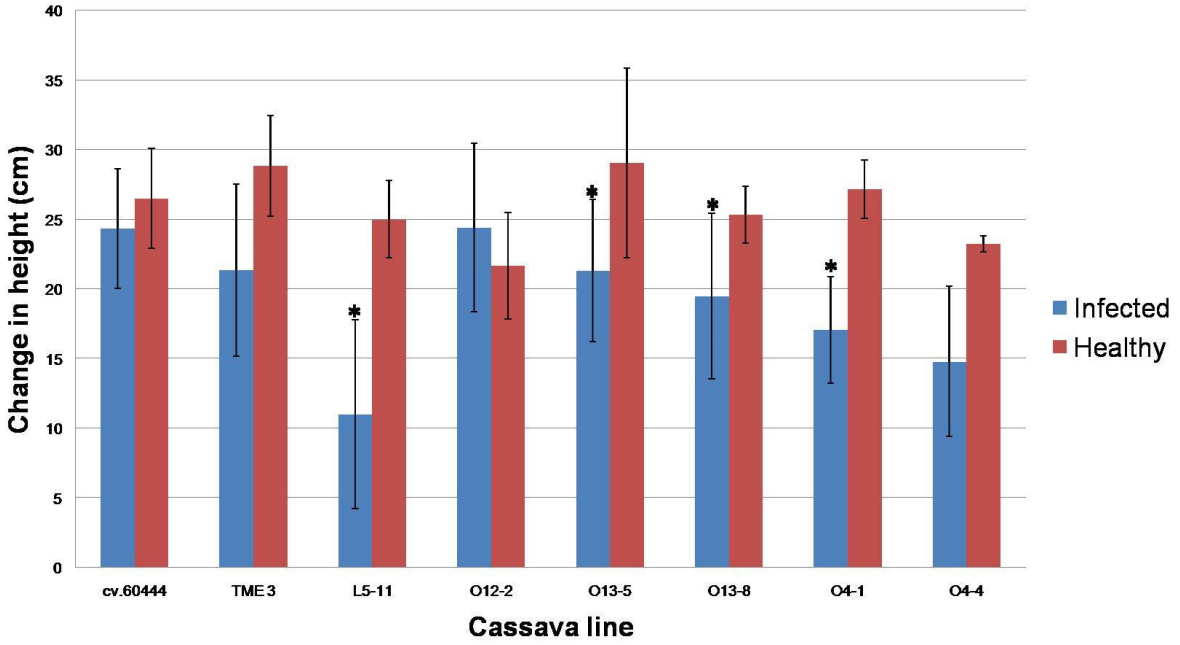
Appendix



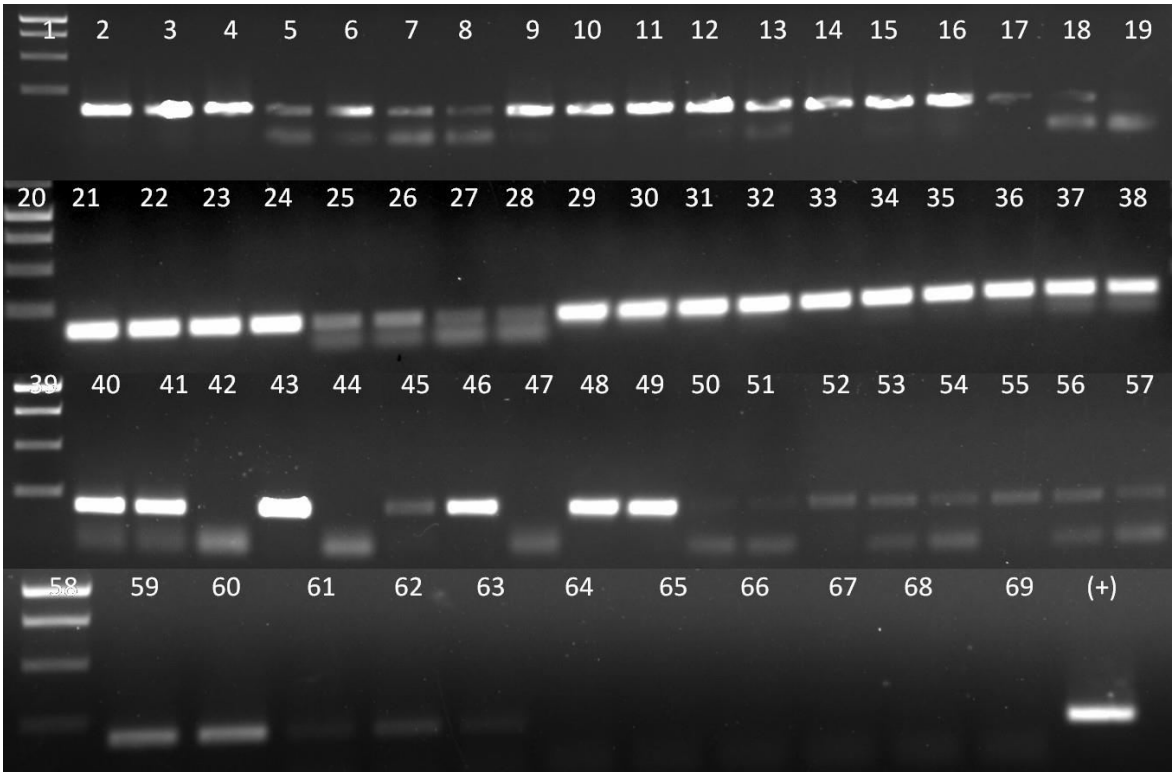
Appendix A: Average symptom severity scores in SACMV AC1/4 transgenic (O14-1, O12-2, O13-5, O13-8, O4-1 and O4-4) as well as tolerant TME3 and untransformed cv.60444 control plants, agro-inoculated with South African cassava mosaic virus at 32 and 65DPI. Scores represent the average symptom severity in 6 plants, and with error bars representing the standard deviation. Asterisks represent statistically significant results.



Appendix B: Average relative fold-change of viral load, in relation to reference gene *UBQ*10 in SACMV AC1/4 transgenic (O14-1, O12-2, O13-5, O13-8, O4-1 and O4-4) as well as tolerant TME3 and untransformed cv.60444 control plants, agro-inoculated with South African cassava mosaic virus, at 32 and 65DPI. Values represent the average of 6 plants, and with error bars representing the standard deviation. Asterisks represent statistically significant results.



Appendix C: Change in height in SACMV AC1/4 transgenic (O14-1, O12-2, O13-5, O13-8, O4-1 and O4-4) as well as tolerant TME3 and untransformed cv.60444 control plants, agro-inoculated with South African cassava mosaic virus at 32 and 65DPI. Scores represent the average symptom severity in 6 plants, and with error bars representing the standard deviation. Asterisks represent statistically significant results.



Appendix D: PCR amplification of SACMV Coat protein in SACMV AC1/AC4 transgenic lines (O12-2, O14-5and O14-8) as well as untransformed cv.60444 and TME3 infected with SACMV, at 14DPI. Lane 1, 20, 39 and 58: 1 Kb plus Molecular weight marker (Thermoscientific). Lane 2-13, O12-2 plants, lane 14-26: O13-5, lane 27-37: O13-8, 38-50: cv.604444 and 51-63: TME3, lane 64: Healthy O12-2, lane 65: healthy O13-5, lane 66: healthy O13-8, lane 67: healthy cv.60444, lane 68: healthy TME3, lane 69: NTC. Purified pBIN DNA-A was included as a Positive control (+).

Appendix E: Student's t-test accessing the mean statistical difference between SACMV AC1/AC4 transgenic lines O12-2, O13-5 and O13-8 transgenic line and cv.60444 symptom severity score index, plant height and SACMV viral load, at 32 and 65dpi

	Mean symptom severity		Viral load		Height	
	32 DPI	65 DPI	32 DPI	65 DPI	Infected	Healthy
TME 3	0.03	0.03	0.05	0.02	0.08	0.77
O12-2	0.02	0.05	0.03	0.02	0.60	0.21
O13-5	0.04	0.02	0.06	0.03	0.04	0.89
O13-8	0.02	0.01	0.09	0.03	0.04	0.74

Appendix F: Student t-testing assessing mean of statistical difference in relative expression of WRKY 14, mes microRNA 319, AMP deaminase, GEM-4 like protein, FPS1, FPS2 and Prxs in FECs transformed with pC-AES, pC-E-ACMV and pC-EACMV.

Gene	AES	E-ACMV	EACMV
WRKY 14	0.01	0.26	0.49
mes-MicroRNA319	0.01	0.085	0.083
AMP deaminase	0.04	0.030	0.080
GEM-4	0.001	0.013	0.001
FPS 1	0.002	0.051	0.067
FPS 2	0.003	6.1E-05	0.031
Prxs	0.002	3.43E-05	0.534

Appendix G: Student's t-test accessing the relative expression of *UMP* synthase or *ADK* in storage roots collected from UMPS and ADK transgenic cassava cv.60444 lines

Line	
UMPS1	0.145
UMPS2	0.196
UMPS13	0.083
UMPS17	0.059
ADK1	0.497
ADK2	0.220