PROTEIN-PROTEIN INTERACTIONS IN A GEMINIVIRUS-CASSAVA SYSTEM

by

BUSISIWE NCUBE KANYIKA

(1279052)

Submitted in fulfilment of the requirements for the degree

Doctor of Philosophy

in

Molecular and Cell Biology

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

Supervisor: Professor Chrissie Marie Rey

Advisor: Dr Nikki Gentle

January 2022

DECLARATION

I, Busisiwe Ncube Kanyika (1279052), hereby declare that this thesis is my own, unaided work. It is being submitted for the Degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

Signed:  Date: 6th day of January 2022

ABSTRACT

Cassava as a hardy drought tolerant crop that offers economic value in poor rural communities as a low input cost crop to farm for rural development. It provides nutritional needs as a gluten-free carbohydrate source is an essential raw material in industrial applications to produce renewable energy, starch, paper, glue, and is relevant to pharmaceuticals and mining. However, cassava is subject to disease, especially those caused by geminiviruses. Geminiviruses can devastate yields of cassava and thus negatively impact the cassava value chain. South African cassava mosaic virus (SACMV) (family *Geminiviridae*) is an economically significant plant virus that hijacks the cassava (host) genome replication machinery for its own multiplication and inevitable plant deterioration of plant health. SACMV systemic movement results in global subcellular changes that alter host gene expression and inhibit plant defence mechanisms. The ability of SACMV to replicate and proliferate is determined by the host availability of a matrix of proteins and metabolites that upon direct/indirect interaction promote virus movement and evasion of plant defences. Transcriptomic studies and virus-host interaction assays have generated a plethora of biologically significant data on tolerant and susceptible host-virus responses within cassava and other species. However, a complete understanding of viral protein mechanisms and clear solutions towards resistance and increased yield remain elusive. Therefore, the objectives of the research reported here were to identify resistance associated genes from transcriptomic data obtained from plants infected with geminiviruses and to gain an understanding of geminiviral systemic movement via SACMV movement protein interactions.

Yeast two-hybrid (Y2H) assays were utilized in the identification of cassava host proteins interacting with the MP cell-cell movement protein of SACMV. Monoclonal antibodies were raised up in rabbit against MP antigenic fragments to co-immunoprecipitate MP-infected host protein (MP-IHP) complexes. *In planta* CRISPR-mediated gene editing in cassava protoplasts was carried out to ascertain the role of three identified host interactors from Y2H assays, namely, nuclear pore complex protein (NUP98), histone-lysine N-methyltransferase ATX4-like (SDG16) and casein kinase 1-like protein HD16 (CKL1) in SACMV accumulation. From a total of 150 million analysed Y2H interactions, 12 novel interactors were identified and ranked according to their predicted biological scores (PBS) from very high to moderate. These were: CKL1 and NUP98A (Very high), lysine--tRNA ligase, cytoplasmic-like (KRS1) (High), photosystem I reaction centre subunit II, chloroplastic-like (PsaD), probable galactinol-sucrose galactosyltransferase 2 (SIP2) (Good) and SDG16, purine uptake permease 1-like (PUP1), homoserine kinase (HSK), phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase and protein-tyrosine-phosphatase (PTEN2A), acyl-CoA N-acyltransferase (NAT), ethylene-responsive transcription factor

ABR1-like (ABR1-like) and protein LIGHT-DEPENDENT SHORT HYPOCOTYLS 4-like (LSH) proteins (Moderate). Co-immunoprecipitation identified seven highly significant proteins, namely gylycine decarboxylase P-protein (GLDP2), heat shock protein 90 (HSP90), histone superfamily protein (H4), ATP citrate lyase subunit B (ACLB-2), sucrose synthase activity (SUS4), ATPase, AAA-type protein (CDC48C) and mitochondrial malate dehydrogenase (mMDH1). Gene editing of NUP98, SDG16 and CKL1 resulted in a down regulation of SACMV accumulation. The protein network of interacting proteins showed that BC1 encoded movement protein (MP) relies on the host for post-translational modifications for its functionality. It also plays a multifunctional role by manipulating host proteins to potentially mediate viral genome movement and activation, and further interferes with global host metabolism contributing to disease symptoms.

DEDICATION

I dedicate this work to my darling mother who will remain loved and appreciated for generations to come.

THEOPHISTA MWANAMBALE

(23 May 1952 - 30 January 2017)

She was always filled with **laughter**, words of **wisdom,** and a **strength** and **peace** beyond understanding for every situation.

Not every day is good, but there's something good in every day.

(The wise words of ***THEOPHISTA MWANAMBALE***)

Praise the Lord, my soul; all my inmost being, praise His holy name. Praise the Lord, my soul, and forget not all His benefits— ***Psalm 103 vs 1-2***

ACKNOWLEDGEMENTS

Dear Prof Rey, thank you for the opportunity to become a better scientist within your laboratory. May the Lord bless and keep you in the years ahead.

I am grateful to the Faculty for the Future (Schlumberger Foundation) for fully funding my research and providing a rare and treasured opportunity to engage with other wonderful women in science.

All my colleagues (past and present) in the Plant Biotechnology lab, that provided amazing support and diverse relational opportunities. The world awaits your goodness.

Lee Ann Soobramoney and Kutlwano Xulu, thank you for filling that part in my heart that needed to finish strong and with the right self-perspective.

Elelwani Ramulifho and Rorisang Seutane, the late fatiguing nights and the long hard days in the laboratory were so much more exciting as you filled them with laughter, compassion, and honest interactions.

Dearest Monica Chansa Mulenga-Gordon, thank you for cheering me on and patiently listening as I went through the most dramatic and heart-wrenching phases of losing my dreams and reclaiming my purpose.

My Alissa Kanyika, a gem beyond measure. Thank you for selflessly praying for me for the past year to finish my PhD well. You have had the opportunity to present any request to your connect group and you chose my success as your prayer request. Thank you for being so selfless and generous. You have shown what it means to be *blessed to be a blessing*.

My Liam Kanyika, thank you for reminding me there was more to my life than a PhD. Thank you for the daily hugs and affection. Thank you for being considerate towards my lack of sleep and late nights.

My Rex Kanyika, I love you. Thank you for the many pots of yummy food and the vacuumed carpets. Thank you for being an available husband and father. Thank you for being the Kanyika that God blessed us with.

Most significant of all, thank You Lord for providing all the love, resources, clarity, strength, and wisdom to be who You created me to be. Thank You for sustaining me in all ways, at all times and in all things.

Even to your old age and gray hairs I am He, I am He who will sustain you. I have made you and I will carry you; I will sustain you and I will rescue you. ***Isaiah 46 vs 4***

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

aa	Amino acid
AC4	RNA-silencing suppressor protein
ACMV	African cassava mosaic virus
AD	Activation domain
ARC	Agriculture Research Council
AtPID	The *Arabidopsis thaliana* Protein Interactome Database
AV2	Pre-coat protein
BCTV	Beet curly top virus
BD	Binding domain
BioGRID	Biological General Repository for Interaction Datasets
BLAST	Basic local alignment search tool
BMRF	Bayesian Markov random field method
BP	Biological process
Ca^{2+}	Calcium ion
CaLCuV	Cabbage leaf curl virus
CC	Cellular component
CMD	Cassava mosaic disease
CP	Coat protein
DTI	Department of Trade and Industry
EACMV	East African cassava mosaic virus
ETI	Effector-triggered immunity
GFP	Green fluorescent protein
GO	Gene ontology
H_2O_2	Hydrogen peroxide
HSP	Heat shock proteins

ICMV	Indian cassava mosaic virus
IR	Intergenic region
IRG	Immunity related genes
JA	Jasmonic acid
LRR	Leucine rich repeats
MAPK	Mitogen activated protein kinase
MF	Molecular function
BC1	Movement protein
MRF	Markov random fields
NBS	Nucleotide-binding site
NMR	Nuclear magnetic resonance
NSP	Nuclear shuttle protein
O^{2-}	Oxide
$ONOO^{-}$	Peroxinitrite
PABC1S	Pathogen-associated molecular patterns
PCR	Polymerase chain reaction
PR	Pathogenesis-related
PTGS	Post-transcriptional gene silencing
RAR1	Mla12 resistance
RBR	Retinoblastoma-related protein
RCR	Rolling circle replication
REN	Replication enhancer protein
REP	Replication-associated/- initiator protein
RNA	Ribonucleic acid
ROS	Reactive oxygen species
SA	Salicylic acid

SA	South Africa
SACMV	South African cassava mosaic virus
siRNA	Small interfering RNA
SLCMV	Sri Lankan cassava mosaic virus
SNF1	Sucrose nonfermenting1
SO	Sequence ontology
SSA	Sub-Saharan Africa
STRING	Search tool for the retrieval of interacting genes/proteins
TAIR	The *Arabidopsis* information resource
TGMV	Tomato golden mosaic virus
TIA	Technical Innovation Agency
TRAP	Transcriptional activator protein
TYLCV	Tomato yellow leaf curl virus
VIGS	Virus-induced gene silencing
Y2H	Yeast two-hybrid

RESEARCH OUTPUTS

Ncube Kanyika, B., Freeborough, W., Gentle, N., Rey, C (2018). A meta-analysis of global host-virus responses in *Solanaceae, Euphorbiaceae* and *Brassicaceae* plant families during geminivirus and ipomovirus infection. Plenary session. *South African Genetics Society / South African Genetics Society (SAGS/SASBi).* 16-18 October 2018. Golden Gate Highlands National Park.

Busisiwe Ncube-Kanyika, Valeria Velásquez-Zapata, Roger P. Wise, Marie Emma Christine Rey. Cassava proteins interacting with the movement protein of South African cassava mosaic virus reveal roles in pathogenicity. **Unpublished.**

Chapter 1

LITERATURE REVIEW

1.1 CASSAVA: CULTIVATION AND RELEVANCE

Manihot esculenta Crantz, commonly called cassava is a tropical and subtropical perennial crop in the family *Euphorbiaceae* (Guo et al. 2014). It is highly heterozygous (Xia et al. 2014; Wang et al. 2014) with a genome spanning an approximate length of 770 Mb for 30,666 genes and a chromosome number of n=18 (Guo et al. 2014). Chromosome sizes across different accessions vary from 1.23 to 2.41 µm per karyotype, with an average chromosome length of 1.74 µm (De Carvalho & Guerra, 2002). It's wild progenitor *M.esculenta* ssp.*flabellifolia* arose from domestication in the Central Brazilian Savannah by propagation of stem cuttings (Allem, 2001). It is tolerant to harsh Sub-Saharan abiotic stress conditions and so its cultivation is wide spread across Sub-Saharan Africa (SSA), Asia and South America (Guo et al., 2014) constituting over 45 accessions (Montero-Rojas et al., 2011) that average over 300 million tonnes per year (FAOSTAT, 2020). Sixty percent is harvested for nutritional needs as a gluten free carbohydrate source and the 40 % as a raw material in industrial applications for the production of renewable energy, starch, paper, glue, and is relevant to pharmaceuticals as an excipient and as a depressant during mineral ore flotation in mining (FAOSTAT, 2020; Guo et al. 2014; Plant village 2015; Bokanga 1995; Poole et al. 2010). Cassava as a hardy drought tolerant crop offers economic value in poor rural communities as a low input cost crop to farm for rural development (FAOSTAT, 2020). Yield in Africa accounts for 58 % of total world yield (FAOSTAT, 2020). Therefore cassava has the potential to play an integral role in African economies (Legg et al., 2014).

1.2 CASSAVA VIRUSES: CASSAVA MOSAIC GEMINIVIRUSES AND CASSAVA BROWN STREAK VIRUS

The study of the impact of plant viruses by morphological descriptions of virus symptoms dates as early as Japan, A.D. 752 in *Eupatorium lindleyanum*, later identified as being caused by eupatorium yellow vein virus (Saunders et al., 2003). An increased understanding of virus transmission was demonstrated by tobacco mosaic virus infection in the late 19th century. The 20th century may be considered the emergence of true virology allowing for isolation, characterisation, sequencing and mechanistic studies of viral pathogens in plants (Hull, 2009). Shortfalls in cassava yield potential are due to diseases caused by two virus families, namely, *Geminiviridae* (genus, *Begomovirus*) and *Potyviridae* (genus, *Ipomovirus*) virus families that comprise cassava mosaic geminiviruses (CMGs) and cassava brown streak virus (CBSV) respectively (Legg & Thresh, 2001; Zinga et al.,

2013; Legg et al., 2015). Both CBSV and CMGs are transmitted by the insect vector *Bemisia tabaci* resulting in cassava yields in certain instances being burdened by both viruses (Brown et al., 2015; Wylie et al., 2018; Zerbini et al., 2017)

Cassava brown streak virus is an RNA virus that belongs to the genus *Ipomovirus*. Ipomoviruses comprise four species that infect in addition to cassava, *Cucurbita, Cucumis* and *Ipomoea* spp. (Wylie et al., 2018). Cassava brown streak virus was initially identified in the first half of the 1900s (Storey & Nichols, 1938) and for seventy-five years was geographically restricted to coastal regions of East Africa (Nichols, 1950). It was considered a virus of low impact to food security and economy until its re-emergence and spread post 2006 (Alicai et al., 2007) encompassing inland regions of Central and North-East Africa (Mulimbi et al., 2012; Munganyinka et al., 2018) and the Comoros islands (Azali et al., 2017) and with the potential to spread to the rest of the tropical and sub-tropical regions that cultivate cassava. Infection results in quality and yield losses ranging from fifty-five to sixty-nine percent of root weight with an additional twenty-four percent loss at market (Hillocks et al., 2001).

Cassava mosaic geminiviruses are DNA viruses with circular single-stranded genomes that belong to the genus *Begomovirus* (Zerbini et al., 2017). Begomoviruses comprise over 300 species of bi- and mono-partite genomes that infect a diverse genus host range. This includes trees such as *Malus*, *Citrus* and *Jatropha* species (Liang et al., 2015; Loconsole et al., 2012; Kashina et al., 2013; Srivastava et al., 2015) and in addition to cassava, other herbaceous species of *Solanum*, *Ipomoea*, *Capiscum*, *Brassica* and multiple legumes (Czosnek & Laterrot, 1997; Wasswa et al., 2011; Srivastava et al., 2017; Akinbade et al., 2010; Mandal et al., 2001; Qazi et al., 2007). Geminiviruses in weed species provide a reservoir for future infection of susceptible, economically relevant host species complicating the control of the viruses (Tahir et al., 2015; Marwal et al., 2014). Therefore, cassava mosaic disease (CMD) is more wide spread than CBSV (Berry & Rey, 2001; Akinbade et al., 2010; Emily et al., 2016; Ntui et al., 2015) resulting in greater economic yield losses. Yield losses are in the range of 65 % to 74 % and are aggravated by excessive vector infestation (Legg et al. 2015; Legg et al. 2014). Disease incidence ranges from 20 % to 80 % and can reach 100 % in individual fields, with high severity (Toualy et al., 2014; Legg et al., 2015). In addition, the presence of more than one strain of virus also results in a synergistic relationship between the virus strains resulting in greater symptom severity and thus detrimental to yield turnover (Zinga et al., 2013; Naseem & Winter, 2016). Yield losses from CMD have had an economic impact in the tropical and Sub-Saharan regions exceeding $2500 million in Africa, India and the United States of America (Varma & Malathi, 2003; Scholthof et al., 2011).

1.3 CASSAVA MOSAIC GEMINIVIRUS

1.3.1 TAXONOMY AND DIVERSITY

The virus family *Geminiviridae* comprises seven genera namely, *Becurtovirus, Grablovirus, Begomovirus, Curtovirus, Eragrovirus, Mastrevirus, Turncurtovirus, Capulavirus* and *Topocuvirus* (Zerbini et al., 2017). The latter two are transmitted by a treehopper and aphid respectively (Zerbini et al., 2017). The largest genus across all viruses is the *Begomovirus* genus with 322 species of bi- and mono-partite genomes that are recognized by the International Committee on Taxonomy of Viruses (Brown et al., 2015). Brown and colleagues carried out species demarcation criteria for the begomoviruses emanating from true pairwise sequence alignments, with the exclusion of sites with gap characters using SDT software for comparisons of 1,826 full-length begomovirus genomes from the NCBI-GenBank database at a sequence identity cut off value of 91 % to separate isolates from different species (Brown et al., 2015).

Cassava mosaic geminiviruses (CMGs) namely, African cassava mosaic virus (ACMV), cassava mosaic Madagascar virus (CMMGV), East African cassava mosaic virus (EACMV) (5 isolates), East African cassava mosaic Kenya virus (EACMKV), East African cassava mosaic Malawi virus (EACMMV), East African cassava mosaic Zanzibar virus (EACMZV), Indian cassava mosaic virus (ICMV) (4 isolates), South African cassava mosaic virus (SACMV) and Sri Lankan cassava mosaic virus (SLCMV) (2 isolates). The diversity of virus variants is strongly attributed to genome re-assortment/ inter- and intra-specific recombination and pseudo-recombination (the acquisition of genes from other genera) (Lefeuvre et al., 2007; Padidam et al., 1999; Silva et al., 2014). Lefeuvre and colleagues statistically ascertained that recombination is a controlled evolutionary mechanism that occurs in genome locations that when encoded do not alter the three dimensional structures of chimeric proteins. The presence of recombination hotspots further adds to the principle of recombination and pseudo recombination functioning primarily in suppression of the host plants post-transcriptional gene silencing (PTGS) defence mechanism thereby increasing virulence and host range (Seal et al., 2006; Lefeuvre et al., 2007). Pseudo recombination results in viruses of low biological fitness, the high recombination rate of geminiviruses may alleviate this decline in virus ability by transfer of the intergenic region, permitting efficient replication and virulence (Zhou et al., 1997). Geminiviruses exhibit an interspecific recombination affinity within the region of the origin of replication as was noted in studies of, cotton leaf curl virus (CLCuV-PK) interspecific recombinants with okra yellow vein mosaic virus (OYVMV); EACMV with an unknown begomovirus that carries similar CP and AV2 as that of tomato yellow leaf curl virus-Israel (TYLCV-Is) (Zhou et al. 1998; Zhou et al. 1998). Recombination functions largely in adaption of the virus to new hosts, increased virulence and vector compatibility (Zhou et al., 1997). EACMV presents diversity and frequent recombination whilst ACMV exhibits fixed genome sequence clusters

(Naseem & Winter, 2016). Viral genome replication occurs in meristematic tissues and the phloem; the latter being less efficient is initiated by translesion synthesis polymerases. Lower efficiency may attribute to a high frequency of point mutations and new disease strains (Richter et al., 2016). Replication across geminiviruses is by three methods; complementary strand replication (CSR), recombination dependent replication (RDR) and rolling circle replication (RCR) and this diversity in replication further increases the genetic flexibility of geminiviruses that adds to new epidemics (Saunders et al., 1991; Maghuly et al., 2014).

1.3.2 TRANSMISSION AND SYMPTOMS

There are eleven species of CMGs (Family *Geminiviridae*, Genus *Begomovirus*), primarily transmitted by arthropod silver leaf whitefly, *Bemisia tabaci Gennadius* (B biotype) (Family *Aleyrodidae*) (Patil & Fauquet, 2015). Whitefly infestation results in leaf chlorosis and a secondary effect of the whitefly population is the accumulation of honeydew which promotes mould growth and further limits the photosynthetic ability of the plant, promoting early senescence (CABI and EPPO, 2016). *Bemisia tabaci* populations, have been predicted to thrive and spread to areas previously uninhabited by the insect, and with the onset of increasing global temperatures and atmospheric carbon dioxide intensifying their economic impact on existing cassava and other cash crops globally (Curnutte et al., 2014; Gilioli et al., 2014). The association of *B. tabaci* with the virus is categorised as persistent/circulative in nature (Opara, 2015; Liu et al., 1997). The acquisition of the virus is for the vector lifespan and the virus once taken up goes through a ‘silent’ avirulent phase within the vector before the vector may become an active transmitter and, successful transmission requires that the virus is taken into the vector homeocel (mediated by virus DNA B movement proteins) and into the salivary organs for secretion into the plant (mediated by DNA A coat protein) (Opara, 2015; Liu et al., 1997). Secondary transmission occurs through contaminated cutting tools used for preparation and propagation of pre-infected cuttings (Toualy et al., 2014; CABI and EPPO, 2016). The control of whiteflies would positively impact cassava geminivirus control. Studies on natural enemies for *Bemisia tabaci* have identified at least nine parasites from the genera *Amitus*, *Encarsia* and *Eretmocerus*, four pathogens from the genera *Aschersonia*, *Beauveria*, *Isaria* and *Paecilomyces* and lastly five predators from the genera *Ceraeochrysa*, *Chrysoperla*, *Delphastus*, and *Nephaspis* (CABI and EPPO, 2016). Mycoinsecticide products have been derived from *Verticillium lecanii*, *Paecilomyces fumosoroseaus* and *Beauveria bassiana* however usage has been limited by the cost of purchase, shelf life and dependence on humid and warm environmental conditions as opposed to the relatively easier acquisition and application of more toxic non-biological insecticides preferred by the farmer (Faria & Wraight, 2001; Mascarin et al., 2013).

The phenotypic characteristics of the CMD post infection are a symptomatic effect of altered phytohormone signalling in infected plant cells (Allie et al., 2014) and are generally noted on the leaves. Symptoms may be broadly categorised from post infection as follows: (A) asymptomatic leaves, (B) mild mosaic, (C) severe mosaic, (D) severe mosaic, distortion, leaf curl, (E) severe mosaic, severe distortion, severe reduction of leaf area, stunting. Severe damage of the cassava roots is also readily noted at harvesting of the tuber (Plant village, 2015; Emily et al., 2016). Disease scoring is based on a standard scale of 1 to 5 based on leaf chlorosis, distortions and plant growth or stunting. Significant to the severity of infection are episomal SEGS1 and SEGS2 (sequences enhancing geminivirus symptoms 1 and 2) that are beta- and alpha-satellite sequences respectively that show 23 % sequence homology and are derived from the cassava genome (Ndunguru et al., 2016). The presence of SEGS1 and SEGS2 interferes with disease tolerance and aggravates the rate and overall symptom severity within seven days of infection (Ndunguru et al., 2016). Small RNA-Seq data revealed the presence of iSEGS-derived small RNAs targeting SEGS1 and SEGS2 and deductions were made that iSEGS are redolent of transposons and influence differential gene expression largely related to biological processes (Maredza et al., 2016). Singularly, iSEGS-derived small RNAs do not induce any infection or negatively alter plant physiology (Ndunguru et al., 2016).

Furthermore, synergy between CMGs aggravates symptom progression (Vanitharani et al., 2004). In infections involving two CMGs, ACMV and EACMV, it was noted that the targeting of the plant silencing pathway via the AC4 and AC2 proteins respectively, varies in time and location, synergistically facilitating the accumulation of viral DNA for more severe symptoms (Vanitharani et al., 2004). Epidemiologically, the increased viral titre from synergistic infections increases the probability of virus transmission by whitefly vectors and further increases the risk of the spread of a more insistent CMD (Fondong et al., 2000).

1.3.3 GENOME STRUCTURE

Cassava mosaic geminiviruses are bipartite, single-stranded, closed circular DNA viruses in which the DNA components, DNA-A and DNA-B molecules are ~2800 and ~2760 nucleotides long, respectively and encapsidated in 30.2 kDa protein subunits (Böttcher et al., 2004). Subunits are arranged as twinned particles of 22.5 nm in diameter comprising two incomplete icosahedra of eleven capsomers each **(Figure 1.1)** (Böttcher et al., 2004; Patil & Fauquet, 2009). Both DNAs contain an intergenic region (IR) which also contains a common promoter region (~200 bp) which is essential for the expression of proteins key in viral replication and systemic invasion of the host (Legg et al., 2015; Bisaro, 1996). It is assumed that the DNA-B component of CMGs may have been a circular ssDNA satellite which through recombination became a part of the geminivirus genome (Nawaz-ul-Rehman & Fauquet, 2009).

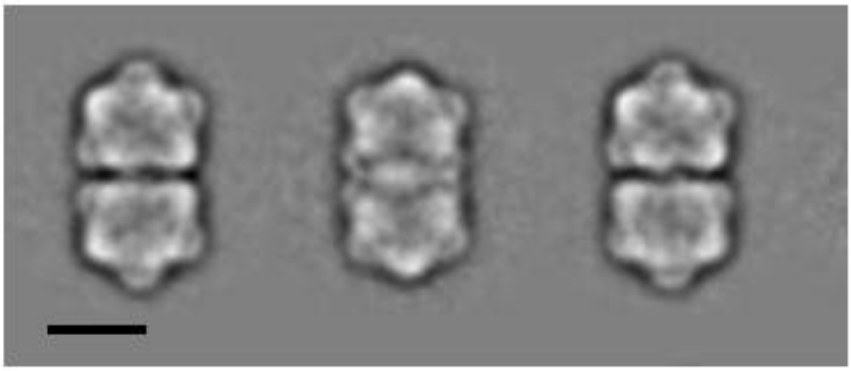


Figure 1.1 Two-dimensional projection maps of cassava geminivirus ACMV particles. Bar 20 nm. Courtesy Böttcher et al., 2004.

DNA-A and DNA-B have six and two open reading frames (ORFs) respectively encoded in the v and c sense respectively **(Figure 1.2)** (Gutierrez, 2000). The proteins coded for in DNA-A are as follows: AV1, coat protein (CP); AV2, pre-coat protein; AC1, replication-associated/-initiator protein (Rep); AC2, transcriptional activator protein (TrAP); AC3, replication enhancer protein (REn) and AC4, RNA-silencing suppressor protein. DNA-B codes for proteins that facilitate plant virus movement as follows: BC1, cell to cell movement protein (MP); BV1, nuclear shuttle protein (NSP) from nucleus to cytoplasm (Zerbini et al., 2017). SACMV specifically constitutes, DNA-A: AV1 (258 aa) and AV2 (116 aa) in the virion-sense direction and overlapping reading frames AC1 (354 aa), AC2 (135 aa) and AC3 (134 aa) and AC4 (98 aa) in the complementary-sense direction; DNA-B: BV1 (258 aa) and BC1 (307 aa) (Berrie et al., 2001).

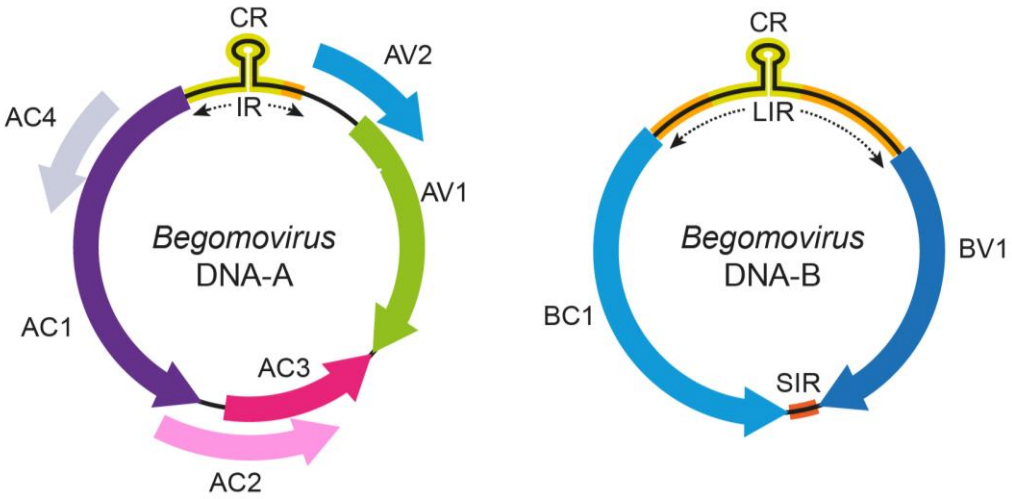


Figure 1.2 Schematic diagram of geminivirus genome organization. Open reading frames (ORFs) and their direction of transcription in the virion-sense (V) or complementary-sense (C) strand, and corresponding protein products are coded by colour. The position of the stem-loop containing the

conserved 5′-TAAGATTCC-3′ sequence located in the long intergenic region (LIR) is shown. CR denotes common regions of DNA A and B; AV1 (CP), coat protein; IR, intergenic region; MP, movement protein; BV1 (NSP), nuclear shuttle protein; AC3 (Ren), replication enhancer protein; AC1 (Rep), replication-associated protein; SIR, short intergenic region; TrAP, transcriptional activator protein. Image and caption adapted from Zerbini et al., 2017 and International Committee on Taxonomy of Viruses (ICTV)

1.3.4 GEMINIVIRUS REPLICATION

Geminiviruses, SACMV included, transcriptionally alter host cell cycle genes to induce conditions that favour the expression of host DNA synthesising machinery for virus replication (Hanley-Bowdoin et al. 2013; Pierce & Rey 2013). At least 72 % of core cell cycle regulators have been shown to be altered in gene expression during SACMV infection (Pierce & Rey, 2013; Allie et al., 2014). Replication of cassava geminiviruses is by host DNA polymerase via rolling circle replication (RCR) initiated from the IR region at a nonananucleotide (TAATATTAC) region of replication and primarily initiated by the Rep protein (Saunders et al., 1991; Fondong, 2013). Viral replication may also transition to recombination-dependent replication, which is initiated by homologous recombination between a partially replicated ssDNA and a closed, circular dsDNA to form a looped molecule that serves as a template for both ssDNA and dsDNA strand synthesis (Hanley-Bowdoin et al., 2013).

Rep introduces a nick generating a free 3′-hydroxyl end and covalently binds to the 5 ' terminus via a phosphotyrosine linkage. The nicked strand is converted to dsDNA by the host replisome by RNA-primed synthesis of a complementary (minus) strand (Bisaro, 1996). This regenerates the origin of replication, which again is nicked by Rep, this time acting as a terminase to release the displaced unit-length plus strand, which is simultaneously ligated to circular form by the closing activity of Rep. Viral replication transitions to recombination-dependent replication, which is initiated by homologous recombination between a partially replicated ssDNA and a closed, circular dsDNA to form a looped molecule that serves as a template for both ssDNA and dsDNA strand synthesis. Thereafter, Rep represses its own transcription, for the accumulation of viral genomes for encapsidation by activation of TrAP expression, which in turn activates CP and NSP expression. Circular ssDNA can then be encapsidated by CP into virions, which are then available for whitefly acquisition and conveyance (Bisaro, 1996; Hanley-Bowdoin et al., 2013) via a six-residue insertion near the tip of the capsomer (Böttcher et al., 2004). NSP binds to viral DNA and moves it across the nuclear envelope, where movement protein (MP) traffics it across plasmodesmata (Bisaro, 1996; Hanley-Bowdoin et al., 2013).

1.4 GEMINIVIRUS HOST MANIPULATION

The understanding of plant-virus relationships helps in the understanding of the underlying molecular aspects of functional genomics of host-virus relationships. Efforts have been made to understand cassava-geminivirus interactions at a transcriptional level (Ascencio-Ibanez et al., 2008; Góngora-Castillo et al., 2012; Pierce & Rey, 2013; Allie et al., 2014; Kushwaha et al., 2015; Liu et al., 2014). The differential up or down expression of genes (Allie et al., 2014) and specifically miRNA expression patterns in the host are indicative of CMG manipulation of the host genome to counteract host defences and consequently result in physiological and morphological abnormalities in the host plant (Bizabani et al., 2021; Rogans & Rey, 2016). In susceptible cassava T200, the synthesis of cassava defence proteins may be reduced by 68 % hindering effective mechanisms of phytohormone signalling, regulated gene expression and pathogen-alert signal transduction (calmodulin-binding proteins); defence related transcription factors (WRKY); heat shock proteins (HSPs); and regulatory and defence proteins (mitogen activated protein kinase-MAPK) (Allie et al., 2014). The upregulation of host genes *pectin lyase* and *plant invertase/pectin methylesterase inhibitor superfamily* members have been proposed to facilitate viral movement across cell walls by cell wall degradation (Allie et al., 2014). Cell wall degradation increases plasmodesmata size exclusion limits to allow cell-to-cell movement, and long distance movement via the phloem to other organs (Allie et al., 2014; Uhrig & MacFarlane, 2008).

Protein-protein interactions are categorised into *in vitro*, *in vivo* and *in silico* and a full protein interaction study may incorporate two or more of the methods for validation of observed interactions. *In vivo* yeast two-hybrid (Y2H) assays and bimolecular immunofluorescence assays (BiFC), complemented with *in vitro* co-immunoprecipitation (Co-IP), *in planta* gene editing and *in silico* computational analysis, are favourable approaches to consider in the demonstration and understanding of the complexity of geminivirus-host interactions (Teixeira et al., 2021; Czosnek et al., 2013; Kumar, 2019; Guerrero et al., 2020; Bosque et al., 2014; Deblasio et al., 2015). Information on host transcriptome responses and protein-protein interaction studies complement an understanding of host-virus responses (Ludvigsen & Honoré, 2018). Geminiviruses interact with host proteins to manipulate the plant host into facilitating their replication and movement (Culver & Padmanabhan, 2007; Sunter & Bisaro, 1992; Pascal et al., 1993; Aberle et al., 2002; Baliji et al., 2010; Gouveia-mageste et al., 2020). Rep protein interacts with retinoblastoma-related (RBR) protein via its conserved alpha helical region and in a study on maize streak virus in cereals, Rep-RBR binding was essential in the proliferation of mesophyll cells indicated by symptom severity (McGivern et al., 2005). The RBR-Rep interaction inhibits the interaction of phosphorylated RBR with E2F transcription factors allowing the latter to activate transcription of polymerases and other replication associated

proteins in differentiated cells (Hanley-Bowdoin et al., 2013). Furthermore, Rep of ACMV harbours a putative cyclin interaction motif (RXL) to facilitate reconditioning of the cell cycle and host replication factors (Hipp et al., 2014). The upregulation of cell cycle genes that promote S-phase of the cell cycle have been identified as key features of geminivirus infection (Allie et al., 2014; Pierce & Rey, 2013; Ascencio-Ibanez et al., 2008). Rep protein further interacts with ubiquitin-conjugating enzyme2 (UBC2) and histone monoubiquitination1 (HUB1) for the monoubiquitination of histone 2B that subsequently promotes trimethylation of histone 3 at lysine 4 on ChiLCV mini-chromosomes (Kushwaha et al., 2017).

African cassava mosaic virus (ACMV) and EACMV both encode post-transcriptional gene silencing (PTGS) suppressors AC4 and AC2 (TrAP) respectively (Vanitharani et al., 2004). The AC4 of ACMV Cameroon strain in *Arabidopsis* interferes with the miRNA duplex unwinding process and inactivates mature miRNAs inhibiting regulation of target mRNAs and inevitably destabilising cell development and regulation of biochemical processes (Chellappan et al., 2005). Transcriptional activator protein (TrAP) additionally also functions in PTGS of the host by interfering with metabolic pathway regulators adenosine kinase (ADK) (Maghuly et al. 2014; Hao et al. 2003; Soosaar 2005) and sucrose nonfermenting1 (SNF1) kinases (Fontes et al., 2004; Hao et al., 2003). Adenosine kinase (ADK) activity in an immune response methylates the 24 nt intergenic region of the virus genome (Rodríguez-Negrete et al., 2009) whilst SNF1 kinases show evidence of conserved function to regulate plant metabolism during stress (Hao et al., 2003). AC4 suppresses miR159 which is crucial in regulation of vegetative growth, flowering time, anther development, seed shape, germination plant growth, morphogenesis and reproduction (Maghuly et al., 2014; Alonso-Peral et al., 2012).

The interaction of NSP with acetyltransferase is essential to geminivirus infection and pathogenicity (Carvalho & Lazarowitz, 2004). It inhibits its interaction with histones and therefore maintains the cell in a dedifferentiated state suitable for virus proliferation (McGarry et al., 2003). NSP interacts with ssDNA and MP (in the cytoplasm) via its N-terminal and C-terminal respectively and upstream of the C-terminal is a leucine-rich nuclear export sequence (NES) that facilitates shuttling of nuclear proteins (Carvalho & Lazarowitz, 2004). The NSP interferes with primary immunity related kinase protein NIK1, hindering its function as an upstream immune signalling component. This is by NSP interaction with an 80 amino acid area of NIK1 that bears the putative active site and activation loop that is crucial for control, substrate recognition and overall catalytic activity of the kinase (Fontes et al., 2004). NSP binds to the kinase domain blocking RPL10 phosphorylation. Phosphorylated RPL10 hinders virus activity in the nucleus (Hanley-Bowdoin et al., 2013).

1.5 GEMINIVIRUS CELL-TO-CELL MOVEMENT PROTEIN (MP): MECHANISMS AND INTERACTIONS

1.5.1 PLANT GEMINIVIRUS SYSTEMIC INFECTION

In order for plant viruses to systemically infect their plant hosts, movement proteins are encoded to facilitate the translocation of viral DNA (Kawakami et al., 2004; Wei et al., 2010; Reagan & Burch-Smith, 2020; Ding et al., 1995; Su et al., 2010). Studies on plant viruses have shown that viruses encode movement proteins that are phosphorylated by host proteins (Kawakami et al., 2003; Kleinow et al., 2009) conferring functionality to move via ER-associated microtubules and filaments (Boyko et al., 2000; Christensen et al., 2009). The movement of viral proteins and viral DNA relies on the use of plasmodesmata by at least three possible mechanisms. These are, (i) direct interaction with the plasmodesmata to increase the size exclusion limit (SEL) enabling viral passage (Ding et al., 1995); (ii) interaction with host factors that mediate the formation of structures anchored to and through plasmodesmata (Wei et al., 2010); (iii) sever plasmodesmata-associated actin filaments to increase SEL (Su et al., 2010).

1.5.2 GEMINIVIRUS MOVEMENT PROTEIN (MP) FUNCTIONALITY

Phosphorylation of Abutilon mosaic virus (AbMV) MP in *E. coli* (Wege & Jeske, 1998) is an essential modification by host proteins towards AbMV MP functionality and in facilitating ubiquitination by the host (Kleinow et al., 2009). AbMV MP site directed mutagenesis in the C-terminal domain at Ser-250 (replaced with alanine or aspartate) phosphorylation site, compromised viral titre and symptoms (Kleinow et al., 2009). However, symptom severity and virus accumulation were enhanced in mutagenesis of C-terminal domain at Thr-221 and Ser-223 (replaced with alanine or aspartate) (Kleinow et al., 2009). Variable responses from amino acid specific mutagenesis on disease progression shown by Kleinow and colleagues infers to a complex multiple MP-host protein interaction (Kleinow et al., 2009). Symptom progression in *N. benthamiana* by tomato golden mosaic virus (TGMV) strains was further identified as an MP dependent mechanism (Stanley & Albrecht, 1992). The geminivirus MP was further identified as a host determinant in SqLCV (Ingham et al., 1995). Studies over the past thirty years have substantiated MP functionality in symptom development in TGMV (Stanley & Albrecht, 1992) and SqLCV MP expressed by transgenic *N. benthamiana* (Pascal et al., 1993). This was further established by BDMV MP eliciting a viral phenotype in non-host tomato cultivars (Hou et al., 2000). The predominantly disordered molecular structure of SACMV MP (Nankoo, 2018) further showed its potential capability in a role influencing the progression of SACMV infection

1.5.3 INTER- AND INTRA-CELLULAR LOCALISATION OF GEMINIVIRUS MP

The role of MP in cell-to-cell migration via plasmodesmata was established in bean dwarf mosaic virus (BDMV) (Noueiry et al., 1994). This was confirmed by squash leaf curl virus (SqLCV) studies in Sf9 insect cells and *N. benthamiana* protoplasts in which MP (BL1) localised NSP (BR1) to the peripheral membrane (Sanderfoot & Lazarowitz, 1995) and cell wall and plasma membrane fractions (Erica Pascal et al., 1993). The facilitated movement of MP within the cytosol via interaction with the cytoskeleton has been elucidated in AbMV studies (Frischmuth et al., 2007; Aberle et al., 2002; Krapp et al., 2017). AbMV-MP reorganised microtubules in mammalian cells (Krapp et al., 2017). In freeze-fracture immuno-labelling, AbMV MP was localised to microsomal vesicles and the plasma membrane in yeast (Aberle et al., 2002; Frischmuth et al., 2004). During SACMV infection, genes that encode for plasmodesmata-associated proteins and beta-1,3-glucanase enzymes were up regulated further supporting viral movement via plasmodesmata (Allie et al., 2014; Pierce & Rey, 2013).

1.5.4 MOVEMENT OF GEMINIVIRAL DNA

The begomoviral coat protein (CP) has been shown through ACMV CP studies to possess three phosphorylation sites in the N-terminus, threonine 12, serine 25 and serine 62 (McGarry et al., 2003; Hipp et al., 2016) and domain activities are assigned functions for viral DNA movement across nuclear membranes (Unseld et al., 2001) by compacting viral DNA (Hipp et al., 2016). The binding of CP to DNA has been assumed to be nonspecific and influenced by the opposite isoelectric points of DNA and CP, however it was revealed that structural changes in the viral DNA may be because of specific binding. The CP does not act alone and additional viral proteins are necessary to mediate infectivity (Carvalho & Lazarowitz, 2004). *Arabidopsis* infectivity studies with cabbage leaf curl virus (CLCV) and squash leaf curl virus (SqLCV) deduced from reverse Y2H and selective mutation that NSP recruits acetyltransferase (NSI) a nuclear protein, to the ssDNA-CP complex, displacing the CP to form an NSP-DNA complex that facilitates channelling across the nuclear membrane and thereafter interacts with the MP for systemic proliferation (Carvalho & Lazarowitz, 2004) through plasmodesmata (Hehnle et al., 2004) by a transport mechanism termed as the “couple-skating model”. In competition binding assays, NSP exhibited an affinity for circular ss- and ds-DNA and MP for open circular ssDNA binding (Rojas et al., 1998).

The couple skating model for viral DNA movement has been hypothesised to involve a complex of ssDNA/dsDNA-NSP to compact viral DNA. The NSP-DNA complex, interacts with MP and MP interacting host proteins to shuttle viral DNA from cell to cell via plasmodesmata (Carvalho & Lazarowitz, 2004). This model has been further legitimised

with regards to the AbMV NSP-MP-DNA complex by electron microscopy (Hehnle et al., 2004). Other studies have identified host protein interactors common to both NP and MP (Zhou et al., 2011; Gouveia-mageste et al., 2020) and are indicative of multiple cytosol localised interactions (Carvalho et al., 2008) to facilitate viral DNA movement.

1.5.5 MOVEMENT PROTEIN-HOST PROTEIN INTERACTIONS

Most host proteins that interact with BC1 encoded MP in begomoviruses have only been discovered recently **(Table 1.1)**. In the past ten years, studies have revealed MP to have at least forty-five host protein interactors (Gouveia-mageste et al., 2021; Uchiyama et al., 2014; Krenz et al., 2010; Lewis & Lazarowitz, 2010). The most recently identified interactors by Gouveia-mageste and colleagues in cabbage leaf curl virus (CabLCV) expanded on the roles of geminivirus MP and its cooperative binding with NSP in key aspects of auxin response, potentially negating the ubiquitin-proteasome pathway and interactions that would enable movement via nucleotide binding (Gouveia-mageste et al., 2021). Specific to the study by Gouveia-Mageste, were interactors distinct to MP with roles involved in gene regulation (LA RNA-binding protein, Small nuclear ribonucleoprotein family protein), cellular signalling and regulation (tetratricopeptide repeat (TPR)-like superfamily protein), microtubule mediated mechanisms (FKBP-like peptidyl-prolyl cis-trans isomerase family protein), and DNA compaction (KTI12-like, chromatin associated protein) (Gouveia-mageste et al., 2021) **(Table 1.1)**.The identification of common interactions is further indicative of the NSP-MP protein complex within the cytosol to cooperatively facilitate the movement of viral DNA. AbMV MP interaction with pin4 and SCD2 proteins further associated viral movement with the cytoskeleton in both mammalian and plant cells (Krapp et al., 2017; Gouveia-mageste et al., 2021). NSP and MP interact with histone H3 which plays a role in compacting viral DNA into nucleosomes, involved in both replication and movement of geminiviruses (Hanley-Bowdoin et al., 2013). Histone H3 binds via its N-terminal region to BDMV (Zhou et al., 2011). Gel overlay assays revealed the migration of H3 from the nucleolus to the plasmodesmata in the presence of MP via the H3-NSP-MP complex (Zhou et al., 2011). Yeast two-hybrid assays and BiFC assays respectively showed AbMV MP homo-oligomerisation via the C terminal domain and interaction with chloroplastic HSC70 via the N terminal (Krenz et al., 2010). Krenz and colleagues, further silenced HSC70 by virus induced gene silencing and noted a reduction in virus spread and an attenuation of virus symptoms (Krenz et al., 2010). The MP-HSC70 interaction in enabling virus proliferation may be attributed to HSC70 import role into the endoplasmic reticulum (ER) (Tsai et al., 2000; Abell et al., 2007). Thus, the HSC70-MP interaction identified as essential to successful AbMV infection (Krenz et al., 2010). Heat shock proteins form motile granules that are associated with actin microfilaments that are

Table 1.1 MP and NSP host interacting proteins. Host proteins identified specific to MP highlighted in blue. Source studies' and their DOIs as follows: Krenz et al. 2010, 10.1016/j.virol.2010.02.011; Lewis and Lazarowitz 2010, 10.1073/pnas.0909080107; Zhou et al., 2011, 10.1128/JVI.00082-11; Uchiyama et al., 2014, 10.3389/fpls.2014.00584; Krapp et al. 2017, 10.3390/v9110334 and Gouveia-Mageste et al., 2021, 10.1101/2020.01.10.901496.

INTERACTING PROTEIN	DESCRIPTION	SOURCES
AT5G02500	HSC70	Krenz et al. 2010
AT1G09200	histone H3	Zhou et al., 2011
AT2G20990	Arabidopsis synaptotagmin SYTA	Uchiyama et al., 2014; Lewis and Lazarowitz, 2010
AT2G01420	pin-formed 4	Krapp et al. 2017
AT3G48860	stomatal cytokinesis defective 2	Krapp et al. 2017
AT2G17670	tetratricopeptide repeat (TPR)-like superfamily protein	Gouveia-Mageste et al., 2021
AT4G28190	developmental regulator, ULTRAPETALA	Gouveia-Mageste et al., 2021
AT5G21160	LA RNA-binding protein	Gouveia-Mageste et al., 2021
AT5G14000	NAC domain containing protein 84	Gouveia-Mageste et al., 2021
AT5G54580	RNA-binding (RRM/RBD/RNP motifs) family protein	Gouveia-Mageste et al., 2021
AT1G73655	FKBP-like peptidyl-prolyl cis-trans isomerase family protein	Gouveia-Mageste et al., 2021
AT1G13870	KTI12-like, chromatin associated protein	Gouveia-Mageste et al., 2021
AT4G14560	indole-3-acetic acid inducible	Gouveia-Mageste et al., 2021
AT2G42390	kinase C substrate, heavy chain-like protein	Gouveia-Mageste et al., 2021
AT3G46900	copper transporter 2	Gouveia-Mageste et al., 2021
AT1G28230	purine permease 1	Gouveia-Mageste et al., 2021
AT2G43810	small nuclear ribonucleoprotein family protein	Gouveia-Mageste et al., 2021
AT1G68185	ubiquitin-like superfamily protein	Gouveia-Mageste et al., 2021
AT1G79930	heat shock protein 91	Gouveia-Mageste et al., 2021
AT1G52730	transducin/WD40 repeat-like superfamily protein	Gouveia-Mageste et al., 2021
AT4G17230	SCARECROW-like 13	Gouveia-Mageste et al., 2021
AT5G06970	plant/protein (DUF810)	Gouveia-Mageste et al., 2021
AT3G26710	cofactor assembly of complex C	Gouveia-Mageste et al., 2021
AT1G29250	alba DNA/RNA-binding protein	Gouveia-Mageste et al., 2021
AT2G44020	mitochondrial transcription termination factor family protein	Gouveia-Mageste et al., 2021
AT1G27930	glucuronoxylan 4-O-methyltransferase-like protein (DUF579)	Gouveia-Mageste et al., 2021
AT5G38610	plant invertase/pectin methylesterase inhibitor superfamily protein	Gouveia-Mageste et al., 2021
AT3G51590	lipid transfer protein 12	Gouveia-Mageste et al., 2021
AT2G37640	barwin-like endoglucanases superfamily protein	Gouveia-Mageste et al., 2021
AT5G66920	SKU5 similar 17	Gouveia-Mageste et al., 2021
AT1G03440	leucine-rich repeat (LRR) family protein	Gouveia-Mageste et al., 2021
AT3G20280	RING/FYVE/PHD zinc finger superfamily protein	Gouveia-Mageste et al., 2021
AT4G27657	hypothetical protein	Gouveia-Mageste et al., 2021
AT3G16100	RAB GTPase homolog G3C	Gouveia-Mageste et al., 2021
AT1G22900	disease resistance-responsive (dirigent-like protein) family protein	Gouveia-Mageste et al., 2021
AT2G36970	UDP-Glycosyltransferase superfamily protein	Gouveia-Mageste et al., 2021
AT1G78460	SOUL heme-binding family protein	Gouveia-Mageste et al., 2021
AT1G68440	transmembrane protein	Gouveia-Mageste et al., 2021
AT4G05060	papD-like superfamily protein	Gouveia-Mageste et al., 2021
AT4G27860	vacuolar iron transporter (VIT) family protein	Gouveia-Mageste et al., 2021
AT4G12890	gamma interferon responsive lysosomal thiol (GILT) reductase family protein	Gouveia-Mageste et al., 2021
AT3G62600	DNAJ heat shock family protein	Gouveia-Mageste et al., 2021
AT4G00370	major facilitator superfamily protein	Gouveia-Mageste et al., 2021
AT1G33810	zinc finger/BTB domain protein	Gouveia-Mageste et al., 2021
AT5G20730; AT4G14560	interacting indole-3-acetic acid inducible 28	Gouveia-Mageste et al., 2021

essential in translocation to plasmodesmata and generally HSP play a role in viral protein unfolding which may permit passage through plasmodesmata (Uhrig & MacFarlane, 2008). The interaction of MP with SYTA provides proof of endocytic movement from the plasma membrane to plasmodesmata (Uchiyama et al., 2014; Lewis & Lazarowitz, 2010).

The identified interactors localise to multiple Cellular Component categories confirmative of the role of MP in systemic movement of the geminiviral genome **(Figure 1.3)**. Geminivirus MP in its function to translocate viral DNA acts co-operatively with geminiviral

protein NSP, and forms complexes with host proteins that possess the ability to bind to nucleotides, proteins, lipids, and enzymatic activity **(Figure 1.4)**. Therefore, the interactions of the geminivirus movement protein show an interference with proteins involved in multiple cellular processes.

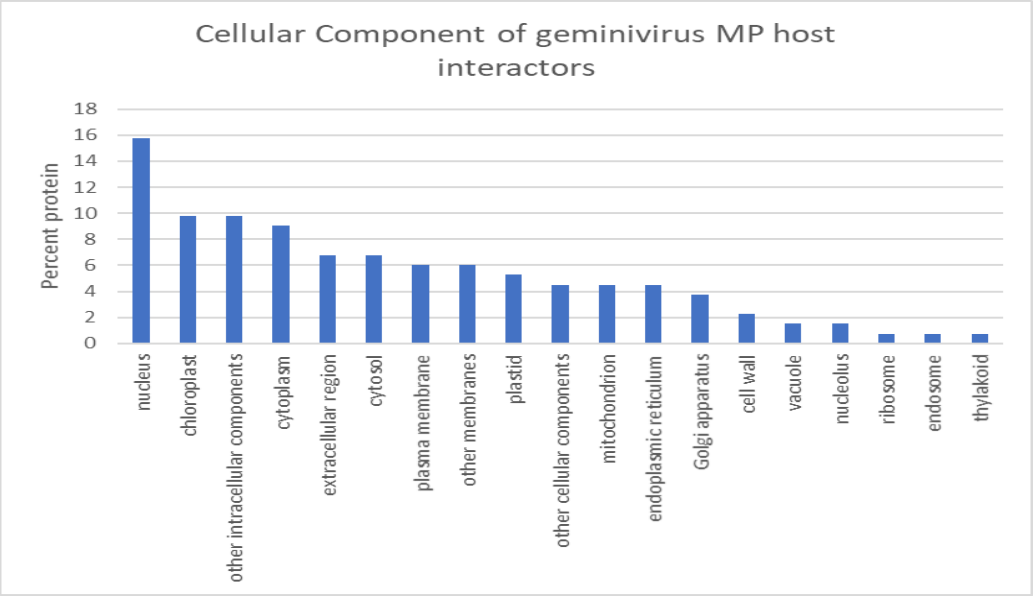


Figure 1.3 The GO Cellular Component categories of geminivirus MP interacting proteins identified to date. Sourced from TAIR.

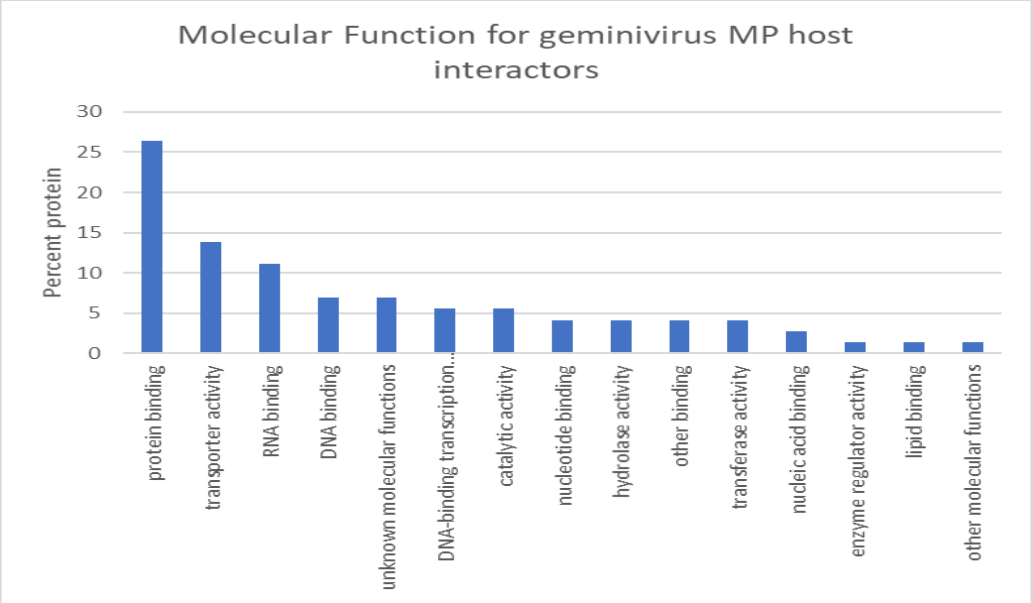


Figure 1.4 The GO Molecular Function categories of geminivirus MP interacting proteins identified to date. Sourced from TAIR.

1.6 PLANT IMMUNITY

1.6.1 TOLERANCE AND RESISTANCE TO CASSAVA GEMINIVIRUSES

With regards to resistance and tolerance in cassava, the terms have been used interchangeably despite only tolerance being a distinct quality of cassava-virus immunity. For clarity, both terms resistance and tolerance will be used with respect to articles cited. Resistant and tolerant (Adriko et al. 2011; Naseem & Winter 2016; Dias do Carmo et al. 2015) genotypes are available and attributed to a single dominant resistance CMD2 gene, a polygenic and recessive CMD1 gene and a CMD3 gene (Okogbenin et al. 2012; Wolfe et al. 2015). A WRKY transcription factor is associated with the CMD2 gene (Fregene et al., 2004). The CMD2 has been incorporated into susceptible genotypes and is better understood and its mechanism of coordinated association with avirulence (Avr) genes and resistance cofactors studied (Allie et al. 2014; Dias do Carmo et al. 2015; Wang et al. 2014). Genome-wide association analysis (GWAS) revealed at least four (4) SSR markers that are associated with SSRY40 (associated with quantitative resistance) and 13 regions of small effect (Wolfe et al. 2015). Multiple interacting loci in tolerance to CMD involving both CMD2 and CMD3 have been postulated (Wolfe et al. 2015). Breeding efforts have led to the identification of molecular markers for resistance gene selection i.e. SSR Y28 and RFLP GY1 markers flanking resistance gene CMD2 (Parkes et al., 2015). Rabbi et al. 2014 constructed a genetic map using single nucleotide polymorphisms (SNPs) to construct a high resolution genetic map of resistance to consolidate resistance markers associated with resistance. The study deduced that resistance has a single source (narrow genetic base) and may be controlled by a single gene or a cluster of resistance genes. Molecular markers from previous studies were further shown by Rabbi et al. 2014 to be inherited as an interdependent group and linked to CMD2 resistance across scaffolds, 05214 (S5214_780931, S5214_30911 and SSRY28); 06906 (SSR NS158 and SSR NS169), 04175 (SSR NS198); 07933 (SSRY106); with no matching scaffold RFLP RME-1. RME-1 has been associated by with CMD2 resistance markers NS169, NS158 and SSRY28 respectively (Dias do Carmo *et al*. 2015). Scaffolds 05214, 06906 and 07933 demonstrated bias to being the main regions of resistance (Allie et al. 2014; Rabbi et al. 2014; Wolfe et al. 2015; Dias do Carmo et al. 2015). More recently, long read-based de novo assembled genomes of CMD-susceptible and CMD2-resistant African cultivars localised CMD2 to a single scaffold (Scaffold 7) (Kuon et al., 2019). CMD2 has broad specificity against geminiviruses with broad sense heritability. Thus, revealing it is not significantly altered by the environment; however, unlike narrow sense heritability (additive gene action) it cannot be readily assumed that it will be persistent in subsequent generations. It is however, currently still the most reliable source of resistance as in addition to high heritability, it is dominant and has broad specificity and the identification of multiple markers in close proximity has further increased the selection and improvement of phenotypes with

tolerance (Dias do Carmo et al. 2015). Exploration of the CMD2 locus and surrounds identified resistance associated genes that encode for peroxidases (MeTME3_00015769 and MeTME3_00015798) and are involved in protein folding (MeTME3_00015870) and posttranscriptional gene silencing (PTGS)/ RNA silencing (SGS3, MeTME3_00015743) (Kuon et al., 2019). The association of SGS3 and peroxidases to the CMD2 locus was significant considering the role of reactive oxygen species in innate and effector triggered immune signalling and the roles of PTGS in defence against viruses.

1.6.2 HOST DEFENCE RESPONSES AGAINST GEMINIVIRUS INFECTION

1.6.2.1 RNA SILENCING

Post transcriptional gene silencing is a highly conserved systemic mechanism present in plants, animals and fungi and are mediated by Small interfering RNAs (siRNA) and microRNAs (miRNAs) (Mourrain et al., 2000). Small interfering RNAs (siRNAs) are derived from dsRNA synthesised by RDR6/RDR2/RDR1 and these are cleaved by RNASe III type enzyme DICER-like endonucleases (DCL) into 21–22 (by DCL1, DCL2 and DCL4) and 24–26 nt (by DCL3) lengths with the assistance of HYL1 and SERRATE (Bologna & Voinnet 2014; Pooggin 2013; Li et al. 2011; Hamilton et al. 2002). DNA virus derived siRNAs are identified by host cell RDR2 (Bologna & Voinnet, 2014) and RDR1 (Wang et al. 2012). Cleaved dsRNA fragments are incorporated into Argonaute forming an RNA-induced silencing complex (RISC). Virus proteins inhibit virus targeted host RNA silencing by interfering with the interaction of 21 to 24 nt interaction with the RISC complex (Soosaar 2005). Argonaute (AGO) proteins are contained in cytoplasmic processing bodies (P-bodies) (Kumakura et al., 2009) and are the main elicitor of RNA silencing (Bologna & Voinnet, 2014). Argonaute 1 (AGO1) is acts primarily against virus invasion and different AGO proteins elicit a second layer of defence (Wang *et al*. 2012). A single guide strand is retained for sequence specific targeting of messenger RNA (mRNA) which is cleaved by AGO1/4 for PTGS. The resulting siRNA lengths play a specific functional role in that the targeting of the silencing assembly and pathway by viruses is size selective and occurs by direct interaction with silencing components and sequestering of siRNAs (Vargason et al., 2003; Xiong et al., 2009). The longer 22–24 nt siRNA function is biased to systemic silencing and methylation of homologous DNA whilst the shorter 21–22 nt siRNA functions exclusively for sequence-specific mRNA degradation (PTGS) (Hamilton et al. 2002; Zhang et al. 2005) targeting virus transcripts (Wang et al. 2012).

MicroRNAs are derived from single strand hairpins and are 19–24 single nucleotide single strand non coding sequences (Rhoades et al., 2002; Vaucheret, 2006). Patanun et al. 2012 screened 169 cassava miRNAs and 21 nt strands were most abundant at 75% followed by 20 nt (14%), 22 nt (7%), 23 nt (2%), and 19 nt (<1%). They function as developmental

regulators targeting proteins with roles in stress responses, metabolism, structural integrity and biosynthesis acting as hetero-silencers and are derived from short stem loop precursors (Vaucheret, 2006; Rhoades et al., 2002; Patanun et al., 2012) that guide the RISC to target mRNA and partially bind for mRNA repression or sequence specific cleavage (Hutvágner 2002). MicroRNA biogenesis resembles that of siRNA, primarily distinguished by the absence of RDR. The initial DCL1 cleavage of pre-miRNA determines sequence and target specificity, and the maturation process of miRNAs are influenced by DCL1 cleavage. Double strand RNA-binding (DRB) proteins, hyponastic leaves 1 (HYL1) and serrate (SE) improve the efficacy and precision of DCL1 whilst tough (TGH) and modifier of SNC1 2 (MOS2) respectively stabilise the DCL1 complex and recruit pri-miRNAs (Bologna & Voinnet, 2014). Cyclophilin 40 (CyP40) interacts with molecular chaperone heat shock protein 90 (Hsp90) to promote AGO1 activity in miRNA activity (Earley & Poethig, 2011) and passenger-strand removal from AGO1 is by disassociation from HSP90 by conformational modification (Bologna & Voinnet, 2014).

RNA silencing mechanisms are mandatory to host anti geminiviral responses that confer resistance or tolerance (Kushwaha et al., 2015; Raja, 2010; Bizabani et al., 2021; Blevins et al., 2006). Variations in host specific responses occur in the production of siRNAs to combat RNA and DNA viruses (Blevins et al., 2006). In infection of *Arabidopsis*, geminivirus CaLCuV and DNA virus cauliflower mosaic virus (CaMV) the former was mainly affected by DCL4, RDR6, HEN1 and the latter by DCL4 respectively (Blevins et al., 2006). In the same study, RNA virus oilseed rape mosaic virus (ORMV), hindered HEN1 mediated mechanisms whilst CaMV impaired processing of endogenous RDR6-derived double-stranded RNA (Blevins et al., 2006). In SACMV infection, AGO1, DCL1 (Bizabani et al., 2021) and DCL3 (Allie et al., 2014) were associated with tolerance in cassava whilst in ToLCNDV infection RDR6, AGO1, DCL1 and SGS31 were attributed to resistance in pepper (Kushwaha et al., 2015). Geminivirus proteins V2, AL2 and L2 of TYLCV, TGMV and BCTV respectively target host RNA silencing mechanisms by interference of SGS3 binding (Wang et al., 2014), inactivation of adenosine kinase (Wang et al., 2003) and histone deacetylase to hinder host silencing (Wang et al., 2017).

1.6.2.2 INNATE AND EFFECTOR TRIGGERED IMMUNITY

The consensus is that plants possess a primary or basal pattern-triggered immunity against viral proteins and other foreign molecules upon detection of pathogen-associated molecular patterns (PAMPs). Recognition occurs at the host cell periphery by surface-localized pattern recognition receptors (PRRs) that activate PAMPs triggered immunity (PTI) (Monaghan & Zipfel, 2012; Bart et al., 2012; Jones & Dangl, 2006). The second line of defence involves more aggressive defence mechanisms associated with Effector-Triggered Immunity (ETI) inside the host cell, often resulting in a hypersensitive response

(HR) to localise the pathogen and restrict systemic movement (Lozano et al., 2011; Bart et al., 2012). ETI is triggered by recognition of pathogen molecules by resistance (R) gene proteins known as receptor or surveillance proteins (Gururani et al., 2012). Resistance gene proteins are classified among NBS-LRR and other immunity related genes (Choi et al., 2016; Gururani et al., 2012). The understanding of NBS-LRR and other immunity related gene (IRG) interactions during virus invasion provide clues to unravelling the molecular mechanisms of tolerance and susceptibility (Choi et al. 2016). There are nine distinct classes of immunity related genes (IRGs) based on the arrangement of the functional domains (Gururani et al., 2012). With regards to SACMV, the upregulation of several NBS domain-LRR genes in response to SACMV, and their potential influence on plant tolerance have been identified (Louis & Rey, 2015).

The HR triggered by ETI leads to activation of a prolonged memory-dependant Systemic Acquired Resistance (SAR) in distal tissues (Bart et al., 2012). The latter is directed to uninfected plant tissues (Kranthi et al., 2013) and provides a broad-spectrum, long duration resistance against pathogenic fungi, oomycetes, viruses, and bacteria (Lu et al., 2015; Fu & Dong, 2013). In both responses, early signalling activities are mediated by calcium, activation of mitogen activated protein kinase (MAPK) cascades and ROS (reactive oxygen species), for the release of oxidising molecules (ONOO-, O^{2}-and H_2O_2), mediation of cell wall strengthening and the subsequent induction of defence phytohormone pathways for salicylic acid (SA), ethylene and jasmonic acid (JA) (Stael et al., 2015; Kruijt & Kock, 2005). Light perception via phytochromes A and B plays a key role in mediating a light dependent HR and resistance signalling immune response (Chandra-Shekara et al., 2006) that results in an influx in ROS, resistance-associated hormones, disease resistance proteins of the NBS-LRR class and increased activity of HSPs (Sano et al., 2014). Oxidising molecules are modulated by the mitogen-activated protein kinase (MAPK) cascades MEK2-SIPK/NTF4 and MEK1-NTF6 (Bellin et al., 2013) and NO may be crucial in the coordination and regulation of calcium and other ROS signalling in plants (Trapet et al., 2015). Cytosolic calcium activates calcium-dependent protein kinase/NADPH oxidase, CPK5 which phosphorylates respiratory burst oxidase homologues D (RBOHD) via its N-terminal domain releasing superoxide (O^{2-}) in the apoplast and conversion to hydrogen peroxide (H_2O_2) spontaneously or catalytically by the action of superoxide dismutase (SOD) increasing cell permeability (Dubiella et al., 2013). Hydrogen peroxide (H_2O_2) in the HR response has also been studied as an activator of salicylic acid (SA) accumulation (Alvarez et al., 1998). The SA, JA and ethylene are defence related signalling hormones involved in perception and are induced and further regulated by an oxidative burst that contributes to the amplification of defence related biomolecules, WRKY and CC–NBS–LRR/TIR–NBS–LRR proteins in a local and systemic manner (Lu et al., 2015; Chanda et al., 2011; Truman et al., 2007; Ochsenbein et al., 2006; Lu et al., 2013; Bernoux et al.,

2011; Day et al., 2006; Belkhadir et al., 2004; Maier et al., 2011; Fu & Dong, 2013; Zhu & Lee, 2015). The relevance of NBS-LRR and posttranscriptional gene silencing machinery in geminiviral infection, was demonstrated in their abundance being inversely proportional to virus accumulation in studies of tomato leaf curl New Delhi virus (ToLCNDV) in resistant host *Capsicum annuum* and SACMV in tolerant cassava TME3 (Bizabani et al., 2021). Therefore, in susceptible hosts *Nicotiana benthamiana* and *Solanum lycopersicum* (Kushwaha et al., 2015), cassava T200 (Bizabani et al., 2021) NBS-LRR and gene silencing transcripts were quantitatively lower.

The induction of a hypersensitive response in geminivirus has been linked mainly to the Rep, NSP and V2 in overexpression studies however, other geminiviral proteins counteract the HR response (Guerrero et al., 2020). Tomato yellow leaf curl virus (TYLCV) suppressed plant cell death in an HSP90 inhibited background in tomato plants (Moshe et al., 2016). Therefore, in whole geminivirus genome infection of plants a HR response against geminivirus infection has not been identified (Guerrero et al., 2020). The upregulation and inhibition of HR inducing host factors has been demonstrated in geminivirus infection (Hu et al., 2019; Mei et al., 2020). HIR proteins facilitate HR mediated cell death and in geminiviruses, the C4 protein of tomato leaf curl Yunnan virus (TLCYnV) induces the upregulation of *hypersensitive induced reaction 1 (HIR1*) (Mei et al., 2020). However, C4 by direct interaction, inhibits HIR1 self-interaction and promotes LRR1-mediated degradation of HIR1 (Mei et al., 2020). Tomato yellow leaf curl China virus (TYLCCNV) upregulates the MAPK pathways and its βC1 satellite interferes with the MAPK pathways by inhibiting the kinase activity of MKK2 and MPK4 (Hu et al., 2019).

1.6.3 RESISTANCE BREEDING EFFORTS IN CASSAVA

The general efforts made to approach the problem of cassava geminiviruses have been to work with communities to create phytosanitary programmes and early warning networks. Additionally, increasing research capacity and feedback in efforts to use simple efficient diagnostic tools, introduce novel molecular techniques for characterisation and resistance breeding and essentially the wide spread provision of quality vegetative propagules (Legg et al., 2015). Conventional breeding from seed to improve both the disease and pest resistance and agronomic properties (i.e. efficiency, nutritional elements, cyanogenic glucosides, root quality and yield) has been successful, with field trials approximating six to seven years (Jennings & Iglesias, 2001). Transcriptional control by RNA interference (RNAi) mainly inhibits the viral mRNA from being encoded and has been ascertained as an effective tool in functional genomics studies in transgenic crops resistant to CBSV (Odipio et al. 2014), SLCMV (Ntui et al., 2015) and ACMV (Zhang et al., 2005) for heritable crop improvement. Introgression of RNAi into farmer preferred varieties as opposed to conventional breeding methods ensures that favourable traits such as maturity, root size,

flavour, starch content and hardiness are maintained (Patil et al., 2015). It also reduces the risk of incorporating other disease susceptibilities previously of no economic importance. The transgenic expression of ribosome-inactivating proteins (RIPs) mediated by an AC2 promoter conferred disease recovery in plants infected with ACMV and its isolates (Hong et al., 1997). The Rep protein ceases its own transcription to enable viral DNA accumulation (Bisaro, 1996; Hanley-Bowdoin et al., 2013). Therefore, the uncontrolled or constitutive expression of Rep proteins in host cells hinders viral proliferation (Chatterji et al., 2001; Brunetti et al., 1997). The expression of virus specific truncated Rep protein encoding the full N-terminal DNA and oligomerisation domains effectively hindered geminivirus proliferation in tomato leaf curl New Delhi virus (ToLCND) (Chatterji et al., 2001) and tomato yellow leaf curl virus (TYLCV) (Brunetti et al., 1997) infected plants and further delayed the onset of symptoms in ACMV and other geminiviruses (Chatterji et al., 2001).

Interestingly, Beyene et al. 2016, demonstrated the loss of heritable CMD2 resistance in somatic embryos of transgenic RNAi cassava TME 204 and in natural resistance cassava TME 7 and TME 3, this has however not been shown with regards to CMD1 and CMD3 resistance and so pooling of resistance is a key concept to consider in CMD resistance breeding/transformation. The use of conventional somatic embryogenesis has however been shown to yield virus free plants from infected mother plants of varieties such as *Nguwo* (Nkaa et al., 2013). A study by Bi *et al*. 2010, confirmed the need for Asian cassava to be introgressed with African germplasm CMD2 resistance as a primary solution to CMD. Okogbenin et al. 2007, incorporated CMD2 resistance by use of molecular marker assisted breeding into elite Latin American (ELA) cassava germplasm that hold the potential to improve the genetic diversity in Africa as ELA germplasm encompass enhanced nutrient, yield, and pest resistance. The introduction of Latin American clones into Africa introgressed with CMD resistance gave yield in the range of 20 to 46 tons per hectare at 8 months, thus exceeding African tolerant variety TMS 30572 which averages yields of 24 tons per hectare and is used as a yield standard in field trials (Okogbenin et al., 2007).

1.7 GENE EDITING

A systems biology approach that follows up results from transcriptome data and/or virus-host interaction studies to determine the influence of host gene knockout or knockdown on virus persistence provides a clear understanding of viral protein mechanisms and confirms genes associated with host susceptibility or resistance (Miles et al., 2016; Czosnek et al., 2013). Thus, gene editing is a key aspect in reverse genetics expression studies to assess specific gene function. Functional genetic screening provides information on host cell factors essential for viral persistence and suppression and protein cellular pathways (DeJesus et al., 2016). The utilisation of virus induced gene silencing (VIGS) in elucidating host genes relevant to viral infection has been prominent in virus studies (Buchmann et al.,

2009; Mwaba, 2011; Mwaba & Rey, 2017; Czosnek et al., 2013; Liu et al., 2002). VIGS has been effective in highlighting the significance of defence genes in *Nicotiana benthamiana;* NbSGT1 and NbSKP1 in tobacco mosaic virus (TMV) resistance (Liu et al., 2002). With regards to understanding geminivirus-host relationships most work in host gene editing has been carried out by use of VIGS (Mwaba & Rey, 2017; Czosnek et al., 2013; Beyene et al., 2017). In SACMV infection, nitric oxide associated protein 1 (NOA1) was associated with symptom severity (Mwaba & Rey, 2017). In TYLCSV infection, eighteen host genes were identified that enable viral accumulation and were associated with post translational modifications (Czosnek et al., 2013). Furthermore, five genes were identified to be significant to conferring resistance to TYLCSV (Czosnek et al., 2013). In EACMV infection of *Arabidopsis*, the silencing of SPINDLY (SPY) was used to rapidly discriminate between EACMV resistant and susceptible varieties (Beyene et al., 2017). VIGS was also used to identify the link between ACMV Rep and recruitment of host factors such as retinoblastoma-related 1 (RBR1) in the reprogramming of the host cell cycle in a manner conducive for viral replication (Bruce et al., 2011).

Clustered regularly interspaced short palindromic repeats/Cas9 (CRISPR/Cas9) has gained precedence after first being introduced as a gene editing tool (Jinek et al., 2012) that functions on cell double strand repair mechanisms that result in homologous recombination (HR), nonhomologous end joining (NHEJ), alternative end joining (alt-EJ) and single-strand annealing (SSA) (Ceccaldi et al., 2016). The CRISPR/Cas9 system has been significant towards targeting the geminivirus genome **(Figure 1.5)** (Ali et al., 2016; Zaidi et al., 2016).

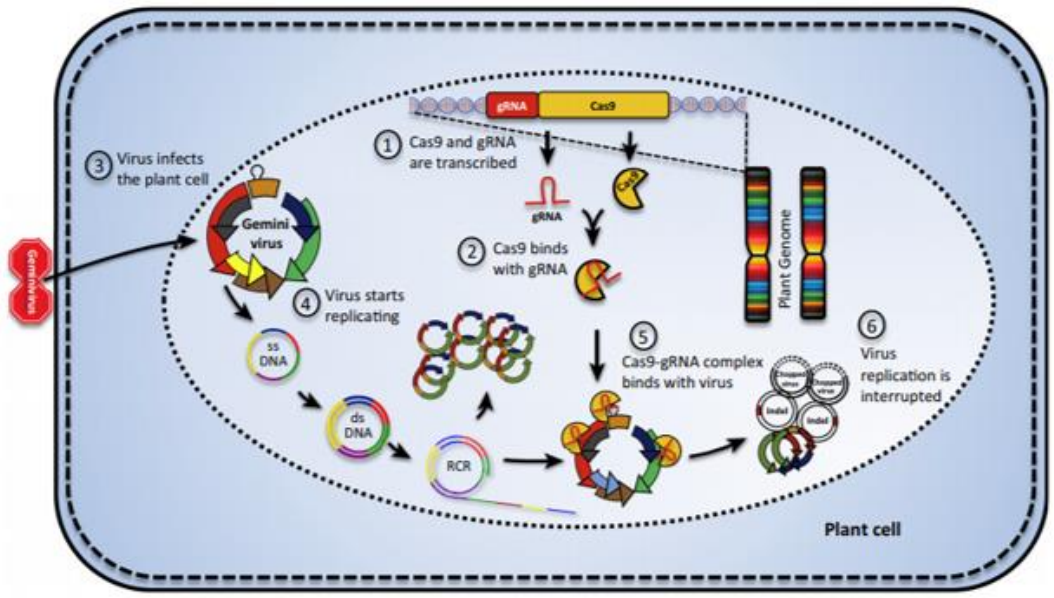


Figure 1.5 Engineering Plant Resistance against Viral Pathogens via the CRISPR/Cas9 System. A stepwise depiction of CRISPR/Cas9-mediated virus interference in the plant cell nucleus. (1)

Components of the CRISPR/Cas9 machinery, gRNA and Cas9, are expressed from the plant genome. (2) Assembly of the gRNA–Cas9 complex. (3) Upon virus challenge, the geminivirus infects the plant cell. (4) The single-stranded DNA (ssDNA) geminivirus genome replicates through the double-stranded DNA (dsDNA) replicative form followed by rolling circle replication (RCR) to produce multiple viral copies. (5) The gRNA–Cas9 complex targets the viral dsDNA at complementary target sites and (6) cleaves the viral genome via double-strand break (DSB) formation, which can be repaired by non-homologous end joining (NHEJ) repair. Alternatively, the formation of DSBs can lead to degradation of the viral genome. Source: Zaidi et al., 2016

In a study by Ali and colleagues, the targeting of noncoding, intergenic regions of geminiviruses, cotton leaf curl Kokhran virus (CLCuKov), merremia mosaic virus (MeMV) and tomato yellow leaf curl Sardinia virus (TYLCSV) was effective in hinderance of viral replication and systemic movement whilst in targeting of coding regions replication and movement were retained (Ali et al., 2016). Targeting of geminiviruses has been shown to be specific (Ali et al., 2015). A confirmation of specificity provides plausible certainty that plant genome editing does not occur. Ali and colleagues showed that, viral targeting was specific to TYLCV and did not target beet curly top virus (BCTV) (Ali et al., 2015). Geminiviruses have a small genome with a high frequency of point mutations (Richter et al., 2016) and therefore may be capable of evading CRISPR/Cas9 targeting (Mehta et al., 2019) most especially within gene editing targeted to coding regions (Ali et al., 2016; Mehta et al., 2018). In the editing of ACMV 33 % to 48 % of targeted and edited viral genomes within the TrAP and Ren coding regions evolved resistance to CRISPR/Cas9 cleavage (Mehta et al., 2019). A single CRISPR/Cas9 array can be encoded with multiple targeting sequences (multiplexed) to enable the editing of multiple sites in the genome (Cong et al., 2013) and this concept has been adopted in plant studies. Multiplexed targeting of chilli leaf curl virus (ChiLCV) and cotton leaf curl disease (CLCuD) in the non-coding and coding regions of the virus genome yielded significant reductions in viral genome accumulation and therefore multiplexing a solution towards negating viral genome evolutionary mechanisms against CRISPR/Cas9 editing (Roy et al., 2019; Mubarik et al., 2021). Multiplexing with CRISPR/Cas9 further provides antivirus mechanisms towards mixed infections (Mubarik et al., 2021).

The use of CRISPR/Cas9 in studying of host gene function towards virus accumulation has mainly been studied in RNA viruses and identification of geminivirus specific work was limited. CRISPR/Cas targeting of the *eIF4E* gene in RNA viruses of tomato bushy stunt virus (TBSV), tobamovirus turnip vein clearing (TVCV), turnip mosaic virus (TuMV), rice yellow mottle virus (RYMV) and rice dwarf virus (RDV) have conferred resistance to multiple plant hosts (Cao et al., 2020; Pyott et al., 2016; Lellis et al., 2002; Robaglia & Caranta, 2006; Wang & Krishnaswamy, 2012). With regards to cassava, the CRISPR/Cas9 targeted editing of *eIF4E* isoforms rendered host tolerance confirmed by a reduction in

symptom severity against RNA viruses, cassava brown streak virus (CBSV) and Ugandan cassava brown streak virus (UCBSV) (Gomez et al., 2019). With regards to geminiviruses and cassava, the targeted knockdown of the E3 ubiquitin ligase gene in cassava revealed an increase in SACMV accumulation, indicative of its role in suppression of viral infection in protoplasts (Chatukuta & Rey, 2020). Work in cassava further provided a rapid screening technique for evaluation of host genes pertinent to geminiviral accumulation (Chatukuta & Rey, 2020).

1.8 NETWORKING TO UNDERSTAND PLANT-VIRUS MECHANISMS

Network biology is a combination of systems biology, graph theory and computational and statistical analyses in which the topology of the graphs represent molecular interaction networks (Winterbach et al. 2013). Networks are illustrated as interaction graphs, where nodes (proteins or genes) and edges or arcs (directed or undirected) represent pairwise interactions. Graph theory approaches have been applied to describe the topological properties of the network: distribution of node degree (number of incoming and outgoing edges per node), network diameter (average of the shortest distance between pairs of nodes), and clustering coefficient (proportion of the potential edges between the neighbours of a node that are effectively observed in the graph). These analyses have led to the observation of some apparently recurrent properties of biological networks: power-law degree distribution, small world, high clustering coefficients, and modularity (Sylvain and van Helden 2006; Zhu et al. 2007).

Interaction networks are based on the assumptions that evolutionarily-conserved proteins tend to have conserved interaction and this information is inferred in prediction of networks (Hao et al., 2016). Interaction orthologue (Interolog) mapping is a based on this assumption and a study in *Arabidopsis* identified 19979 predicted interactions for 3617 *Arabidopsis* proteins (Geisler-Lee et al. 2007). Proteins that exhibit a high number of interactors tend to evolve slowly because a greater number of sites on the protein must be maintained for functionality with diverse proteins i.e. rate of evolution is inversely proportional to number of interactors (Chen & Xu 2003). Genetic algorithms rely on predictive pairwise assumptions based on the patterns of natural evolution, (conserved protein residue profiles-primary and secondary structures) are essential in predicting locality, function and interacting partners (Hao et al., 2016). Iterative methods improve existing algorithms and consequently increase confidence values and give more accurate network relationships interactions (Shih & Parthasarathy, 2012). Docking networks are employed in predictions for non-conserved regions/domains (Hao et al. 2016; Liu et al. 2008).

Leal et al. 2013 constructed a network by utilising a rigorous mathematical data mining method, kernel canonical correlation analysis (KCCA), which is ideal in predicting functional

relationships and allows the integration of existing/ prior knowledge data for biological inference. New immunity related genes (IRGs) in cassava and *Arabidopsis* and their functional relationships between new genes and known genes were identified using this method. The interactions between and *Arabidopsis* were constructed into a network exhibiting potyvirus proteins interacting among themselves, directly with plant proteins, and further extended the interactome to secondary host protein interactors (Elena & Rodrigo, 2012). Interaction networks demonstrate the functional relationships between proteins/genes and therefore have the capacity to reveal new and unexplored mechanisms. Geminivirus network analysis has provided a comprehensive approach towards studying virus-host interactions and revealed that geminiviral proteins have a tendency to target highly connected host proteins (Wang et al., 2017; Gouveia-mageste et al., 2020). The identification of host hub proteins provides clues to potential targets for engineering resistance to viral infection (Wang et al., 2017). The network topology of geminivirus MP and NSP host proteins identified key hub proteins involved in auxin-mediated responses, a finding novel to MP and NSP targeted interactions (Gouveia-mageste et al., 2021).

There are five main molecular interaction network models (Winterbach et al., 2013). These are as follows:

1) Association networks. These model relationships between molecules whether it be by binding, co-expression and structural similarities (Winterbach et al., 2013). They depict direct and indirect interactions that are inferred from one gene directly interacting with two genes and from this the two genes are inferred to interact indirectly (Li et al., 2015). Protein structural similarity networks provide information on amino acid, peptide, 3-dimensional and functional similarities.
2) Functional networks. These model gene and protein functional relations (Winterbach et al., 2013). The link length is directly proportional to function and the function being related to surrounding protein/ gene neighbours (Sharan et al., 2007). These are represented by genetic interaction networks used to evaluate the expression of the phenotype as a result of epistasis, and are represented as links between nodes (genes) and demonstrate the direct relationship between genes and the similarity to genes that express a similar pattern of interactions (Jacobs & Segrè, 2012). In this sense, epistatic networks can be clustered into within and between module relationships (Segrè et al., 2004).
3) Protein-protein Interaction Networks (PPI Networks) are undirected networks that model protein binding. In being undirected, the degree of node is a one dimensional distribution considering incoming and outgoing edges as a single set of connections (Winterbach et al., 2013). Online databases are available with STRING being the most concise (Franceschini et al., 2013). Datasets providing experimental data from GFP

fusion experiments and with AmiGO annotation, Swiss-Prot annotation, and localization based on TAIR gene descriptions further allow subcellular localisation of predicted interaction networks (Geisler-Lee et al. 2007). Visualisation of protein networks is facilitated by programs such as CYTOSCAPE (Geisler-Lee et al. 2007).

4) Transcription-regulatory Networks (TR Networks) are bipartite networks with one set of nodes representing genes and the other representing transcription factors (TFs). They are directional and indicate causal relationships between genes and transcription factors (Li et al., 2015).
5) Metabolic Networks are bipartite networks that model the relationships between the biochemical reactions that occur in cells and the substrates (amino acids, sugars and lipids) and energy involved in the reactions therefore presenting protein and metabolite information that incorporates transcription factor binding and phosphorylation (Zhu et al., 2007; Winterbach et al., 2013).

Data collection and network analysis goes through two main steps:

1) Normalisation

The data obtained is sourced from literature and from different databases and experiments must be pre-processed to remove experimental bias and allows data from different units to be comparable. Algorithms used are Discretization Method-I/II (Chandrasekhar et al., 2011) and may be specific to the experimental method used i.e. in microarrays; MAS5 (Li et al., 2015), Lowess normalization methods (Kim et al. 2006) and Discretization Method-III (Chandrasekhar et al., 2011) and all normalised data further undergoes cross-platform normalisation (Li et al., 2015).

2) Statistical Reconstruction

Methods for reconstruction of interaction networks are correlation-based methods (for co-expression networks), Supervised learning (for gene association or regulatory networks), Probabilistic Graphical Models (combine probability theory and graph theory), Meta-prediction (for network topology) and the resultant network may be further evaluated by existing high confidence networks, cross validation and functional coherent module assessment (Li et al. 2015).

1.9 ASSIGNMENT OF GENE FUNCTION FOR PROTEIN INTERACTION STUDIES

In the establishment of PPI networks, the assignment of gene function plays a critical role in assigning value to protein expression and interaction studies for the construction of reliable and referable biological network models. MapMan (http://gabi.rzpd.de/projects/MapMan/) and Gene Ontology (GO)

(http://www.geneontology.org/) (Klie & Nikoloski, 2012) are gene function tools commonly used. The former is specifically designed for plant specific pathways whilst the latter was initially fashioned for microbial systems and pathways (Klie & Nikoloski, 2012) and has been extended to plant studies.

MapMan has three modules, a Scavenger module that organises the measured parameters into hierarchical functional categories, an ImageAnnotator module that visualizes the data on a gene-by-gene basis on schematic diagrams ('maps') of biological processes and a PageMan module uses the same ontologies to statistically evaluate responses at the pathway or processes level (Usadel et al., 2009; Thimm et al., 2004). MapMan may be used for transcripts, proteins, metabolites and enzymes and was developed for use with *Arabidopsis*, but has already been extended for use with several other species (Usadel et al., 2009). It has been extended to cross species studies between maize and *Arabidopsis* (Usadel et al., 2009), the functional mapping and characterisation of biotic stress responses in solanaceous species has been especially relevant (Rotter et al., 2007), grapevine auto- and hetero-graft studies (Cookson et al., 2014), tobacco (Ling et al., 2013), categorizing chromoplast proteins in various fruits and vegetables (Wang et al., 2013) and in cassava; shade avoidance responses (Ding et al., 2016), polyethylene glycol (PEG)-Induced dehydration stress (Fu et al., 2016) and heat-girdling and it's inhibiting effect on starch biosynthesis (Zhang et al., 2015).

Gene Ontology is based on four functional categories: Molecular Function (MF) describes activities, such as catalytic or binding activities, at the molecular level; Biological Process (BP) describes biological goals accomplished by one or more ordered assemblies of molecular functions; Cellular Component (CC) describes locations, at the levels of subcellular structures and macromolecular complexes (Klie & Nikoloski, 2012); and Sequence Ontology (SO) a more recent ontology which permits the classification and standard representation of sequence features (such as "polycistronic" and "maternally imprinted") (Gene Ontology Consortium (GOC), 2004). The GOC comprises specific research groups for compilation of GO biological databases such as TAIR (Lamesch et al., 2012).

There are two main methods of predicting protein function, namely direct and indirect methods. The direct methods are based on predictions with respect to a proteins' neighbourhood of proteins by 'guilt by association' and 'kin liability' (Pinkert et al., 2010). The former infers that if gene A is involved in function X and we obtain evidence that gene B functionally associates with A, then B is also involved in X (Koonin 2003). This concept extends pathways by assigning protein function to unannotated proteins within protein complexes and pathways (Geisler-Lee et al. 2007). The algorithms commonly used are neighbourhood based, graph theoretic, probabilistic and integrating multiple data, the first

three use interaction methods based on confidence levels (Sharan et al., 2007). 'Kin liability' is module assisted and clusters the network into groups and assigns function to those groups based on the majority of known annotations within the identified groups/ modules (Pinkert et al., 2010; Sharan et al., 2007). The detection of modules within a network is by unsupervised clustering algorithms (Chandrasekhar et al., 2011), expanding seed complex, integrating gene expression data and integrating diverse genomic data methods that vary in presentation of a graphical interface, overlapping modules and the use of interaction weights based on confidence levels (Sharan et al., 2007).

The binary Markov Random Fields-Deng (MRF-Deng) is a commonly used direct method of annotation based on the probability that's protein functionality depends on the number of its direct neighbours in the network that perform the function and the number of those that do not. This method introduces uncertainty due to unannotated proteins and is confounded with increased numbers of unannotated proteins (Kourmpetis et al., 2010). The unannotated proteins are not regarded in the initial logistic regression; however Gibbs sampling is employed for functional inference of the proteins with unknown function ("unannotated proteins") and therefore the unannotated proteins are considered in a model that initially disregards unannotated proteins (Kourmpetis et al., 2010). Kourmpetis et al. 2010 improved on the MRF and developed a probabilistic approach for protein function prediction i.e., a Bayesian approach to an existing Markov Random Field method (BMRF method) by performing simultaneous estimation of the model parameters and prediction of protein functions. Additionally, the use of an adaptive Markov Chain Monte Carlo algorithm led to more accurate parameter estimates and improved prediction performance. This was a direct method that which when applied to protein interaction data from S. *cerevisiae* provided novel predictions, using 340 Gene Ontology terms, for 1170 unannotated proteins (Kourmpetis et al., 2010).

1.10 RATIONALE

Cassava is a New Age crop encompassing many benefits (Frow & Yearley, 2010) as a low input crop that is, environmentally sustainable and well adapted to the erratic Sub-Saharan climate. Nutritionally, in Sub-Saharan Africa the production cost of the main commercial staple crop, maize, has escalated over the past three years. Cassava is harvested for its roots and leaves and would serve as an affordable high carbohydrate alternative with substantial vitamin and mineral content. Industrially, it offers a local reservoir source for bioenergy production and a renewable source for petrochemicals and industrial materials. In Sub-Saharan Africa, cassava has been acknowledged as not only a source of nutrition but essential to industries and therefore a crop with great potential for economic upliftment of poor communities.

However, the obtaining high yields is hindered by the prevalence of cassava mosaic disease (CMD) caused by cassava mosaic geminiviruses (CMGs). This limits economic returns to less than 25 % and world yield losses currently exceed US$1.9 billion (Alene et al. 2013). Cassava resistance mechanisms and host responses against CMGs have been studied (Allie et al. 2014; Soosaar, 2005; Pierce & Rey 2013; Wang et al. 2014). Genetic engineering has also been a useful tool in conferring plant resistance (Ntui et al. 2015; Vanderschuren et al. 2012). Plant genomics provides a platform to study geminivirus interactions and identify possible viral protein interference with cassava metabolic processes (Soosaar, 2005; Maghuly et al. 2014; Pierce & Rey 2013; Pooggin, 2013; Soitamo et al. 2012). Transcriptomic studies have generated a plethora of biologically significant data on tolerant and susceptible host-SACMV responses (Allie et al., 2014; Pierce & Rey, 2013).The virus elicits a cascade of well-coordinated manipulation of the susceptible cassava host altering the cell to favour its own persistence by the ability to effectively suppress the translation of cassava RNA silencing (Allie et al., 2014; Pierce & Rey, 2013). Tolerant cassava varieties effectively counteract virus proteins by specific resistant mechanisms targeting viral proteins (Allie et al., 2014; Pierce & Rey, 2013; Louis & Rey, 2015).

In vivo Y2H assays have been employed in identifying binary interactions of cassava-geminiviral MP and NSP protein (Frischmuth et al., 2007; Carvalho & Lazarowitz, 2004) and this shows their relevance to be used to clarify cassava-geminivirus interactions and discover novel interactions. Furthermore, viral proteins tend to also involve weak and transient interactions (Maio et al., 2020; Pascal et al., 1994) that are detectable in Y2H assays (Meyer & Schomburg, 2008). Also, due to the complexity of protein interactions and the ability of viruses to perturb multiple cellular processes, it is important to interactions by *in vitro* Co-IP assays. Co-immunoprecipitation assays present a simple but robust technique to reveal the complex interactome networks of viral proteins (DeBlasio et al., 2017; Hoyle, 2018). Identified interactors may further be elucidated for conferring resistance or susceptibility by gene editing techniques such as CRISPR/Cas9 (Chatukuta & Rey, 2020). The current information on cassava-CMG interactions is limited and has still not resulted in effective crop improvement.

In this research cassava-MP protein interactions were studied by Y2H and Co-IP assays to develop an understanding of South Africa cassava mosaic virus (SACMV) MP mechanisms. Three significant interactors identified from interaction assays were knocked down to observe influence on SACMV replication.

This data provided a valuable resource for novel proteins involved in the cassava geminivirus infection response and their role in potentiation of either resistance or susceptibility for an informed approach towards combating geminivirus infection.

1.11 OBJECTIVES AND AIMS

1.11.1 MAIN OBJECTIVE

The main objective of this research was as follows:

To study the SACMV movement protein (MP) by use of Y2H and Co-IP assays and CRISPR/Cas9 mediated gene knockdown to identify the specific pathways that it interferes with and thereby construct an elaborate model of cassava-MP protein interactions. Subsequently identify possible molecular processes that can be targeted to develop strategic disease resistance to minimise yield losses.

1.11.2 SPECIFIC AIMS

There were three specific aims to this research. These were as follows:

1. To identify novel interactions of the CMD-tolerant cassava TME3 landrace host proteins with SACMV MP at 32 dpi using a high-throughput robust Y2H screening and postulate on their relevance to the functional mechanisms of SACMV MP during SACMV infection of cassava TME3.
2. To co-immunoprecipitate native and expressed SACMV MP with CMD-tolerant cassava TME3 landrace host proteins and therefore identify SACMV MP affiliated host proteins by mass spectrometry and postulate on their relevance to the functional mechanisms of SACMV MP during SACMV infection of cassava TME3.
3. To evaluate the impact of CRISPR mediated gene knockdown of CMD-tolerant cassava TME3 host genes previously identified to interact with SACMV MP by Y2H screening on SACMV replication in cassava TME3 protoplasts.

Chapter 2

THE IDENTIFICATION OF DIRECT INTERACTIONS BETWEEN CASSAVA TME3 HOST PROTEINS AND SACMV MP PROTEINS BY Y2H ASSAYS

2.1 INTRODUCTION

The Y2H system has over the past 30 years become established as an effective protein-protein interaction system. It is based on the principle of the *Saccharomyces cerevisiae* GAL4 protein which comprises a DNA-binding domain that activates transcription upon physical interaction with an activation domain containing acidic sequences necessary for interaction to occur. This principle was first made known by the activation of the GAL4 protein after interaction with GAL80, an inhibitor of the yeast transcriptional activator bound to an activation domain (Ma & Ptashne, 1988). Fields and Song later established the Y2H system as a genetic system to detect protein-protein interactions by binding of two yeast proteins designated as bait and prey to the N-terminal GAL4 DNA binding domain (BD) and the C-terminal GAL4 activating domain (AD), respectively (Fields & Song, 1989). Positive interactions result in the expression of reporter genes (Fields & Song, 1989). The reporter gene often used for convenient visibility is the *lacZ* gene which encodes for β-galactosidase enzyme that when bacteria grown on media containing X-gal results in blue colonies for positive selection (Serebriiskii & Golemis, 2000; Fields & Song, 1989). Colonies of low blue colour intensity are indicative of weak interactions and may be explored or excluded as false positives (Mehla et al., 2015; Luque et al., 2002). The *Escherichia coli* derived LexA ($LexA_{1-202}$ amino acid fragment) Y2H system, provides an alternative to the BD of the GAL4 Y2H system (Golemis & Brent, 1992). Two systems were designed by Fashena and colleagues in which the *LexA* gene is either upstream of *HIS3* gene or both *HIS3* and *lacZ* genes (Fashena et al., 2000). **Figure 2.1** illustrates the Y2H principle in which interaction leads to transcriptional activation of reporter genes/auxotrophic markers for selection and identification of positive clones.

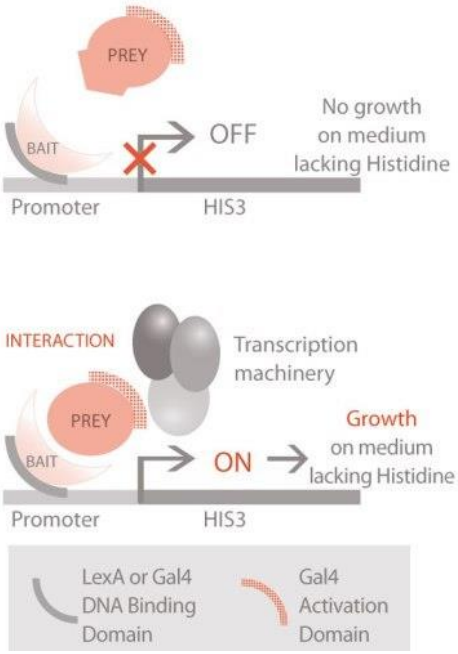


Figure 2.1 A schematic representation of the Y2H system. The interaction of two proteins reconstitutes an active transcription factor and enables yeast growth on selective media. BAIT, protein of interest. PREY, protein partner of the bait. Source, Hybrigenics services.

Selection of positive clones may be enhanced by complementation with one (Artigas-Jerónimo et al., 2020) or more (Fields & Song, 1989; Durfee et al., 1993; He et al., 2019; Alcântara et al., 2019; Li et al., 2018) auxotrophic gene markers such as *LEU2*, *HIS3*, *ADE2*, *URA3, TRP* and *LYS2* that allow for growth on amino acid deficient media, in order to increase the reliability of identified clones as positive. Binary interactions systems such as the Y2H system are prone to false positives (Brückner et al., 2009; Braun et al., 2009). Auto-activators that lead to reporter gene transcription without a physical protein interaction can account for up to 95 % of identified results (Trigg et al., 2017; Shivhare et al., 2021). To increase stringency in selection, a Y2H system, in which there is a reciprocal transfer of proteins fused to the AD and BD domains, enables the confirmatory identification of positive clones (Román et al., 2014). Carrying out Y2H assays in three biological replicates and only selecting those interactions observed in triplicate is another strategy to increase confidence in observed interactions (Alcântara et al., 2019). Stellberger and colleagues further designed a more accurate system that increased coverage and reduced false positives by use of vector constructs that made both C-terminal fusion proteins of DNA-binding and activation domains. These vectors when combined with existing vectors for N-terminal fusions allowed four bait-prey combinations (NN, CC, NC, and CN) (Stellberger et al., 2010). Most recently, to reduce the carry-over of false positives arising from auto-activation, an elimination strategy that integrated *URA3* into the BD domain was designed

to inhibit autoactivated clones from growing in media supplemented with 5-Fluoroorotic acid (5-FOA) (Shivhare et al., 2021).

Yeast two-hybrid systems were initially used to study a single protein-protein interaction pair at a time and have since evolved into identifying interactions in a high throughput manner resulting in complex interaction maps for various networks, for example, *Helicobacter pylori* (Rain et al., 2001), *Saccharomyces cerevisiae* (Cagney & Uetz, 2001) and varicella-Zoster Virus (VZV) (Stellberger et al., 2010). In the case of high throughput Y2H assays that look at identifying bait interactors from a wide array of prey proteins, normalised and non-normalised cDNA libraries can both be used to give a more representative transcriptome (Henry et al., 2019). In the study by Henry and colleagues, the normalised library improved access to transcripts of low abundance and increased the overall representation of transcripts whilst the latter non-normalised library, generated longer transcripts and more transcript isoforms (Henry et al., 2019). High throughput assays identify more interactions, and it has thus become necessary to improve on the qualitative selection of positive clones based on the *lacZ* gene β-galactosidase assay in a quantitative manner. Quantitative pellet X-gal (PXG) (Möckli & Auerbach, 2004) and microplate assay colorimetric reporters (Serebriiskii et al., 2000) have been developed to both improve clone selection in high throughput assays and provide an additional tier to rank interactions from high to low colour intensity as a measure of interaction strength. In order to complement the emergence of high throughput Y2H data within research labs, and provide opportunity for re-analysis of data, statistical pipelines such as DEEPN (Pashkova et al., 2016), 2HybridTools (Cauchy et al., 2019), Y2H Scores (Velásquez-Zapata et al., 2020; Velásquez-Zapata et al., 2021) and other computational prediction methods (Pan et al., 2021) have been designed for retrieval of high confidence interactors. Access to a comprehensive Y2H library for biotic and abiotic stress studies, for *Arabidopsis thaliana* challenged with multiple bacterial pathogens and chemicals associated with environmental stress responses have also been made accessible (Matiolli & Melotto, 2018).

The understanding of disease mechanisms in plant-pathogen interactions has been better elucidated by Y2H protein-protein interaction studies. In fungi pathogen studies the complexity of the interaction among 126 *Ustilago maydis* effectors (Alcântara et al., 2019), and multiple host interactors of effector proteins of *Bremia lactucae* (Pelgrom et al., 2020), have been elucidated by Y2H systems. In bacterial pathogen studies, defence factors against bacterial invasion for *Pseudomonas* have been identified (Oblessuc et al., 2020). There exist several studies that have utilised the Y2H system to investigate plant geminivirus interactions. The Rep protein of tomato golden mosaic virus (TGMV) and tomato yellow leaf curl Sardinia virus (TYLCSV) was found to interact with the host cell sumoylation enzyme, NbSCE1 (Castillo et al., 2004). Both Ren and Rep proteins of

TYLCSV interact with proliferating cell nuclear antigen (PCNA) involved in DNA replication (Castillo et al., 2003). Proof of homo-oligomerisation of geminivirus proteins, MP of abutilon mosaic virus (AbMV) (Frischmuth et al., 2004) and Rep of wheat dwarf virus (WDV) (Missich et al., 2000), has been demonstrated. The AL2 (TrAP) of TGMV interacts with calmodulin-like protein 39, a regulator of RNA silencing (Yong Chung et al., 2014) and DNA-binding protein PEAPOD2 (PPD2) (Lacatus & Sunter, 2009). The latter further interacts with the CP promoter of TGMV and cabbage leaf curl virus (CabLCV) and may be indicative of AL2 involvement in CP activation (Lacatus & Sunter, 2009). Binary interactions within yeast have also allowed a better understanding of viral strategies, most notably in interaction with kinases (Fontes et al., 2004; Hao et al., 2003; Jia et al., 2016; Wang et al., 2003; Shen et al., 2011). TrAP of TGMV and BCTV have been shown to interact with sucrose nonfermenting1 (SNF1) kinase (Hao et al., 2003) and adenosine kinase (ADK) (Wang et al., 2003). The βC1 protein of tomato yellow leaf curl China β-satellite (TYLCCNB) interacts with the sucrose-nonfermenting1-related kinase (SnRK1) (Shen et al., 2011). The cotton leaf curl Multan virus (CLCuMuV) betasatellite (CLCuMuB) βC1 interacts with S-phase kinase-associated protein (SKP1/ASK1) involved in ubiquitination, and thus enhances viral symptom accumulation and progression (Jia et al., 2016). The NSP of TGMV and tomato crinkle leaf yellows virus (TCrLYV) interacts with the NSP-interacting kinase (NIK) potentially negating the viral defence activity of NIK (Fontes et al., 2004). Geminivirus protein interactions with host proteins have also provided information on anti-viral mechanisms in the plant. The Rep (C1) of tomato leaf curl Yunnan virus (TLCYnV) interacts with the core autophagy-related protein (ATG8h), an interaction that restricts viral progression (F. Li et al., 2020). The N-terminal region of viral protein, Ren (C3) of TYLCV, beet curly top virus (BCTV) and TGMV interacts with DNA polymerase α subunit 2 (POLA2) as an essential exploitative mechanism for viral DNA replication (Wu et al., 2021). The C4 of beet severe curly top virus (BSCTV) interacts with receptor kinase CLAVATA 1 (CLV1), an important developmental receptor, that when hindered by interaction results in developmental abnormalities noted in geminivirus infection (Li et al., 2018). The wheat dwarf geminivirus (WDV) RepA protein interacts with Geminivirus Rep A-binding (GRAB) proteins of the NAC domain family of proteins involved in growth, organ development and senescence (Xie et al., 1999).

Yeast two-hybrid systems have gained popularity due to their reliability in that they have been able to verify interactions that have been affirmed by prior studies and are therefore repeatable (Fields & Song, 1989; Möckli & Auerbach, 2004). They have also shown consistency in results obtained from similar virus proteins screened with different host prey proteins. For example, this is noted in geminivirus studies that have been able to confirm interactions in two or more geminivirus species with prey proteins from different plant hosts (Li et al., 2020; Wang et al., 2003). Results obtained from Y2H assays also allow for the

identification of novel interacting proteins (Xie et al., 1999). Albeit many false positives results present, it is still possible to put in control measures that identify true positive interactions (Stellberger et al., 2010; Brückner et al., 2009; Shivhare et al., 2021). The proximity of interaction of proteins may also play a significant role in encouraging the occurrence of an interaction that within the original cell would never occur and bears no biological relevance. For example, within the original host, consider hypothetical proteins A and B that are separated by sub cellular compartmentalisation that create a physical barrier that isolates protein A from protein B. Both protein A and B function independently of each other and there is no biological need for the proteins to interact within the cell. The yeast cellular environment is also suited for proteins to fold correctly and thus have binding sites within the correct conformation for interaction. However, proteins that require post-translational modifications for functionality to enable an interaction within the original cell may not be facilitated within the yeast cellular environment resulting in false negatives. Fastidious interactions that require specific conditions pertaining to temperature, pH, steric hinderances and post-translational modifications may also be classified as false negatives (Brückner et al., 2009; Huang & Bader, 2009). Furthermore, weak and transient protein-protein interactions can be observed in Y2H assays (Meyer & Schomburg, 2008). Viral proteins tend to also involve weak and transient interactions (Maio et al., 2020; Pascal et al., 1994) making this technology ideal to identify and investigate viral-host protein interactions.

2.2 AIM AND OBJECTIVES

Previous work with Y2H assays to identify single binary interactions with SACMV MP protein by the Rey laboratory proved to be unsuccessful (Nankoo, 2018). Therefore, a high throughput system was preferred for the identification of novel interactions of SACMV MP within cassava host proteins.

The aim of this study was to identify novel interactions of the CMD-tolerant cassava TME3 landrace host proteins with SACMV MP at 32 dpi using a high-throughput robust Y2H screening outsourced to Hybrigenics Services, S.A.S., Evry, France (http://www.hybrigenics-services.com)

The main objectives were as follows:

1. Propagation and evaluation of SACMV infection of cassava TME3, for harvesting of leaf tissue at 32 days post infection (dpi).
2. Extraction of cassava TME3 RNA and SACMV MP codon optimised expression vector to enable cloning into appropriate vectors for Y2H screening.
3. To evaluate the relevance of identified interactions using existing literature with regards to understand SACMV MP interactions during SACMV infection.

2.3 METHODS

The cassava TME3 prey cDNA library synthesis and MP bait cloning and Y2H screening were carried out by Hybrigenics services (S.A.S., Evry, France). The samples were prepared as follows prior to further processing by Hybrigenics services:

2.3.1 PREY SAMPLE PREPARATION

2.3.1.1 MICRO-PROPAGATION AND POTTING OF CASSAVA TME3 PLANTLETS (70 DAYS)

This was done according to established protocol with modifications (Allie et al., 2014). Sterile nodal cultures of cassava TME3 landrace were micro-propagated on Murashige and Skoog (MS) medium (Murashige & Skoog, 1962) supplemented with 2 % sucrose, 0.002 mM $CuSO_4$, 0.78 % plant tissue culture agar, pH 5.8, for 3 weeks at 28 °C with a 16h photoperiod under 8000–10000 lux. Rooted plantlets were transferred into Jiffy® pellets (Jiffy Products International) under clingwrap in a controlled growth chamber at 28 °C with a 16h photoperiod under 8000–10000 lux. Jiffies were acclimatised to the growth chamber environment. This was carried out over a 3 week period with three holes made in the clingwrap at the end of each week and thereafter, the cling wrap was completely removed.

2.3.1.2 INFECTION OF CASSAVA TME3 PLANTLETS WITH SACMV INFECTIOUS CLONES (32 DAYS) AND SYMPTOM EVALUATION

Plantlets were infected with equal quantities of SACMV pBIN-DNA-A and pBIN-DNA-B infectious, head-to-tail, dimers transformed into *Agrobacterium tumefaciens* strain Agl1 (Agl1) (Berrie et al., 2001). Infectious clones were cultured in YEP broth (100 µg/ml carbenicillin, 100 µg/ml kanamycin, 30 °C, 200 rpm) to an OD_{600} of 1.8-2.0. The cultures were pelleted by centrifugation for 20 min at 4200 xg. The pellet was thereafter washed in an equal volume of sterile distilled water by centrifugation for 20 min at 4200 xg and finally resuspended in an equal volume of fresh YEP broth. Equal volumes of SACMV DNA-A and DNA-B were mixed. A 200 µl syringe with a 5 mm needle was used to inoculate the plantlets through the petioles and stems of cassava TME3. Wild type untransformed Agl1 was cultured and inoculated in the same manner. Plant height and symptom severity was recorded for 32 days. Symptoms were scored according to previous studies (Allie et al., 2014; Hahn et al., 1980) on a scale of 0 to 3;0 (no symptoms/healthy), 1 (mild leaf distortion or mild chlorosis), 2 (moderate leaf distortion and chlorosis) and 3 (severe leaf distortion and chlorosis with reduction in leaf area). Harvesting of leaf tissue was carried out at 32 days post infection (32 dpi) from SACMV-infected cassava TME3 plants. Only symptomatic

leaves were harvested (symptom score 0-3). Confirmation of SACMV by was by PCR amplification with coat protein primers (Fwd: '5-CGTGGCTGTGAA-3') and (Rev: '5-CGCGTGACATCA-3'). Harvested samples were immediately frozen in liquid nitrogen and stored at −80 °C prior to RNA extraction.

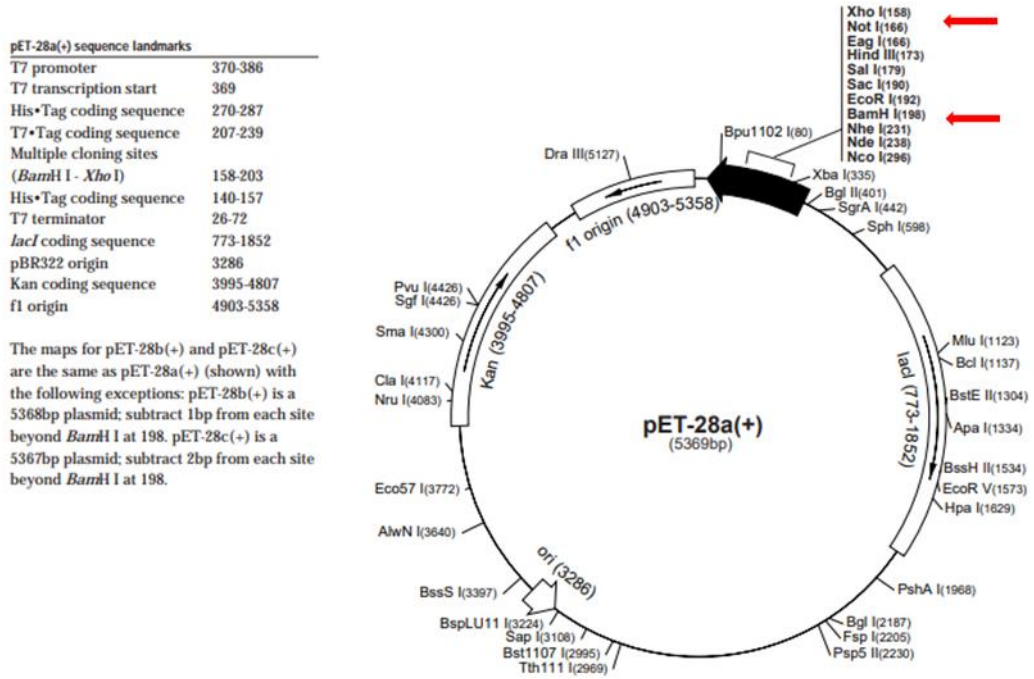
2.3.1.3 RNA EXTRACTION

A CTAB RNA extraction method (Chang et al., 1993) effective in pine tissue was modified to be ideal for RNA extraction from cassava. Two hundred milligram of leaf tissue in 2 ml tubes was ground in liquid nitrogen. One ml of CTAB buffer (2 % (w/v) CTAB, 2 % (w/v) (polyvinylpolypyrrolidone-40) PVP-40, 100 mM (hydroxymethyl aminomethane HCl (Tris-HCl)), pH 8.0, 25 mM ethylenediaminetetraacetic acid (EDTA), pH 8.0, 11.7 % (w/v) NaCl, 10 % 1 M (hydroxymethyl aminomethane HCl (Tris-HCl)), pH 8.0) supplemented with 0.1 % (v/v) β- mercaptoethanol and with 0.1 % (v/v) spermidine trihydrochloride was added to the ground leaf tissue. The sample was incubated at 65 °C for 12 min with shaking every 2 min. Thereafter, an equal volume (1 ml) of chloroform: isoamyl alcohol (CIA) (24:1) was added, the sample was vortexed and centrifuged at 4 °C for 10 min at 12000 xg. The supernatant was transferred to a new pre-chilled (1-4 °C) tube and the addition of CIA, centrifugation, and transfer of supernatant repeated 4 times. The fourth supernatant was transferred to a new tube and a 25 % volume of 10 M Lithium chloride (LiCl) was added and mixed by gentle inversion. The sample was left to precipitate overnight at -20 °C. Thereafter, the LiCl RNA solution was centrifuged at 4 °C for 65 min at 12000 xg and the LiCl pellet decanted carefully. The pellet was resuspended with 1 ml pre-chilled (-20 °C) 70 % ethanol and thereafter centrifuged at 4 °C for 10 min at 12000 xg. The supernatant was discarded, and the ethanol wash repeated. The pellet was air dried at room temperature for 10 min and resuspended in 50 µl of RNAse free water and stored at -80 °C. RNA samples were pooled into a total of four samples and integrity assessed. RNA quantity and purity were assessed by A260/A280 and A260/A230 ratios. RNA integrity was further assessed on a 1 % TAE bleach gel (Aranda et al., 2012). Additionally, RIN score determination (Schroeder et al., 2006) was outsourced to Inqaba Biotec (Pretoria, South Africa).

2.3.2 BAIT SAMPLE PREPARATION

The BC1 gene (923 bp) (codon-optimised to GenBank ID: NP_620668.1 BC1 [South African cassava mosaic virus] SACMVsBgp2-BC1) was previously cloned into pET-28a (+) (Novagen) for protein expression by *E. coli* BL21(DE3) using NdeI and NotI restriction sites **(Figure 2.2)** (Nankoo, 2018). *E. coli* BL21(DE3) transformed with BC1- pET-28a (+) was cultured in 10 ml SOB media (0.5 % (w/v) yeast extract, 2 % (w/v) tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgSO4, 10 mM $MgCL_2$, 50 µg/ml kanamycin and incubated for 16 h

at 37 °C on a shaker at 200 rpm. The resulting bacterial culture was centrifuged at room temperature for 10 min at 12000 xg to pellet the bacteria. Plasmid extraction on the bacterial pellet was carried out according to the GeneJET Plasmid Miniprep Kit (ThermoFisher Scientific, Massachusetts, USA) and the plasmid confirmed by restriction digest and PCR. Restriction enzymes NdeI and NotI digest were visualised by agarose gel electrophoresis. PCR amplification (annealing at 58 °C and elongation for 30 seconds for 30 cycles resulting in a 900 bp amplicon) with primers, (Fwd: '5-TATAGGATCCATGGACGCCCAATTTACCCG-3') and (Rev: '5-ATGCGGCCGCTTACAATTGCCTTGGTG-3'). The undigested plasmid was also confirmed by sequencing (Inqaba Biotec, Pretoria, South Africa). The whole undigested plasmid was stored at -20 °C and provided to Hybrigenics for bait cloning.



pET-28a(+) sequence landmarks	
T7 promoter	370-386
T7 transcription start	369
His•Tag coding sequence	270-287
T7•Tag coding sequence	207-239
Multiple cloning sites (*Bam*H I - *Xho* I)	158-203
His•Tag coding sequence	140-157
T7 terminator	26-72
lacI coding sequence	773-1852
pBR322 origin	3286
Kan coding sequence	3995-4807
f1 origin	4903-5358

The maps for pET-28b(+) and pET-28c(+) are the same as pET-28a(+) (shown) with the following exceptions: pET-28b(+) is a 5368bp plasmid; subtract 1bp from each site beyond *Bam*H I at 198. pET-28c(+) is a 5367bp plasmid; subtract 2bp from each site beyond *Bam*H I at 198.

Figure 2.2 The TB074VM pET-28a-c (+) vector map construct provided by Novagen. The BC1 gene (923 bp) inserted between NdeI and NotI restriction sites (red arrows) on the map.

2.3.3 Y2H SCREENING (ULTImate Y2H screen)

Yeast two-hybrid screening (*ULTImate* Y2H screen) was performed by Hybrigenics Services, S.A.S., Evry, France (http://www.hybrigenics-services.com). The 923 bp codon-optimised coding sequence for MP was PCR-amplified from pET-28a and cloned into pB27 as a C-terminal fusion to LexA (LexA-BC1). The MP bait was neither toxic to the yeast nor auto-activating. The construct was confirmed by sequencing of the entire insert and used

as a bait to screen a randomly primed cassava TME3 SACMV infected (32 dpi) cDNA library constructed into pP6 prey vector. pB27 and pP6 vectors were derived from the original pBTM116 (Vojtek & Hollenberg, 1995) and pGADGH (Bartel et al., 1993) plasmids, respectively. One hundred and fifty million clones (15-fold the complexity of the library) were screened using a mating approach with YHGX13 (Y187 ade2-101:loxP-kanMX-loxP, matα) and L40ΔGal4 (mata) yeast strains as previously described (Fromont-racine et al., 1997). Seventy His+ colonies were selected on a medium lacking tryptophan, leucine, and histidine. The prey fragments of the positive clones were amplified by PCR and sequenced at their 5' and 3' junctions. The resulting sequences were used to identify the corresponding interacting proteins in the GenBank database (NCBI) using a fully automated procedure. A confidence score (PBS, for Predicted Biological Score) was attributed to each interaction as previously described (Formstecher et al., 2005).

2.3.4 CONFIDENCE SCORE

The PBS relies on two different levels of analysis. Firstly, a local score that considers the redundancy and independence of prey fragments, as well as the distribution of reading frames and stop codons in overlapping fragments. Secondly, a global score that considers the interactions found in all the screens performed at Hybrigenics using the same library. This global score represents the probability of an interaction being nonspecific. For practical use, the scores were divided into four categories: from A (highest confidence) to D (lowest confidence). A fifth category (E) specifically flags interactions involving highly connected prey domains previously found several times in screens performed on libraries derived from the same organism. Finally, several of these highly connected domains were confirmed as false positives of the technique and tagged as F. The PBS scores have been shown to positively correlate with the biological significance of interactions (Rain et al., 2001; Wojcik et al., 2002).

2.3.5 BIOINFORMATICS ANALYSIS

Gene ontology was carried out using the TAIR GO Slim/Functional Classification online function (https://www.*Arabidopsis*.org/tools/bulk/go/index.jsp). The list of MP interacting proteins identified from Y2H screening were extended by inclusion of interactors identified from an *Arabidopsis* interactome (Velásquez-Zapata et al., 2021) that comprises only experimentally determined interactions from BioGRID 4.1.190 version (Stark et al., 2006), and literature review (Mukhtar et al., 2011; Weßling et al., 2014; Smakowska-Luzan et al., 2018; Trigg et al., 2017; Mcwhite et al., 2021).

2.3.5.1 POOLED BIOINFORMATICS ANALYSIS

The list of identified MP interacting proteins and their interactors from the *Arabidopsis* interactome was further extended in STRING v11.0 (https://string-db.org/). Active interaction sources were retrieved via the textmining, experiments and databases options with a minimum interaction score of 0.400. Network extension was carried out to obtain a sample size that would be statistically ideal for pathway enrichment analysis. Pathway enrichment was determined via web based tool ShinyGO v0.60 (Ge et al., 2020). ShinyGO pathway enrichment analysis was based on hypergeometric distribution followed by FDR correction (FDR < 0.05).

2.3.5.2 SELECTIVE BIOINFORMATICS ANALYSIS

Interactors were categorised into 3 sets based on the ability to link each identified interactor to experimentally determined interactors from either the *Arabidopsis* interactome or STRING (Set X and Set Y, respectively) and those that could not be linked to experimental interactors but could be linked to interactors in STRING from databases and textmining (Set Z). Out of the 12 identified MP interactors, Set X (AT5G64750, AT1G03130, AT1G10390 and AT3G03940) were linked to interactors from the *Arabidopsis* interactome (experimentally determined interactions). Of the remaining identified interactors without any identified interactors from the *Arabidopsis* interactome, Set Y (AT3G11710, AT3G57520, AT4G27910 and AT3G19420) were assigned interactors via STRING (Active interaction sources were restricted to experimental data). Set Z (AT1G28230, AT2G17265, AT2G27980 and AT1G07090) were not linked to any experimentally determined interactors in both the *Arabidopsis* interactome and STRING. These were therefore assigned interactors with Active interaction sources of textmining, experiments and databases. All data extended in STRING was analysed for KEGG pathway enrichment in the ShinyGO web tool. A comparison of enriched KEGG pathways was conducted.

2.4 RESULTS

2.4.1 SACMV-INFECTED CASSAVA TME3 MORPHOLOGICAL FEATURES

Of the SACMV-infected cassava TME3 plants, 17 % did not present symptoms and SACMV was not detected by PCR (data not shown). Only infected and symptomatic plants were considered for sampling **(Appendix 2.1)**. Initial symptom progression was noted from 12 to 22 dpi. Plants that did not show symptoms post 22 dpi were not successfully infected. The characteristic symptoms of SACMV infection recorded herein were comparable to prior studies (leaf curling, chlorosis, leaf distortion and reduced photosynthetic leaf area) (Allie et al., 2014; Bengyella et al., 2015; Bizabani et al., 2021) **(Figure 2.3)**. Out of 30 infected plants, scoring was noted as follows, 4 %, 50 % and 10 % of plants had a symptom severity score of 3, 2 and 1 respectively giving an overall mean symptom score of 2.3. All plants were notably symptomatic at 32 days post inoculation. Increase in plant height ranged from 7 % - 267 % from the initial plant height at 0 dpi **(Appendix 2.1)**. PCR amplification with SACMV coat protein primers further confirmed sampled plants were infected by yielding a single 95 bp band **(Appendix 2.2)**.

Figure 2.3 Visual representation of symptom score (0-3) of cassava TME3 infected with SACMV infectious clones. 0 (no symptoms/healthy), 1 (mild leaf distortion and mild chlorosis), 2 (moderate leaf distortion and chlorosis), 3 (severe leaf distortion and chlorosis with reduction in leaf area).

2.4.1 RNA INTEGRITY: NON-DENATURING GEL AND RIN VALUES

Total RNA on non-denaturing agarose gel run in 1 % TAE buffer gave distinct bands for 18S and 28S RNA fragments **(Figure 2.4A)**. Degradation of RNA presented by a smearing of bands was not noted. The four pools of RNA spectrophotometric values were as follows: A260/A280 ratios of 2.16, 2.18, 2.23 and 2.05 and concentrations of 194, 183, 347 and 405 ng/µl, respectively. Total volume concentrations were 159, 151, 622, 417 µg, respectively. These parameters and the total concentration approximating 1 mg fulfilled the requirements for library construction by Hybrigenics services. RIN determination by the Agilent 2100 bioanalyzer was set for evaluation of eukaryotic RNA. RIN values range from 1 to 10, the latter denoting RNA of high quality. Extracted RNA RIN values were 8.00, 8.30, 7.10 and 7.90, respectively **(Figure 2.4C)**. Capillary gel electrophoresis images showed an

abundance of high molecular weight fragments, 28S and 18S RNA below the 4 kb and 2 kb positions, respectively **(Figure 2.4B)**. The rRNA ratio (28S/18S) ranges from 0 to 2, the latter denoting RNA of the highest quality. rRNA ratios were 1.6, 1.8, 1.8 and 1.5, respectively.

2.4.2 CONFIRMATION OF BC1 PLASMID CLONED INTO pET-28a (+) BY RESTRICTION DIGEST AND PCR AMPLIFICATION

The restriction enzyme (NdeI/NotI) digest of BC1-pET-28a (+) resulted in only two expected size fragments of 923 bp for the BC1 coding sequence insert and ~5400 bp for the pET-28a(+) fragment **(Figure 2.5A)**. Amplicons of the BC1 insert were confirmed by a single band below the 1000 kb region (~923 bp fragment) **(Figure 2.5B)**.

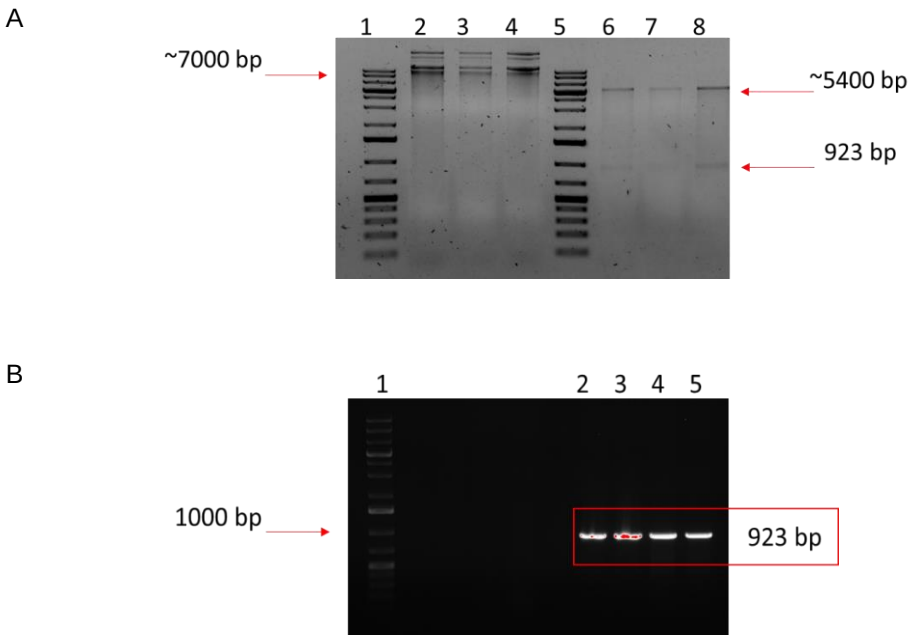


Figure 2.5 Confirmation of BC1 insert in pET-28a(+). **A**. 1 % agarose gel showing undigested BC1-pET-28a (+) plasmid (2-4) and BC1- pET-28a (+) digested with NdeI/NotI to confirm the presence of the BC1 insert resulting in two fragments of 923 bp BC1 insert (3-4) bp and 5400 bp pET28a(+) (6-7) fragments, GeneRuler^{TM} 1kb Plus DNA ladder (1, 5). **B**. 1 % agarose gel showing BC1- pET-28a (+) plasmid PCR products to confirm the presence of the BC1 insert resulting in a 900 bp fragments (2-5), GeneRulerTM 1kb Plus DNA ladder (1).

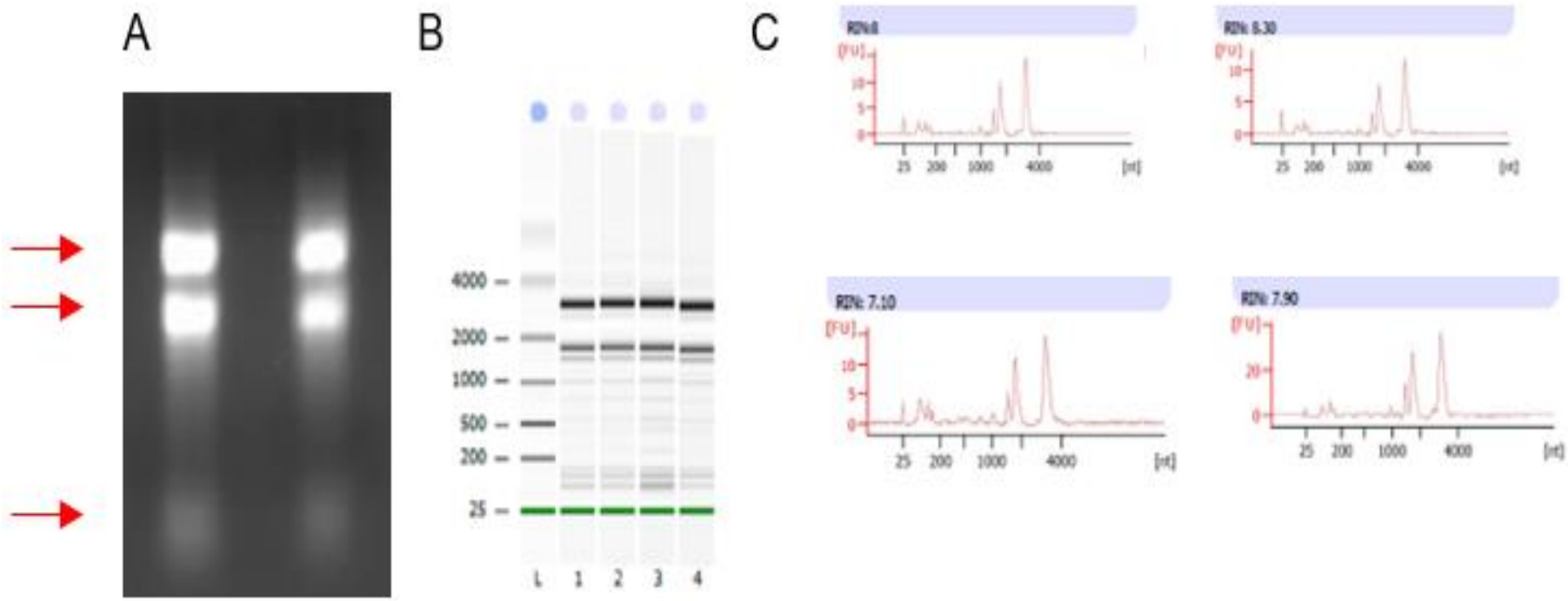


Figure 2.4 RNA Integrity determination. **A**. 1.3 % non-denaturing agarose gel showing total RNA extracted from cassava TME3 leaf tissue. Intact RNA separated into 28S, 18S and 5S bands indicated by red arrows from top to bottom. **B**. RIN capillary gel electrophoresis image obtained for four individual pools of total RNA sample relative to Ladder (L). **C**. Electropherogram and RIN value for each sample.

2.4.4 INTERACTORS IDENTIFIED BY Y2H SCREENING

A Y2H screen of SACMV MP as bait against a cassava TME3-SACMV-infected library identified interacting proteins classified according to their PBS categories (A-D) **(Figure 2.6, Table 2.1)**. These were identified by their *Manihot esculenta* gene identities and for further analysis were annotated according to their *Arabidopsis* homologue identities. Identified interactors were, casein kinase 1-like protein HD16 and nuclear pore complex protein NUP98A (Very high confidence-A), lysine--tRNA ligase, cytoplasmic-like (high confidence-B), photosystem I reaction centre subunit II, chloroplastic-like and probable galactinol-sucrose galactosyltransferase 2 (Good confidence-C) and histone-lysine N-methyltransferase ATX4-like, purine permease 1-like, homoserine kinase, phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase and protein-tyrosine-phosphatase PTEN2A, uncharacterized protein, ethylene-responsive transcription factor ABR1-like, protein light-dependent short hypocotyls 4-like and chloroplast (Moderate confidence-D). Chloroplasts , due to their abundance in plant leaf tissue were not regarded for further bioinformatics analysis.

Screen Parameters

Nature	cDNA
Reference Bait Fragment	South African cassava mosaic virus - SACMVsBgp2-BC1 (aa 1-307) ; hgx5000v1
Prey Library	Manihot esculenta virus infected_RP1
Vector(s)	pB27 (N-LexA-bait-C fusion)
Processed Clones	70 (pB27_A)
Analyzed Interactions	150 millions (pB27_A)
3AT Concentration	0.0 mM (pB27_A)

Global PBS®

Global PBS (for Interactions represented in the Screen)		Nb	%
A	Very high confidence in the interaction	2	14.3%
B	High confidence in the interaction	1	7.1%
C	Good confidence in the interaction	2	14.3%
D	Moderate confidence in the interaction This category is the most difficult to interpret because it mixes two classes of interactions : - False-positive interactions - Interactions hardly detectable by the Y2H technique (low representation of the mRNA in the library, prey folding, prey toxicity in yeast)	9	64.3%
E	Interactions involving highly connected (or relatively highly connected) prey domains, warning of non-specific interaction. The total number of screens performed on each organism is taken into account to set this connectivity threshold: 20 interactions to different bait proteins in our entire database for Human, 10 for Mouse, Drosophila and Arabidopsis and 6 for all other organisms. They can be classified in different categories: - Prey proteins that are known to be highly connected due to their biological function - Proteins with a prey interacting domain that contains a known protein interaction motif or a biochemically promiscuous motif	0	0.0%
F	Experimentally proven technical artifacts	0	0.0%
Non Appliable			
N/A	The PBS is a score that is automatically computed through algorithms and cannot be attributed for the following reasons : - All the fragments of the same reference CDS are antisense - The 5p sequence is missing - All the fragments of the same reference CDS are either all OOF1 or all OOF2 - All the fragments of the same reference CDS lie in the 5' or 3' UTR		

Figure 2.6 ULTImate Y2H screen summary of SACMV MP vs infected cassava TME3. Global PBS scores are rated from highest confidence to lowest confidence score (A to F) and non-applicable scores.

2.4.5 ANNOTATION OF INTERACTORS IDENTIFIED BY Y2H SCREENING

Table 2.1 List of SACMV MP interacting candidates from Y2H screen annotated according to *Arabidopsis thaliana* homologues and functional domain (Pfam ID)

PBS category	Prey ID (NCBI)	Prey annotation	Alternative ID	*Arabidopsis* ID
A	LOC110603635	casein kinase 1-like protein HD16 (CKL1)	Manes.16G117000	AT3G03940
A	LOC110624105	nuclear pore complex protein (NUP98A)	Manes.01G002100	AT1G10390
B	LOC110619246	lysine--tRNA ligase, cytoplasmic-like (KRS1)	Manes.07G091500	AT3G11710
C	LOC110609589	photosystem I reaction centre subunit II, chloroplastic-like (PsaD)	Manes.02G143000	AT1G03130
C	LOC110622020	probable galactinol-sucrose galactosyltransferase 2 (SIP2)	Manes.09G072200	AT3G57520
D	LOC110599949	histone-lysine N-methyltransferase ATX4-like (SDG16)	Manes.14G006400	AT4G27910
D	LOC110606559	purine permease 1-like (PUP1)	Manes.18G062600	AT1G28230
D	LOC110621944	homoserine kinase (HSK)	Manes.01G017200	AT2G17265
D	LOC110624740	phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase and protein-tyrosine-phosphatase (PTEN2A)	Manes.10G077300	AT3G19420
D	LOC110626469	acyl-CoA N-acyltransferase (NAT)	Manes.11G153000	AT2G27980
D	LOC110627812	ethylene-responsive transcription factor ABR1-like (ABR1)	Manes.12G146000	AT5G64750
D	LOC110630773	protein LIGHT-DEPENDENT SHORT HYPOCOTYLS 4-like (LSH)	Manes.14G126500	AT1G07090
D	EU117376.1	chloroplast	NA	NA

2.4.6 GENE ONTOLOGY OF IDENTIFIED MP INTERACTORS

Gene ontology categorisation of the MP interactors revealed the cellular localisation and associated biological processes impacted. A percentage frequency was assigned to denote the categories represented by the interactors relative to the whole *Arabidopsis* genome **(Figure 2.7,Table 2.2)**. The topmost GO Slim category for "Cellular Components" represented were those for "chloroplast" 19 %, "nucleus" 16 % and, "plastid", "thylakoid", "other membranes", "cytoplasm" at (9 %). With regards to GO Slim category, "Biological Process", the most highly represented were other " "cellular processes" (21 %), "other metabolic processes" (16 %) and, "cellular protein modification process", "response to chemical", "transport" (6 %). GO Slim category, "Molecular Function" was most represented by "transferase activity", "catalytic activity", "hydrolase activity", "protein binding" (13 %), followed by "kinase activity" (10 %) and "DNA binding", "transporter activity", "RNA binding" (6 %).

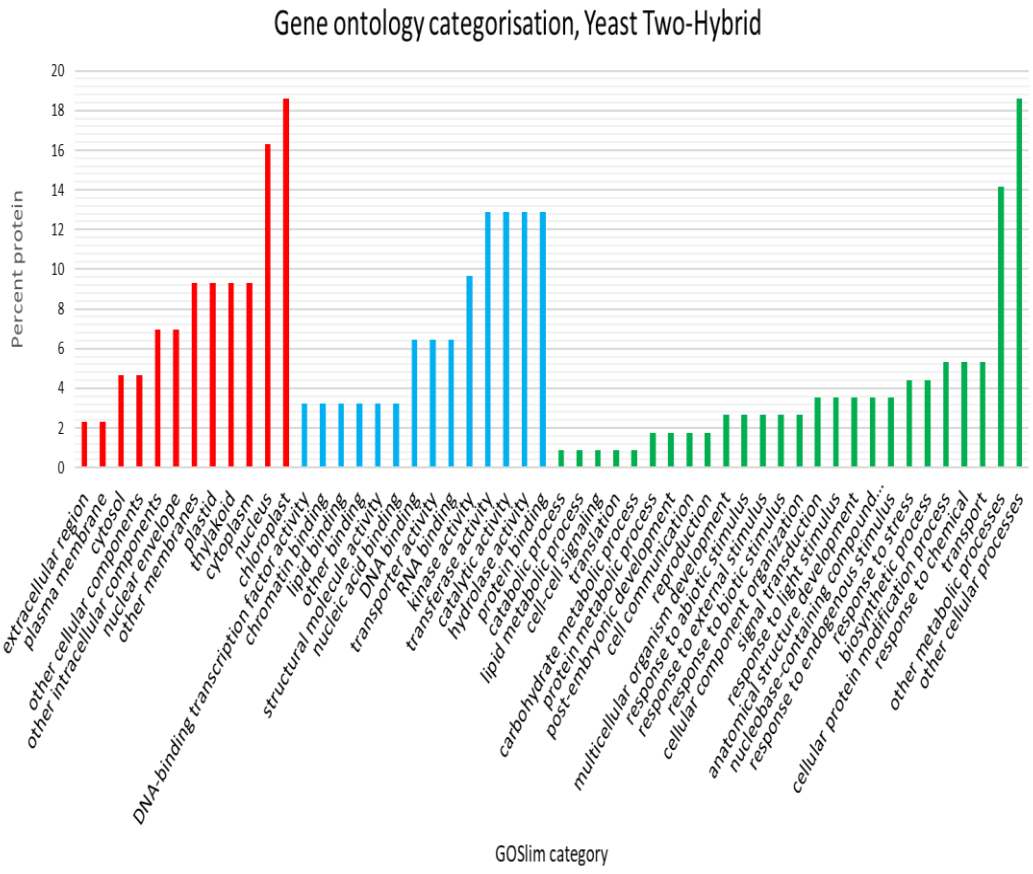


Figure 2.7 The Gene Ontology of MP interacting proteins identified from Y2H assays, Cellular component (Red), Biological process (Green), Molecular function (Blue).

Table 2.2 Gene ontology categorisation of interactors identified by Y2H assays

GO Category	Functional Category	Percentage
GO Cellular Component	extracellular region	2
	plasma membrane	2
	cytosol	5
	other cellular components	5
	other intracellular components	7
	nuclear envelope	7
	other membranes	9
	plastid	9
	thylakoid	9
	cytoplasm	9
	nucleus	16
	chloroplast	19
GO Molecular Function	DNA-binding transcription factor activity	3
	chromatin binding	3
	lipid binding	3
	other binding	3
	structural molecule activity	3
	nucleic acid binding	3
	DNA binding	6
	transporter activity	6
	RNA binding	6
	kinase activity	10
	transferase activity	13
	catalytic activity	13
	hydrolase activity	13
	protein binding	13
GO Biological Process	catabolic process	1
	lipid metabolic process	1
	cell-cell signalling	1
	translation	1
	carbohydrate metabolic process	1
	protein metabolic process	2
	post-embryonic development	2
	cell communication	2
	reproduction	2
	multicellular organism development	3
	response to abiotic stimulus	3
	response to external stimulus	3
	response to biotic stimulus	3
	cellular component organization	3
	signal transduction	4
	response to light stimulus	4
	anatomical structure development	4
	nucleobase-containing compound metabolic process	4
	response to endogenous stimulus	4
	response to stress	4
	biosynthetic process	4
	cellular protein modification process	5
	response to chemical	5
	transport	5
	other metabolic processes	14
	other cellular processes	19

2.4.7 KEGG PATHWAY ENRICHMENT OF EXTENDED INTERACTION NETWORK (POOLED)

To understand the extent of impact on biological pathways the interaction network of identified interactors was evaluated for KEGG pathway enrichment **(Figure 2.8A-B)**. The topmost enriched KEGG pathway was, “Proteasome”. The data was represented in two ways. In the first **(Figure 2.8A)**, the percentage of proteins (x axis) was plotted against the enriched KEGG pathway (y axis). In the second analysis, a hierarchical clustering tree based on the enriched KEGG pathways was created **(Figure 2.8B)**. Enrichment analysis was based on hypergeometric distribution followed by FDR correction (<0.05). The hierarchical clustering tree summarised the correlation among significantly enriched KEGG pathways.

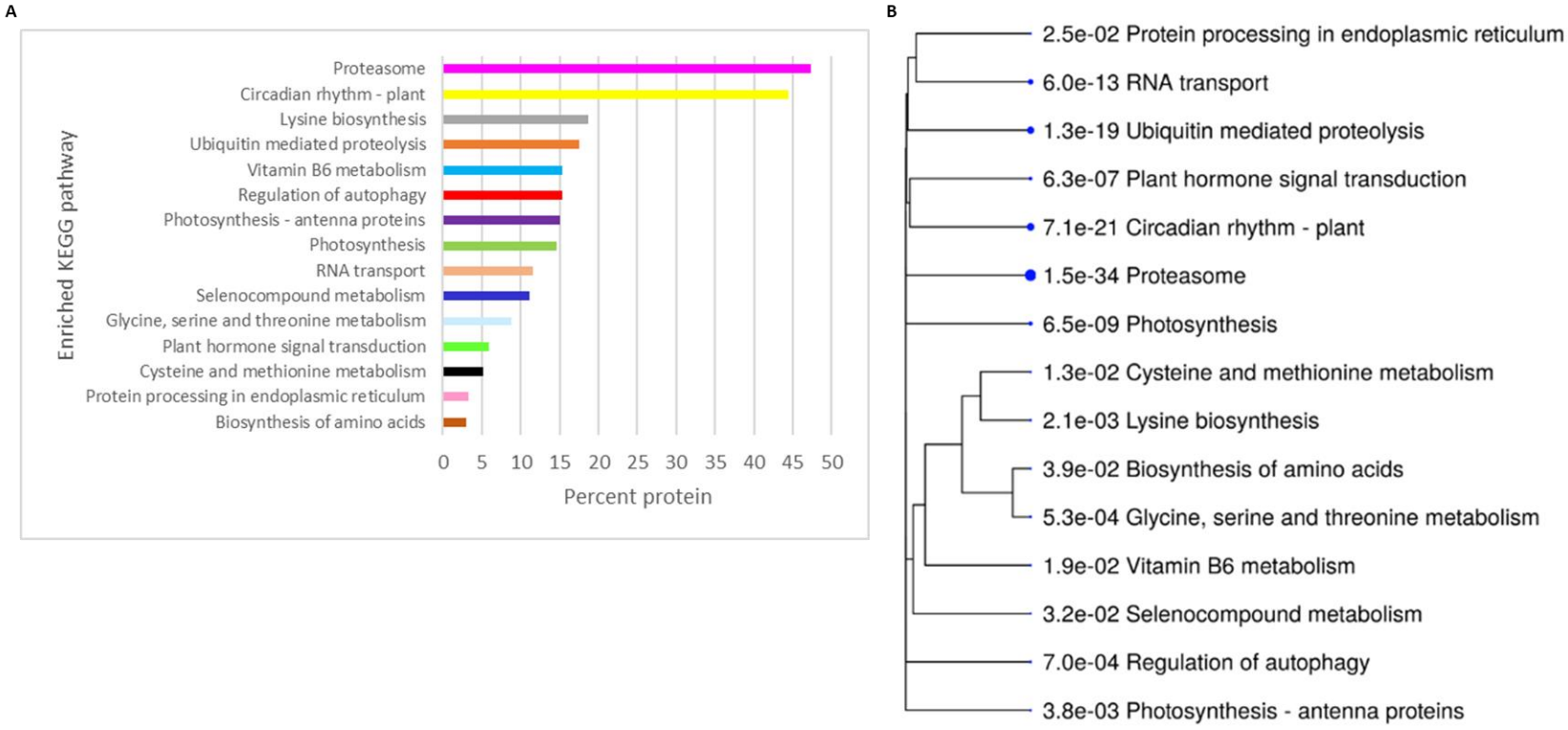


Figure 2.8 Enriched KEGG pathways identified from *Y2H* assays. *A*. Bar chart plotted against percentage of associated proteins. The enriched KEGG pathways from the interaction network of MP interacting proteins (pooled) identified by Y2H assays. Enrichment analysis based on hypergeometric distribution followed by FDR correction (<0.05). *B* Hierarchical clustering tree based on the enriched KEGG pathways from interaction network of MP interacting proteins (pooled) identified by Y2H assays. Pathways with many shared genes are clustered together. Bigger dots indicate more significant p-values (<0.05).

2.4.8 ENRICHED KEGG PATHWAY COMPARISONS OF INTERACTING PROTEINS

Twenty six enriched KEGG pathways were common to at least two of the identified interactors associated with MP **(Figure 2.9, Table 2.3)**.

Table 2.3 Enriched KEGG pathways common to two or more proteins identified by Y2H assays

KEGG pathway	Number of proteins	Arabidopsis ID				
Ribosome	2	AT3G03940	AT1G28230			
Circadian rhythm - plant	2	AT3G03940	AT5G64750			
RNA transport	2	AT1G03130	AT1G10390			
Proteasome	2	AT1G03130	AT1G10390			
Ubiquitin mediated proteolysis	4	AT1G03130	AT1G10390	AT5G64750	AT3G11710	
mRNA surveillance pathway	3	AT1G03130	AT1G10390	AT4G27910		
Spliceosome	3	AT1G03130	AT1G10390	AT3G19420		
DNA replication	5	AT1G10390	AT5G64750	AT4G27910	AT3G11710	AT2G17265
Nucleotide excision repair	4	AT1G10390	AT5G64750		AT2G17265	AT3G11710
Pyrimidine metabolism	3	AT1G10390	AT5G64750	AT1G07090		
Purine metabolism	3	AT1G07090	AT1G10390	AT5G64750		
Plant hormone signal transduction	2	AT1G28230	AT5G64750			
Base excision repair	4	AT3G11710	AT5G64750		AT4G27910	AT2G17265
Protein processing in endoplasmic reticulum	3	AT3G19420	AT1G07090	AT5G64750		
Mismatch repair	4	AT3G11710	AT2G17265	AT3G19420	AT4G27910	
Galactose metabolism	2	AT3G57520	AT2G17265			
Metabolic pathways	4	AT3G57520	AT2G17265	AT3G19420	AT1G07090	
Lysine degradation	2	AT2G27980	AT4G27910			
Inositol phosphate metabolism	2	AT1G07090	AT3G19420			
Phosphatidylinositol signalling system	2	AT1G07090	AT3G19420			
Regulation of autophagy	2	AT1G28230	AT3G19420			
Carbon metabolism	2	AT2G17265	AT3G19420			
Vitamin B6 metabolism	2	AT1G28230	AT2G17265			
Biosynthesis of secondary metabolites	3	AT2G17265	AT2G27980	AT1G07090		
Cysteine and methionine metabolism	2	AT2G17265	AT2G27980			
Alanine, aspartate, and glutamate metabolism	2	AT2G17265	AT1G07090			

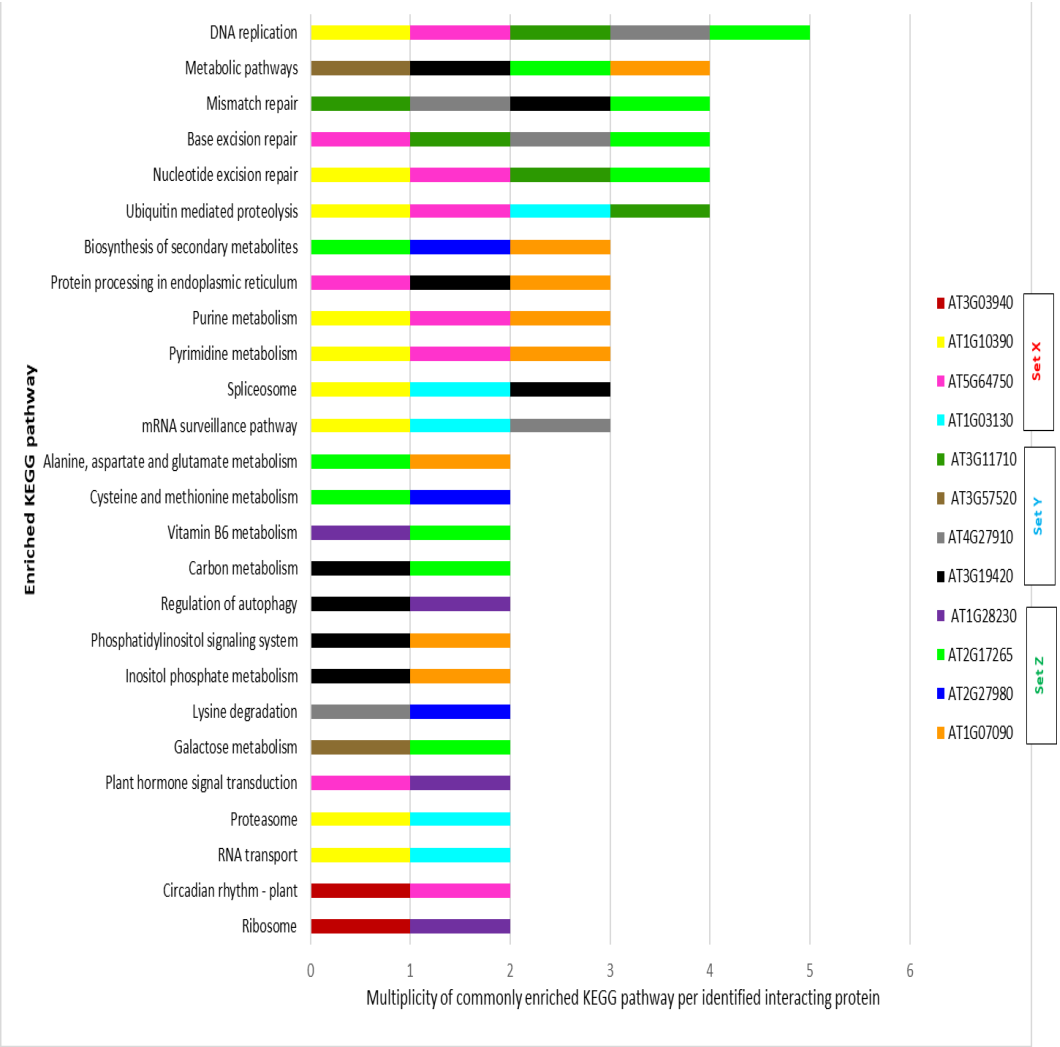


Figure 2.9 The multiplicity of enriched KEGG pathways common to at least two-of the identified MP interacting proteins. The interaction network of each MP interacting protein (pooled) was extended in STRING and enrichment analysis based on hypergeometric distribution followed by FDR correction (<0.05) in ShinyGO.

2.5 DISCUSSION

2.5.1 SYMPTOM SEVERITY AND TOLERANCE FEATURES OF SACMV-INFECTED CASSAVA TME3

Cassava TME3 has been established as an SACMV tolerant land race that albeit infection and symptom progression, there is a period post-infection in which asymptomatic leaves begin to emerge from the apical meristem (Allie et al., 2014; Bengyella et al., 2015; Bizabani et al., 2021). These later develop symptoms. The phenomenon of symptom recovery in RNA and DNA virus infected plants has been attributed to post-transcriptional and transcriptional gene silencing characterised by the accumulation of 21–22 nt vsiRNAs and 24 nt vsiRNAs, respectively (Ghoshal & Sanfaçon, 2015). In a separate study, RNA silencing pathway enzymes (AGO1 and DCL1) regulated by microRNAs (miR162 and miR168) were crucially linked to cassava TME3 tolerance (Bizabani et al., 2021). In *Arabidopsis*, infected with geminivirus BCTV, tolerance was linked to AGO4 mediated methylation (Raja et al., 2008). Prior studies (Allie et al., 2014; Bengyella et al., 2015; Bizabani et al., 2021) scored symptoms in susceptible T200 landrace and tolerant TME3 using a 5-tier index, with 5 exhibiting overtly gross reduction in leaf area. The cassava TME3 plants infected, in this study showed leaf chlorosis, distortion, and a reduction in leaf area. However, by comparison to previous studies, the culminating symptoms observed were not as severe. Hence the modification of the symptom severity score was adjusted to accommodate a 4-tiered severity score index.

2.5.2 RNA INTEGRITY FOR A HIGH-QUALITY cDNA LIBRARY

High RNA quality is a prerequisite for high quality cDNA library synthesis. Cassava leaf tissue is high in fats, polysaccharides and polyphenolics (Laya & Koubala, 2020; Staack et al., 2019) and this can be problematic in extraction of good quality RNA. The base method (Chang et al., 1993) initially designed for extraction from pine and modified for cassava in this study proved to be ideal for this purpose. Spectrophotometric values (A260/A280 ratios 2.05-2.23, 183-405 ng/ul) were comparable to other studies focused on the construction of high quality cDNA libraries for Y2H screening in *Curvularia* and *Cymbidium* sp. (A260/A280 ratios of 1.90-2.00, 100-700 ng/ul) (Xu et al., 2020; Gao et al., 2015). Extracted RNA RIN values were acceptable (7.1-8.3) for a superior cDNA library synthesis, and other studies have reported similar values (7.3-8.1) in *Actinidia* (kiwifruit) (Dharmaraj et al., 2019) and higher values (>8.0) from *Marchantia (bryophytes)*, tumour cells (mammalian) *and Arabidopsis* (Shivhare et al., 2021; Dharmaraj et al., 2019; Barreiro-Alonso et al., 2019; Rosas-Diaz et al., 2018). In the aforementioned studies, *Actinidia* RIN much like cassava TME3 values were lower and this may be attributed to sample sources being rich in polysaccharides and phenols (Afshar-Mohammadian et al., 2018). In cassava, RIN values

ranging between 6-8.1 (root and leaf) (Behnam et al., 2019), 7.2-8.0 (root) (Guan et al., 2019) and >6.0 (leaf) (Irigoyen et al., 2020) were sufficient for downstream enzymatic and transcriptomic analysis. It may be inferred that RIN values are species specific or inherently limited by the chemical composition of the sample source. Capillary gel electrophoresis images showed an abundance of high molecular weight fragments, 28S and 18S RNA **(Figure 2.4 A/B)**. The presence of an abundance of high molecular weight fragments is directly proportional to RNA integrity (Schroeder et al., 2006).

2.5.3 MP INTERACTED WITH PROTEINS ASSOCIATED TO MULTIPLE CELLULAR LOCATIONS

Gene ontology categorisation of the MP interactors revealed the cellular localisation associated with the interacting proteins **(Figure 2.7)**. MP is expected to traverse the nuclear and cell membranes and intercellularly. Intercellular movement is via the ER associated membrane system which connects cells via the plasmodesmata which crosses cell walls. The CC GO was comparable to prior literature in the localisation of geminiviral movement proteins to various intracellular organelles and membranes **(Figure 2.7)**. The C-terminal *anchor* and N-terminal *pilot* domains of AbMV MP amino acid domains have been respectively linked to intracellular localisation and, plasma membrane and nuclear membrane targeting (Zhang et al., 2002). Immuno-labelling techniques have detected AbMV MP in the plasma membranes (Aberle et al., 2002; Kleinow et al., 2009; Frischmuth et al., 2007), microsomal vesicles (Aberle et al., 2002) and endoplasmic reticulum membranes in the case of squash leaf curl virus (SqLCV) (Ward et al., 1997; Zhang et al., 2002) independent of any facilitation by plant host proteins. However other studies have shown facilitation by interaction with host HSC-70 (Krenz et al., 2010). Krenz et al. (2010) presented evidence for MP oligomerisation at the cell periphery and within chloroplasts *in planta* (Krenz et al., 2010). Furthermore, MP dependent re-localisation of the NSP from the nucleus to the plasma membrane has been shown (Frischmuth et al., 2007). In transient expression assays, BDMV-MP localised to plasmodesmata mediated by histone H3 (Zhou et al., 2011) and AbMV MP co-localised with HSC70-1 at the cell periphery and within chloroplasts (Krenz et al., 2010). Work by Zhao and colleagues intricately elaborated on the role of chloroplasts in plant virus infection (Zhao et al., 2016). Chloroplasts play a highly functional defence role in the provision of enzymes crucial to the biosynthesis of defence hormones; SA and JA, release of ROS and stromule transport of defence factors (Zhao et al., 2016) and therefore, the ability of MP to interact with chloroplast proteins may also allow for an interference of proteins crucial to anti-viral defence. The role of chloroplastic enzymes in antiviral defence was shown by PLRV infection in a phytoene desaturase (PDS) silenced background characterised by unhindered viral progression (DeBlasio et al., 2018). From the GO Slim terms for CC, BP, and MF, within this objective, it may be postulated

that perturbations noted in geminivirus studies may be well attributed to MP interaction with the host proteome. The noted GO Slim terms related to MP interacting proteins within this objective have been noted in geminivirus-host whole transcriptome studies for example with Pepper golden mosaic virus (PepGMV) and SACMV (Góngora-Castillo et al., 2012; Allie et al., 2013). *Nicotiana benthamiana* that is susceptible to SACMV, showed an overrepresentation of GO Slim categories, “nucleus, “other cellular components” (“Cellular Component”), “transferase activity, “other binding” (“Molecular Function”) and “other cellular processes”, “other metabolic processes” (“Biological Process”) (Allie et al., 2013). Within the “Biological processes”, GO Slim terms to mention a few, “transport, “response to stress”, transport”, “biosynthetic process”, “cellular protein modification process”, “signal transduction”, “anatomical structure development”, “nucleobase-containing compound metabolic process” were perturbed in transcriptome analysis of tolerant *Capiscum* sp. infected with PepGMV (Góngora-Castillo et al., 2012).

2.5.4 KEGG PATHWAY ANALYSIS REVEALED THAT MP INTERACTS WITH PROTEINS INVOLVED IN VARIOUS PLANT METABOLIC PATHWAYS

To have an overview understanding of how all the noted interactions of MP may impact morphological pathways, it was necessary to identify enriched KEGG pathways **(Figure 2.8A, 2.9)**. Most especially those that were common to the interacting proteins. It was noted that several KEGG terms were enriched and therefore, to get a general overview of enriched KEGG pathways, a word cloud (https://worditout.com/) was used to identify the term most prominent to identified interactors. The major term identified with MP interactions in pathway enrichment was the word “metabolism”. Seventy five percent (9/12) of the identified interactors were associated with metabolic pathways or metabolism **(Figure 2.8A, 2.9)**. The disruption and hijacking of metabolism and alteration of protein abundance in virus infection has long been established (Willets & Wong, 1980; Wang & Maule, 1995; Thaker et al., 2019). In a recent study, geminivirus jatropha leaf curl Gujarat virus (JLCuGV), caused metabolic alterations in amino acid, lipid and carbohydrate metabolism (More et al., 2021). Cassava TME3 was sampled at 32 dpi, a timepoint of peak SACMV infection and symptom progression (Allie et al., 2014). Therefore, the RNA extracted for Y2H assays may have been enriched with transcripts expressed in favour of host susceptibility and host metabolic dysregulation.

2.5.4.1 CARBOHYDRATE METABOLISM

In this study, leaf tissue sampled for Y2H screening presented symptoms consistent with an alteration in metabolic dysregulation with features of chlorosis and leaf curling. The symptoms noted and the interaction of MP with proteins (PTEN2A, HSK, LSH, SIP2) crucial to metabolism provides evidence of both MP and SACMV perturbation of carbohydrate

metabolism. Alteration of carbohydrate metabolism is a feature noted in leaf morphology and photosynthetic efficiency as it impacts the biochemical composition of leaf constituents by an abnormal accumulation of starch, carotene and chlorophyllase and a depletion of chlorophyll and xanthophyll components (Willets & Wong, 1980). Liu and colleagues further confirmed the increase in chlorophyllase transcripts in cassava ACMV infection and its link to chlorophyll degradation and compromised photosynthetic efficiency (Liu et al., 2014). A down regulation of the photosynthesis pathway has been noted in other virus studies of TYLCV (Wu et al., 2019), tobacco rattle virus (TRV) (Fernández-Calvino et al., 2014) and rice stripe virus (RSV) (Zhao et al., 2016). *Cucurbita* sp. infected with zucchini yellow mosaic virus (ZYMV) presented a reduction of total chloroplasts and an irregular chloroplast morphology linked to elevated plastoglobuli and starch and reduced thylakoids (Zechmann et al., 2003). These variations were also noted in *Nicotiana* sp. and attributed to the betasatellite (βC1) of radish leaf curl disease (RaLCB) (Bhattacharyya et al., 2015). In the latter study, photosynthetic efficiency of chloroplasts was reduced by 60 % (Bhattacharyya et al., 2015).

Geminivirus MP is a symptom determinant (Hou et al., 2000; Kleinow et al., 2009) and SACMV MP association with chloroplast protein (PsaD) and proteins involved in carbon metabolism (PTEN2A, HSK) and translocation of sugar compounds (SIP2) would have an impact on symptom severity **Figure 2.9)**. In *Arabidopsis*, turnip vein-clearing virus (TVCV), cucumber mosaic virus (CMV) and cauliflower mosaic virus (CaMV) the alteration of starch biosynthesis by the silencing of chloroplastic phosphoglucomutase enzyme led to an amelioration of symptom severity but not virus proliferation (Handford & Carr, 2007).

2.5.4.2 PROTEIN METABOLISM

In this study, MP was associated with proteins (NUP98, SDG16, PsaD, PTEN2A, HSK, CKL1, PUP1) involved in amino acid biosynthesis, amino acid metabolism, spliceosome, ribosome, RNA transport, mRNA surveillance and the ubiquitin mediated proteolysis which would impact host transcription, host protein profile and the abundance of amino acids within the host. An alteration of protein metabolism was noted by pathways linked to host gene expression and ubiquitin mediated proteolysis. For plant viruses to be able to thrive within the host, they have been shown to down regulate specific host gene expression. In pea seed-borne mosaic virus (PSbMV) localised down regulation of starch metabolic enzymes was consistent with an upregulation of viral RNA (Wang & Maule, 1995). In TRV infection genes associated with ribosomal function were down regulated whilst those linked to the ubiquitin-proteasome pathway were upregulated (Fernández-Calvino et al., 2014). The latter deliberately resulting in an influx of free amino acids for the TCA cycle (Fernández-Calvino et al., 2014). A feeding mechanism into the TCA cycle is favoured in virus infection in order to meet energy demands for viral replication and mobilisation in

mammalian cells (Mayer et al., 2019). Alteration of the amino acid composition of infected plants also provides a nutritional benefit to insect vectors for continual transmission of virus particles (Guo et al., 2019). TYLCV in infected tomato plants alters the free amino acid composition of phloem sap inadvertently favouring the insect vector (Guo et al., 2019). Alteration of amino acid metabolism may hinder effective immunity and interacting proteins associated with amino acid metabolism, degradation and biosynthesis were CKL, HSK, SDG16 and NAT. Aspartate is a precursor for multiple amino acids (lysine, threonine, isoleucine and methionine) crucial for plant immunity (Zeier, 2013). Lysine degradation results in the formation of pipecolic acid (Broquist, 1991) which confers systemic acquired resistance (SAR) by increasing NO and ROS species (Wang et al., 2018). Vitamin B6 induces autophagy and is typically identified as a significant cofactor for metabolic enzymes involved in diverse biochemical reactions essential to growth, development, immunity and stress tolerance and is responsible for oxidative stress tolerance (Vanderschuren et al., 2013).

2.5.5 THE PPI NETWORK OF MP INTERACTING PROTEINS REVEAL INTERSECTING PATHWAYS TARGETED BY MP DURING SACMV INFECTION OF CASSAVA TME3

It was noted that interacting proteins associated MP to four major categories of common KEGG pathways. To date, these associations have not been identified with regards to SACMV MP function. These were, **Nucleotide associated** including DNA replication, Nucleotide excision repair, Base excision repair, Mismatch repair, Pyrimidine metabolism, Purine metabolism. **Protein associated** including Ubiquitin mediated proteolysis, Proteasome, Spliceosome, mRNA surveillance pathway, Protein processing in endoplasmic reticulum, Ribosome, RNA transport, Amino acid metabolism (Cysteine and methionine metabolism, Alanine, aspartate and glutamate metabolism, Vitamin B6 metabolism, Lysine degradation, Biosynthesis of secondary metabolites). **Growth, cellular repair, and development associated** (Circadian rhythm – plant, Regulation of autophagy) and **Sugar and lipid signalling associated** (Plant hormone signal transduction, Inositol phosphate metabolism, Phosphatidylinositol signalling system, Galactose metabolism).

2.5.6 MP INVOLVEMENT IN VIRAL REPLICATION

The involvement of MP by interaction with proteins (NUP98, HSK, SDG16, PTEN2A, LSH, SIP2, KRS1) involved in viral replication was evident in the predominance of nucleotide associated pathways (DNA replication, Nucleotide excision repair, Base excision repair, Mismatch repair, Pyrimidine metabolism, Purine metabolism). Of further interest was the interference of the mRNA surveillance pathway which is crucial to identifying and obliterating aberrant and/or non-host DNA. Interference of this pathway has also been shown by other mammalian viruses (Wada et al., 2018; Nakano et al., 2013). An effective

suppression of PVX invasion in plants requires a collective contribution by both the nonsense-mediated decay (NMD) and RNA silencing pathway (Garcia et al., 2014). A down regulation of NMD genes occurs in PVX infection (Garcia et al., 2014). An affiliation to cell cycle related functions was noted It is well established that geminiviral proteins such as Rep can hijack the cell cycle to manipulate the host replication machinery for its own DNA replication. Nuclear shuttle protein (NSP) responsible for cell to cell movement has been shown to interfere with the cell cycle by interaction with *Arabidopsis* acetyltransferase which results in maintaining the cell in a dedifferentiated state (McGarry et al., 2003). While binding of MP has not been shown to date to bind to cell cycle components, MP also may interfere with gene regulation and inevitably growth, development and immune and adaptive stress responses by interaction with proteins within this category (Oliveira et al., 2017; Lorković, 2009).

2.5.7 MP INVOLVEMENT IN CIRCADIAN RYTHM

The circadian rhythm plays a critical role in growth, development and biologically priming the plant for defence against invading pathogens (Wang et al., 2011). The interacting proteins (ABR1, CKL1) associated with this pathway further highlight the need for post translational modifications and hormone signalling as a necessity for maintenance of circadian rhythm. An inhibition or interference of these proteins in their function would not only contribute to growth and developmental aberrations but also alter the plant's capacity to elicit an effective anti-viral response that comprises of defence hormone signalling strategies and post translational activation of immune responsive proteins.

2.5.8 MP INVOLVEMENT IN CELL SIGNALLING

SACMV MP was associated with proteins (PTEN2A, LSH, ABR1, PUP1) involved in plant signalling essential to cell function and defence. Sugar and lipid signalling pathways that incorporate phosphatidyl inositol, raffinose and galactose are derived from myo-inositol and are involved in lipid biosynthesis essential to signalling, transport of proteins and mRNA, abiotic and biotic stress, membrane trafficking, mRNA export, stress tolerance and storage of phosphorus (Valluru & Van den Ende, 2011). Infection of *Jatropha* sp. with JLCuGV, had a significant reduction of galactinol and myo-inositol (More et al., 2021).

2.5.9 INTERACTORS IDENTIFIED BY Y2H SCREENING

The Y2H screening resulted in the identification of twelve proteins. Each interacting protein was a binary interaction with its own specific influence within the cell system. It was important to fully scrutinise each interacting partner and the predictive impact on specific biological processes within the cell upon interaction with MP. Each interactor was further elaborated for its potential physiological role in the SACMV MP-cassava relationship. It

must be noted that complete verification and appreciation of each interaction would require further *in planta* and mutational studies.

2.5.9.1 INTERACTION OF MP WITH **CKL1** MAY ALTER ABA AND GA SIGNALLING

Casein kinase belongs to a class of casein kinases also identified as EL1-like (AEL) proteins and MUT9-like kinases (MLK) and is also homologous to casein kinase like (CKL) proteins (Chen et al., 2018; Zheng et al., 2018). It functions as an ATP and DNA binding serine/threonine kinase (Zheng et al., 2018; Gaudet et al., 2011) and is localised to the cytoplasm and nucleus (Gaudet et al., 2011). It is involved in processes of endocytosis, gibberellic acid (GA) mediated signalling pathway, histone phosphorylation, ubiquitation, negative regulation of abscisic acid-activated signalling pathway, peptidyl-serine phosphorylation, protein phosphorylation and regulation of protein phosphorylation (Zheng et al., 2018; Gaudet et al., 2011; Chen et al., 2018).

Casein kinases influences cell growth and development by regulating ABA and GA hormone signalling pathways (Zheng et al., 2018; Chen et al., 2018). Therefore, the interaction of SACMV MP with CKL1 potentially influences hormone dysregulation altering plant growth and development. Transcriptomic analysis in SACMV infected susceptible cassava identified an up regulation of gibberellin-regulated family proteins (Allie et al., 2014), to date, the role of gibberellins in geminivirus infection is not well elucidated. Albeit altered modulation of GA leads to a retarded growth and abnormal male reproductive growth (Hu et al., 2017) features noted in SACMV infection. In *Arabidopsis*, PYR/PYL, perceives and mediates ABA signalling, albeit when phosphorylated by casein kinase results in PYR/PYL ubiquitination and degradation thus suppressing the ABA signalling pathway (Chen et al., 2018). Prior studies have linked plant pathogen susceptibility to ABA (Alazem et al., 2014; Mohr & Cahill, 2007; Qiu et al., 2015). MP may interact with casein kinase and limit its ability to phosphorylate PYR/PYL and thus allow ABA signalling and enhance plant vulnerability to SACMV. The interaction of MP with casein kinase when linked to a study by Sun and colleagues (Sun et al., 2020) may elude to MP playing an indirect role in facilitating NSP transcription by enabling the ABA pathway. In mungbean yellow mosaic virus (MYMV) the NSP promoter (P_{BV1}) houses three ABA-responsive elements required for AC2-mediated transcriptional activation of BV1 (Sun et al., 2020). The perception of ABA leads to the activation of sucrose non-fermenting 1-related subfamily 2 protein kinases (SnRK2s) and the subsequent expression of ABA-responsive elements and ABA responses (Sun et al., 2020; Chen et al., 2018). SnRK2 acts antagonistically to serine/threonine kinase sucrose-nonfermenting1-related kinase (SnRK1), in that SnRK1 has been identified in contributing to host anti-viral activity (Shen et al., 2009; Shen et al., 2011; Testerink et al., 2004) whilst activation SnRK2 induces ABA signalling and disintegration of SnRK1 (Coello et al., 2012). ABA confers abiotic stress

tolerance (Cutler et al., 2010; Sah et al., 2016; Kim et al., 2010). Plants infected with RNA viruses, cucumber mosaic virus (CMV), tobacco mosaic virus (TMV), tobacco rattle virus (TRV) viruses and brome mosaic virus (BMV) respectively showed improved drought tolerance potentially linked to an accumulation of ABA (Xu et al., 2008). Whilst the mode of action within the study by Xu and colleagues speculated various reasons for this, it would thus be essential to investigate the role of SACMV MP in interference of ABA signalling pathway regulation by over expression of MP and CKL1 mutants to access the impact of overexpression on virus accumulation and proliferation.

2.5.9.2 INTERACTION OF **CKL1** WITH MP MAY BE ESSENTIAL TO MOVEMENT AND PHOSPHORYLATION OF MP

Phosphorylation is a key aspect related to the function of viral proteins. In *Arabidopsis* a member of the CK1 family of proteins, casein kinase 1-like 6 (CKL6) contains a C-terminal domain signal that localises to and binds cortical microtubules (Ben-Nissan et al., 2008). The interaction of MP to a member of the CK1 family of proteins and the understanding that the identified interactor (CKL1) is a homolog to CKL6, it may be relevant to propose that the interaction may be closely related to facilitate viral DNA movement. In *Nicotiana benthamiana* GFP tagged MP of tobacco mosaic virus (TMV) in protoplasts and leaf tissue localised from the endoplasmic reticulum to microtubular filaments (Heinlein et al., 1998) and the disruption of microtubules in protoplasts impeded the ability of MP to localise viral RNA to the cell periphery (Más & Beachy, 2000). Squash leaf curl virus (SqLCV) BL1 (MP) has also been shown to localise to the endoplasmic reticulum (Ward et al., 1997). The MP of abutilon mosaic virus (AbMV) and cleome leaf crumple virus (ClLCrV) showed a manipulation of the cytoskeletal system by a reorganisation of microtubules into bundles (Krapp et al., 2017).

2.5.9.3 INTERACTION OF **CKL1** WITH MP MAY HINDER HOST PROTEIN POST TRANSLATIONAL MODIFICATIONS AND PLAY A ROLE IN PHOSPHORYLATION OF MP

Phosphorylation of movement proteins (MPs) has been studied in AbMV (Kleinow et al., 2009; Bisaro, 1996), tomato mosaic virus (ToMV) (Matsushita et al., 2000), potato virus A (PVA) (Ivanov et al., 2003), CMV (Nemes et al., 2017) and TMV (Citovsky et al., 1993). Ninety-five percent of New World geminivirus MP sequences contained a conserved tyrosine phosphorylation site (Ho et al., 2014) plainly highlighting the significance of phosphorylation to MP functionality. The mobility of tobacco mosaic virus (TMV) through plasmodesmata is relegated by the phosphorylation of its P30 MP by a cell wall-associated protein kinase that interacts with and sequesters P30 in mature leaf tissue with secondary plasmodesmata ideal for cell to cell trafficking (Citovsky et al., 1993). The 2b protein of

cucumber mosaic virus has two conserved phosphorylation sites essential to movement between the nucleus and cytoplasm and suppressor activity (Nemes et al., 2017). Post translational phosphorylation has been employed as a host defence mechanism, in tomato yellow leaf curl China β-satellite (TYLCCNB) which plays a role similar to that of MPs as a symptom determinant was phosphorylated by a defence related (SnRK1) reducing viral load and intensity of symptoms (Shen et al., 2011; Shen et al., 2012). Phosphorylation of AL2 and Rep of CaLCuV and TGMV respectively, by SNRK1 delayed symptom progression and severity and interfered with viral DNA anchorage and consequently, reduced viral replication (Shen et al., 2018; W. Shen et al., 2014). The phosphorylation of AbMV-MP controls AbMV host specificity and viral DNA mobility (Kleinow et al., 2020). The role of casein kinases in PVA revealed that PVA-CP is phosphorylated at Thr-242 by casein kinase 2 (CK2), a modification essential to virion assembly and systemic infection and potentially regulated the RNA binding ability of CP for assembly and migration (Ivanov et al., 2003). Considering the multiplicity of the role noted in this single interaction (PVA-CP vs CK2) it opens the assumption that the MP of SACMV interacts with CKL1 to facilitate attachment, migration and release of viral DNA whilst enhancing plant cell susceptibility locally by the dysregulation of ABA. It could be ideal to identify specific phosphorylation sites within MP and determine the functional relevance of each phosphorylation site by mutational and *in planta* expression studies and identify target regions for MP repression within the susceptible host.

2.5.9.4 THE INTERACTION OF MP WITH **NAT** AND **ABR1** PROTEINS HAS IMPLICATIONS FOR JASMONATE BIOSYNTHESIS PATHWAY

Jasmonate biosynthesis is essential for the plants primary response to wound healing and increases the production of defence molecules against biotic stress in *Manihot* and other plant species (Allie et al., 2014; Jeevalatha et al., 2017; Gupta et al., 2021; Rosas-Díaz et al., 2016). Yeast two-hybrid screening resulted in the identification of two proteins involved in the regulation of the jasmonate biosynthesis pathway. These were an acyl-CoA N-acyltransferase (NAT) and ethylene-responsive transcription factor ABR1-like (ABR1).

The NAT protein contains a PHD-type zinc finger domain and a TPL binding domain. Multiple proteins with the aforementioned domains are collectively referred to as JAZ proteins and play a regulatory role in repressing the jasmonate biosynthesis pathway (Ali & Baek, 2020; Ruan et al., 2019). The bHLH transcription factors (MYC) regulate jasmonate mediated defence and response to abiotic stress and wound healing (Lorenzo et al., 2004; Cheng et al., 2011; Ali & Baek, 2020). JAZ proteins interact via the C-terminus with co-repressors TOPLESS (TPL), TPL-related proteins (TRPs) and, via the N- terminus with novel interactor of JAZ (NINJA) to inhibit MYC in the absence of wound induced jasmonyl-L-isoleucine (JA-Ile) (Pauwels et al., 2020). With regard to geminiviruses, wounding as a

result of whitefly inoculation and with respect to this study, needle inoculation has the capacity to activate JA-Ile, resulting in the ubiquitination of JAZ proteins and the release of MYC transcription factors for jasmonate mediated defence response factors (Pauwels et al., 2020). TME3 infected with SACMV exhibits tolerance that may be attributed to an early jasmonate signalling response (Pierce & Rey, 2013). The interaction with SACMV MP may favour both viral proliferation but activate jasmonate biosynthesis in tolerant and resistant plant hosts. Jasmonate responses are absent in susceptible host responses to geminiviruses (Allie et al., 2014; Pierce & Rey, 2013; Rosas-Díaz et al., 2016) and therefore the noted interactions of SACMV MP may be specific to tolerant and resistant species. Future investigation of this interaction in both susceptible and resistant species would be ideal.

ABR1 also referred to as ethylene response factor 111 (ERF111) belongs to the ERF/AP2 transcription factor family and was initially affiliated with repression of ABA responses (Pandey et al., 2005), but also regulates JA, auxin and ethylene signalling (Cai et al., 2014; Ye et al., 2020; Zhang et al., 2011). ERFs have a GCC rich region for pathogen resistance and DRE element regions for abiotic stress tolerance respectively (Müller & Munné-Bosch, 2015). Characteristic to this family of proteins, is a high responsiveness to methyl jasmonate (MeJA) and wounding, involvement in stress responses and wound repair mechanisms (Bäumler et al., 2019). Within this family of proteins ERFs function within the jasmonate signalling pathway to mediate, plant defence mechanisms against pathogens (Lorenzo et al., 2004) and mechanisms of JA and auxin in lateral root formation and regeneration (Cai et al., 2014; Ye et al., 2020). Therefore, a binding of MP would be inhibitory towards ABR1 roles crucial to host defence.

Noted interaction with the NAT and ABR1 hormone regulating proteins may result in variable susceptibility and defence related scenarios. (1) The interaction of MP with ABR1-like protein in the C terminal region may mimic either of the two co-repressors resulting in the repression of MYC and inhibition of JA responsive genes. (2) The interaction of MP with ABR1 may competitively inhibit MYC and co-repressors leading to the transcription of JA responsive genes. (3) The interaction of MP with ABR1 hinders its repressive interaction with MYC leading to JA responsive genes. (4) The interaction of MP with ABR1 and casein kinase positively regulates ABA signalling leading to susceptibility or a pathogen and abiotic stress response positively regulating the transcription of JA responsive genes. *In planta* evaluations and mutations within MP to identify specific regions of interactions that could be targeted to hinder MP interference with proteins of the JA biosynthetic pathway would be ideal. (5) The DNA binding transcription factor activity of ABR1 may also play a role in transcriptional activation of the viral genome. Geminiviruses have transcriptional activating ORFs TrAP/AC2/AL2 (Yang et al., 2007; Jeske, 2009) and the interaction with

ABR1 may implicate the manipulation of additional host factors to manipulate for this purpose.

2.5.9.5 THE INTERACTION OF MP WITH **PTEN2A** MAY PROMOTE MOVEMENT OF VIRAL DNA AND DISRUPT CELLULAR PROCESSES, AND IMMUNE SIGNALLING

PTEN2A is a lipid signalling molecule that influences plant defence, abiotic stress tolerance, growth and reproduction (Pribat et al., 2012; Wang & Chapman, 2013; Yunus et al., 2015; McLoughlin & Testerink, 2013). PTEN2A localises to the nucleus and within the cytoplasm is anchored to microtubular vesicles via phosphatidylinositol 3 phosphate (PI3P) (Naguib et al., 2015). An electrostatic cell membrane binding mechanism facilitates its translocation (Chaozu et al., 2018). In *Arabidopsis*, PTEN dephosphorylates the 3′ phosphate group of phosphatidylinositol phosphates (PI(3,5)P, PI3P, PI(3,4)P2) and has a strong binding affinity for phosphatidic acid (Pribat et al., 2012). Phosphatidic acid is a crucial lipid signalling molecule crucial to water stress signalling (McLoughlin & Testerink, 2013), reproductive development (Yunus et al., 2015) and ABA signal transduction in germinating tissues (Katagiri et al., 2005). In plant-virus relationships, it has been found to promote red clover necrotic mosaic virus (RCNMV) viral RNA replication (Hyodo et al., 2015). Thus, the interaction of MP with PTEN2A potentially facilitates microtubular migration of viral DNA, disrupts cell signalling and regulation of cellular processes, and hinders immune signalling.

2.5.9.6 THE INTERACTION OF MP WITH **PTEN2A** MAY PLAY A ROLE IN REGULATION OF MP FUNCTION

The identification of CKL1 as a potential facilitator of MP phosphorylation was interesting considering that post-translational phosphorylation of virus proteins is essential for functionality (Ivanov et al., 2003; Nemes et al., 2017; Citovsky et al., 1993). PTEN2A plays a key role in dephosphorylation (Pribat et al., 2012). It would be worth considering an interplay of host proteins, CKL1 and PTEN2A inclusive, in the phosphorylation and dephosphorylation of MP. A concept that has also been postulated in a recent AbMV-MP study (Kleinow et al., 2020). Triple phosphorylation mutants of AbMV-MP resulted in one-third more severe symptoms (Kleinow et al., 2020). It may be postulated that PTEN2A may play a key role in site specific dephosphorylation of MP that would either negate the function of MP to translocate viral DNA and induce symptoms or dephosphorylation would result in the amelioration of pathogenicity conferred by MP. The latter enabling host longevity and favouring virus persistence within a living albeit infected host.

The interaction of PTEN2A with phosphatidic acid which further interacts with HSP90 and SnRK2 serine/threonine protein kinase (Testerink et al., 2004; Pribat et al., 2012) and prior

evidence of geminivirus association with heat shock proteins HSP90 (Moshe et al., 2016), and HSC70 (Gorovits et al., 2013), βC1 satellite interaction with kinase SnRK1 (Shen et al., 2011) and additionally the potential upregulation of ABA by MP interaction with casein kinase and PTEN2A may provide a preliminary backbone for an investigation into SACMV MP host interactions and mechanisms related to plant susceptibility and/or tolerance to SACMV.

2.5.9.7 INTERACTION OF MP WITH **LSH** MAY FAVOUR VIRAL REPLICATION BY ALLOWING CELL DIFFERENTIATION

LSH is a DNA binding developmental regulator located in the nucleus and mitochondrion (Gaudet et al., 2011) that plays a role in the suppression of organ differentiation in the boundary cells of meristematic organs (Takeda et al., 2011) and mediates phototropism (Gaudet et al., 2011). Geminivirus TGMV replication occurs in differentiated cells (Nagar et al., 2002) and therefore, the hinderance of LSH would favour geminivirus replication. It is possible that, the interaction with MP may hinder its ability to mediate cell differentiation and therefore favour viral replication (Nagar et al., 2002).

2.5.9.8 THE INTERACTION OF MP WITH **LSH** MAY CONTRIBUTE TO MICROTUBULAR ABNORMALITIES, TRANSLOCATION OF VIRAL DNA AND INDUCE VIRAL REPLICATION

No studies were identified that associated LSH to directly targeting the nuclear membrane or associating with the cytoskeletal system. LSH has, however, been shown by Y2H screening to interact with AT3G27960 (BioGRID:7240), a kinesin that contributes to lateral stability of cortical microtubules (Ganguly et al., 2020). Phenotypic anomalies that arise from defects in cortical microtubules are in shoot and root misdirection and defective cell walls (Mathur & Hülskamp, 2002) and in *Drosophila* a disruption of cortical microtubule stability resulted in elevated mitosis (Bossing et al., 2012) and inhibited normal growth towards light. The binding of MP with LSH may interfere with LSH binding to AT3G27960 leading to an increase in mitosis and a weakened cell wall integrity, both factors that would favour viral replication and proliferation (Kong & Hanley-Bowdoin, 2002). LSH protein has been shown to interact with the Rep SACMV by Y2H screening (Rey lab, unpublished data) and this albeit requiring further investigation, highlights a potentially legitimate dual role of SACMV proteins in facilitating viral DNA replication via LSH. Association of MP with LSH may also contribute association with microtubules which is relevant to geminiviral MP and Rep movements (Fondong, 2013).

2.5.9.9 THE INTERACTION OF MP WITH **SIP2** MAY HINDER CELL SIGNALLING AND REGULATED GROWTH AND DEVELOPMENT

SIP2 is an alpha-galactosidase that catabolises raffinose to sucrose and galactose and may be involved in the release of raffinose in sink tissues (Peters et al., 2010). Orthologues of this enzyme have been shown to transport carbohydrate sugars across the phloem (Sprenger & Keller, 2000). With regards to transportation, SIP2 localises to both the cytoplasm and plasmodesma. Plasmodesma are channels within the cell wall that enable cell to cell communication by passage of molecules (Haywood et al., 2002). The interaction of MP with SIP2 therefore further implicates MP in the systemic localisation of geminiviruses across the cell walls and MP would thus interact with proteins that are localised to the cell wall or proteins that may enable passage across cell walls. Hence an interaction with SIP2 carries relevance with regards to the function of MP in movement. In *Arabidopsis*, SIP2 is induced by heat shock transcription factor A2 (HSFA2) (Nishizawa-Yokoi et al., 2008), which is elevated in response to temperature and salt tolerance (Ogawa et al., 2007). In TYLCV infection of tomato, a repression of heat shock proteins, HSFA2 inclusive was inversely proportional to viral load and plant susceptibility (Ghandi et al., 2016). The binding of HSFA2 with TYLCV CP, V2, C1, C2, C3 and C4 proteins whilst the interaction with CP was further confirmed to occur within the nuclear fractions thus indicating the suppression of HSFA2 activity on transcription (Ghandi et al., 2016). HSFA2 is involved in the coordinated light responses during germination and seedling development (Nelson et al., 2010) and interacts with ASYMMETRIC LEAVES1 (AS1) in co-ordinating leaf development (Theodoris et al., 2003). The interaction of MP with SIP2 provides evidence into the targeted inhibition of stress and developmental proteins induced by HSFA2 resulting in a compromised stress response, developmental anomalies, and leaf malformations. The potential role of SIP2 in transfer of raffinose to sink organs, shows MP may interfere with allocation of raffinose. Raffinose, galactinol and myo-inositol are interconnected. The accumulation of galactinol is directly linked to the availability of precursor molecules myo-inositol and galactose (Karner et al., 2004). Whilst myo-inositol and sucrose concentrations affect the accumulation of raffinose (Karner et al., 2004). Infection of *Jatropha* with JLCuGV was characterised by a down regulation of galactinol and myo-inositol (More et al., 2021). Therefore, studies of the host metabolome during SACMV MP host expression in a SIP2 mutational background could aid in providing a conclusive evaluation of its impact on the accumulation of inositol and its derivatives.

2.5.9.10 THE INTERACTION OF MP WITH **NUP98A** MAY FACILITATE VIRAL DNA MOVEMENT

NUP98A is a mobile nucleoporin colocalised to the nuclear pore complex and within the nucleus and its translocating ability is dependent on continuous transcription by RNA

polymerase I/II (Griffis et al., 2002). Its biological processes involve export of mRNA and RNA across the nuclear membrane into the cytoplasm, import of proteins into the nucleus and regulation of shade avoidance (Gaudet et al., 2011; Gallemí et al., 2016; Powers et al., 1997). Various studies have highlighted the manipulation of nuclear pore complex proteins by mammalian viruses to promote viral replication and translocation of viral genomes and proteins (De Jesús-González et al., 2021). In animal studies,NUP98 has been implicated in the transport of virulence factors in Rift Valley fever virus infection (Lau & Weber, 2020). Furthermore, in HIV infection, NUP98 and the nuclear pore complex (NPC) have been implicated in the trafficking of HIV-1 cDNA across the nuclear membrane (Ebina et al., 2004). Therefore, the ability of NUP98 to migrate DNA and proteins proposes the feasibility of interaction of MP facilitating movement of MP anchored SACMV DNA and MP independently.

2.5.9.11 THE INTERACTION OF MP WITH **NUP98A** MAY HINDER HOST GENE EXPRESSION AND HINDER THE MITIGATION OF AN IMMUNE RESPONSE

NUP98 has also been associated with gene activation (Zhang et al., 2020; Sump & Brickner, 2017; Cruz et al., 2018; Franks et al., 2017). In terms of antagonising host immunity, influenza A virus and vesicular stomatitis virus matrix protein (VSV M) target NUP98 transcription and activity in an effort to block mRNA export as a means to cripple an effective immune response (Satterly et al., 2007; Von Kobbe et al., 2000). Therefore, an interference with NUP98 by interaction with MP potentially provides a platform for virus genome translocation and ablates the recruitment of effective immune signalling by the plant for defence.

2.5.9.1 THE INTERACTION OF MP WITH **SDG16** INHIBITS HOST GENE EXPRESSION

SDG16 is crucial to the post translational methylation of H3 and H4 histones of thousands of genes linked to vegetative and reproductive development (Chen et al., 2017). Methylation of H3K4 is associated with gene activation and consequent transcription and translation (Black et al., 2012). Rep protein of ChiLCV has been implicated in the interaction with host proteins ubiquitin-conjugating enzyme2 (UBC2) and histone monoubiquitination1 (HUB1) for the monoubiquitination and trimethylation of H3K4 of viral mini-chromosomes (Kushwaha et al., 2017). Geminivirus, Indian cassava mosaic virus (ICMV), targeted silencing of H3K9 histone methyltransferase KRYPTONITE (NbKYP), (a methyltransferase associated with gene repression/chromatin silencing (Black et al., 2012)) rendered increased hypersensitivity to ICMV in *Nicotiana* sp. (Sun et al., 2015). In geminiviruses, the AC2/TrAP/L2 of ICMV/TGMV/BCTV interferes with histone methylation of host DNA and cytosine methylation of the invading virus by host factors (Castillo-González et al., 2015; Sun et al., 2015; Raja et al., 2008). AC2 of ICMV transcriptionally activates related to

abscisic acid insensitive 3 (ABI3) and VP1 (related to ABI3/VP1 2 (RAV2)), a transcriptional repressor of H3K9 (Sun et al., 2015) whilst the TrAP of TGMV, interacted with H3K9, inhibiting its activity (Castillo-González et al., 2015). Individual *Arabidopsis* mutants (of cytosine methyltransferases, H3K9 methyltransferase, methylation pathway, or methyl cycle enzymes) exhibited an increased susceptibility to CaLCuV and BCTV respectively (Raja et al., 2008). Methylation and RNA silencing pathway enzymes AGO1 and DCL1 and, DRM, CMT3 and AGO4 have been linked to geminivirus infected host recovery in SACMV and BCTV infection respectively (Bizabani et al., 2021; Raja et al., 2008) whilst the latter, domains rearranged methylase (DRM), chromomethylase3 (CMT3) and argonaute4 (AGO4) were shown to be directly involved in viral genome cytosine methylation by bisulphite sequencing (Raja et al., 2008). The tomato yellow leaf curl China virus betasatellite (TYLCCNB) interacts with and inhibits S-adenosyl homocysteine hydrolase (SAHH), a methyl cycle enzyme limiting its ability to methylate host and viral DNA (Yang et al., 2011). The AC2 of mungbean yellow mosaic Indian virus (MYMIV) by immunoprecipitation showed that AC2, binds to both AGO1 and RDR6 of inhibiting host TGS (Kumar et al., 2015). These studies only briefly highlight the targeting of host methylation factors by geminivirus proteins. It may be inferred that the interaction of MP with H3K4 would notably allow for the inhibition of the methyltransferase role in gene activation and the relegation of transcripts essential to immune responses

.The interaction of SACMV MP with host H3K4 may also serve as a viral mechanism to activate viral DNA transcription. The nature of this interaction was identified in chilli leaf curl virus (ChiLCV) with respect to the Rep protein (Kushwaha et al., 2017). Viral mini-chromosomes associated with either H3K4 or H3K9 and exhibited lower and higher viral DNA methylation respectively (Ceniceros-Ojeda et al., 2016). The latter was characteristic to the recovered phenotype of tolerant plants (Ceniceros-Ojeda et al., 2016). The down regulation of histone acetyl transferase an activator of transcription in the recovery phenotype of cassava TME3 (Allie et al., 2014), would limit viral transcription and further adds credence to viral utilisation of host transcription activators for viral transcription and consequent viral protein translation. The mechanism in ChiLCV was later elaborated to involve the recruitment of ubiquitination enzymes by Rep protein for monoubiquitination of Histone 2B and the consequent trimethylation of H3K4 on the ChiLCV mini-chromosomes for transcription of the viral genes in *Nicotiana benthamiana* (Kushwaha et al., 2017). Of further inference to the interference of methylation and gene expression is the prior noted interaction of MP with NUP98 of this objective. NUP98 in mammalian cancer studies has been implicated in gene activation and by inducing trimethylation of H3K4 methyltransferase thus resulting in increased gene expression (Zhang et al., 2020; Sump & Brickner, 2017; Cruz et al., 2018; Franks et al., 2017). The interaction of MP with NUP98

and H3K4 collectively alludes to a mechanism by MP in host gene suppression for further investigation.

2.5.9.13 THE INTERACTION OF MP WITH **PUP1** MAY SERVE AS AN ANCHOR FOR MP MEDIATED VIRAL DNA PROLIFERATION AND DYSREGULATE THE ACCUMULATION OF CYTOKININS

PUP1 is a transmembrane protein that functions in proton coupled transport of purine nucleobases and their derivatives (adenine, trans-zeatin, pyridoxine (vitamin B6), caffeine, adenosine, cytokinin) via the xylem and phloem to and from shoot tissues (Bürkle et al., 2003; Szydlowski et al., 2013; Gillissen et al., 2000). The interaction of MP with SACMV MP may inhibit PUP1 activity resulting in an accumulation of bioactive cytokinins. Purine permease 14 (PUP14) a homologue to PUP1 in *Arabidopsis*, depletes bioactive cytokinin resulting in a suppression of the cytokinin response (Zürcher et al., 2016). In TGMV and SCTV infection, exogenous application of cytokinin favoured viral replication (Baliji et al., 2010). An upregulation of cytokinin responsive genes is induced by pathogenicity proteins (AL2/C2) of begomo- and curto-viruses via interaction with adenosine kinase (ADK), impeding its ability to phosphorylate cytokinin and thus increase bioavailable cytokinin (Baliji et al., 2010). PUP1 was recently identified as an NSP and MP interacting protein in geminivirus cabbage leaf curl virus (CabLCV) (Gouveia-Mageste et al., 2020).

2.5.9.14 THE INTERACTION OF MP WITH **PUP1** MAY DYSREGULATE THE ACCUMULATION OF VITAMIN B6

Vitamin B6 (pyroxidine) albeit not greatly studied in geminiviruses, is well identified for amelioration of oxidative stress (Havaux et al., 2009), is an enzyme cofactor (Mooney et al., 2009) and in a non-virus study conferred ROS quenching during *Rhizoctonia* disease of potato (Samsatly et al., 2020). Therefore, vitamin B6 in amelioration of oxidative stress, maintains integrity of DNA and limits autophagy (Filomeni et al., 2015; Haxim et al., 2017). Cancer studies have, however, shown that autophagy sustains cell energy for the recruitment and action of DNA repair mechanisms (Filomeni et al., 2015; Haxim et al., 2017). Most recently, autophagy in plant virus interactions was reviewed (Huang et al., 2020). Geminivirus, C1 of tomato leaf curl Yunnan virus (TLCYnV) interacts directly with the core autophagy-related protein ATG8h translocating C1 to autophagosomes for degradation by nuclear autophagy, limiting virus infection (Li et al., 2020). Geminivirus, TYLCCNV and its betasatellite (TYLCCNB) up-regulate a PTGS suppressor, calmodulin-like protein which interacts with suppressor of gene silencing 3 (SGS3) inducing the degradation of SGS3 via the autophagy pathway thereby eliminating SGS3 for virus targeted RNA silencing (Li et al., 2017). In cotton leaf curl Multan virus (CLCuMuB), its betasatellite (CLCuMuB) was also found to favour autophagy by interaction with

glyceraldehyde-3-phosphate dehydrogenase (GAPC), a negative regulator of autophagy (Ismayil et al., 2020).

PUP1 interaction with MP may play a key role in the anchorage of MP for viral DNA long distance movement. It may also impede the translocation of cytokinin from infected cells and therefore, promote viral replication by induction of cytokinin responsive factors. Albeit the interaction with MP may also hinder its ability to transport purine nucleobases resulting in a down regulation of nucleobase transport that may have an inhibitory effect on virus proliferation. A suggestion would be that the limitation in the translocation of purine nucleobases such as vitamin B6, may create a pool of B6 locally available to the infected cell for ROS amelioration and thereby conserve viral DNA and inhibit autophagy. Studies that have shown autophagy to create a microenvironment suitable for DNA repair and viral proliferation suggest a need for specific *in planta* investigations into monitoring vitamin B6 within infected cells and its impact during SACMV infection.

2.5.9.15 THE INTERACTION OF MP WITH **HSK** MAY INTERFERE WITH AMINO ACID BIOSYNTHESIS PATHWAYS AND BIOLOGICAL PROCESSES THAT REQUIRE METHYLATION

HSK is a chloroplastic, ATP binding kinase that catalyses the phosphorylation of L-homoserine to L-homoserine phosphate as an intermediate step to threonine and methionine biosynthesis (Lee et al., 2005; Lee & Leustek, 1999) and further functions as a promoter of the circadian responsive autonomous flowering pathway (Zhu et al., 2005). The interaction with MP may much like CKL1 and PTEN2A may both play an essential role in post translational modifications of MP within the chloroplasts. However, HSK is specific to homoserine.

Homoserine kinase functions in methionine and threonine biosynthesis, hence an interaction with MP that interferes with these pathways would compromise protein synthesis negatively impacting defence, developmental and regulatory processes. To be considered is S-adenosyl methionine a product of ATP and L-methionine essential as a methyl donor resulting in methylation of DNA, various RNAs, proteins and phospholipids (Cantoni, 1952). Methylated phospholipids such as phosphatidylcholine is hydrolysed to intermediate choline and phosphatidic acid which are critical to defence responses that involve lipid signalling for rapid response mechanisms, hormone signalling, vesicular trafficking and membrane-cytoskeleton interactions (Zhao, 2015; Xing et al., 2021; Abd-El-Haliem & Joosten, 2017). Protein methylation occurs largely on arginine and lysine residues (Lee et al., 2005). Arginine methyltransferases are elevated in nitric oxide (NO) stress signalling to relegate mRNA splicing of stress-related genes (Hu et al., 2017). Histone lysine methyltransferases (SDG8, SDG25) methylate histones H3K4 and H3K36,

that regulate systemic acquired resistance and are critical to accumulation of lipids and cuticle integrity (Lee et al., 2016). Lysine methylation of non-histone proteins modulates response to cadmium stress (Serre et al., 2020). In the chloroplast, methionine and threonine are catalysed to 2-ketobutyrate, a precursor for isoleucine biosynthesis and thus methionine and threonine biosynthesis influence the composition of branched chain amino acids (Singh & Shaner, 1995). The significance of the isoleucine pathways in wound and pathogen response may be appreciated by the predominant and most active form of jasmonate being conjugated to isoleucine (12-hydroxy-jasmonoyl-L-isoleucine) (Poudel et al., 2019). The interaction of MP with HSK, therefore, may interfere with amino acid biosynthesis pathways that are crucial to modulation of plant defence responses.

2.5.9.15 THE INTERACTION OF MP WITH **KRS1** MAY INTERFERE WITH THE TRANSLATION OF HOST PROTEINS

KRS1 is a nucleotide binding ligase involved mainly in the ATP dependent transfer of a cognate and non-cognate transfer of amino acids for translation of mRNA for protein synthesis (Wu et al., 2007; Gaudet et al., 2011). The non-cognate nature of other tRNA ligases is a common occurrence (Plant et al., 2007; Wu et al., 2007). An impediment in the function of KRS1 may attribute to the occurrence of mutations in host proteins arising from SACMV infection. Geminivirus SACMV induces mutations in the ubiquitin E3 ligase gene, MeE3L (Chatukuta & Rey, 2020). Non-standard Watson-Crick base pairs arising from non-cognate and near-cognate tRNA prevail if there are sufficient hydrogen bonds for stability resulting in ‘’geometrical mimicry’’ that allows non-standard base pairing to occur (Roy et al., 2015; Westhof et al., 2014). Therefore, an interaction of MP with this ligase may slow down protein synthesis or hinder accurate protein synthesis due to a lowered availability of tRNAs for cognate transfer and result in error prone protein synthesis. The Gag protein of human immunodeficiency virus type 1 (HIV-1) interacts via its nucleocapsid domain with tRNA-ligase facilitating viral transcription (Javanbakht et al., 2003) and complexes with viral tRNA(Lys3) which contains anti-codon-like element (TLE) to facilitates viral protein translation (Jones et al., 2013). At least nine plant RNA virus genera have been identified with Transfer RNA-like structures (TLSs) to mimic at their 3’ ends binding valine, histidine and tyrosine (Dreher, 2009). DNA viruses have also been shown to encode multiple tRNAs (Albers & Czech, 2016). Geminiviruses are reliant on host machinery for their replication and translation (Hanley-Bowdoin et al., 2013; Ramesh et al., 2017) therefore a mechanism that involves MP to manipulate host tRNA ligase is plausible. The prospect of MP playing a mimicry role in interaction with the ligase to support viral translation would require further investigation. The mimicry may allow for a translational complex to form with the ribosome and limit ribosomal function for host proteins and mediate translation of viral proteins.

2.5.9.17 THE INTERACTION OF MP WITH **PsaD** MAY ALTER CHLOROPLAST STRUCTURE AND FUNCTION

PsaD is a significant photosynthetic protein that forms complexes with ferredoxin a key electron acceptor in photosynthesis and as a key component in the biosynthesis of pigments and assimilation of sulphur and nitrogen (Hanke & Mulo, 2013). PsaD is both a protein and mRNA binding protein (Dutta et al., 2014; Bach-Pages et al., 2020). The latter implicates RNA in regulating PsaD function and organisation with other proteins (Bach-Pages et al., 2020). The gene knockout of the PsAD has been shown to be lethal in *Arabidopsis thaliana* (Haldrup et al., 2003; Ihnatowicz et al., 2004). In plant virus studies, virus localisation to chloroplasts and interaction with chloroplastic proteins is not uncommon (Zhao et al., 2016). Recently in cassava, the cassava common mosaic virus (CsCMV) is a potexvirus that leads to aberrations in chloroplast shape and reduced function, efficiency and an increase of the sucrose/starch ratio indicating an alteration in carbon storage and assimilation (Zanini et al., 2021). PsaD complexes with ferredoxin and therefore provides insight into an effect arising from the interference of MP may be from a study of PVX (Yang et al., 2020). Virus protein PVX-p25 protein interacts with chloroplast protein ferredoxin 1 (FD1) correlating to reduced callose deposition at plasmodesmata arising from reductions in abscisic acid (ABA) and salicylic acid (SA) (Yang et al., 2020). In geminivirus studies, the MP of AbMV interaction with chloroplastic HSC-70 induces stromules ideal for systemic viral trafficking (Krenz et al., 2010). Geminivirus radish leaf curl betasatellite (RaLCB) βC1 targets the chloroplast and associates with the photosystem II reaction centre impairing chloroplast ultrastructure and function (Bhattacharyya et al., 2015). Further studies elucidated that photosystem II oxygen-evolving enhancer protein 2 (PsbP) involved in plant defence binds viral DNA negating viral accumulation and migration, however, the interaction of PsbP with βC1 interferes with its ability to counter virus invasion (Gnanasekaran et al., 2019). The interaction with MP may therefore directly contribute to phenotypic deviations (stunting, size reduction and leaf abnormalities) manifested during SACMV infection because of alterations in carbon metabolism.

2.6 SUMMARY

Figure 2.10 provides a schematic summary of the proposed model that highlights the consequences of MP interacting with the identified host proteins that may favour viral translocation and replication or result in an antiviral response.

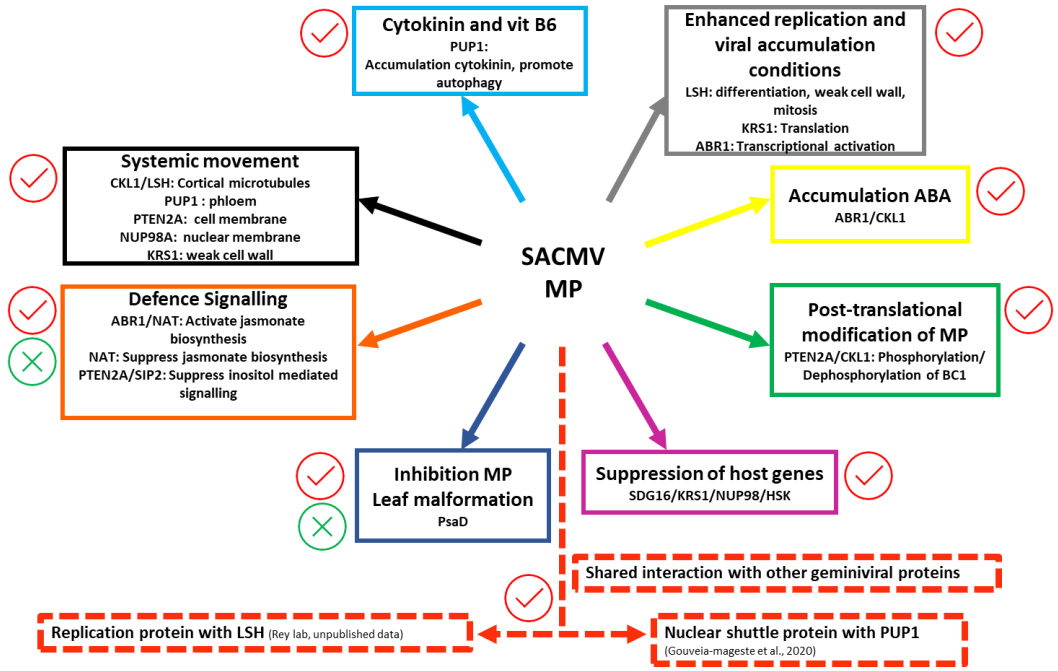


Figure 2.10 The proposed model of the interaction of MP with host proteins (coloured arrows and polygons) resulting in alteration of the host proteome and metabolome that results in either a viral advantage (red tick) or an anti-viral host response (green cross). The interaction of MP with cassava TME3 host proteins revealed that MP interaction is multifaceted and predominantly favours viral replication, proliferation, and translation. The post translational modification of MP by phosphorylation and dephosphorylation via interaction with PTEN2A and CKL1 confers functionality to MP in either the nucleus or cytoplasm (green). MP interacts with proteins for translocation: within the nuclear environment via cortical microtubules (CKL1/LSH), traverses the nuclear membrane (NUP98), cell to cell trafficking and within the cytoplasm (PTEN2A), phloem trafficking (PUP1) and contributes to cell wall weakening (LSH/KRS1) (black). Interaction of MP with KRS1 and LSH enhances conditions for viral replication and viral protein translation by inhibition of LSH promoting cell differentiation and recruitment of KRS1 for ribosomal binding and translation of viral protein (grey). A hinderance of host gene expression essential to eliciting an effective defence response by MP involves binding of NUP98 and SDG16/H3K4 to limit gene activation and the recruitment of KRS1 retarding host protein translation (purple). The accumulation of hormones ABA and cytokinin in favour of viral replication by inhibitory interaction of MP with ABA suppressors CKL1 and ABR1 (yellow) and PUP1 for accumulation of bioavailable cytokinin (sky blue). Interaction of MP with PUP1 further limits translocation of bioavailable vitamin B6 resulting in amelioration of ROS within cells and the promotion of autophagy by inhibitory interaction with MP (sky blue). Interaction of MP with NAT competitively hinders the suppression MYC2 activating jasmonate biosynthesis and leading to an anti-viral host response (orange). Interaction of MP with NAT would however suppress jasmonate signalling and release ABA responses favouring viral replication (orange). Interaction of PTEN2A with MP would suppress binding to and phosphorylation of defence signalling derivatives of inositol whilst interaction with SIP2 would hinder the translocation of raffinose and other sugars essential to defence signalling (orange). Interaction of MP with PsaD inhibits photosynthetic efficiency and amplifies viral symptoms and should PsaD play an antiviral role within chloroplasts would bind to MP haltering viral replication and accumulation (navy blue). PUP1 and LSH from previous work have been shown to interact with Rep and NSP geminiviral proteins therefore indicative of co-operative interaction of MP with other geminiviral proteins (red dashes).

Chapter 3

THE IDENTIFICATION OF HOST PROTEINS INTERACTING WITH MP MOVEMENT PROTEIN OF GEMINIVIRUS SACMV IN ORDER TO UNDERSTAND THE SYSTEMIC INFECTION OF *MANIHOT ESCULENTA* CRANTZ.

3.1 INTRODUCTION

In order to facilitate biochemical changes in cellular pathways, proteins interact as stable mono- or hetero-dimers or oligomers that function as a single macromolecular complex unit (Jones & Thornton, 1996). The assembly of ribosomes and spliceosomes requires over 300 and 100 proteins respectively (Staley & Woolford, 2009). Most regulatory proteins such as transcription factors and those involved in post-translational modifications such as kinases and methylases tend to interact transiently with other proteins with variable binding strengths (Alvarez et al., 2020). Plant viruses and viruses in general have been shown to induce complex cellular responses within the host (Willets & Wong, 1980; Wang & Maule, 1995; Thaker et al., 2019). Viral proteins act as avirulent or virulent factors that interact with host proteins to establish infection. *In vivo* techniques such as Y2H (Fields & Song, 1989) and bimolecular immunofluorescence (BiFC) (Ghosh et al., 2000) have been used in the identification of binary protein interactions. However, to fully identify the complexity of *in planta* interactions, often multiple binary interactions are required. Therefore, due to the complexity of protein interactions and virus alteration of cellular processes, it is important to identify complex interactions via Co-IP. Multiple Y2H and BiFC studies, for example, showed that potyvirus proteins form a highly connected interactome arising from direct virus protein interactions with 2-10 host proteins (Bosque et al., 2014).

Protein microarrays are a favourable tool to capture protein-protein interactions, albeit complex. More recently, a single study utilising a complex protein microarray generated an interactome with MP and NSP of cabbage leaf curl virus (CabLCV) comprising 343 host proteins represented by 20 hub networks (Gouveia-Mageste et al., 2020). They demonstrated *in planta* that viral NSP, a facilitator of the nucleocytoplasmic trafficking of viral (v)DNA, interacts with a new endosomal vesicle-localized plant-specific syntaxin-6 protein. Immunoprecipitation assays present a simple but robust technique to reveal the complex interactome networks of viral proteins. Virus-host interaction studies have shown that *in vitro* isolation and identification of primary and secondary virus-host protein interactions is best achieved by the technique of Co-IP (DeBlasio et al., 2017; Hoyle, 2018). In a study by Wang et al. (2017), a geminivirus-host protein interactome was inferred using affinity purification and mass spectrometry. Plant interactors with each of the encoded proteins of tomato leaf curl virus (TYLCV) were identified in *Nicotiana benthamiana* and interaction networks inferred (Wang et al., 2017). The interactome network of potato leafroll

virus (PLRV) (genus, *Polerovirus*) was elucidated by Co-IP from coat protein (CP) and readthrough protein (RTP) binding proteins that spanned a total of 746 annotated host proteins (Deblasio et al., 2015).

Co-immunoprecipitation involves the anchorage of an antigen-specific antibody to a solid substrate (e.g. agarose, magnetic beads, microtiter plate) within a protein lysate or extract to enable the antibody binding of the antigen together with its physiological protein binding partners (DeBlasio et al., 2017; Iqbal et al., 1960; Deblasio et al., 2015). Four mandatory aspects to Co-IP are;1) a protein (or extract) that contains non-denatured proteins that maintain their physiological conformation and are attached to their physiological binding partners (interaction complex), 2) an antibody specific to a target antigen within the protein lysate (or extract) mixture, 3) multiple negative controls to eliminate proteins that bind non-specifically to the antibody substrate (false positives) and 4) an accurate and high throughput identification of binding proteins by a technique such as mass spectrometry (Hoyle, 2018; Deblasio et al., 2015).

False positives in Co-IP can account for 45-94 % of identified proteins (Hoyle, 2018; Deblasio et al., 2015; Huang & Kim, 2013). In addition to multiple negative controls, different strategies have been adopted to reduce the occurrence of false positives/ contaminating proteins. Antigens of low abundance tend to result in increased false positives, therefore, the use of a two-step chemical crosslinking strategy that utilises disuccinimidyl suberate and dithiobis[succinimidylpropionate] may be employed for accurate identification of true interactors (Huang & Kim, 2013). Sciuto and colleagues, further developed a two-step Co-IP (TIP) method that consists of two serial Co-IPs using a biotinylated antibody (Ab) and an anti-biotin resin that enables the gentle elution of the protein complex prior to re-precipitation with beads (Sciuto et al., 2018). False positives may also be reduced by pre-clearing of lysate, a process that involves the prior incubation of extract with solid substrate prior to anchorage with antibody, to reduce the percentage of proteins with a high binding affinity for the solid substrate (Kushwaha et al., 2017; Lario et al., 2013).

Co-immunoprecipitation is advantageous in that it provides a simplistic and cost effective mechanism that does not require complex expression cloning techniques associated with Y2H and BiFC assays (Ghosh et al., 2000; Fields & Song, 1989). The resulting protein extract from lysed cells is generally representative of the host proteome, albeit extraction techniques can be optimised for identification and characterisation of a target protein and its interactors relative to specific subcellular fractions (Zheng et al., 2016; Zhu et al., 2017). It also allows for a high throughput manner in which to identify more accurately the immediate interactome network of the target protein as opposed to identifying a single binary interaction and extrapolating via computational resources (Wang et al., 2017;

Deblasio et al., 2015). Disadvantages of Co-IP are the requirements for large volumes of sample in the evaluation of low abundance proteins (Zhu et al., 2017) and the need to optimise protein extraction protocols for target protein enrichment (Deblasio et al., 2015). Albeit, a large number of false positives, exclusion of contaminating proteins may be avoided by the use of multiple negative controls during experimentation (Hoyle, 2018; Deblasio et al., 2015; Huang & Kim, 2013).

In efforts to understand the role of SACMV in cassava, transcriptomic studies (Allie et al., 2014; Pierce & Rey, 2013; Allie et al, 2013) and efforts to characterise SACMV MP (Nankoo, 2018) and SACMV- replication-associated protein (Rep) (Ayres, 2018) have been carried out. Information on the structural characterisation of SACMV MP have advanced, but an understanding of the functional mechanisms of SACMV MP is still limited to extrapolation from other geminivirus species studies in non-cassava hosts.

3.2 AIM AND OBJECTIVES

The aim of this study was to co-immunoprecipitate native and expressed SACMV MP with cassava TME3 host proteins and therefore identify SACMV MP affiliated host proteins by mass spectrometry and postulate on their relevance to the functional mechanisms of SACMV MP during SACMV infection of cassava TME3.

The main objectives were:

1. To evaluate cassava TME3 protein lysate (extract) and antibody for suitability in Co-IP experiments.
2. To evaluate proteins identified from Co-IP experiments for physiological applicability to SACMV MP functional roles during SACMV infection of cassava TME3 .

3.3 METHODS

In brief, the experimental methods were carried out as illustrated below:

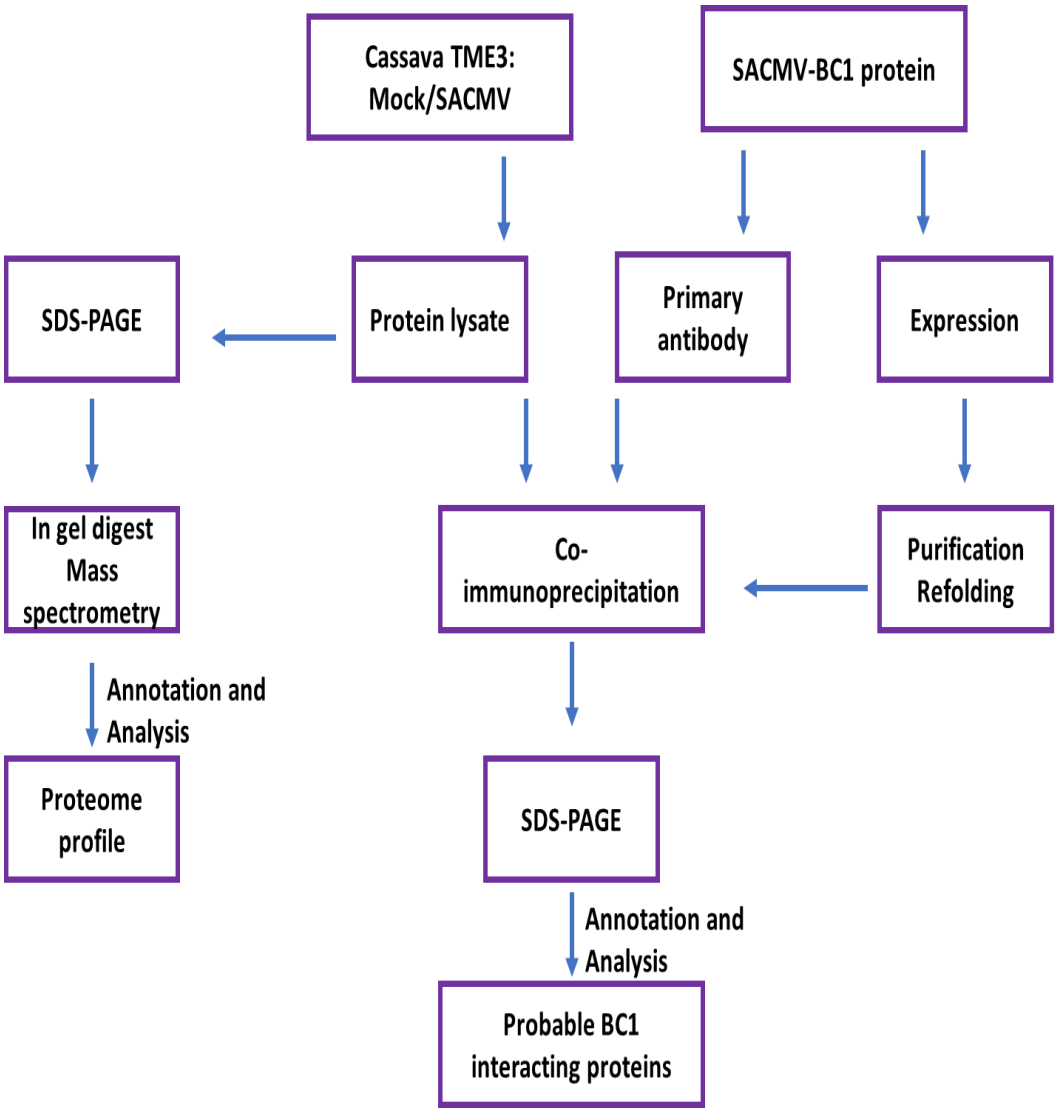


Figure 3.1 The Co-IP workflow to identify probable MP/BC1:cassava interacting proteins from cassava landrace TME3 infected with infectious clones of SACMV. Plants were mock-inoculated with untransformed *Agrobacterium* (negative control). Leaf tissue was assessed for a preliminary proteomic profile (red arrows). The MP amino acid sequence was used to design polyclonal antibodies specific to MP (antigen) (green arrows). MP was expressed in *E. coli* and purified and refolded (orange arrows). Co-IP (blue arrows) was carried out with protein extract, primary antibody, and MP by use of Protein A magnetic beads. Eluted fractions from Co-IP were run on an SDS-PAGE gel and analysed by in-gel digest followed by mass spectrometry. Peptides identified from mass spectrometry were annotated. Annotated proteins were analysed by simplistic and computational methods to identify host interactors of MP.

3.3.1 MICRO-PROPAGATION OF CASSAVA TME3

This was done according to an established protocol with modifications (Allie et al., 2014). Sterile nodal cultures of cassava TME3 landrace were micro-propagated on Murashige and Skoog (MS) medium (Murashige & Skoog, 1962) supplemented with 2 % sucrose, 0.002 mM $CuSO_4$, 0.78 % plant tissue culture agar, pH 5.8, for 3 weeks (28 °C, 16h photoperiod, 8000–10000 lux). Rooted plantlets were transferred into Jiffy® pellets (Jiffy Products International) under clingwrap in a controlled growth chamber (28 °C; 16h photoperiod, 8000–10000 lux). Jiffies were acclimatised to the growth chamber environment. This was carried out over a 3-week period with three holes made in the clingwrap at the end of each week and thereafter, the clingwrap was completely removed.

3.3.2 INFECTION OF CASSAVA TME3 PLANTLETS WITH SACMV INFECTIOUS CLONES

Plantlets were infected with equal quantities of SACMV pBIN-DNA-A and pBIN-DNA-B infectious, head-to-tail, dimers transformed into *Agrobacterium tumefaciens* strain Agl1 (Berrie et al., 2001). Infectious clones were cultured in YEP broth (100 µg/ml carbenicillin, 100 µg/ml kanamycin, 30 °C, 200 rpm) to an OD_{600} of 1.8-2.0. The cultures were pelleted by centrifugation for 20 min at 4200 xg, washed in an equal volume of sterile distilled water by centrifugation for 20 min at 4200 xg and finally resuspended in an equal volume of fresh YEP broth. Equal volumes of SACMV DNA-A and DNA-B were mixed. A 200 µl syringe with a 5 mm needle were used to inoculate the plantlets through the petioles and stems of cassava TME3. Wild type untransformed Agl1 was cultured, and plants inoculated in the same manner (mock-infected control). Plant height and symptom severity were observed over 32 days. Symptoms were scored at 32 days post inoculation (dpi) according to previous studies (Allie et al., 2014; Hahn et al., 1980) on a scale of 0 to 3; 0 (no symptoms/healthy), 1 (mild leaf distortion or mild chlorosis), 2 (moderate leaf distortion and chlorosis) and 3 (severe leaf distortion and chlorosis with reduction in leaf area). Harvesting of leaf tissue was carried out at 32 dpi from SACMV-infected cassava TME3 plants. Only symptomatic leaves were harvested (symptom score 1-3), and confirmation of infection established via PCR amplification of the viral coat protein (AV1) primer set (Fwd: '5-CGTGGCTGTGAA-3') and (Rev: '5-CGCGTGACATCA-3') **(Appendix 3.1, Appendix 2.1)**. From mock-inoculated cassava TME3 plants, new leaves from the apical meristem were also harvested for the Co-IP experiments. Harvested infected and mock samples were immediately frozen in liquid nitrogen and protein extraction carried out.

3.3.3 PROTEIN EXTRACTION

This was carried out according to DeBlasio et al. 2015 with modifications (Deblasio et al., 2015). Three ground leaf tissue samples were combined and divided into equal portions for evaluation of five lysis buffers (**Buffer 1**: 20 mM HEPES-KOH (pH 7.4), 110 mM K_2PO_4, 0.1 % Tween, 0.1 % Triton X-100, 150 mM NaCl, 2 mM $MgCl_2$, 1/100 PI, 0.5 mM PMSF; **Buffer 2**: 50 mM HEPES-KOH (pH 7.4), 100 mM KOAc, 2 mM MgCl2, 1/100 PI, 0.5 mM PMSF; **Buffer 3**: Buffer 2 + 0.4 % Triton X-100; **Buffer 4**: 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.4 % Triton X-100, 1/100 PI, 0.5 mM PMSF; **Buffer 5**: 25 mM Tris-HCl (pH 7.4), 150 mM NaCl, 1 % Triton X-100, 1 mM EDTA, 1/100 PI, 0.5mM PMSF). Cryo-lysis of leaf tissue was carried out with a mortar and pestle into a fine powder. Two grams of finely powdered leaf tissue was added to 12 ml of each buffer. The resulting mixture was incubated on ice for 30 min and briefly vortexed. The resulting extract was centrifuged for 10 min at 3000 rpm. The supernatant was filtered through mira-cloth to obtain a clear pale green extract prior to assessment of protein extract concentration and quality.

Protein concentration was determined by use of the Pierce™ BCA Protein Assay kit (ThermoFisher Scientific, Massachusetts, USA). Estimated protein concentrations were calculated from standard curves plotted from the absorbance values obtained at OD_{562} of extract against albumin standards **(Table 3.1)**.

Extract concentration and quality were further assessed on 12 % SDS-PAGE mini-PROTEAN gels (TGX™ Fastcast™ acrylamide kit Bio-Rad Laboratories, Inc) at a concentration of 15 ug per sample. Buffer 3 **(Table 3.1)** was chosen for downstream application and stored at -20 °C.

3.3.4 EVALUATION OF CASSAVA TME3 LYSIS BUFFERS AND EXPRESSED MP PROTEIN LYSATE

From the resulting gels for SACMV-infected and mock infected leaf tissue protein extracts, individual sample lanes were sliced out and stored in sterile distilled water at 2-4°C. The slices were evaluated for the presence of native MP protein and general protein composition by in-gel digest and mass spectrometry (AB Sciex 6600 TripleTOF mass spectrometer) by the CSIR-Proteomics laboratory team (Pretoria, South Africa). Annotation of protein fragments identified by mass spectrometry was conducted through Protein pilot software. Protein pilot software uses a standard target decoy database searching method and a proteome database was used to identify and validate proteins from spectral data. The Universal Protein Resource Knowledgebase (UniProtKB) database (http://www.uniprot.org) for *Manihot esculenta* Crantz (cassava) comprising 38,326

computationally analysed proteins was used to assign peptides to proteins. These were annotated to their closest *Arabidopsis* homologues.

The main parameters considered were the Unused ProtScore, Peptides (95%) and the False discovery rate (FDR < 0.05). The Unused ProtScore denotes the uniqueness of a peptide to a single protein and thus reduces inaccuracies from homology. The Unused ProtScore threshold was for values greater than 0.05 (95 % confidence). Peptides (95%) denotes the number of distinct peptides with at least 95% confidence attribution to a protein. The Peptides (95%) threshold was for values greater than 0. All reversed proteins from the decoy database were disregarded.

The identified proteins for mock and SACMV-infected samples were analysed for KEGG pathway enrichment by use of web based tool ShinyGO v0.60 (http://bioinformatics.sdstate.edu/go/) (FDR < 0.05) **(Appendix 3.1, Appendix 3.2)**. The identified proteins for mock and infected samples were analysed for KEGG pathway enrichment by use of web based tool ShinyGO v0.60 (FDR < 0.05). Distinguishing pathways were noted **(Figure 3.3)**. These samples served as controls for the presence and absence of SACMV MP.

3.3.5 MP OVER-EXPRESSION, DENATURATION, PURIFICATION, AND REFOLDING

This was carried out according to an established protocol (Nankoo, 2018). The expressed and purified MP was used as a positive antigen control for antibody specificity and was also utilised as an antigen for Co-IP of MP interactors in mock samples. In brief, protein overexpression was induced with 0.25 mM IPTG. The resulting protein was extracted with Tris based lysis buffer and dialysis with decreasing concentrations of urea solution. Thereafter, protein purification was carried out using a nickel charged gravity flow column with increasing concentrations of imidazole solution to wash and elute for expressed MP. Protein refolding was conducted through two cycles of dialysis in refolding buffer and concentration using sucrose. Well isolated eluates of denatured MP protein were pooled together and stored in multiple aliquots (4 °C and -20 °C). Pooling was also carried out for denatured samples for refolding of the MP protein. SDS-PAGE was carried out to ascertain the purity and presence of denatured and refolded MP obtained. The concentration of the pooled samples was determined by absorbance values at $OD_{340,}$ and protein fractions visualised on 12 % SDS-PAGE mini-PROTEAN gels (TGX™ Fastcast™ acrylamide kit Bio-Rad Laboratories, Inc).

3.3.5.1 MP PROTEIN OVEREXPRESSION

MP- pET-28a (+) transformed *E. coli* BL21(DE3) was cultured in 10 ml SOB media (0.5 % (w/v) yeast extract, 2 % (w/v) tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgSO_4$, 10 mM

$MgCL_2$, 50 µg.ml^{-1} kanamycin) and incubated overnight at 37 °C for 16 h on a shaker at 200 rpm. The resulting bacterial culture was centrifuged for 10 min at 12000 xg and room temperature to pellet the bacteria. The pellet was re-suspended and cultured in 100 ml fresh SOB media (50 µg.ml^{-1} kanamycin) and the initial absorbance at OD_{600} recorded (OD_{600} = 0.1-0.2). The fresh culture was incubated for approximately 1h at 37 °C on a shaker at 200 rpm to an OD_{600} of 0.4-0.6 and induced with 0.25 mM of IPTG. The induced culture was incubated for 4 h at 37 °C on a shaker at 200 rpm to allow for protein expression, thereafter, centrifuged for 15 min at 15 000 xg. The supernatant was discarded, and the pellet stored at -80 °C.

3.3.5.2 MP PROTEIN DENATURATION

The resulting pellet from protein overexpression was thawed on ice for 25 min. The thawed pellet was briefly vortexed and sonicated on ice (Microson Ultrasonic Cell Disruptor, Lasec SA) with lysis buffer (50 mM Tris, 5 % glycerol, 50 mM NaCl, pH 7-8) added in a 1:10 ratio. Lysozyme was added to 1 % and buffered solution briefly vortexed and sonicated on ice. The solution was centrifuged for 25 min at 25 000 xg and the pellet retained and resuspended in urea solution (8 M urea, 50 mM Tris- HCl pH 8.5, 0.05 % NaN_3 pH 7-7.5). The resuspended solution was sonicated and centrifuged for 25 min at 25 000 xg. The supernatant was retained and placed in dialysis tubing and dialysed overnight in a 4 M urea solution (4 M urea, 20 mM sodium phosphate, 0.5 M NaCl, 10 mM imidazole, pH 7-7.5).

3.3.5.3 MP PROTEIN PURIFICATION

Protein purification was carried out by use of a gravity flow column packed with a nickel charged NTA Profinity™ IMAC resin (Bio-rad). A volume of 2 ml of 50 % nickel resin slurry was washed with 5 ml of sterile distilled water and the flow through discarded. The washed resin was equilibrated with 3 ml of 4 M urea solution and the flow through discarded. To the equilibrated resin, dialysed protein was added, and the protein-resin mixture incubated at 4 °C for 30 min on rotation to enable the binding of protein molecules to the resin. The flow-through was collected in 3 x 2 ml fractions. The resin with bound protein was washed with 5 ml of 4 M urea solution (4 M urea, 25 mM imidazole, pH 7-7.5) and the flow-through collected in ~ 1.7 ml fractions. To the resin, protein was eluted using 6 ml of elution buffer (4 M urea, 250 mM imidazole, pH 7-7.5) and the flow-through collected in ~1.7 ml fractions. All fractions were stored at 4 °C.

All the eluted protein fractions were prepared with 2X SDS buffer (added in a 1:1 ratio) and run on a 12 % SDS-PAGE mini-PROTEAN gels (TGX™ Fastcast™ acrylamide kit Bio-Rad Laboratories, Inc). Based on the SDS-PAGE results, column fractions containing purified MP protein (35 kDa) were pooled together **(Figure 3.4)**.

3.3.5.4 MP PROTEIN REFOLDING

Pooled fractions of purified MP protein were placed in dialysis tubing and dialysed overnight in 2 L of refolding buffer (50 mM Tris-HCl pH 10, 0.5 M L-Arginine, 2 % (v/v) Triton X-100, 10 mM DTT). The pooled fractions in dialysis tubing were transferred into fresh refolding buffer (50 mM Tris-HCl pH 10, 0.5 M L-Arginine, 10 mM DTT) for two overnight dialysis cycles. Refolded protein was run on 12 % SDS-PAGE mini-PROTEAN gels (TGX™ Fastcast™ acrylamide kit Bio-Rad Laboratories, Inc) to check for aggregation **(Figure 3.5)**.

3.3.6 MP ANTIBODY PRODUCTION AND ASSESSMENT

The MP amino acid sequence (NP_620668.1 BC1 [South African cassava mosaic virus]) was submitted to Abmart Inc. (China) for polyclonal antibody production. The protein sequence was subject to a detailed analysis using the Abmart propriety algorithm. An "epitope score" was assigned to each residue based on structural features, sequence conservation, hydrophobicity, solvent exposure, and numerous other criteria. Twelve-residue fragments with highest overall scores were chosen as epitopes to generate antibodies for the MP protein. Eight fragments were identified spanning from the N- to C-terminus and a single antigenic fragment was selected for antibody production in rabbits **(Figure 3.6)**. Affinity column purified rabbit polyclonal antibody was used to purify the MP antibody (Abmart). Western blotting was carried out to ascertain specific antibody binding to purified SACMV MP. Determination of detectable MP in lysate at various dilutions (1:10, 1:100, 1:200, 1:500) of a concentrated solution (13.5 µg) in lysis buffer was performed **(Figure 3.7)**.

3.3.7 WESTERN BLOTTING TO IDENTIFY MP

Western blots were carried out to ascertain antibody-MP antigen binding in (i) SACMV infected plant protein extract, (ii) purified MP protein and (iii) co-immunoprecipitated protein complexes. Proteins separated by SDS-PAGE were transferred onto a Trans-Blot Turbo™ PVDF membrane (Bio-Rad, Hercules, USA) via a Trans-Blot Turbo Transfer System (Bio-Rad, Hercules, USA) for 3 min. The membranes were incubated for 30 min with continuous agitation in 5 % skim milk in TBSTT (10 mM Tris-HCl, 150 mM NaCl, 0.1 % Tween 20, pH 7.5). This was followed by membrane washing with TBSTT (X3) for 5 min. The membrane was incubated at 4 °C, for 16 h in primary polyclonal antibody (Abmart, China), diluted 1 $\mu g.ml^{-1}$, in 5 % skim milk in TBSTT with continuous agitation. The membranes were washed with a double volume of TBSTT buffer (X3) for 5 min. After washing, the membrane was incubated at room temperature with constant agitation for 45 min with goat anti-mouse HRP conjugated secondary antibody (Bio-Rad, Hercules, USA) diluted at 1:3000 in blocking buffer and Streptactin-HRP (Bio-Rad, Hercules, USA) at a 1:9000 dilution with the blocking

buffer. The Streptactin-HRP was used to detect the Precision Plus Protein™ Unstained Protein Standards (Bio-Rad, Hercules, USA). The membrane was washed vigorously with two TBSTT 2x for 10 min and TBS (10 mM Tris-HCl, 150 mM NaCl, pH 7.5) 1x for 10 min to remove non-specifically bound secondary antibody (X3) for 10 min. The protein blots were developed using chemiluminescence substrate, Clarity™ Western ECL Substrate (Bio-Rad, Hercules, USA) and viewed with the Gel Doc™ EZ imager (Bio-Rad, Hercules, USA). Twelve percent SDS-PAGE gels for MP, mock-inoculated and SACMV-infected lysates were processed by in-gel digest and mass spectrometry analysis to determine the presence or absence of MP.

3.3.8 CO-IMMUNOPRECIPITATION

Co-immunoprecipitation was carried out to identify protein complexes associated with MP binding. This was done for samples in triplicate comprising, three negative controls, SACMV infected-no antibody (SACMV-nAb), mock-no antibody (nSACMV-nAb), mock infected-antibody (nSACMV-Ab), one positive control/treatment, expressed MP-antibody (BC1-Ab) and two treatments, SACMV infected-antibody (SACMV-Ab), mock-expressed MP-antibody (nSACMV-BC1-Ab). The negative controls allowed for non-specific binding of proteins to magnetic beads to be identified and excluded from the treatments.

Co-immunoprecipitation was by use of the MagReSyn® Protein A beads (Resyn Biosciences, Pretoria, SA) and the DynaMag™-2 Magnet (ThermoFisher Scientific, Massachusetts, USA) magnetic separator in accordance with the MagReSyn® online protocol. All washing steps were conducted with TBST buffer (50 mM Tris pH 7.5, 150 mM NaCl with 0.025 % Tween 20). A minimum volume of 50 µl MagReSyn® Protein A was used to bind, 60 µg of MP antibody and incubated at 2-4 °C for 16 h with sample to a total volume of 650 µl with continuous rotation. Antibody-MP-protein complexes were eluted with 2.5 % acetic acid and the eluates pooled. The pH of the pooled eluates was neutralised with 5 M NaOH. The eluted protein samples were run on 12 % SDS-PAGE mini-PROTEAN gels (TGX™ Fastcast™ acrylamide kit Bio-Rad Laboratories, Inc) and the sample lanes sliced out and stored in sterile distilled water at 2-4°C for subsequent in-gel digest and mass spectrometry.

3.3.9 IN-GEL DIGEST AND MASS SPECTROMETRY OF CO-IMMUNOPRECIPITATES

Both the in-gel digest and mass spectrometry were outsourced to the CSIR-Proteomics Laboratory (Pretoria, South Africa) according to their optimised protocols. Data was acquired on a TripleTOF® 6600 System mass spectrometer (Sciex, USA) in data-dependent acquisition mode. The raw files (.wiff) were processed using Protein Pilot V5 (Sciex, USA) software which uses a standard target decoy database searching method and

a proteome database to identify and validate proteins from spectral data. Data was matched against a SWISS-PROT database of *Arabidopsis thaliana* sequences from the Universal Protein Resource Knowledgebase (UniProtKB) database (http://www.uniprot.org).

3.3.10 IDENTIFICATION OF THE MOST SIGNIFICANT HOST-COMPLEX INTERACTORS

Two methods were utilised to identify the most significant MP interactors. The first was a simplistic method that utilised an exclusion concept based on the elimination of all proteins identified in the negative controls from all sample results. The second method mathematically assigned specificity and confidence values to results computationally using high throughput software. Collaborating partners from Iowa State University carried out the latter method through Significance Analysis of INTeractome Express (SAINTexpress) (Teo et al., 2014) and Y2H Scores (Velásquez-Zapata et al. 2020) software to assign confidence scores to observed interactors.

SAINTexpress was run using the samples without antibody as controls and with antibody as treatment. Two pseudo replicates were generated to run the analysis. Results were compiled in a unique file presenting, (i) bait-prey interactions, (ii) the average counts, (iii) the Saint score, (iv) fold change and (v) FDR. Protein annotations were also included, containing gene names and IDs. To run Y2H-SCORES (Velásquez-Zapata, Elmore, Banerjee, et al., 2021), the samples were classified as non-selected (no antibody) and selected (antibody). Two scores were calculated including the enrichment and the specificity score. To calculate the specificity score the mock-inoculated sample with antibody was also included in the selected group, resembling the effect of an empty bait in the design. Results were summarized reporting bait, prey, enrichment, specificity score, a Borda ensemble and a Borda quantile with the percentile that the Borda score occupies in the density of values.

3.3.10.1 A SIMPLISTIC EXCLUSION METHOD

A simplistic approach was taken to identify proteins exclusive to treatment samples. To achieve this, all proteins identified from the negative controls were excluded from the results obtained from the treatment samples. ‘Sticky’ proteins and highly similar proteins (isozymes) were identified in control and treatment samples and excluded.

3.3.10.2 COMPUTATIONAL METHODS: Y2H SCORES AND SAINT ANALYSIS

Data assigned scores was obtained from collaborating partners (University of Iowa). A stringency in selection of significant interactors was applied to the data obtained from

SAINT and Y2H Scores software analyses, respectively, taking into consideration that the results obtained had a low number of interactors.

Parameters for SAINT were set as follows: SAINT score < 0.72, Bonferroni FDR (BFDR) ≤ 0.1); Y2H Scores (Enrichment > 0.57, Specificity > 0.17, sum score > 0.96, Borda score > 3.27 **(Appendix 3.3)**. All significantly identified proteins in the negative controls and the treatments were excluded from the data.

For Y2H Scores analysis, identified interactors with significance from within the top ranked 27th percentile (Borda_quantile 0.27) based on the Borda counts ranked using the Wilcoxon signed rank test **(Appendix 3.4)**.

Results from both software tools were collated. 'Sticky' proteins and highly similar proteins (isozymes) were identified in control and treatment samples and excluded from the total of highly significant results. The results were carried forward for further characterisation.

3.3.11 GENE ONTOLOGY CATEGORISATION AND KEGG PATHWAY ENRICHMENT ANALYSIS

Gene ontology was carried out using the TAIR GO Slim/Functional Classification online function (https://www.*Arabidopsis*.org/tools/bulk/go/index.jsp) **(Figure 3.8)**.

To further elaborate the biological significance of identified cassava TME3 interactors, their interactors were identified from an *Arabidopsis* interactome (Velásquez-Zapata et al., 2021) with only experimentally determined interactions from BioGRID 4.1.190 version (Stark, 2005), and literature review (Mukhtar et al., 2011; Weßling et al., 2014; Smakowska-Luzan et al., 2018; Trigg et al., 2017; McWhite et al., 2020).

The list obtained was analysed for pathway enrichment by use of web based tool ShinyGO v0.60 (Ge et al. 2020, http://bioinformatics.sdstate.edu/go60/) **(Figures 3.9, 3.10)**. ShinyGO pathway enrichment analysis was based on hypergeometric distribution followed by FDR correction (FDR < 0.05).

3.4 RESULTS

3.4.1 SACMV INFECTED CASSAVA TME3 MORPHOLOGICAL FEATURES

For Co-IP only those with a symptom score greater than or equal to 2 were considered. Therefore, scoring was noted as follows 40 % (Score 3), 60 % (Score 2) giving an overall mean symptom score of 2.4. All plants were notably symptomatic at 32 days post inoculation **(Appendix 3.1)**. Amplification with SACMV (coat protein) CP primers further confirmed infection with a single 95 bp band **(Appendix 2.2)**. Visual representation of symptoms as depicted in Chapter 2 (Figure 2.3).

3.4.2 SELECTION OF IDEAL LYSIS BUFFER FOR CO-IMMUNOPRECIPITATION

Protein concentrations ranged from 2508 to 3967 $\mu g.ml^{-1}$ and 3161 to 4731 $\mu g.ml^{-1}$ in SACMV and mock-inoculated samples, respectively. Mean protein concentrations were 3135.8 and 3328.2 $\mu g.ml^{-1}$ in SACMV and mock-inoculated samples, respectively. Buffers 3 and 5 were comparable in protein concentration however, Buffer 5 had a deep green colour and was excluded because of colour intensity. Buffer 3 had a clearer extract compared to all buffers evaluated. Buffer 2 and 3 showed a high concentration of plant proteins distributed across a wide molecular weight (MW) range (250-15 kDa) **(Figure 3.2)**. A high concentration of protein bands within the 55 kDa and 15 kDa range may have been indicative of Rubisco (hexadecameric) subunits (Parry et al., 1987), which are often found in protein extraction from leaves due to their high abundance. MP migrated below the 37 kDa position confirming its ~35 kDa molecular weight **(Figure 3.2)**. Buffer 3 was selected based on protein concentration **(Table 3.1)**, a visually pale green extract (not shown) and clear SDS-PAGE bands **(Figure 3.2, Figure 3.3)**.

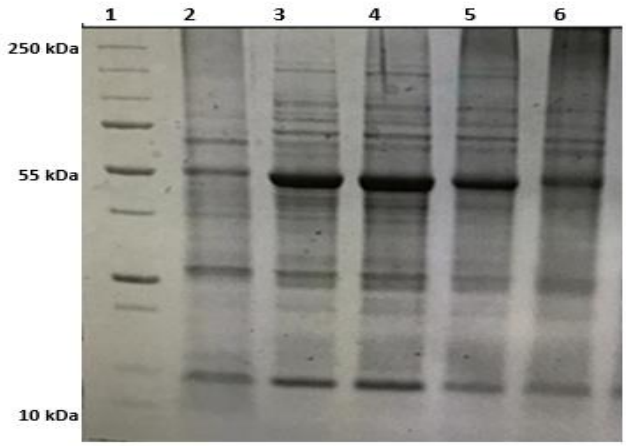


Figure 3.2 A 12% SDS-PAGE gel to evaluate buffer suitability for Co-IP experiments. SDS-PAGE gel showed protein distribution of 15 µg of cassava TME3 extract proteins per sample from SACMV infected leaf tissue. Buffer 1-5 (lanes 2-6). Precision Plus Protein™ Standards (lane 1).

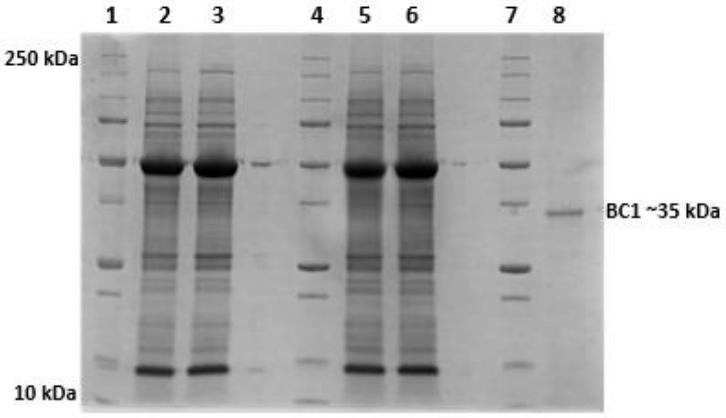


Figure 3.3 A 12% SDS-PAGE gel to evaluate buffer suitability for Co-IP experiments. Protein concentration loaded was 15 µg of cassava TME3 extracted proteins per sample from SACMV infected leaf tissue. Buffers 1-5 are represented in (lanes 2-6, respectively. Precision Plus Protein™ Standards (lane 1).

Table 3.1 Protein extract quantification. Estimated protein concentrations were calculated according to the Pierce BCA kit from standard curves plotted from the absorbance values obtained at OD_{562} of SACMV infected and non-infected leaf tissue samples against albumin standards. Sample buffer 3 (yellow highlight) was chosen based on a high concentration, clear SDS-PAGE bands and a clear extract (not shown)

SAMPLE	BUFFER	VALUE A	VALUE B	AVERAGE	CORRECTED	GRADIENT	INTERCEPT	**ug/ml
SACMV INFECTED	1	1.104	1.096	1.1	0.969	0.0004	0.0851	2635
	2	1.146	1.027	1.0865	0.9625	0.0004	0.0407	2508
	*3	1.349	1.301	1.325	1.196	0.0003	0.0752	*3736
	4	1.242	1.287	1.2645	1.1375	0.0004	0.0043	2833
	5	1.251	1.401	1.326	1.199	0.0003	0.0088	3967
MOCK	1	1.542	1.622	1.582	1.451	0.0004	0.0851	3840
	2	1.323	1.372	1.3475	1.2235	0.0004	0.0407	3161
	*3	1.641	1.606	1.6235	1.4945	0.0003	0.0752	*4731
	4	1.425	1.638	1.5315	1.4045	0.0004	0.0043	3501
	5	1.545	1.462	1.5035	1.3765	0.0003	0.0088	4559

**Estimated protein concentrations were calculated according to the Pierce BCA kit from standard curves plotted from the absorbance values obtained at OD_{562} of SACMV infected and non-infected leaf tissue samples against albumin standards. *Sample buffer 3 (yellow highlight) was chosen based on a high concentration, clear SDS-PAGE bands and a clear extract (not shown).

3.4.3 EXPRESSED MP PROTEIN

With reference to SDS-PAGE gel **(Figure 3.4)** the non-induced and IPTG induced (lane 2 and 3) *E. coli* BL21 expression vector both showed significant expression of MP protein at the 35 kDa molecular weight region. Soluble MP was present in the denatured fraction (lane 4). Purification washes using nickel charged gravity flow column presented prominent bands in the 35 kDa region (lanes 5, 6 and 7). The pooled elution of MP presented a single clear band (lane 8) with no visible contaminating proteins. Therefore, purification was considered successful and MP refolded. The presence of MP in the refolded fractions was confirmed by the presence of a single band within the 35 kDa region. Additional work within the laboratory has confirmed the refolding of MP by fluorescence and native gels have shown refolded MP to remain in the well due to oligomerisation via the C-terminal (Nankoo & Rey, unpublished).

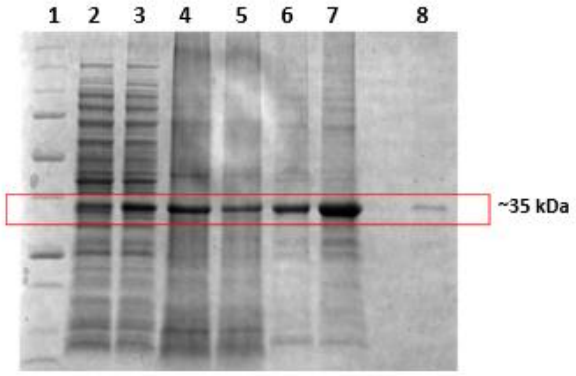


Figure 3.4 A 12% SDS-PAGE gel showing the step-wise expression and purification of MP protein (~35 kDa) from E. coli with respective lanes denoted in brackets; Non-induced-no IPTG (2), Induced-IPTG (3), Solubilised (4), Wash 1 (5), Wash 2 (6), Wash 3 (7), Pooled elution (8), Precision Plus Protein™ Standards (1)

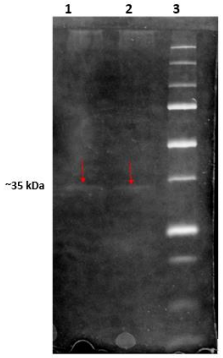


Figure 3.5 A 12% SDS-PAGE gel confirming the presence of MP protein (~35 kDa) in the refolded fractions from purified denatured elution fractions (red arrows). Refolded MP (1-2), Precision Plus Protein™ Standards (3).

3.4.4 MP ANTIBODY: PRODUCTION AND ASSESSMENT

The fragment with the highest epitope score towards the C-terminal region of MP (amino acid position 187-198, sequence TPTARSQQPITH, epitope score 232.5) was chosen as an ideal antigenic site for antibody production **(Figure 3.6)**.

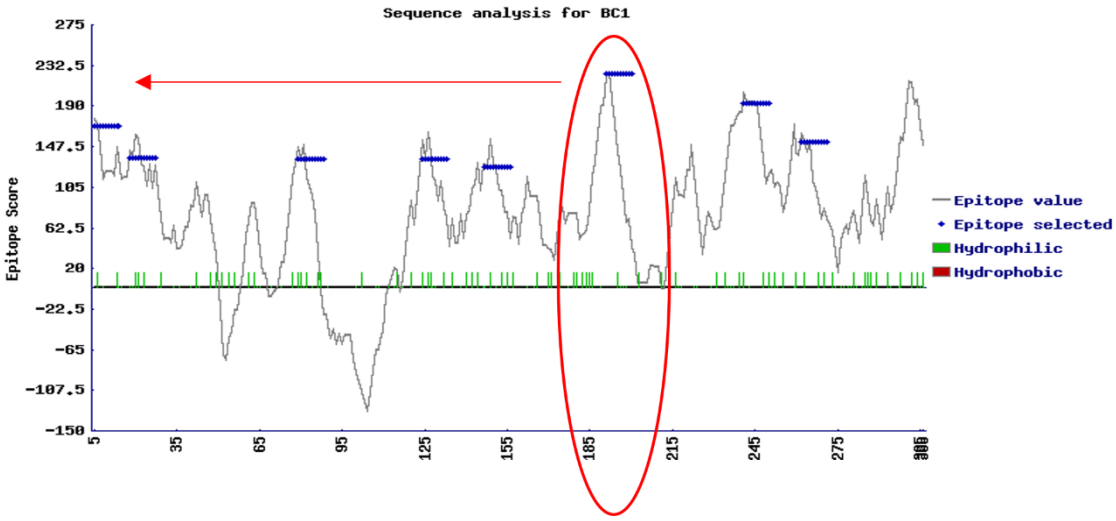


Figure 3.6 A propriety algorithm (Abmart, China) assigned epitope scores (based on structural features, sequence conservation, hydrophobicity, solvent exposure, and numerous other criteria) to regions of amino acid positions; (1-12), (14-25), (75-86), (120-131), (143-154), (187-198), (237-248) and (258-269); Circled in red, the region (187-198) with the highest epitope score (232.5) was chosen as the antigenic site for antibody production.

Western blot detection of native MP was not possible. *E. coli* expressed MP was detectable at dilutions of 1:10 and 1:100 (maximum threshold 0.135 µg) but not 1:200 and 1:500 **(Figure 3.7)**. Detectable concentrations of native MP in SACMV infected leaf tissue extract by western blotting was not achieved.

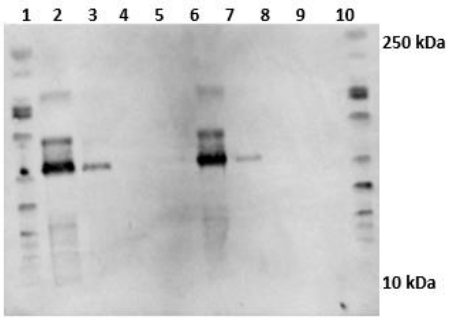


Figure 3.7 Western blot of MP to determine antigen binding and at different protein dilutions to determine sensitivity; Dilutions of 1:10 (2,6) and 1:100 (3, 7) showed positive detection of the MP

antigen. Dilutions of 1:10 indicated multiple fragment size detection due to MP degradation(lower bands) and oligomerization (higher bands). Dilutions of 1:200 (4, 8) and 1:500 (5, 9) did not detect the MP antigen. Precision Plus Protein™ Standards (1, 10).

3.4.5 PRELIMINARY EVALUATION OF PROTEIN EXTRACTSS: MASS SPECTROMETRY

In-gel digest and mass spectrometry analysis of purified ~35 kDa MP sample revealed the presence of 21 contaminating proteins in addition to SACMV MP which was identified with high confidence (Unused ProtScore: 56; Peptides (95%): 50). Identified MP peptides spanned 71.99 % coverage over the length of the MP amino acid sequence. In-gel digest and mass spectrometry analysis of the extract from the healthy sample was carried out on protein bands around 55-25 kDa regions to ascertain the absence of MP. Analysis identified the presence of 16 contaminating proteins (mammalian derived proteins) in the MP sample. MP was not identified in healthy samples. In total 247 proteins from cassava TME3 were significantly identified. Full lane in-gel digest and mass spectrometry analysis of the infected extract revealed the presence of 2206 significant proteins of which 2158 were attributed to *Manihot esculenta*. In the infected extract MP was identified with high confidence (Unused ProtScore: 10; Peptides (95%): 5). Identified MP peptides spanned a 27.04 % coverage over the length of the MP amino acid sequence. This confirmed the presence of SACMV derived MP (native MP) in infected samples. Of the 2158 *Manihot esculenta* proteins 1905 were cross mapped to *Arabidopsis* homologues. Single lane in-gel digest and mass spectrometry analysis of the mock-inoculated sample revealed the presence of 2099 were significant proteins of which 2088 were attributed to *Manihot esculenta*. Of the 2088 *Manihot esculenta* genes 1837 were cross mapped to *Arabidopsis* homologues. In the mock-inoculated MP was not identified.

3.4.6 PATHWAY ENRICHMENT ANALYSIS OF ANNOTATED PROTEINS IDENTIFIED

Figure 3.8 depicts the KEGG pathways that were identified from whole lane in-gel digest of SACMV-infected and mock-inoculated samples.

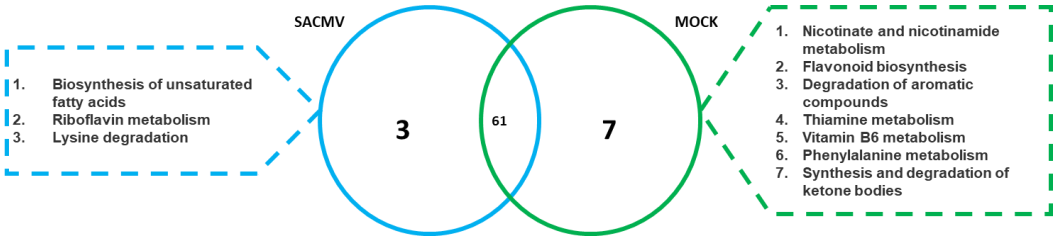


Figure 3.8 A Venn diagram distinguishing KEGG pathways enriched in mock (green) and SACMV-infected (blue) samples. Pathways exclusively enriched to each sample have been shown in pentagons.

3.4.7 IDENTIFICATION OF THE MOST SIGNIFICANT INTERACTORS

3.4.7.1 A SIMPLISTIC EXCLUSION METHOD

From the simplistic approach, a total of 70 (SACMV-nAb, nSACMV-nAb, nSACMV-Ab), 32 (BC1-Ab), 44 (SACMV-Ab) and 21 (nSACMV-BC1-Ab) proteins were identified. A total of 17 interactors were initially identified as unique to treatment samples (expressed MP-antibody, SACMV infected-antibody, mock-expressed MP protein-antibody) prior to exclusion of sticky proteins and their isoforms.

3.4.7.2 COMPREHENSIVE METHODS: Y2H SCORES AND SAINT ANALYSIS

In SAINT analysis, 28 proteins had a SAINT score > 0.72 and considered significant interactors identified from SAINT analysis **(Appendix 3.3)**. Eleven were common to negative controls. SAINT analysis identified less significant interactors than Y2H Scores. SAINTexpress is computationally designed to identify high confidence interactions from mass spectrometry data (Teo et al., 2014). SAINTexpress includes external interaction data in computing reliable scores in a probabilistically objective manner (Teo et al., 2014).For Y2H Scores analysis, 56 interactors were identified with significance within the top ranked 27th percentile (Borda_quantile 0.27) based on the Borda counts **(Appendix 3.4)**. Thirty seven were common to negative controls . Two proteins (Tubulin alpha-3 chain and Actin-11) exclusive to treatments were identified as isozymes of tubulin and actin proteins identified in negative controls. A larger number of significant interactors was identified than SAINTexpress.Y2H Scores is a statistical framework initially designed for selection of positive interactors within the Y2H framework (Velásquez-Zapata et al., 2020).

Thirty-nine proteins in treatments analysed by Y2H Scores and SAINTexpress were common to negative controls. From both Y2H Scores and SAINTexpress sixteen proteins were common to both analyses and not identified in negative controls. From these proteins, a total of 9 proteins were identified as isozymes of proteins identified in negative controls and excluded. These were: Ribulose bisphosphate carboxylase large chain, ATP synthase subunit, Ribosomal proteins (40s, 50s, 60s), Chaperonin 60 subunits , Glyceraldehyde-3-phosphate dehydrogenase andCatalase-2.A total of 7 protein interactors remained.

Of all the putative interactors identified, seven final protein interactors were identified as **highly significant MP protein complex partners.** These proteins were not identified in Chapter 2 (Y2H assays):

1. AT2G26080 (Glycine dehydrogenase (decarboxylating) 2 (GLDP2))
2. AT5G56030 (Heat shock protein 90-2 (HSP90.1))
3. AT1G07660 (Histone H4 (H4))

4. AT5G49460 (ATP-citrate synthase beta chain protein 2 (ACLB-2))
5. AT3G43190 (Sucrose synthase 4 (SUS4))
6. AT5G03340 (Cell division control protein 48 homolog E (CDC48C))
7. AT1G53240 (Malate dehydrogenase 1 (Mmdh1))

MP protein was identified in the positive control/treatment expressed MP-antibody (BC1-Ab) and two treatments, SACMV infected-antibody (SACMV-Ab) and mock-inoculated-purified MP protein-antibody (nBC1-Ab). There was no MP protein identified in the three negative controls, SACMV infected-no antibody (SACMV-nAb), mock infected-no antibody (nSACMV-nAb) and mock infected-antibody (nSACMV-Ab) samples.

3.4.8 GENE ONTOLOGY ANALYSIS OF THE MOST SIGNIFICANT IDENTIFIED INTERACTORS

Gene ontology (GO) categorisation of the MP interactors revealed the cellular localisation and processes potentially impacted by SACMV MP during infection. A percentage frequency was assigned to denote the categories represented by the interactors in relation to the whole *Arabidopsis* genome **(Figure 3.9, Table 3.2)**. The topmost GO Slim category for Cellular Components (CC) represented were those for cytoplasm (12 %), “chloroplast”, “cytosol”(10 %) and “nucleus”, “other cellular components”, “plasma membrane”, “cell wall”, “mitochondrion”, “vacuole” (7 %). With regards to GO Slim category, Biological Process (BP), the most represented were “other cellular processes” (20 %), “response to stress” (16 %), “other metabolic processes” (11 %) and, “response to abiotic stimulus” (10 %). GO Slim category, Molecular Function (MF) was most represented by “other binding”, “protein binding” (24 %), followed by “transferase activity” (14 %) and “RNA binding”, “nucleotide binding”, “hydrolase activity”, “catalytic activity” (10 %).

Table 3.2 Gene ontology categorisation of interactors identified by Co-IP assays

GO Category	Functional Category	Percent	GO Category	Functional Category	Percent
GO Cellular Component	endoplasmic reticulum	2	GO Biological Process	other biological processes	1
	other membranes	2		post-embryonic development	1
	plastid	5		lipid metabolic process	1
	Golgi apparatus	5		nucleobase-containing compound metabolic process	1
	other intracellular components	5		response to external stimulus	1
	extracellular region	5		cellular protein modification process	1
	nucleus	7		anatomical structure development	1
	other cellular components	7		flower development	1
	plasma membrane	7		transport	1
	cell wall	7		cell cycle	1
	mitochondrion	7		response to biotic stimulus	1
	vacuole	7		carbohydrate metabolic process	3
	chloroplast	10		biosynthetic process	3
	cytosol	10		multicellular organism development	3
	cytoplasm	12		protein metabolic process	4
GO Molecular Function	RNA binding	10		cellular component organization	4
	nucleotide binding	10		catabolic process	6
	hydrolase activity	10		response to chemical	6
	catalytic activity	10		response to abiotic stimulus	10
	transferase activity	14		other metabolic processes	11
	other binding	24		response to stress	16
	protein binding	24		other cellular processes	20

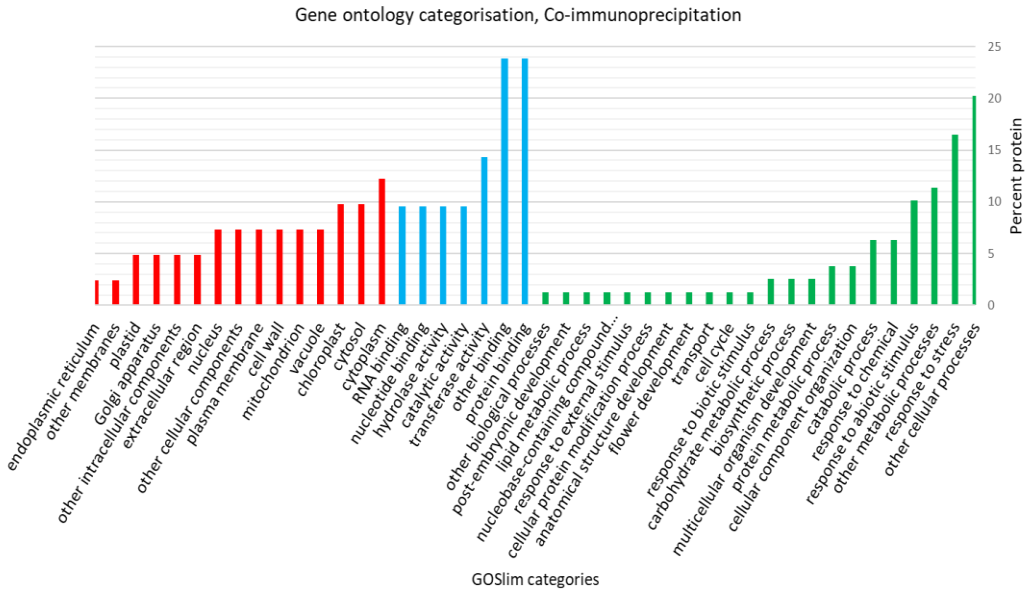


Figure 3.9 The Gene Ontology of the most highly significant seven MP interacting proteins identified from Co-IP mass spectrometry analysis, Cellular Component (Red), Biological Process (Green), Molecular Function (Blue).

3.4.9 KEGG PATHWAY ENRICHMENT ANALYSIS

The interactome for MP binding partners revealed twenty-one enriched KEGG pathways (**Table 3.3, Figures 3.10, 3.11**). The most enriched pathways and the percentage of proteins identified to each pathway was the Proteasome pathway at 80 %.

Table 3.3 Enriched KEGG pathways of co-immunoprecipitated proteins

KEGG pathway	Percentage
Proteasome	80
RNA polymerase	43
RNA degradation	18
Nucleotide metabolism (Pyrimidine, Purine)	17, 15
Carbon fixation in photosynthetic organisms	13
Citrate cycle (TCA cycle)	11
Glyoxylate and dicarboxylate metabolism	9
Galactose metabolism, Circadian rhythm – plant, Protein processing in endoplasmic reticulum	8
Nucleotide excision repair, RNA transport,	6
Plant-pathogen interaction, Cysteine and methionine metabolism, Carbon metabolism, Glycolysis / Gluconeogenesis, Ubiquitin mediated proteolysis	5
Ribosome, Microbial metabolism in diverse environments	3
Metabolic pathways	2

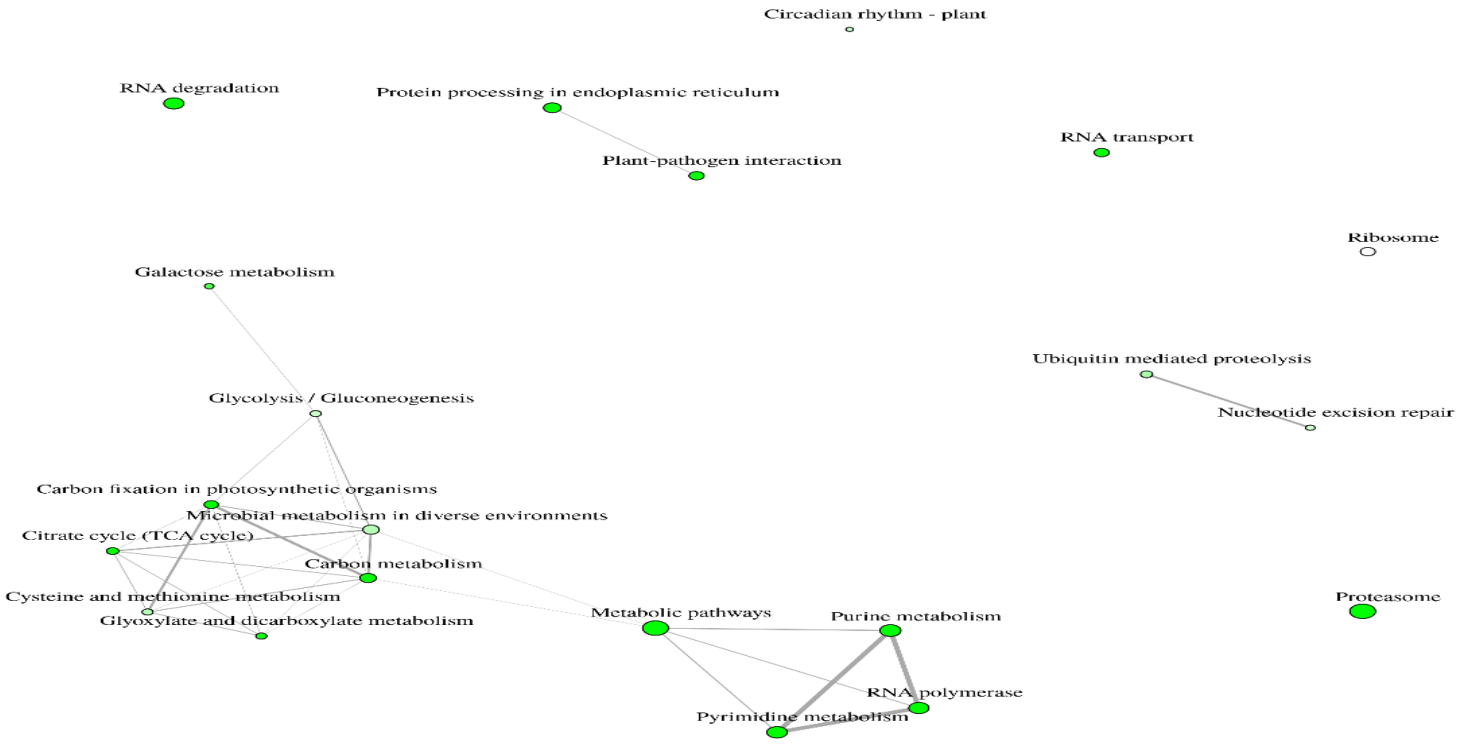


Figure 3.10 Protein interaction network derived from the enriched KEGG pathways of co-immunoprecipitated MP host proteins visualised in ShinyGO. The network shows the relationship between enriched KEGG pathways. Two pathways (nodes) are connected if they share 20% (default) or more genes. Darker nodes are more significantly enriched gene sets. Bigger nodes represent larger gene sets. Thicker edges represent more overlapped genes.

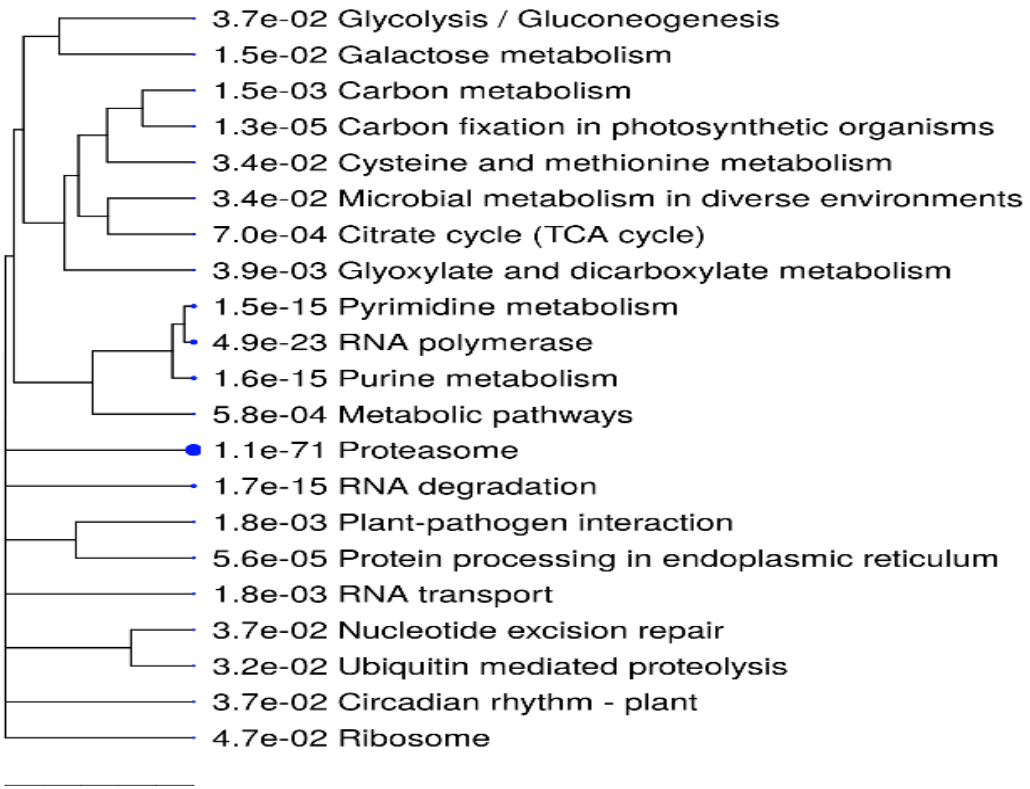


Figure 3.11 Hierarchical clustering tree derived from enriched KEGG pathways of co-immunoprecipitated MP-host proteins interaction network visualised in ShinyGO. The hierarchical clustering tree summarises the correlation among significantly enriched KEGG pathways. Pathways with many shared genes are clustered together. Bigger dots indicate more significant P-values (<0.05).

3.5 DISCUSSION

3.5.1 SELECTION OF LEAF SAMPLES WAS DETERMINED BY SYMPTOM SEVERITY

The BC1-encoded movement protein (MP) of geminiviruses is present in low concentrations relative to plant biomass (Kleinow et al., 2008). Viral load is directly proportional to symptom severity in cassava TME3 infected with SACMV (Allie et al., 2014). Therefore, only plants that presented a symptom score ≥ 2 were considered to have increased relative viral concentrations present in sampled tissue **(Appendix 3.1)**.

3.5.2 QUALITY OF CLARIFIED LEAF EXTRACT

A down regulation of ribosomal proteins, metabolic enzymes and an inhibition of gene activation is a common feature of virus infection (Satterly et al., 2007; Von Kobbe et al., 2000; Fernández-Calvino et al., 2014; Wang & Maule, 1995). This may explain the reason why, SACMV infected leaf tissues had a mean protein concentration ~6 % less than mock-inoculated samples across the five buffers evaluated **(Table 3.1)**. Plant pigments, lipids, secondary metabolites and polysaccharides have a tendency to bind to proteins and reduce SDS-PAGE quality and so it is essential to reduce these components during extraction (Wu et al., 2014). HEPES based buffer 3 was ideal for Co-IP experiments in cassava TME3 and has been used in prior virus-host studies in *Nicotiana benthamiana* (Deblasio et al., 2015). The extract obtained from Buffer 3 was an almost clear pale green colour indicative of reduced chlorophyll pigments and proteins. This was considered ideal in that chlorophyll components such as Rubisco account for up to 76 % of leaf proteins and may play a major contaminating role due to the relative abundance in plant leaf tissues (Evans & Caemmerer, 1987). The native form of proteins and intact physiological protein complexes are pre-requisites for the identification of binding partners in Co-IP experiments. Therefore, more stringent cleaning protocols that incorporate the use of TCA/Acetone and numerous wash steps were not considered as these would have interfered with protein conformation and result in protein degradation (Niu et al., 2018).

3.5.3 ANTIBODY DETECTION OF MP

At higher concentrations of MP, the banding pattern detected by western blots was indicative of possible oligomerisation, degradation, or post-translational modifications (phosphorylation and ubiquitination) of MP. Oligomerisation of MP in AbMV by binary interaction experiments was determined to occur via the C-terminal region (Krenz et al., 2010). Multiple post-translational phosphorylation events of AbMV-MP expressed in *E. coli*, *Saccharomyces* and plants has been shown to result in multiple banding patterns due to the altered migration arising from either a lower positive charge or an increase in the total molecular weight (Kleinow et al., 2008). Much like phosphorylation, the polyubiquitination

of TMV-MP and subsequent degradation by ubiquitin mediated proteolysis in the early and later infection stages respectively, resulted in the presentation of western blots with multiple band fragments (Reichel & Beachy, 2000).

It was not possible to identify native MP using western blots. Therefore, more sensitive mass spectrometry was ideal in detection of MP. Physiological concentrations of pathogenic proteins such as fungal effector molecules and the AbMV-MP by western blots is not always effective (Kleinow et al., 2009; Kleinow, Tanwir, et al., 2009; Alcântara et al., 2019). This may be due to higher host to pathogen molecule ratio (Alcântara et al., 2019). In the case of geminivirus AbMV-MP, the transient nature of the protein leads to low concentrations of protein (Kleinow et al., 2008).

3.5.4 IDENTIFICATION OF THE MOST SIGNIFICANT INTERACTORS

SAINT analysis was more stringent in identifying the most significant interactors. SAINTexpress is computationally designed to identify high confidence interactions from mass spectrometry data (Teo et al., 2014). SAINTexpress includes external interaction data in computing reliable scores in a probabilistically objective manner (Teo et al., 2014). The initial result obtained using Y2H Scores was also effective in identifying initial interactors. It would be essential to re-analyse the SAINTexpress data and the Y2H data in a comparable manner. This will be explored further to ensure that both computational tools are fairly evaluated. Assumptions on the most effective computational tool cannot be made at this time.

3.5.5 GENE ONTOLOGY OF IDENTIFIED INTERACTORS

The GO ontology Cellular Component indicated that MP interacts with proteins that are well distributed throughout the intracellular regions of the cell and associates with multiple cell organelles and the cytoskeletal system. Previous studies on geminivirus MPs are in agreement with the recorded localisation from this study (Aberle et al., 2002; Kleinow et al., 2009; Frischmuth et al., 2007; Zhang et al., 2002; Ward et al., 1997; Krenz et al., 2010; Sanderfoot & Lazarowitz, 1995; Stanley & Albrecht, 1992). The localisation of geminivirus MP to the vacuole was of particular interest. The plant vacuole together with the golgi apparatus and endoplasmic reticulum is part of the secretory pathway (Patarroyo et al., 2013) and has a complex endomembrane trafficking system essential to the distribution of lipids, polysaccharides and proteins (Zhang et al., 2014; Patarroyo et al., 2013). It is therefore not surprising to find that MP being a viral DNA translocation protein associates with vacuole localised proteins to enter the ER via the golgi body to facilitate intracellular movement to the cell membrane and plasmodesmata for intercellular movement. Vacuoles also play a role in plant defence by either programmed cell death induced by the release

of vacuolar hydrolytic enzymes to effect virus genome or protein digestion (Hara-Nishimura & Hatsugai, 2011). The geminivirus MP is associated with endoplasmic reticulum-derived tubules in developing phloem cells (Krapp et al., 2017).

In addition to trafficking, the secretory pathway is also involved in modification and transport of proteins (Patarroyo et al., 2013), it was therefore interesting to note that the gene ontology terms cellular protein modification process and transferase activity were identified in the GO categorisation of identified interactors under BP and MF, respectively. This furthers alludes to the manipulation of not only transport, but a cohort of host post translational mechanisms for MP phosphorylation and trafficking, which may play a role in SACMV infection. Post translational modification has been demonstrated to play a critical regulatory role in AbMV MP functionality (Kleinow et al., 2008). Geminivirus MPs also play critical roles as symptom determinants (Stanley & Albrecht, 1992; Pascal et al., 1993; Hou et al., 2000). The affiliation of MP with chloroplast (CC) and BP categories involved in cellular processes, metabolic processes and post-embryonic and reproductive development would contribute to symptoms due to altered chloroplast function and perturbed plant growth noted in SACMV infection (Allie et al., 2014).

The GO categorisation for MF was prominent for ‘other’ binding and protein binding. Proteins with a tendency for multiple binding partners are generally referred to as hub proteins (Patil & Nakamura, 2006). Hub protein structures are also predominantly disordered with a low mean hydrophobicity and high net charge (Patil & Nakamura, 2006). The SACMV MP may be categorised as a hub protein as it has an approximately 65 % highly disordered secondary structure and is ~52 % hydrophilic and ~38 % hydrophobic (Nankoo, 2018).

3.5.6 KEGG PATHWAY ENRICHMENT ANALYSIS OF IDENTIFIED INTERACTORS

The interactome for MP complex partners revealed twenty-one enriched KEGG pathways (**Figures 3.10, 3.11**). Enriched KEGG pathways comprised mainly of those involved in carbon and energy metabolism (Carbon fixation, Citrate cycle (TCA cycle), Glyoxylate and dicarboxylate metabolism, Galactose metabolism, Carbon metabolism, Glycolysis / Gluconeogenesis); Protein and amino acid associated pathways (Proteasome, Protein processing in endoplasmic reticulum, Cysteine and methionine metabolism, Ubiquitin mediated proteolysis, Ribosome), Nucleotide metabolism (Pyrimidine metabolism, Purine metabolism), overall metabolism (Microbial metabolism in diverse environments) and adaptive responses (Plant-pathogen interaction, Circadian rhythm – plant).

Most of these pathways have been discussed prior in Chapter 2 and interference with metabolism, noted in Chapter 2. Plant-pathogen interaction pathway and the proteasome

pathway will be discussed. The latter referred to in Chapter 2, will still be delved into due to its prominence among the enriched pathways presented in this chapter.

3.5.6.1 PLANT-PATHOGEN INTERACTIONS

The plant-pathogen interaction may be characterised by the production of reactive oxygen species and localized programmed cell death/hypersensitive response (HR). It has been shown that in spite of an upregulation of cell death associated host genes during geminivirus, TYLCV and tomato leaf curl New Delhi virus (ToLCNDV) infection of tomato, these viruses suppress the hypersensitive response and subsequently cell death (Moshe et al., 2016). Cassava does not exhibit the HR to pathogen infection, including geminiviruses, and SACMV may also suppress the HR. However oxidative and other biotic stress linked to plant-pathogen responses have been shown in cassava landraces T200 and TME3 in response to SACMV (Allie, et al., 2014).

3.5.6.2 PROTEASOME/UBIQUITINATION

Proteasome and ubiquitin mediated proteolysis pathways associated with MP interactors were enriched. The post translational ubiquitination and proteolysis of viral proteins has been appreciated as a host control response to plant virus infections, namely, TMV-MP (Reichel & Beachy, 2000), βC1 of tomato yellow leaf curl China virus (TYLCCNV) (Shen et al., 2016) and tomato leaf curl New Delhi virus (ToLCNDV) (Sahu et al., 2016), amongst others. An upregulation of 26SP subunit RPT4 and Ubiquitin conjugating enzyme E2 was characteristic of resistance to tomato leaf curl New Delhi virus (ToLCNDV) (Sahu et al., 2016). The MP of RNA virus TMV, undergoes polyubiquitination at the endoplasmic reticulum and followed by ER dislocation and degradation in the 26S proteasome (Reichel & Beachy, 2000). The host RING E3 ligase of *Nicotiana benthamiana* subverts viral proliferation and symptom progression by binding with and mediating ubiquitination and 26S proteasomal degradation of the TYLCCNV-βC1 (Shen et al., 2016). To offset the 26S proteasome degradation pathway, viruses have employed direct targeting of proteasome associated enzymes. Recently, susceptibility during SACMV infection was linked to the alteration of a ubiquitination E3 ligase enzymes by SACMV-induced mutations resulting in loss of function, and compromising defence (Chatukuta & Rey, 2020). The β satellite of cotton leaf curl Multan virus (CLCuMuB-βC1) and the TrAP of tomato yellow leaf curl Sardinia virus (TYLCSV) both interfere with host ubiquitination/proteasome pathway by interaction with SKP1 and COP9 signalosome (CSN5), respectively (Lozano-Durán et al., 2011; Jia et al., 2016). SKP1 and CSN5 are both essential for the formation and activity of a multi-subunit ubiquitin ligase SKP1/CUL1/F-box (SCF) complex responsible for activating JA defence signalling (Jia et al., 2016; Lozano-Durán et al., 2011). MP could target the proteasome to impede antiviral defence responses.

3.5.6.3 CARBON METABOLISM

The protein-protein interaction network generated a network centred around carbon metabolism. Cassava TME3 was sampled at a peak symptomatic time-period (32 dpi). During virus infection, interference with carbon metabolism leads to symptoms presented within the leaves (Willets & Wong, 1980). The dysregulation of carbon metabolism in geminivirus infected plants leads to an accumulation of starch in chloroplasts (Bhattacharyya et al., 2015; Zanini et al., 2021). Starch accumulation is disruptive to chloroplast ultrastructure and consequently function leading to symptoms of leaf maldevelopments noted in geminivirus (Bhattacharyya et al., 2015) and other virus infection (Zanini et al., 2021).

3.5.7 IMMUNOPRECIPITATED INTERACTORS

In mono- and bi-partite geminiviruses, ORFs that are encoded in the complementary-sense function in DNA replication, regulation of transcription and interfere with cellular processes (Gutierrez et al., 2004). MP is encoded in the c-sense in bipartite genomes such as that of SACMV (Berry & Rey, 2001). Furthermore, a multifunctional role among viral proteins is common (Rosas-Díaz et al., 2016). Therefore, it was not surprising to identify MP with proteins not only involved in potential movement but also replication and of influence in various cellular processes **(Figure 3.9, Figure 3.10, Figure 3.11)**. The identified proteins may be ideal candidates for further characterisation of MP functional mechanisms, and potential targets for virus control strategies in future. Studies of this nature typically result in the identification of hundreds of proteins and thus a more generalised interactome approach (Wang et al., 2017; Deblasio et al., 2015). This study resulted in the identification of only seven significant proteins. This low number of identified interactors enables a feasible protein by protein scrutiny of each protein and its relevance to MP function and impact SACMV infection. Of the identified interactors, GLDP2, mMDH1, SUS4 and ACLB-2 were identified as metabolic enzymes. Both GLDP2 and mMDH1 are mitochondrial enzymes. The latter, is involved in methionine metabolism whilst the former with ACLB-2 are involved in the glycolytic pathway (Fatland et al., 2002; Tomaz et al., 2010). SUS4 provides precursor molecules for metabolic pathways including those of the aforementioned enzymes (Bieniawska et al., 2007). Histone H4, HSP90 and CDC48 were identified as proteins that play key roles in DNA organisation, ubiquitination, and chaperone activity.

3.5.7.1 INVOLVEMENT OF HSP90, H4 AND CDC48 IN SACMV ACCUMULATION AND MOVEMENT

The results in this study, demonstrated that HSP90, H4 and CDC48 may be involved in MP-mediated intracellular and intercellular movement of geminiviral DNA. HSP90 is a highly abundant protein, accounting for up to 2 % of total cellular proteins and is

upregulated or down regulated relative to plant conditions (Sable et al., 2018). Also, HSP90 is involved in various cellular processes (Gaudet et al., 2011; Hubert et al., 2009). Therefore, in this study it was anticipated to identify HSPs in all treatments and negative controls. Albeit HSP90 was only significantly identified in the SACMV-Ab treatment. This may be explained as MP capture would be enriched due to antibody binding, and MP is known to bind to heat shock family proteins (Gorovits & Czosnek, 2017). The interactions between the geminivirus TYLCV and HSP70 and HSP90 in plants are necessary for virus movement and infection to proceed (Gorovits & Czosnek, 2017). Both the NSP and MP of bean dwarf mosaic virus (BDMV) also bind to histone H3, forming a highly compact viral DNA-H3-NSP-MP complex in the nucleus that facilitates the transport of the viral movement complex to the cell periphery and plasmodesmata (Zhou et al., 2011). The AbMV-MP viral movement complex associates with a heat shock cognate (70-kDa protein) which facilitates movement through stromules and plastids (Krenz et al., 2010). Studies in BDMV confirmed the interaction and co-localization of H3 and NSP, and H3 and MP to the nuclear environment and cell periphery, respectively, to facilitate systemic movement of viral DNA (Zhou et al., 2011). Geminiviruses form minichromosomes with host histones (Pilartz & Jeske, 1992; Jeske, 2009) and therefore an interaction with histones may act to compact viral DNA for movement through plasmodesmata and nuclear pores (Zhou et al., 2011). Histone H4 is a DNA binding core component of the nucleosome. The nucleosome is a histone octamer that comprises of H2A, H2B, H3 and H4 assembled in one H3-H4 heterotetramer and two H2A-H2B heterodimers (www.uniprot.org/uniprot/P59259). Considering that Histone H3 and H4 form tetramers and comparatively contain regions of statistically significant homology, most specifically in the DNA binding region (Mil' & Biniukov, 1988), and tend to share common plant and viral protein interacting partners (Kumar et al., 2020; Hye et al., 2007; Zhao et al., 2013; Rose et al., 2004), it would be feasible for the interaction of H4 with MP to occur to facilitate SACMV DNA movement similar to that involving H3. The link between H4 to intracellular viral movement and replication was also noted in the helper component-proteinase (HcPro) of potyvirus, bean common mosaic virus (BCMV), binding to H3 and H4, for viral propagation (Kumar et al., 2020). Prior work affiliating geminivirus movement to heat shock proteins, histones and NSP and findings within this study provide evidence that SACMV DNA-H4-HSP90-NSP complex may be involved in movement in infected cassava TME3 plants.

The interaction of MP with H4 within this study adds further credence to the involvement of histones and histone related proteins in geminiviral replication and plant susceptibility. Plants tolerant to geminiviral infection show a trend in this regard. Tolerant cassava TME3 exhibits a SACMV tolerant phenotype at 67 dpi (Allie et al., 2014; Bengyella et al., 2015). Whilst in TME3 there was no differential expression of H4, a down regulation of histone acetyltransferase was recorded in the recovery phenotype (Allie et al., 2014). Histone

acetyltransferases are involved in the acetylation of histones resulting in the gene activation (Black et al., 2012). In cabbage leaf curl virus (CLCV), histone acetyltransferase acetylates H3, H2A and H4 enabling viral replication and NSP-mediated translocation from the nucleus to the cytoplasm (McGarry et al., 2003). Therefore, a downregulation of histone acetyltransferase in the recovery phenotype in TME3 may counteract viral replication and movement. In the recovery phenotype of *Capiscum annum* infected with pepper golden mosaic virus (PepGMV), histone H4 and H3 were down regulated further, providing evidence of recovery linked to inhibition of histones and histone related proteins. Furthermore, histone H4 is a marker for actively mitotic cells in the S-phase and has been used to monitor TYLCV infection in tomato fruit (Just et al., 2017). Geminiviruses manipulate host replication machinery which is accessible in the cell mitotic phase (Kong & Hanley-Bowdoin, 2002). Therefore, an upregulation and availability of histones corresponds to viral replication and susceptibility.

The functional role of histones requires further species-specific investigation in the broader understanding of histone affiliation to geminivirus infection. Studies on SACMV infection in susceptible plants were not clear on the general role of histones in susceptible SACMV infection. Transcriptomic studies of geminivirus infected plants have shown conflicting *histones 3* and *4* expression in susceptible plant species. Susceptible cassava T200 demonstrated an upregulation whilst *Nicotiana benthamiana* plants showed the opposite (Allie et al., 2014; Allie & Rey, 2013). In tolerant phenotypes, cassava TME3 (SACMV infected) and pepper plants (PepGMV infected) showed a repression of *histone acetyltransferase*, and *histones 3* and *4*, respectively (Allie, et al., 2014; Góngora-Castillo et al., 2012) whilst tomato chlorotic mottle virus (ToCMoV) and chilli leaf curl virus (ChiLCV) resistant plants showed an up regulation of H4 (Carmo et al., 2017; Kushwaha et al., 2015) and H1 (Kushwaha et al., 2015).

Viruses manipulate the ubiquitin pathway to create a microenvironment that supports viral replication and persistence (Hanley-Bowdoin et al. 2013; Alcaide-Loridan & Jupin 2012). Ubiquitination has also been shown to play a role in geminivirus infection. In TYLCV infection, HSP90-mediated ubiquitin/proteasome plays a specific role in degradation of viral protein (Moshe et al., 2016). An inhibition of HSP90 in tomato plants resulted in the increased accumulation of TYLCV CP and V2 viral proteins in the nucleus. In another study, chilli leaf curl virus (ChiLCV) Rep protein was shown to bind the geminiviral genome and interacts with NbUBC2 and NbHUB1 for the monoubiquitination of histone 2B that subsequently enhanced transcription of viral genes promoting infection (Kushwaha et al., 2017).

Further support for role of HSPs in SACMV infection comes from transcriptomic studies, where the down regulation of genes encoding for HSP90 in TYLCV (Ghandi et al., 2016)

and SACMV (Allie, et al., 2014) in infected plants were associated with plant susceptibility. In *Nicotiana benthamiana* HSP90 inhibition compromised resistance to PVX,TMV and *Pseudomonas syringae* infection (Lu et al., 2003). The inhibition of HSP90 by interaction with MP may be three-fold, as it can increase the occurrence of misfolded proteins, impact the accumulation of defence proteins and destabilise the 26S proteasome, and hinder the destruction of misfolded and unassembled proteins (Imai et al., 2003; Lu et al., 2003). This could compromise the recruitment and functionality of host defence mechanisms and promote infectivity in SACMV-infected TME3. Geminivirus MPs have been well characterised as symptom determinants (Kleinow et al., 2008), and the silencing of chloroplastic HSP90 in *Nicotiana tabacum* was followed by leaf chlorosis (Bhor et al., 2017). Therefore, the interaction of MP with HSP90 may play a role in contributing to symptoms by competitively binding to and thus inhibiting HSP90 functions.

Of the identified MP-interacting proteins, HSP90 and CDC48, can also both function as molecular chaperones in facilitating ubiquitin mediated proteolysis (Imai et al., 2003; Schuberth & Buchberger, 2005) and play key roles in the accumulation of cyclins for initiation of the cell cycle (Moreno et al., 2019; Parisi et al., 2018). This is essential for geminiviral replication (Bruce et al., 2011; Hipp et al., 2014). In geminivirus infection of potato, CDC48 was 3-fold upregulated in susceptible species in comparison to that of resistance species further signifying its necessitated upregulation to geminiviral accumulation (Jeevalatha et al., 2017). African cassava mosaic virus (ACMV) studies in *Nicotiana benthamiana* showed that reconditioning of the cell cycle and initiation of RCR by Rep is most efficient in young vegetative leaves, whilst null in mature leaves, from 31 and 56 day old plants respectively, highlighting that in mature plants, appropriation of the host for RCR is a complex process that is not solely mediated by Rep and requires additional viral and plant factors (Bruce et al. 2011) and this was later demonstrated by Hipp and colleagues in identification of a putative cyclin interaction motif (RXL) in ACMV Rep protein as a clue to its replication for cells that lack the RB homologue or its expression (Hipp et al., 2014). CDC48 also interacts with H3 and H2B histones, regulating histone monoubiquitination in regulating double strand (DBS) DNA repair, and euchromatination for transcription (Mérai et al., 2014; Bonizec et al., 2014). Interacting partners HSP90, H4, CDC48 alongside SACMV MP appear to be pathogenicity determinant, promoting infection and virus movement.

Furthermore, It may be postulated that MP recruitment of HSP90 may also serve as a molecular chaperone for viral protein folding, stability and nucleocapsid assembly as noted in other plant and animal virus studies (Santoro et al., 2009). MP is a predominantly disordered protein (Nankoo, 2018) much like the hepatitis NS3 virus (Vega et al., 2013). Hepatitis NS3 requires a HSP90 chaperone complex for stabilisation (Ujino et al., 2009).

Recently, the nucleoprotein (NP) of human coronavirus required HSP90 for stability (Li et al., 2020). An interaction of this nature may be relevant to MP functionality. The chemical inhibition of HSP90 within a SACMV replicon cell culture system would be a feasible way to monitor the impact on stability/degradation of SACMV proteins (Ujino et al., 2009). Other studies have shown the manipulation of HSP90 to facilitate replication of bamboo mosaic virus (BaMV) and induction of host factors for viral transcription of human immunodeficiency virus (HIV), ebola virus (EBOV) and SARS-CoV (Huang et al., 2012; Joshi et al., 2016; Smith et al., 2010; C. Li et al., 2020). Concomitantly, from these results herein, and from other studies we propose that BC1-encoded MP recruits HSP90 as a molecular chaperone for the SACMV DNA movement complex in TME3 leading to susceptibility to SACMV by, not only promoting virus movement, but also enhancing viral DNA transcription at the systemic infection stage (32 dpi).

3.5.7.2 CDC48-MP CROSSTALK BETWEEN VIRUS REPLICATION AND SYMPTOM AMELIORATION

CDC48 is localised to the nucleus, golgi apparatus, cytoskeleton and the cytoskeleton system involved in cell wall assembly (phragmoplast). MP is required for intracellular and virus cell-to-cell and long-distance movement, through its cooperative interaction with the nuclear shuttle protein (NSP) in bipartite begomoviruses (Noueiry et al., 1994; Ward et al., 1997). Existing data supports a model where viral DNA is exported from the nucleus by NSP, and the MP located at the cytoplasmic face of nuclear cell membranes or microsomal vesicles binds to NSP–DNA-HSP complexes and enables transfer to adjacent cells via the ER/cytoskeleton network (Kleinow et al., 2009). CDC48 therefore may facilitate cytoplasmic movement of the SACMV DNA- MP-NSP complex.

CDC48 also functions in an ATP dependent manner to fulfil roles involved in biological processes associated with mitosis and protein transport (Gaudet et al., 2011). By similarity, CDC48 has been associated with the AAA-ATPase complex required for the efficient dislocation of ER-lumenal degradation substrates, and their subsequent proteolysis by the proteasome (Schuberth & Buchberger, 2005) and other processes involving autophagosome maturation, mitotic spindle disassembly, retrograde protein transport and ER to cytosol (Gaudet et al., 2011) and plant cytokinesis (Rancour et al., 2002). CDC48 proteins commonly function as a CDC48 protein complex (segregase) with its heterodimeric cofactor UFD1-NPL4 (Mérai et al., 2014; Moreno et al., 2019). The CDC48 complex plays a key role in the nuclear accumulation of cell cycle initiator proteins and ribosomal DNA (rDNA) to the nucleolus for the activation of mitosis and ribosomal RNA genes respectively, and euchromatisation of DNA thus enhancing transcription and protein synthesis (Mérai et al., 2014; Parisi et al., 2018). CDC48 complex releases and prevents degradation of ubiquitinated Cln3 and G1 and D1 cyclins promoting their nuclear

localisation and accumulation to induce the start of the cell cycle (Parisi et al., 2018). In this case, it may be that MP is involved via binding to CDC48 for movement, viral DNA transcription and protein synthesis, interference with the cell cycle and interference with the ubiquination/proteasome pathways contributing to susceptibility/infection.

CDC48 plays a key role as a molecular chaperone for the endoplasmic reticulum (ER)-associated protein degradation (ERAD), a ubiquitin proteasome system that involves the removal of polyubiquinated proteins from the endoplasmic reticulum to the 26S proteasome (Bègue et al., 2019; Schuberth & Buchberger, 2005; Jarosch et al., 2002). The inherent trafficking role of geminivirus MPs and association with the ER, makes MPs such as SACMV MP and/or its chaperones likely targets of ERAD after their transport role is accomplished. The MP of TMV, interacts with CDC48 subsequently leading to its degradation (Niehl et al., 2012). This channelling to the 26S proteasome may also contribute to the low abundance of geminivirus MPs (Kleinow et al., 2009; Kleinow et al., 2009; Alcântara et al., 2019). Geminivirus MPs additionally are symptom determinants (Stanley & Albrecht, 1992; Erica et al., 1993; Hou et al., 2000). The accumulation of symptom determinants is directly proportional to symptom severity and therefore, degradation may therefore play a role in ameliorating symptoms and thus enhancing host longevity for virus survival (Kleinow et al., 2009).

3.5.7.3 GLDP2 INTERACTION WITH MP MAY INTERFERE WITH SULPHUR LINKED HOST DEFENCE MECHANISMS

GLDP2 is a mitochondrial (Heazlewood et al., 2004) RNA-binding enzyme involved in regulating the RNA-binding proteome (RBPome) (Bach-Pages et al., 2020) and is an essential enzyme in photorespiration but has shown its greatest relevance in the non-photorespiratory pathways (Engel et al., 2007). GLDP proteins comprise GLDP1 and GLDP2. GLDP2 shares 91 % sequence identity with GLDP1 and both are abundant in photosynthesising organs and are complementary in function (Engel et al., 2007). It has been shown through lethality of GDLP1/2 double mutants in *Arabidopsis* that GDLP proteins are obligatory and indispensable enzymes for the conversion of glycine to serine in non-photorespiratory plant metabolism in all plant tissues (Engel et al., 2007). GLDP proteins are therefore essential to TCA cycle reactions, the G6P/OPP shunt, fatty acid biosynthesis and starch and sucrose synthesis (Xu et al., 2021). The complementarity of GLDP proteins did not result in any significant changes in single mutants relative to wild type *Arabidopsis* (Engel et al., 2007). Studies that evaluate single mutants for pathogen responses are currently not available. Limited to information on healthy mutant plants, the direct or indirect interaction of MP with GLDP2 significantly altering plant metabolic processes or photosynthesis remains in question.

GLDP2 belongs to the class of methionine cycle (MTC) enzymes which are targeted by viral proteins in the negation of host antiviral responses during geminivirus and potyvirus infection (Mäkinen & Swarnalok, 2019). Multiprotein assemblies of methionine cycle enzymes are formed during biotic and abiotic stress (Mäkinen & Swarnalok, 2019). GLDP2 forms part of MTC enzymes that are connected to glucosinolate biosynthesis (Mäkinen & Swarnalok, 2019). The down regulation of glucosinolate metabolism has been further identified as a repression of sulphur containing compounds that play a pivotal role in host antiviral mechanisms (Lozano-Durán et al., 2012). The over-expression of TYLCCNV βC1 and TYLCV C2 in *Arabidopsis* resulted in the down regulation of MYC2 regulated synthesis of glucosinolates (Li et al., 2014; Lozano-Durán et al., 2012). Therefore, the identification of GLDP2 potentially interferes with plant metabolites relevant to antiviral defence.

3.5.7.4 SUS4 INTERACTION MAY BE AN ANTIVIRAL MECHANISM OR RESULT IN DYSREGULATION OF MULTIPLE METABOLIC PATHWAYS

SUS4 catalyses the cleavage of sucrose to fructose and UDP-glucose (Bieniawska et al., 2007) and these products then feed into cell metabolic pathways (Stein & Granot, 2018). Glucose and fructose are phosphorylated and channelled to mitochondria for glycolysis and into the plastids for glycolysis, shikimate pathway and starch biosynthesis (Stein & Granot, 2018). Thereafter, intermediates and final product (pyruvate) of the glycolytic pathway provide precursor molecules for amino acid and lipid metabolism (Baxter et al., 2007; Stein & Granot, 2018). The metabolome of TYLCV infected tomato revealed a direct relationship between pyruvate availability and its sufficiency in providing energy for plant fitness in resistant genotypes (Moshe et al., 2012).

Inhibition of sucrose synthase by interaction with MP may cause dysregulation of plant functions due to interference in metabolic pathways. Plant geminivirus infection reduces plant yield but does not lead to plant death (Allie et al., 2014; Bisaro, 1996; Moshe et al., 2016). The existing supplementary role of invertases in sucrose cleavage may play a role in minimising the detrimental effect of the MP-SUS4 interaction and allowing some level of plant sustainability (Stein & Granot, 2018). Furthermore, geminiviruses may have adapted to targeting non-lethal targets to benefit their replication and survival within the host. Survival features of geminiviruses have been observed in inhibiting plant cell death responses (Kumar, 2019).

The interaction of SUS4 with MP may also account for poor accumulation of starch in cassava tubers during geminivirus infection (Varma & Malathi, 2003). High sucrose synthase activity in leaf tissue increases the source to sink accumulation of starch in cassava tubers (Li et al., 2016). SnRK1 has been identified as an anti-geminiviral protein which when overexpressed increases plant resistance to geminiviruses (Shen et al., 2011; Shen et al., 2009; Shen et al., 2012) and starch accumulation (Li et al., 2016). Therefore,

starch biosynthesis from this perspective strongly correlates to host resistance and an interference of this pathway much like that of lipid metabolism could render the plant susceptible to attack. Furthermore, efficient starch storage efficiency was identified in cassava plants with a strong xylem architecture (Li et al., 2016) and may be attributed to sucrose synthase role contributing to cellulose biosynthesis in carbon partitioning (Coleman et al., 2009). An interference of cellulose biosynthesis pathways and silencing of SUS4 contributed to features of plant stunting and chlorosis respectively (Choe et al., 2021; Yao et al., 2020). South Africa cassava mosaic virus-MP targeting of SUS4 could contribute to symptoms in TME3 at 32 dpi. A down regulation of cellulose biosynthesis induced by TYLCV infection in tomato has recently been linked to plant stunting and reduced plant size (Choe et al., 2021). SUS4/SUS1 mutants presented stunting and an overt accumulation of carbohydrates resulting in chlorosis (Yao et al., 2020). Accumulation of carbohydrates in leaf tissue occurs due to restriction in carbohydrate translocation (Osmond et al., 1998). Plant viruses have been shown to influence an upsurge of sugars to possibly meet the energy demands required for virus replication and proliferation and an interference with SUS4 suggests a mechanism by which this is achieved (Willets & Wong, 1980; Wang & Maule, 1995; Thaker et al., 2019; Kogovšek et al., 2016). Considering the beneficial role of SnRK1 to the host anti-viral mechanisms, and the SACMV-tolerant phenotype of TME3, it would be valuable to investigate the influence of SUS4 silencing and over expression in SACMV infection.

3.5.7.5 MP-ACLB-2 INTERACTION MAY INTERFERE WITH FATTY ACID MEDIATED DEFENCE RESPONSES

ACLB-2 is a cytosolic enzyme involved in the primary synthesis of acetyl-CoA, a primary precursor for the synthesis of C18-C24 fatty acid biosynthesis (Fatland et al., 2002). The concept of fatty acid biosynthesis in plant defence has been well established (Weber, 2002; Lim et al., 2017; Laxalt & Munnik, 2002). The fatty acid biosynthetic pathways are essential to development and defence as they provide precursors for signalling molecules essential to induction of hormone signalling networks, effector and PAMP-triggered biotic stress responses, multiple abiotic stress responses and feed into cytoskeleton and cell wall integrity (Lim et al., 2017). A down regulation of fatty acid biosynthesis is commonly noted in geminivirus and potyvirus susceptibility (Allie et al., 2014; Pierce & Rey, 2013; Miozzi et al., 2014; Jeevalatha et al., 2017; Ascencio-Ibanez et al., 2008; Amuge et al., 2017; Maruthi et al., 2014). Fatty acid biosynthesis and its influence on plant defence has been strongly linked to the jasmonic (JA) biosynthesis pathway which involves the initial peroxidation of fatty acid precursor molecule linolenic acid followed by multiple beta oxidation steps (Wasternack, 2007; Weber, 2002) and phosphatidyl inositol and its derivatives (Lim et al., 2017). The exogenous application of JA in plants compromised by geminivirus (C2 of TGMV) infection has the capacity to alleviate symptoms (Lozano-Durán et al., 2012).

Cassava TME3 is tolerant to SACMV, therefore, it was interesting to note that in pepper plants that exhibit tolerance to PepGMV, jasmonate related gene expression was noted from the early onset symptomatic phase and the later phase of recovery (Góngora-Castillo et al., 2012). Interestingly, in SACMV-infected TME3 the upregulation of several JA genes was noted at early infection (12 dpi) but was not in the susceptible T200 landrace. At the onset of the symptomatic phase (32 dpi) both T200 and TME3 exhibited an increase in expression of several genes related to biosynthesis and signalling of JA (Rey, unpublished). JA is also induced by oxidative and biotic stress (Allie et al., 2014). Geminiviral interaction with jasmonate biosynthesis and signalling is complex. The C2 of tomato yellow leaf curl Sardinia virus (TYLCSV) hinders selective aspects of JA perception, defence and lipid transport but not response to oxidative stress (Lozano-Durán et al., 2011; Rosas-Díaz et al., 2016). Long chain fatty acid biosynthesis also plays a key role in sphingolipids, cutin and wax biosynthesis essential to structural integrity (Rowland et al., 2006; Lü et al., 2009; Zheng, 2005). Structural integrity is an essential mechanism in plant defence and the down regulation of related genes is a core common feature of susceptibility in geminivirus infection (Ascencio-Ibanez et al., 2008; Miozzi et al., 2014; Jeevalatha et al., 2017; Pierce & Rey, 2013). The interaction of MP with ACLB-2 may interfere with the initial onset of plant defence dependent on fatty acid biosynthesis, cell architecture, energy utilisation and metabolic signalling. The identification of an association that directly targets carbon metabolism and fatty acid biosynthesis is very novel to geminivirus MPs and warrants further investigation.

3.5.7.6 mMDH1 INTERACTION WITH MP MAY ALTER CARBON PARTITIONING, ENERGY METABOLISM AND NUTRIENT AVAILABILITY

mMDH1 catalyses the conversion of NAD-dependent dehydrogenase reactions in a reversible manner and is involved in biological processes of, biotic and abiotic stress response. mMDH1 operates in the TCA cycle and converts malate to oxaloacetate in a reversible manner, thus providing precursors for acetyl CoA, carbon dioxide and pyruvate and regulates carbon partitioning (Yoshida & Hisabori, 2016; Tomaz et al., 2010; Hüdig et al., 2014; Lindén et al., 2016; Tan et al., 2010; Kawamura & Uemura, 2003; Sarry et al., 2006; Grant et al., 2006). mMDH occurs in two homologous forms;mMDH1 and mMDH2 (Tomaz et al., 2010). mMDH2 activity is 28 % and 67 % lower than mMDH2 activity on substrates, oxaloacetate and malate respectively (Hüdig et al., 2014). Therefore, an interference with mMDH1 activity by interaction with MP may have a greater impact. *Arabidopsis* mutants, *mmdh1* and *mmdh2* had reduced growth rate, an elevated respiratory rate and reduced photorespiratory flux due to a reduction in carbon assimilation (Tomaz et al., 2010). Increased respiration rates in early stages of infection and within systemic infection has long been confirmed in plant virus interactions (Loebenstein, 1959; Bates & Chant, 1975). It is proposed that MP binding with mMDH1 in cassava TME3 contributes to

a general biotic stress response, leading to disease symptoms and reduced growth at the onset of the systemic infection stage (32 dpi). In order to reduce the oxidative burst from photorespiration, it has been postulated that plant viruses lower the photorespiratory flux (Souza et al., 2019). TYLCV and tomato chlorosis virus (ToCV) transcriptomics showed that reduced growth and chlorosis were due to the alteration of carbon assimilation by a reduction in multiple metabolic enzymes of the photorespiration, Calvin cycle and photosynthesis pathways (Seo et al., 2018). Furthermore, expression of AC2 of tomato chlorotic mottle virus (ToCMoV), led to down regulation of cytosolic mMDH (cMDH) (Carmo et al., 2013). Therefore, interaction of MP may competitively interfere with mMDH1 function and contribute to symptom progression and a dysregulation of carbon assimilation noted in virus compatible studies. It is proposed that mMDH1 interaction with MP could potentially alter the methionine cycle and alter amino acid biosynthesis, secondary metabolites and methylation products involved in regulation of gene expression (Mäkinen & Swarnalok, 2019). It would be interesting to evaluate mMDH1 and mMDH2 activity in tolerant and susceptible cassava genotypes to fully determine the impact of interfering with either one of these proteins by SACMV in tolerant and susceptible cassava.

3.6 SUMMARY

In summary, MP interaction with cellular components linked to several metabolic pathways and chaperones, may contribute not only to movement of SACMV DNA complex, but disease symptoms in TME3 at 32 dpi. While TME3 is tolerant to SACMV and recovers at 67 dpi, it does nonetheless exhibit symptoms and virus replication at 32 dpi when this study was performed, albeit lower severity and virus load compared to the susceptible T200 landrace. **Figures 3.12 and 3.13** are brief illustrations highlighting the impact and interconnectivity of identified interaction partners. Mitochondrial enzyme GLDP2 has not been included in the schematics. GLDP2 interaction with MP would potentially alter the methionine cycle and alter amino acid biosynthesis, secondary metabolites and methylation products involved in regulation of gene expression (Mäkinen & Swarnalok, 2019).

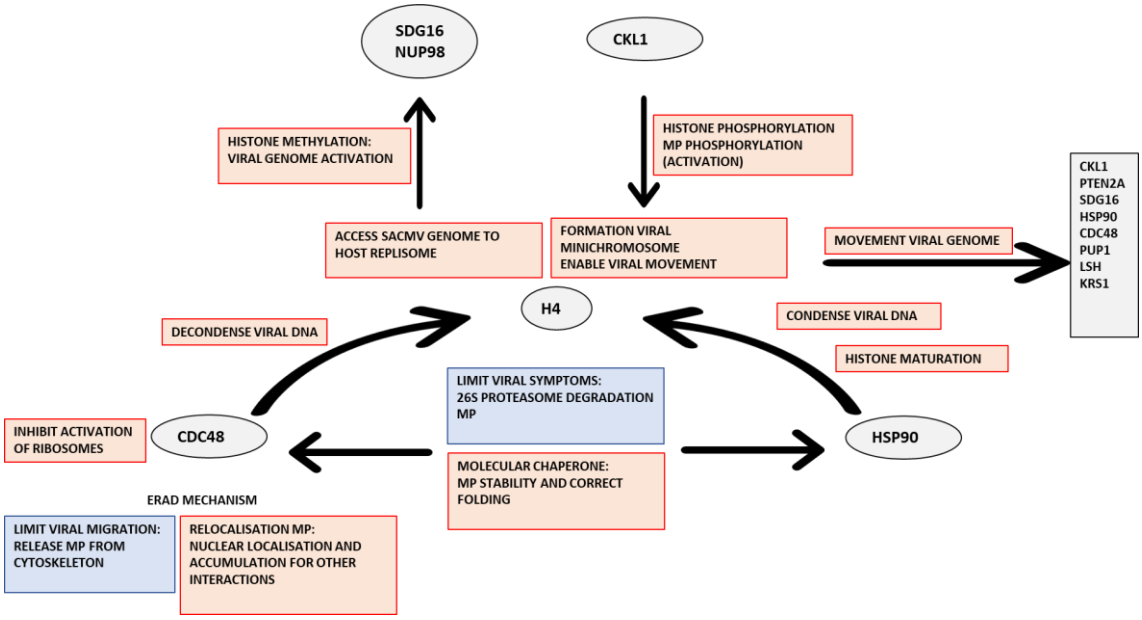


Figure 3.12 A proposed model of the impact from the interaction of MP with HSP, CDC48 and H4 proteins resulting in manipulation/alteration of host protein activity or an anti-viral response. Host proteins are shown in grey, manipulation/alteration in favour of SACMV in red and anti-viral host response in blue. Molecular chaperones HSP90 and CDC48 contribute to the stability and folding of MP and limit viral symptoms by degradation of MP via the 26S proteasome. The ERAD mechanism of CDC48 further limits viral migration by release of MP from the cytoskeleton system and relocates MP to the nucleus and accumulation of MP for other interactions. CDC48 interaction with MP inhibits CDC48 mediated activation of ribosomes essential to relegation of proteins for an antiviral response. Both CDC48 and HSP90 interact with histone H4 to decondense or condense viral DNA respectively. The latter resulting in the access of viral DNA to the host replisome whilst the former enables the formation of viral minichromosomes for viral DNA migration.

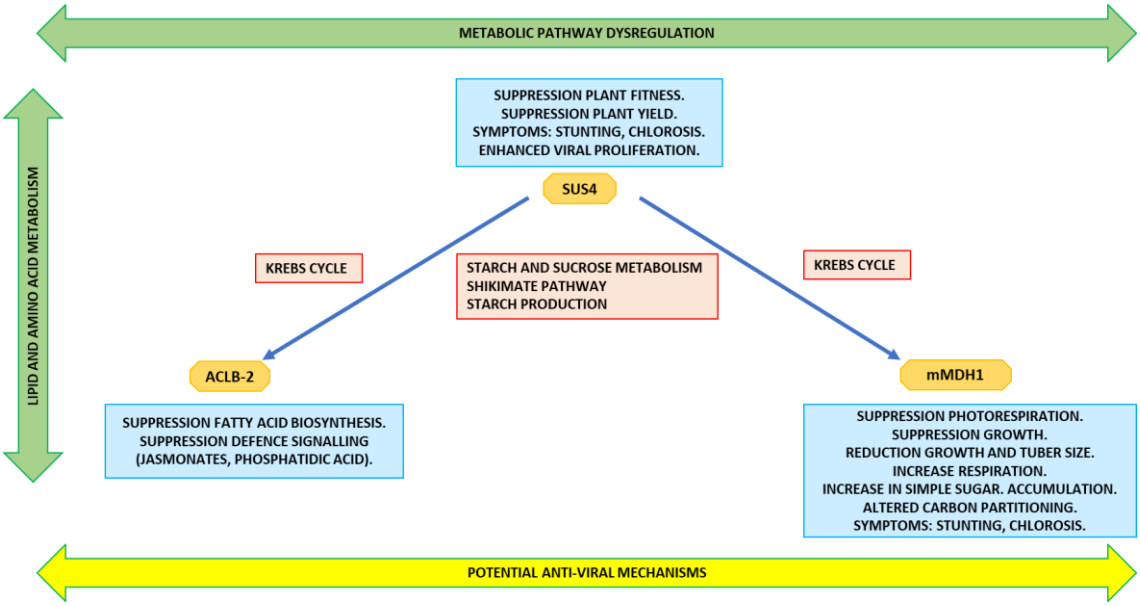


Figure 3.13 A proposed model of the host metabolic enzymes ACLB-2, mMDH1 and SUS4 and the potential impact of interaction with MP and the metabolic pathways impacted. Metabolic enzymes are shown in light orange. Limitations arising from interaction with MP for each enzyme shown in blue. Overall pathway dysregulation shown in green. All enzymes may have the potential for anti-viral mechanisms as discussed and this is shown in yellow. SUS4 provides precursor molecules for metabolic pathways (starch and sucrose metabolism, shikimate pathway, starch production and the KREBS cycle). Interference of SUS4 activity by MP would result in the suppression of plant fitness and yield and would enhance SACMV symptoms of stunting and chlorosis. SUS4 may translocate viral DNA and MP via the cytoskeleton and in an ATP dependent manner and enhance viral proliferation. ACLB-2 and mMDH1 are both associated with the KREBS cycle. ACLB-2 interference by MP would suppress fatty acid biosynthesis and defence signalling mediated by fatty acid derived jasmonate and phosphatidic acid. mMDH1 interference would result in suppression of photorespiration, growth and tuber yield and an increase in simple sugars, respiration, and an alteration of carbon partitioning. mMDH1 interaction with MP would contribute to chlorosis and stunting noted in SACMV infection of cassava.

Chapter 4

THE EVALUATION OF SACMV INFECTION IN CASSAVA TME3 PROTOPLASTS TRANSFORMED BY CRISPR MEDIATED GENE KNOCKDOWN

4.1 INTRODUCTION

Gene editing plays a pivotal role in identifying the impact of silencing or overexpression of a specific gene on a particular feature, for example viral replication. Conventional methods such as virus induced gene silencing (VIGS) delivered via *Agrobacterium* have proved effective in gene silencing but involve the propagation and transformation of plants in a time consuming and laborious manner (Mwaba, 2011; Lentz et al., 2018; Courdavault & Besseau, 2020). Previous studies in our laboratory using cassava and *Nicotiana benthamiana* entailed a minimum of 8 and 16 weeks respectively from propagation to sample collection (Mwaba, 2011; Mwaba & Rey, 2017). A high throughput cassava protoplast system using clustered regularly interspaced short palindromic repeats (CRISPR) Cas9 mediated gene knockdown was recently introduced as an effective and rapid alternative (Chatukuta & Rey, 2020). Chatukuta and Rey showed an increase in SACMV viral load in cassava protoplasts simultaneously transfected with SACMV and CRISPR/Cas mediated silencing of E3 ubiquitin ligase, and also demonstrated induced mutations in the E3L sequence due to gene editing (Chatukuta & Rey, 2020). In plants such as cassava that are recalcitrant to gene manipulation, the pairing of protoplasts with CRISPR/Cas systems in plant-virus interaction studies provides an easy, simplistic and rapid method for gene manipulation and evaluation (Chatukuta & Rey, 2020; Jiang & Doudna, 2017).

The concept of direct transfer of an expression plasmid into protoplasts without the use of an *Agrobacterium* delivery system has made the use of protoplasts a more amenable and preferred solution for transient transformation studies (Potrykus et al., 1985; Krens et al., 1982). Initial studies involved the transformation of monocot and dicot derived protoplasts either by electroporation (Sofiari, 1996; Fromm et al., 1985; Jones et al., 1989) or polyethylene glycol (PEG) mediated delivery of DNA (Maas et al., 1995; Sofiari, 1996). The transformation and regeneration of transformed protoplasts has been successful in many plant species (Hsu et al., 2021; Guo et al., 2005; Jeong et al., 2021; Caisová & Jobe, 2019; Kang et al., 2020; Andersson et al., 2017; Lin et al., 2018) but cassava is highly recalcitrant to transformation which is also cultivar-dependent. Comparatively fewer transformation studies in cassava have been reported (Wen et al., 2020). PEG mediated delivery has taken precedence in recent years due to ease of use, amenability to multiple samples, and efficiency range from 50-90 % (Wu et al., 2017; Shen et al., 2014; Chatukuta & Rey, 2020). The ease of use and capacity to process multiple samples may make amends for studies

that present low efficiencies. In the absence of chemical or mechanical enhancers for delivery, the mechanism of transfection of protoplasts with viral genomes has been postulated to occur by either pinocytosis or damaged regions of the cell membrane (Burgess et al., 1973). Protoplasts have further been appreciated as a means to amplify and harvest viable and packaged plant viruses (Kikkawa et al., 1982; Lee et al., 2001; Rottier, 1980). Viral replication occurs from 6 h post transfection and increases exponentially from 24 to 48 h post transfection with a decline at 72 h post infection (Shadwick, 2007; Kikkawa et al., 1982). The viability and replication of viral DNA in protoplasts presents protoplasts as viable and reliable representatives of plant responses to virus infection (García-Neriaand & Rivera-Bustamante, 2010; Méndez-Lozano et al., 2003; Chatukuta & Rey, 2020).

The use of the CRISPR/Cas9 system in protoplasts to induce gene knockout, via double strand breaks (DSBs) to promote the occurrence of mutations arising from erroneous non-homologous end-joining (NHEJ) has been an effective silencing tool in cassava (Chatukuta & Rey, 2020; Wu et al., 2017), maize (Sant'ana et al., 2020), *Brassica* (Murovec et al., 2018), wheat (Brandt et al., 2020) and tobacco (Lin et al., 2018). The CRISPR/Cas system provides adaptive immunity to bacteria and archaea against viruses and plasmids and was initially derived from *Streptococcus pyogenes* and programmed as a genome editing tool for sequence specific gene targeting (Jinek et al., 2012) **(Figure 4.1)**. In the CRISPR/Cas system a guide RNA (gRNA) sequence that is complementary to a 20 nucleotides (protospacer) for target DNA facilitates complementary binding pre-cleavage by Cas enzyme (Jinek et al., 2012). The sgRNA region is a fusion product of CRISPR RNA (crRNA) and trans-activating CRISPR RNA (tracrRNA) that are involved in complementary binding and recruitment of the Cas enzyme (Jinek et al., 2012). Located downstream of the complementary target region is a mandatory protospacer-associated motif (PAM), which is prerequisite and specific to each Cas for double strand cleavage activity via its RuvC-like nuclease (RuvC-like) domain and the histidine-asparagine-histidine (HNH) motif (Jinek et al., 2012).

There are two classes (class I, class II) of CRISPR/Cas systems further distinguished into six distinct Cas enzyme types (I-VI) (class I: types I, II, IV; class II: types II, V, VI) (Liu et al., 2020). The class I systems functions as protein complexes whilst class II, utilises a single Cas enzyme and is therefore the simplest and ideal for genome editing in the CRISPR/Cas system (Jiang & Doudna, 2017). Cas derived from different species are currently in use and each recognises a specific PAM sequence i.e. *Staphylococcus aureus* Cas9 (SaCas9), *Francisella novicida* Cas12a (FnCas12a), *Streptococcus pyogenes* Cas9 (SpCas9), *Streptococcus thermophilus* (Sth1Cas9) *Lachnospiraceae bacterium* ND2006 (LbCas12a), *Actinomyces naeslundii* Cas9 (AnaCas9) and nCas9-activation-induced

cytidine deaminase (nCas9-Target-AID) (Hsu et al., 2021; Schiml et al., 2014; Schiml et al., 2016; Schindele & Puchta, 2020; Jiang & Doudna, 2017; Steinert et al., 2017). To increase amenability of CRISPR/Cas systems for a plants of variable growth temperature requirements, temperature-tolerant ttLbCas12a was engineered for efficiency at 22 ˚C and 28 ˚C (Schindele & Puchta, 2020). Complementary base pairing induces conformational changes of the Cas activating PAM recognition and unwinding of the target DNA for cleavage (Jinek et al., 2012).

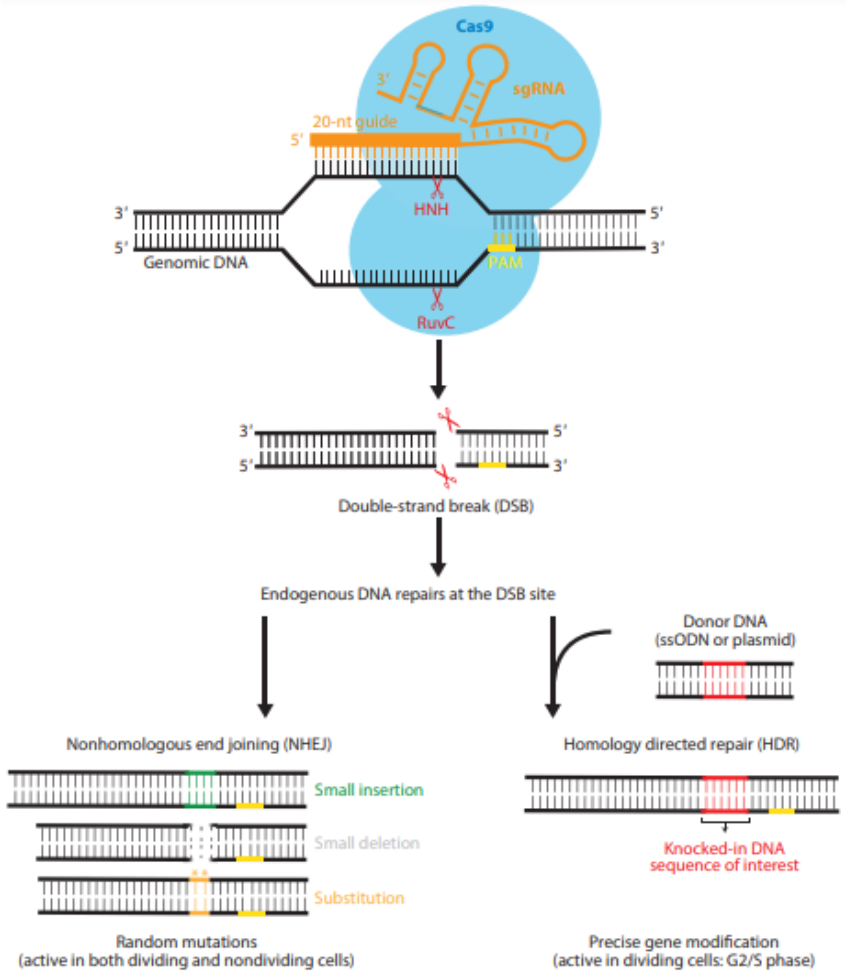


Figure 4.1 A schematic representation of RNA-guided double-stranded DNA cleavage by CRISPR-Cas using programmable guide RNA. In the CRISPR/Cas system a guide RNA (gRNA) sequence that is complementary to 20 nucleotides (protospacer) for target DNA facilitates complementary binding and cleavage by Cas resulting in double strand breaks (DSB) that are repaired by either non-homologous end joining (NHEJ), or homology directed repair (HDR) resulting in genetic transformations. (Adapted from Jiang & Doudna, 2017).

The introduction of double strand breaks (DSBs) by Cas9 leads to one of four outcomes, homologous recombination (HR), nonhomologous end joining (NHEJ), alternative end joining (alt-EJ) and single-strand annealing (SSA) (Ceccaldi et al., 2016). The former allows the inclusion of a decoy template complementary to the target region but with an additional sequence to alter the final replication product of the target region homology-directed repair (HDR) (Chu et al., 2015). The latter three do not require homologous DNA and are thus more error prone producing outcomes that lead to mutational changes such as insertions, deletions, substitutions and chromosomal aberrations and results in gene silencing or loss-of-function and therefore relevant to CRISPR/Cas systems (Ceccaldi et al., 2016).

The applicability of CRISPR/Cas is widely demonstrated from mammalian studies for disease modelling to biomedicine and agricultural enhancements (Zhao et al., 2019). Initial induction of mutations in plants and protoplasts was induced by HDR and NHEJ in *Nicotiana benthamiana* and *Arabidopsis* (Li et al., 2013). NHEJ has been identified as the predominant mechanism resulting in various mutational outcomes in plant studies (Li et al., 2013). The targeting of at least two regions in the target DNA is essential to increase mutation frequencies which may range from 1 % to 50 % in an individual gRNA (Li et al., 2013; Chu et al., 2015; Chatukuta & Rey, 2020; Lin et al., 2018; Brandt et al., 2020) and exceed 80 % in other instances (Lin et al., 2018). The use of two or more different and independent specific Cas enzymes has also been effective in targeting multiple DNA regions (Steinert et al., 2017).

A basic shortfall of CRISPR/Cas systems is the occurrence of off-targets arising from the promiscuity of DNA binding in regions of only five bases in sequence similarity (Jiang & Doudna, 2017; Yanfang et al., 2013). In this case the use of Cas enzymes that identify longer PAM sequences has been identified as a possible way to increase target specificity (Steinert et al., 2017). Traditional SpyCas9 identifies NGG/NAG whilst Sth1Cas9, and SauCas9 PAM sequences target NNAGAA/ NNGGAA and NNGRRT/ NNGRR(N) respectively (Steinert et al., 2017). The use of more complex Cas types that require the formation of a Cas complex may be a way to reduce mismatches and improve a correlation between binding and induction of activity in a target specific manner (Osakabe et al., 2020). Cas type I subtype CRISPR TiD, comprising of five Cas enzymes for recognition, binding and cleavage activity showed a higher stringency in effecting cleavage relative to mismatches in mammalian cells and tomato (Osakabe et al., 2020). A minimum of three mismatches within the sgRNA completely abolished TiD activity (Osakabe et al., 2020). The selection of a promoter to drive sgRNA or Cas expression may play a significant role in increasing transformation and mutation efficiency (Sun et al., 2015). Overexpression of either sgRNA or Cas9 may result in the increased occurrence of off-targets (Sun et al., 2015). Most sgRNA expression vectors are often driven by the U6 RNA polymerase-III

promoters from either *Arabidopsis* or respective to the plant species under consideration (Andersson et al., 2017; Sun et al., 2015; Schindele & Puchta, 2020; Nadakuduti et al., 2019; Chatukuta & Rey, 2020). Cas may be driven typically by a 35S promoter of cauliflower mosaic virus or tobacco mosaic virus promoter or other ubiquitin promoter of plant origin (Schindele & Puchta, 2020; Andersson et al., 2017; Chatukuta & Rey, 2020).

With regards to this study, three genes identified from Y2H screening work reported in Chapter 2 were targeted for CRISPR-mediated gene editing to evaluate response to SACMV in cassava TME3 protoplasts.

4.1.1 AIM AND OBJECTIVES

The main aims of this study were to validate in vivo the function the impact of CRISPR mediated gene knockdown of cassava host genes previously identified to interact with SACMV MP by in vitro Y2H screening (*Manes.16G117000-CKL1, Manes.14G006400-SDG16* and *Manes.01G002100-NUP98*). The influence of gene knockdown was evaluated by the effect on replication of SACMV replication in cassava TME3 protoplasts. If the targeted MP interactors were involved in virus accumulation, then knocking out their function was hypothesised to hamper replication of SACMV *in vivo* in protoplast cells.

The main objectives were:

1. To evaluate CRISPR mediated gene knockdown in cassava TME3 protoplasts based on gene expression and gene editing.

2. To evaluate SACMV replication in CRISPR edited and non-CRISPR edited cassava TME3 protoplasts.

Figure 4.2 gives a schematic overview of the experimental procedure.

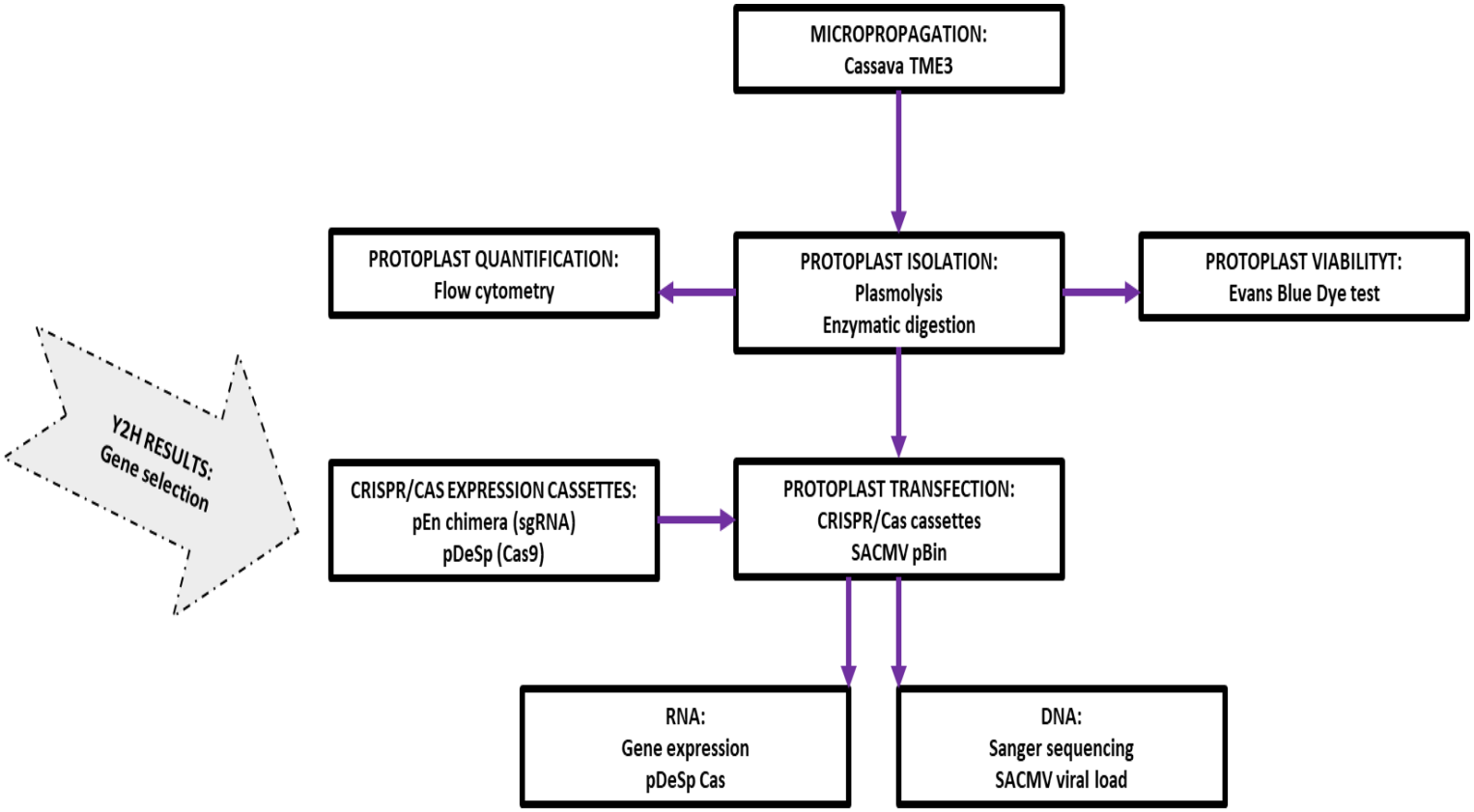


Figure 4.2 A schematic representation of the main steps in the experimental procedure taken to evaluate protoplast transfection with SACMV and CRISPR/Cas expression clones.

4.2.1 CRISPR-CAS9 VECTORS

CRISPR-Cas9 (*Streptococcus pyogenes*) cassettes comprising of Gateway entry pEn-Chimera (Addgene plasmid # 61432; http://n2t.net/addgene:61432; RRID: Addgene_61432) and Gateway destination pDe-Cas9 (pDeSp) vector (Addgene plasmid # 61433; http://n2t.net/addgene:61433; RRID: Addgene_61433) **(Figure 4.3, Figure 4.4)** vectors were donated by Holger Puchta (Schiml et al., 2016). The pEn-Chimera carries the sgRNA expression system flanked by attR sites to allow for Gateway cloning into pDeSp via its *ccdB* gene. All cloning and confirmation of cloned products was carried out according to the established protocol of the Holger Puchta laboratory (Fauser et al., 2014).

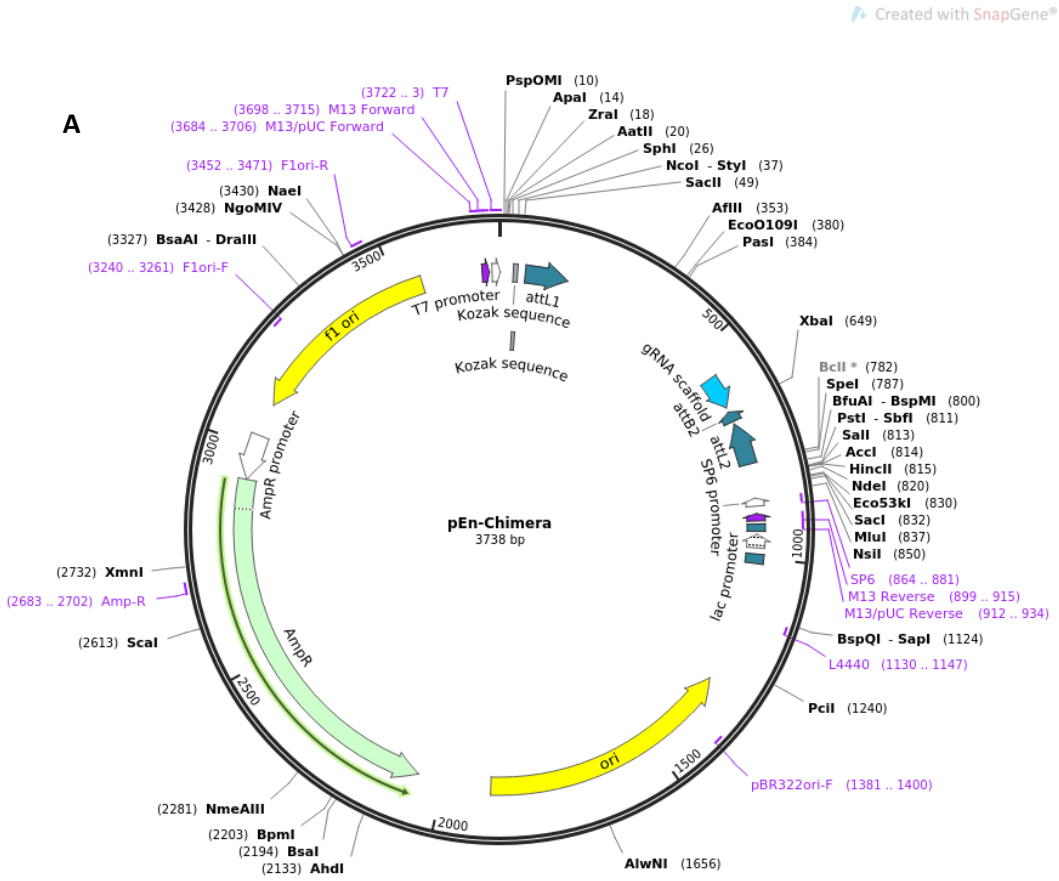


Figure 4.3 The CRISPR/Cas cassette system. The pEn chimera vector construct map provided by the Holger lab.

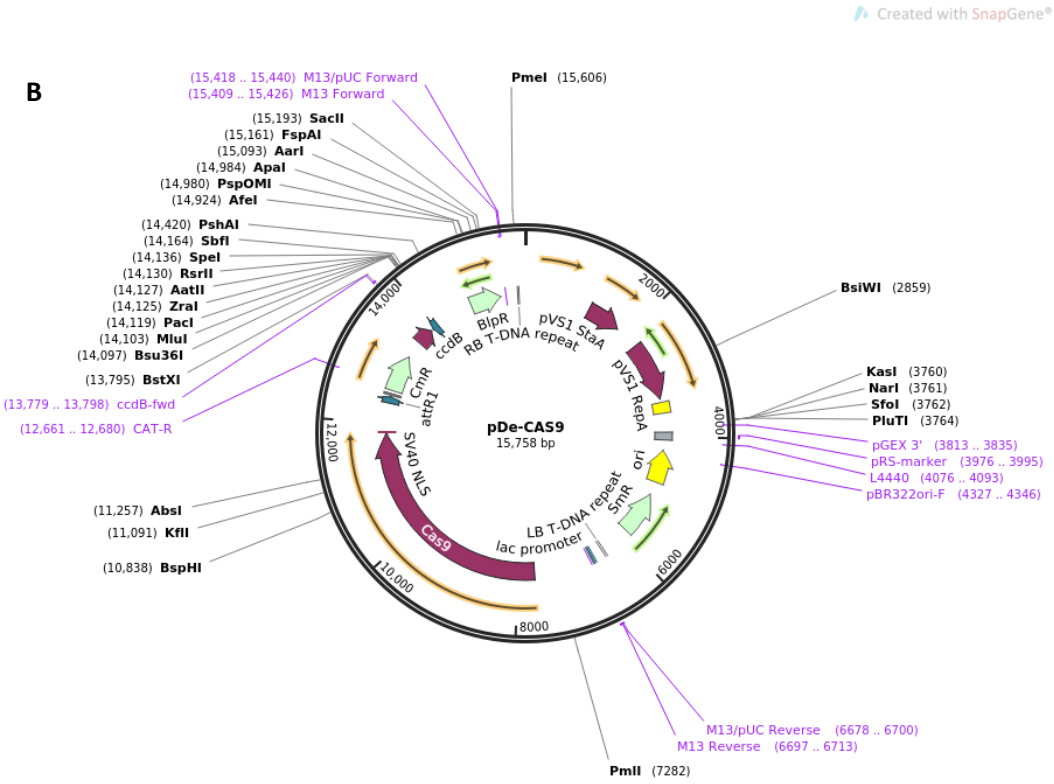


Figure 4.4 The CRISPR/Cas cassette system. The pDeSp vector map construct map provided by the Holger lab.

From the transcripts identified by Y2H screens (Chapter 2), three were selected for single CRISPR-Cas9 gene knockdown. These were ranked with PBS categories A, A (very high confidence) and D (moderate confidence), respectively. These were identified using *Manihot esculenta* at Phytozome (Bredeson et al., 2016) according to the v6.1 genome assembly i.e., *Manes.16G117000.1, casein kinase 1-like protein HD16 (CKL1/CK/C), Manes.14G006400.1, histone-lysine n-methyltransferase atx4-related (SDG16/H) and Manes.01G002100.1, nucleoporin autopeptidase (NUP98/N)* and will be abbreviated as shown in parenthesis options for simplicity. For each interactor, an amino acid sequence commonly shared by all positive clones for each specific identity of prey fragments (Selected Interaction Domains (SID)) was annotated according to the Pfam ID (Mistry et al., 2021) i.e., for SDG16, SET domain (PF0856, SM00317, SM0058), for CKL1, protein kinase domain (SM00220, PF0069) and for NUP98, peptidase S59, nucleoporin domain (PF04096).

For each gene of interest, two 20-nucleotide CRISPR spacer sequences (oligos) were designed targeting separate exonic sites within the gene of interest using CCTop software (Stemmer et al., 2015) with respect to the *S. pyogenes* (PAM) sequences (GGG, AGG, TGG, or CGG) and annotated in the format, fwd-5'-ATTG + protospacer and rev-5'-AAAC

+ rev-com protospacer for forward and reverse, respectively. Protospacer sequences were as follows,

- CKL1 (Fwd: '5-GGTGGTGGAGTACGTGGTCGGGG-3',

  Rev: '5-CCAGTGAGGGGATCCATAGAAGG-3')

- SDG16 (Fwd: '5-CCACTTGTGCGGACGTCCAGAGG,

  Rev: '5-CTTGTCGGAATCCACGTTGGAGG-3')

- NUP98 (Fwd: '5-TTTGGCCAGAAACCTGCTTTTGG-3',

  Rev: '5-AAAGGTGAAGGACTGAATGAAGG-3').

Oligos were diluted to a concentration of 50 µM. A total volume of 4 µl (fwd- and rev-oligo) were added to 46 µl of nuclease free H_2O and incubated for 5 min at 95 °C and 20 min at room temperature. Each oligo sequence pair was restriction cloned into pEn-Chimera cleaved with restriction enzyme BbsI (ThermoFisher Scientific, Massachusetts, USA) (Fauser et al. 2014). For each restriction, 0.5 µl of annealed oligo was incubated with 2 µl (~60 ng) of cleaved pEn-Chimera with 1 µl T4 DNA ligase and 1 µl of buffer in a 10 µl buffered reaction volume at 4 °C for 16 h followed by 2 h at room temperature. The resulting ligation reaction was cloned into competent *E. coli* Dh5α cells and plated onto LB agar (100 µg/ml carbenicillin) at 37 °C. Transformants were confirmed by PCR (annealing at 56°C and elongation for 30 seconds for 30 cycles resulting in a 300 bp amplicon) with SS42-U6 Promoter (Fwd: '5-TCCCAGGATTAGAATGATTAGG-3') and (Rev: 5'-AAAC + rev-com-3') protospacer primers. Amplicons were visualised by agarose gel electrophoresis **(Figure 4.5)**. The Gateway® LR reaction was used to clone the transformed pEn-Chimera into pDeSp (Fauser et al. 2014). pDeSp was prior transformed into DB3.1™ competent cells and the plasmid purified according to the GeneJET Plasmid Miniprep Kit (ThermoFisher Scientific, Massachusetts, USA) protocol. For plasmid transformation, 2 µl (~100 ng) of transformed pEn-Chimera was vortexed with 2 µl (~100 ng) pDeSp, 4 µl of TE buffer and 2 µl of LR clonase II (Invitrogen) and incubated overnight at 25 °C. The reaction products were transformed into competent *E. coli* Dh5α cells and plated onto LB agar (100 µg/ml spectinomycin, 37 °C). Transformants were confirmed by PCR (annealing at 60 °C and elongation for 60 seconds for 35 cycles resulting in 917 bp amplicon) with SS42-U6 Promoter (Fwd: '5-TCCCAGGATTAGAATGATTAGG-3') and SS102-35S-Promotor (Rev: '5-CACCATGTTATCACATCAATCC-3') primers, AflII/BspH1 restriction digest (13600 bp and 1750 bp fragments) **(Figure 4.5** and by sequencing outsourced to Inqaba Biotec (Pretoria, South Africa).

4.2.2 MICRO-PROPAGATION OF CASSAVA TME3 PLANTLETS

This was done according to an established protocol with modifications (Allie et al., 2014). Sterile nodal cultures of cassava TME3 landrace were micro-propagated on Murashige and Skoog (MS) medium (Murashige & Skoog, 1962) supplemented with 2 % sucrose, 0.002 mM CuSO4, 0.78 % plant tissue culture agar, pH 5.8, for 4 weeks at 28 °C with a 16 hour photoperiod under 8000–10000 lux light intensity. Three biological replicates for each experimental gene target were performed. Each biological replicate comprised 300 mg of leaf tissue from four cassava TME3 plantlets **(4.3.3)**. Quantitative PCR for each sample was carried out in three biological and two to three technical replicates **(4.3.6; 4.3.7)**.

4.2.3 PROTOPLAST ISOLATION

This was carried out according to an established laboratory protocol with modifications (Chatukuta and Rey, 2020). All techniques were carried out under aseptic conditions. Approximately 300 mg of leaf tissue from 4 plantlets was cut into 1-3 mm strips and plasmolysed by immersion into 10 ml CPW9M medium (Anthony et al. 1995) for 1 hour. The strips were then vacuum infiltrated in a freshly prepared 10 ml enzyme digestion solution (5 mM morpholinoethanesulphonic acid (MES), 1.2 % cellulase, 0.6 % macerozyme, CPW9M medium) for 30 min and then incubated for 16 h at 25 °C in the dark, on a shaker at 40 rpm. Thereafter, the mixture was shaken for 5 min at 80 rpm to release protoplasts. Released protoplasts were harvested through a 75 µm sieve and the filtrate was centrifuged in a round bottomed tube at 100 xg for 10 min to pellet the protoplasts. The supernatant was discarded, and the pellet washed and centrifuged twice in CPW9M medium. Protoplast integrity, viability and quantification were assessed.

4.2.4 PROTOPLAST EVALUATION

Protoplast integrity was checked using the FLoid™ Cell Imaging Station (Life Technologies) and ZEISS Axioscope **(Figure 4.6, Figure 4.7)**. Protoplast viability was determined by Evans' Blue Dye staining (Huang et al. 1986) **(Figure 4.8, Figure 4.9)**. Quantification was conducted via flow cytometry using the BD Accuri™ C6 flow cytometer (BD Biosciences, New Jersey, USA). Data analysis was performed using the FCS Express 7 Research Edition software (Treestar, Inc, Oregon, USA). Live cells were estimated based on the forward versus side scatter (FSC vs. SSC) (Figure 4.9 A-F). Protoplasts were resuspended in MMg solution (Wu et al. 2017) to a concentration of 10^4-10^6 cells per ml.

4.2.5 PROTOPLAST TRANSFORMATION

For each of the three biological replicates for cassava TME3, (i) Cas9 only [15 µg *Streptococcus pyogenes* plasmid (pDeSp)], (ii) Cas9+sgRNA [(7.5 µg per sgRNA transformed pDeSp)] and (iii) SACMV infectious clones [4 µg of pBIN19-SACMV-DNA-A

and pBIN19-SACMV-DNA-B plasmids each] (Berrie et al. 2001) were gently mixed with 200 µl of protoplasts. Immediately after, 25% polyethylene glycol 4000 (PEG4000) was added and the resulting mixture incubated at room temperature for 20 min. The mixture was gently diluted with three volumes of W5 solution (154 mM NaCl, 125 mM CaCl2, 5 mM KCl, 2 mM MES, pH 5.8). The transfection mixture was centrifuged twice at 100 xg for 2 min and the retained protoplasts resuspended in 300 µL of WI solution (4 mM MES, 0.5 M mannitol, 20 mM KCl, pH 5.8) (Yoo et al. 2007) and incubated for 24 h at 28°C in the dark to induce gene expression. After incubation, protoplasts were centrifuged at 100 xg for 5 min and the pellet frozen in liquid nitrogen and stored at -80°C prior to DNA and RNA extraction.

4.2.6 DNA EXTRACTION AND ANALYSIS

DNA was extracted from the transformed protoplasts using QIAzol (Qiagen, Maryland, USA) according to a user-developed protocol (www.qiagen.com/it/resources/; Qiagen, Maryland, USA). The resulting DNA was DpnI-digested (ThermoFisher Scientific, Massachusetts, USA) to remove input DNA and the presence of SACMV CP determined by PCR amplification and visualisation on agarose gel was done **(Figure 4.11)**. Quantitative PCR (qPCR) for SACMV relative viral load was performed in triplicate using the viral coat protein primer set (Fwd: '5-CGTGGCTGTGAA-3') and (Rev: '5-CGCGTGACATCA-3') amplification relative to the 18S rRNA as the reference gene **(Figure 4.12)**. The Delta-Delta Ct method that quantifies expression based on the assumption that the gene of interest and the reference gene have been amplified by primers of equivalent efficiency (Livak and Schmittgen, 2001). The calculation is represented by the formula $2^{-\Delta\Delta C_t}$ where, C_t represents the cycle threshold and $\Delta\Delta C_t$ represents the differences in C_t values for treated and control samples. That is,

$$\Delta\Delta C_t = \big(Ct(target\ gene, treated) - Ct(reference\ gene, treated)\big) - \big(Ct(target\ gene, control) - Ct(reference\ gene, control)\big)$$

Non-parametric tests suited for small sample sizes (< 30) were used to analyse expression data. Kruskal Wallis H Test (Chan and Walmsley, 1997) to determine differences in CP expression between protoplast DNA samples analysed **(Appendix 4.1)**. The Kruskal Wallis H Test, required the calculation of an H statistic value according to the formula below:

$$H = \frac{12}{n(n+1)} \Sigma \frac{R_i^2}{n_i} - 3(n+1)$$

where n is the total number of values, R is the sum of the ranks for each sample, and n_i is the number in each sample. Each value was ranked from the smallest to the largest and the H value calculated.

This was followed by a post-hoc test using the Mann-Whitney U Test for pairwise comparisons (Mann and Whitney, 1947). The null hypothesis (H_0) assumes that there are no statistically significant differences between to samples. The test calculates the U statistic for two samples under comparison U is calculated for each sample as follows:

$$U_x = R_x - \frac{n_x(n_x + 1)}{2}$$

where, R is the sum of ranks in the sample, n is the number of items in the sample and x is either sample designated as a or b. Thereafter, the z statistic and the corresponding p-value were calculated to determine the validity of the null hypothesis. The standard deviation of the obtained U statistic was calculated, and the z statistic computed i.e.,

$$U_{statistic} = \min(U_a, U_b)$$

$$stdev\ of\ U = \sqrt{\frac{(n_a * n_b) * (n_a + n_b + 1)}{12}},\ z = \frac{U - ((n_a * n_b)/2)}{stdev}$$

The p-value was calculated based on the standard normal distribution of the z statistic. For evaluation of mutation frequency, the two oligo/gRNA regions for each target gene were amplified with appropriate primers designated the gene name abbreviation and the numbers 1 and 2:

For casein kinase 1-like protein HD16:

- **C1**: Fwd:‘5-TCTAGAGCAATGAACCAAGATG-3’,
  Rev:‘5-GTAAAGGGTCAAGGGACAATTA-3’
- **C2**: Fwd:‘5-TTCCTCTATCCCAGTGAGGGG,
  Rev:‘5- GATTTCCGCTTGGGAGTGCT-3’

For histone-lysine n-methyltransferase atx4-related:

- **H1**: Fwd:‘5-AGGACAGTGGAGGTCTCACG-3’,
  Rev:‘5-ACCGTGAAGGGAGTACCTGA-3’
- **H2**: Fwd:‘5-AGGACAGTGGAGGTCTCACG-3’,
  Rev:‘5-TCATACCACAACCTCCACAGG-3’.

For nucleoporin autopeptidase:

- **N1**: Fwd:‘5-CTGGGTTGTTCCTGGTCTTCTA-3’,

  Rev:‘5-TAGGCTGTGATTGCTGACTTGT-3’

- **N2**: Fwd:‘5-ACCTTTTGGAACCTCTAGTACACC-3’,

  Rev:‘5-GGACGGAAAAGCTGAGGTTGA-3’

Amplification was carried out using the Phusion U Green Hot Start DNA Polymerase (ThermoFisher Scientific, Massachusetts, USA). Sequencing of PCR products was carried out on the ABI 3500XL Genetic Analyzer was outsourced to Inqaba Biotec (Pretoria, South Africa). Geneious software (Version 2021.1) created by Biomatters was utilised for sequence alignment and mapping to the genomic sequence (reference) (http://www.geneious.com; Kearse et al. 2012). It was noted that cassava TME3 genome varies from available reference genome for cassava germplasm variety AM560-2 (Bredeson et al., 2016). In the instance that the target region and its vicinity in TME3 varied from that obtained from AM560-2, sequences were aligned to unpublished cassava TME3 sequence data (RefSeq ID: RSFT01000007, GenBank assembly GCA_003957995.1). This was the case for NUP98 which varied by a single nucleotide insertion (adenine) within the sequenced region of interest relative to AM560-2. Each sequencing result was aligned with Geneious alignment (Global alignment with free end gaps, default parameters) to its reference sequence and visually examined within the gRNA region and vicinity for variations (mutations) for transformed and untransformed samples The Geneious mapper was employed in the map to reference tool. Sequence translation was carried out relative to the reference in all three forward frames. Amino acids were coloured according to polarity. Highlighting was carried out for all disagreements of the nucleotide and translated amino acid sequences relative to the reference sequence **(Figure 4.13, Figure 4.14)**. Mutation frequency was determined for each gRNA (gRNA1, gRNA2) **(Figure 4.15)**.

4.2.7 RNA EXTRACTION AND cDNA SYNTHESIS

This was according to the QIAzol (Qiagen, Maryland, USA) manufacturers protocol for transformed protoplasts and a CTAB RNA extraction method for Y2H assays. All surfaces and gloves were sprayed with 70 % ethanol and RNase Away® (Molecular Bio Products Inc., San Diego, USA) to decontaminate them. RNA quantity and purity were assessed by A260/A280 and A260/A230 ratios for both protoplasts and Y2H assays.

First strand cDNA synthesis was performed using the RevertAid First Strand cDNA Synthesis Kit (ThermoFisher Scientific, Massachusetts, USA). Amplification of pDeSp for each target gene was carried out to ascertain activity within protoplasts with SS42-U6 Promoter and SS102-35S-Promotor primers **(Figure 4.16)**. Quantitative PCR (qPCR) using

the Maxima SYBR Green/ROX qPCR Master Mix (2X) (ThermoFisher Scientific, Massachusetts, USA) of the cDNA was carried out in triplicate for each target gene with appropriate primer pairs i.e.,

- **CK**: Fwd:‘5-GGGTTTACATCCCAGGTTTATG-3’,

  Rev:‘5-GCCAGCAACAGAAGTGATATAG-3’.

- **HMT**: Fwd:‘5-GAGCAGTTGATGGAAGATGAG-3’,

  Rev:‘5-CACACTATGTCACCCGAATAAA-3’.

- **NP**: Fwd:‘5-GAGGAACTGAGATGGGAAGA-3’,

  Rev:‘5-CGCTAGAGACACCAAAGTTATT-3’

Target gene expression was determined relative to the 18S rRNA reference gene amplified with primers 18S (Fwd: ‘5-ATGATAACTCGACGGATCGC-3’, Rev: ‘5-CTTGGATGTGGTAGCCGTTT-3’) **(Figure 4.17)**. PCR primer efficiencies were calculated for each primer pair according to the formula,

$$E = \frac{-1}{10^{-slope}}$$

For each qPCR experimental set up, two negative controls were employed. That is, without template and without reverse transcriptase, respectively. The $2^{-\Delta\Delta C_t}$ method (Livak and Schmittgen, 2001) was used to determine relative gene expression and the data obtained was analysed by Kruskal-Wallis (Chan and Walmsley, 1997) and Mann-Whitney U (Mann and Whitney, 1947) tests, respectively.

4.3 RESULTS

4.3.1 CRISPR-CAS9 VECTOR CLONING

Successful ligation of the oligo insert in-frame with the U6 promoter in pEn-chimera (Gateway entry clone) was confirmed by the presence of a 300 bp amplicon **(Figure 4.5A)**. The successful transformation of pDeSp (Gateway destination vector) **(Figure 4.5C)** was confirmed by the presence of both the gRNA and Cas9 promoters presented as clear single bands of the 300 bp amplicon for the oligo insert (gRNA) and a 917 bp amplicon for the Cas promoter (35S promoter) **(Figure 4.5A-B)**. Digestion with AflII/BspH1 resulted in two clear fragments of 13600 bp and 1750 bp fragments for the pDeSp vector and the Cas9 and sgRNA fragment with promoters **(Figure 4.5D)**. pDeSp and pEn-chimera Sanger sequencing confirmed the correct insertion of sgRNA driven by the U6 promoter and the insertion of the sgRNA and promoter region in frame with the Cas9 scaffold of pDeSp.

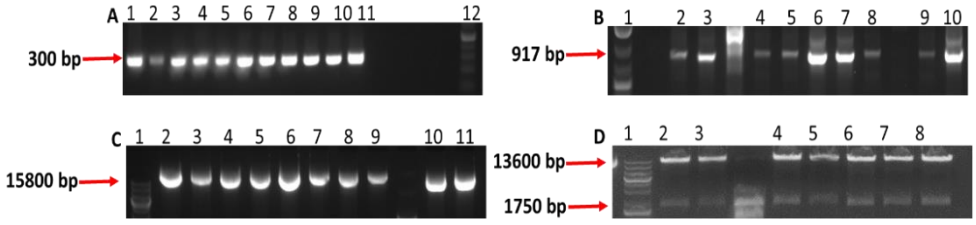


Figure 4.5 1.5 % agarose gel electrophoresis of PCR amplified target regions in Gateway vector transformed pDeSp post LR Clonase reaction **A.** 300 bp SS42-U6 Promoter/rev-oligo; CKL1 (lanes 1-4), NUP98 (lanes 5-7), SDG16 (lanes 8-11) **B.** 917 bp SS42-U6 Promoter/SS102-35S-Promotor CKL1 (lanes 2-4), NUP98 (lanes 5-7), SDG16 (lanes 8-10) **C.** Purified transformed pDeSp vector (pDeSp+sgRNA) CKL1 (lanes 2-5), NUP98 (lanes 6-8), SDG16 (lanes 9-11) **D.** 1 % agarose gel showing pDeSp digested with AflII/BspH1 to screen for successful transformation following LR Clonase reaction resulting in 13600 bp and 1750 bp fragments; CKL1 (lanes 2, 3), NUP98 (lanes 4, 5), SDG16 (lanes 6, 7, 8). GeneRuler^{TM} 1kb Plus DNA ladder for gels **A/C/D** (lane 1) and **B** (lane 12).

4.3.2 PROTOPLAST EVALUATION

Isolated protoplasts were presented as spherical structures with chloroplasts at the membrane periphery and showed high integrity. Protoplast diameter ranged from ~ 10-35 µm **(Figure 4.6, Figure 4.7)**. Evans Blue Dye confirmed protoplast integrity and staining gave an average of 80 % viability from ten fields of view for three samples **(Figure 4.8, Figure 4.9)**. With reference to Figure 4.9 the forward versus side scatter (FSC vs SSC) gating of isolated protoplasts showed a high concentration of protoplasts within the circled region **(Figure 4.10A-C)** and outside of this region was the presence of debris and other irregularly shaped particles. Events measured ranged from ~11 000 to 40 000 events/µl **(Figure 4.10D-F)**. With respect to a viability of 80 %, protoplast yields per gram fresh leaf tissue weight (gFW) ranged from 5.44 x 10^6 to 6.83 x 10^6/gFW.

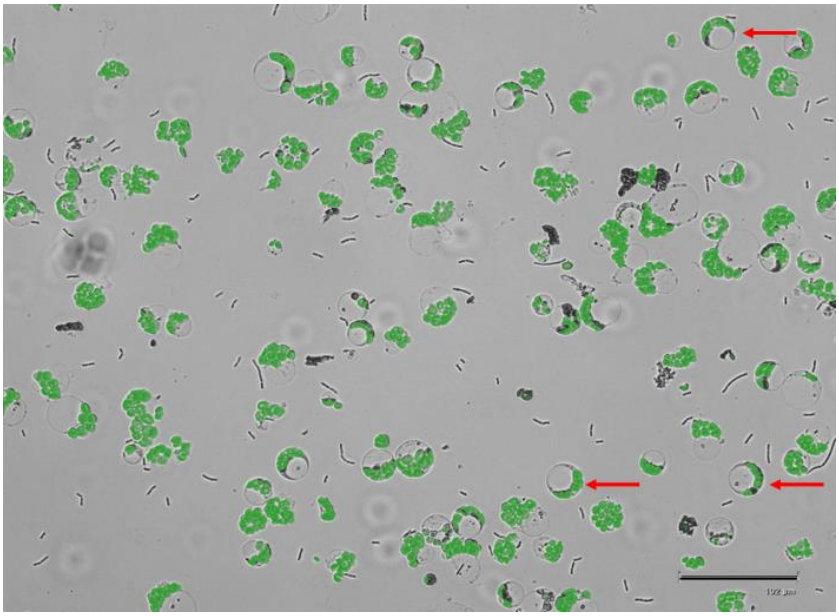


Figure 4.6 Protoplast isolation from cassava TME3 leaf mesophyll viewed with the Floid™ Cell Imaging Station at 20X magnification. Protoplast morphology demonstrated integrity as viewed by the spherical shape and intact round cell membrane with chloroplasts in the periphery (indicated with red arrows). Bar=102 µm.

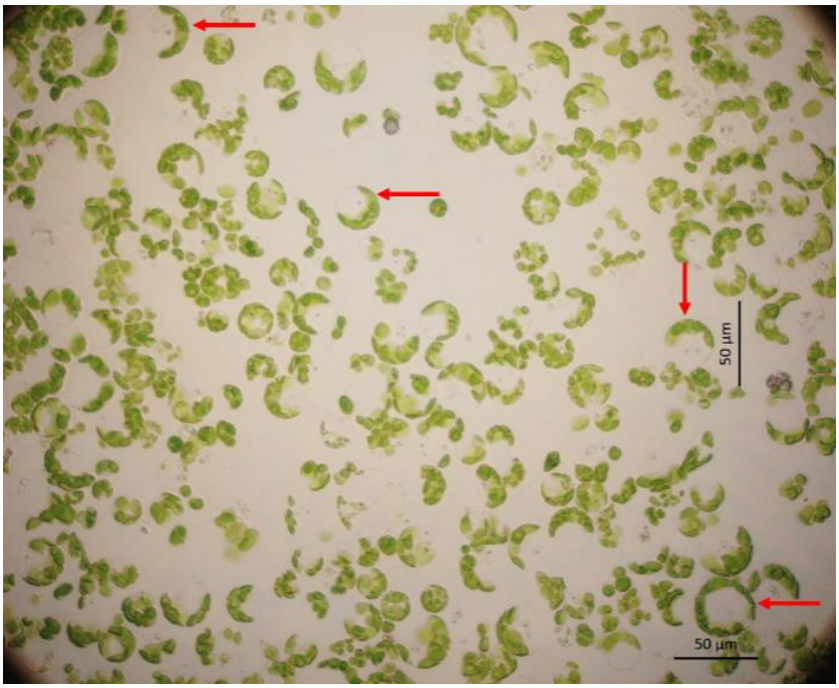


Figure 4.7 Protoplast isolation of cassava TME3 leaf mesophyll viewed with the ZEISS Axioscope microscope. Protoplast morphology presented spherical shape with chloroplasts in the periphery (indicated with red arrows). Bar=50 µm.

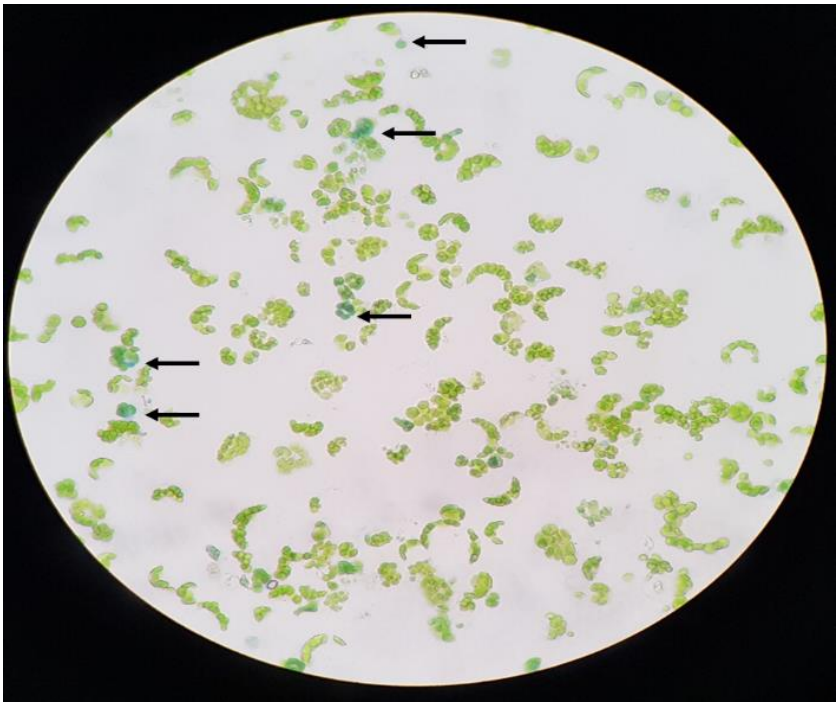


Figure 4.8 Protoplast viability was assessed by Evans Blue dye staining and visualised at 40X magnification. Stained blue cells indicated with black arrows were considered non-viable. Image was taken by use of a Samsung galaxy 8 smartphone.

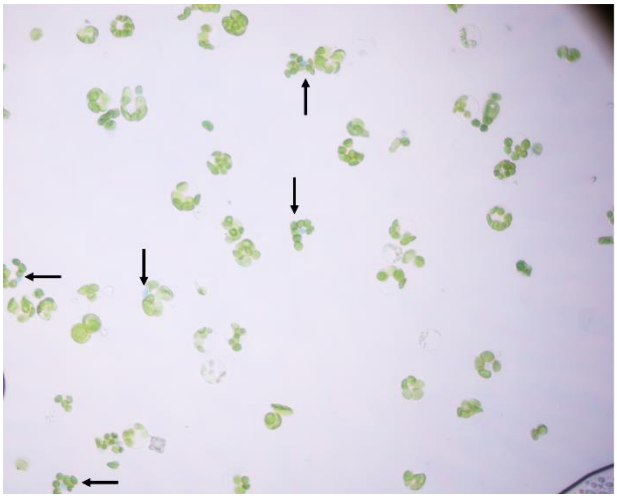


Figure 4.9 Protoplast viability assessed by Evans Blue dye staining and viewed with the ZEISS Axioscope microscope. Stained cells indicated with black arrows were considered non-viable.

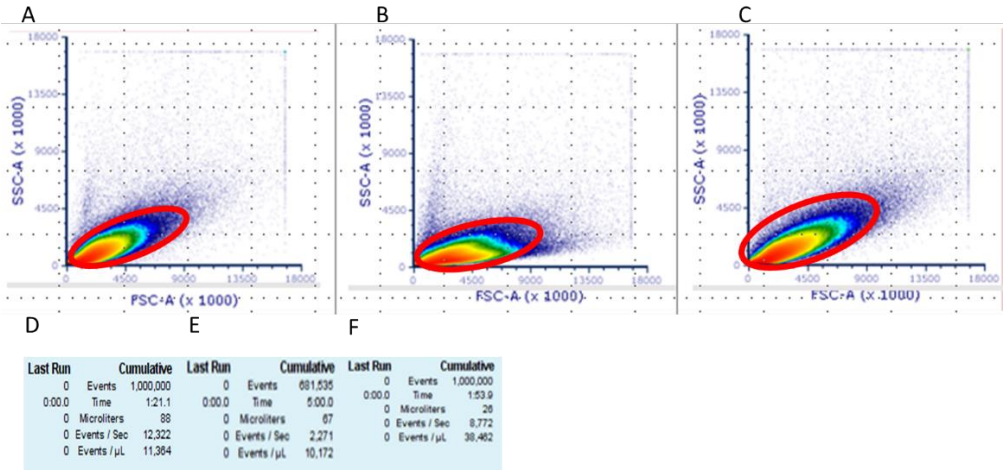


Figure 4.10 Protoplast quantification. BD Accuri™ C6 Flow Cytometer (BD Biosciences, San Jose, CA, USA) measurements. (A-C) Flow cytometric density measurement on a log scale plotted by the effect of protoplast size and shape on the forward scatter (FSC-A) and side scatter (SSC-A) for cassava TME3. Viable protoplasts are from within the red circled regions. (D-F) Events measured within protoplast solutions of cassava TME3 ~11 000-40 000 events/µl.

4.3.3 SACMV DETECTION AND QUANTIFICATION

SACMV CP PCR amplification of DpnI digested protoplast DNA gave clear bands at the 100 bp position for SACMV+pDeSp+sgRNA SDG16/NUP98/CKL1 **(Figure 4.11)** and SACMV (S) (not shown).

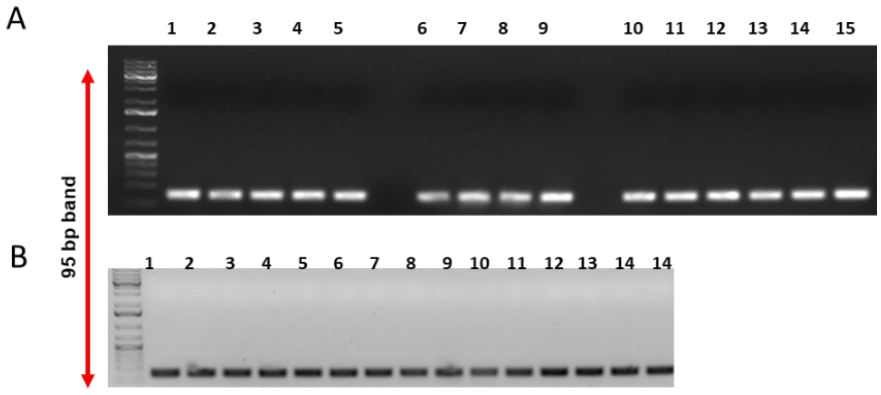


Figure 4.11 DpnI digested protoplast DNA amplified for the presence of SACMV CP presence prior to viral load expression analysis. Expected CP amplicon can be viewed at 95 bp. A and B denoted SDG16 (1-5), NUP98 (6-9), CKL1 (10-15) relative to relative to GeneRulerTM 1kb Plus DNA ladder (left border).

Viral DNA accumulation was assessed in SACMV transfected protoplasts and in SACMV+pDeSp+sgRNA for SDG16/NUP98/CKL1. The viral DNA in treatment samples (SACMV+pDeSp+sgRNA) was visualised in a bar graph standardised against SACMV

transfected protoplasts **(Figure 4.12)**. Viral accumulation in SACMV+pDeSp+sgRNA treated protoplasts was lower than in SACMV transfected protoplast samples.

All Kruskal Wallis Test **(Appendix 4.1)** was performed to determine if CP expression value was the same for four different transfection conditions/treatments (SACMV+pDeSp+sgRNA) of plant protoplasts. The test revealed that the expression values among the four transfection conditions were not the same. There was a statistically significant difference in CP expression among two or more of the transfection treatments ($H_{calculated}$, 15.600 > $H_{tabulated}$, 7.501). Therefore, all medians were not equal, and the null hypothesis was rejected.

The Mann Whitney U test **(Appendix 4.2)** showed that there were no significant differences between SACMV+pDeSp+sgRNA treated samples (p-calculated 0.25 (SDG16), 0.43 (NUP98) and 0.93 (CKL1) > 0.05). However, there were significant differences between each SACMV+pDeSp+sgRNA treated sample and SACMV transfected samples (p-calculated 0.00002 < 0.05). Simultaneous gene targeting and infection with SACMV showed a negative log fold change of 0.55 to 0.62. This represents an average reduction of 0.57 or 57 %.

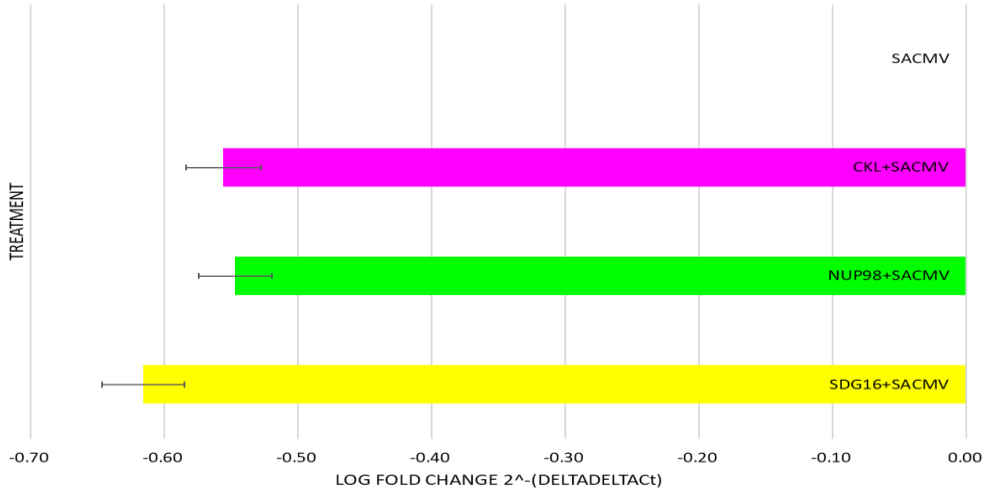


Figure 4.12 Real time PCR was carried out on DpnI digested DNA for evaluation of SACMV DNA viral load based on coat protein (CP) accumulation relative to 18S rRNA as the reference gene. Viral DNA accumulation was assessed in SACMV transfected protoplasts and in and SACMV+pDeSp target gene. The viral DNA in treatment samples (SACMV+pDeSp target gene) was visualised in a bar graph standardised against SACMV transfected protoplasts. SACMV+pDeSp+sgRNA for SDG16 (yellow), SACMV+pDeSp+sgRNA for NUP98 (green), SACMV+pDeSp+sgRNA for CKL1 (pink), SACMV only (S).

4.3.4 MUTAGENESIS

- Sanger sequencing results for CRISPR targeted SDG16 were inconclusive and could not be reported due to the presence of artefacts and amplification of a different region of the genome.

- No mutations were noted in the negative control samples (pDeSp vector only).

- The overall mutation frequencies ranged from 10 % to 60 %. Individual mutation frequencies for CKL1 ranged from 40 % to 60 % and that of NUP98 from 10 % to 20 %. Insertions and deletions accounted for 38 % of observed mutations.

- Results showed that pDeSp+NUP98 for gRNA1 yielded a single insertion after the third adenine within the gRNA1 region and 11 substitutions **(Figure 4.13)**. The region for gRNA2 yielded four substitutions within the gRNA region and 1 substitution in the vicinity. Therefore, pDeSp+NUP98 showed gene modifications frequencies of 20 % and 10 % for gRNA1 and gRNA2 respectively relative to the cassava TME3.

- pDeSp+CKL1 showed gene modifications frequencies of 60 % and 40 % for gRNA1 and gRNA2 respectively **(Figure 4.14)** relative to AM560-2 reference genome which was equivalent to cassava TME3 reference sequence for the targeted region. gRNA1 displayed multiple insertions, deletions and substitutions biased to the gRNA region except for gRNA1e that showed additional disagreements upstream to the 1740 bp position. Disagreements downstream from the gRNA region were also noted. gRNA2 did not show any disagreements within the target gRNA region albeit substitutions and insertions were noted from 6 bp (40 bp position) upstream of the gRNA region to the 80 bp position. Control samples transfected with untransformed pDeSp did not show any variation from the genomic/reference sequence. Mutations within the nucleotide sequences highlighted disagreements in the amino acid sequences within all three translated frames relative to the translated genomic sequences.

- All observed mutations resulted in an altering of the amino acid sequences. Mutations frequencies for each gRNA targeted region were calculated as percentages of total mutations attributed per target gene in both gRNA regions **(Figure 4.15)**. Casein kinase (CKL1) gRNA accounted for the largest number of variables. That is 58 % insertions, 24 % substitutions and 7 % deletions. For NUP98, gRNA1 accounted for 6 % of insertions and 65 % of substitutions. NUP98 gRNA2 showed 29 % of substitutions. No insertions were noted in NUP98 gRNA2 and neither NUP98 gRNA1 nor NUP98 gRNA2 yielded any deletions.

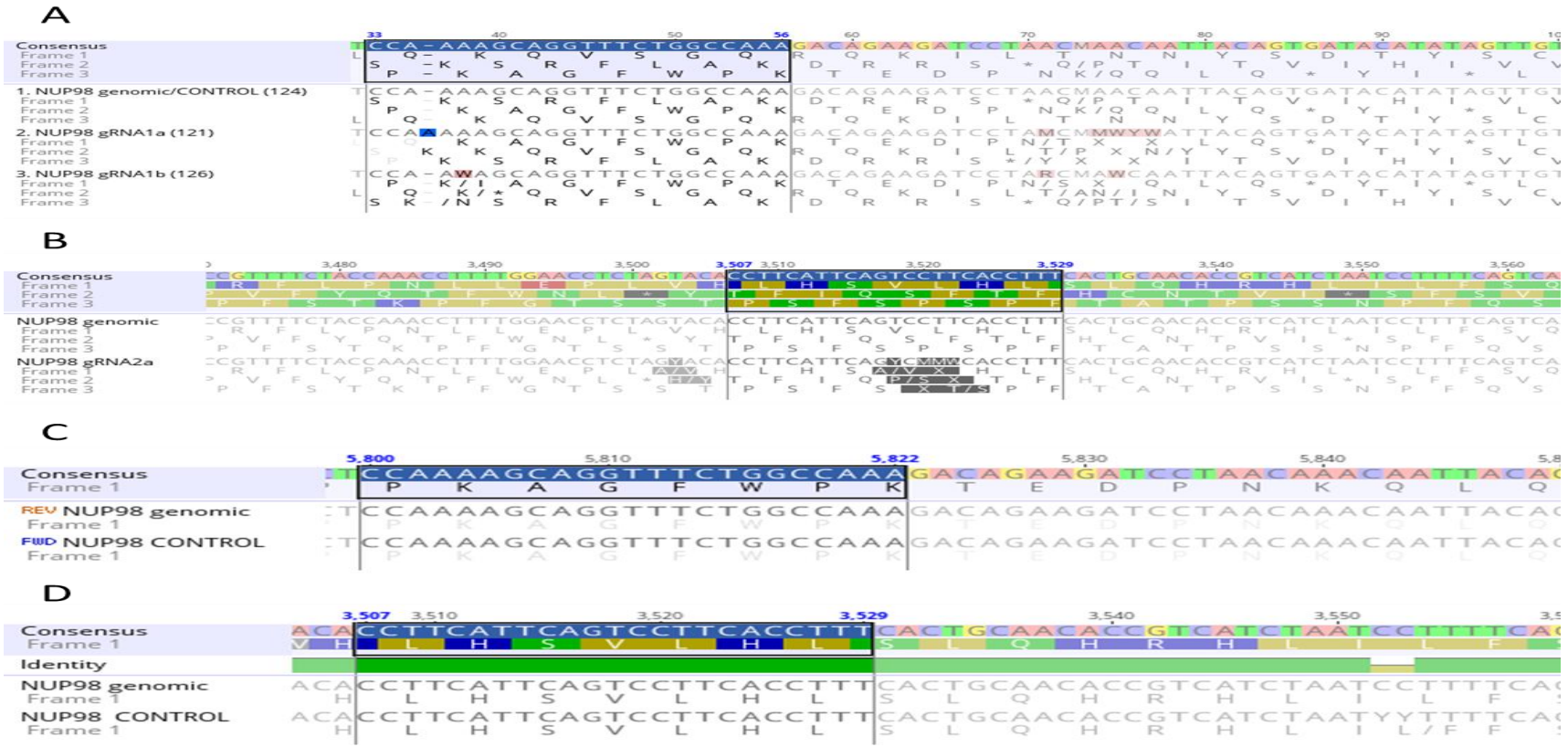


Figure 4.13 *Nucleoporin autopeptidase (NUP98)* transformed and untransformed protoplast DNA genome sequence samples compared to Manes.01G002100 genomic sequence (NUP98 genomic). Alignment with sgRNA regions, gRNA1 (A), gRNA control (B). Targeted sgRNA region shown within vertical borders. A. gRNA1a and gRNA1b transformed B. gRNA2a transformed C. gRNA1 untransformed (Control) D. gRNA2 untransformed (Control). Mutations in nucleotide sequences highlighted accordingly substitutions in grey. All sequences translated in three frames relative to the consensus.

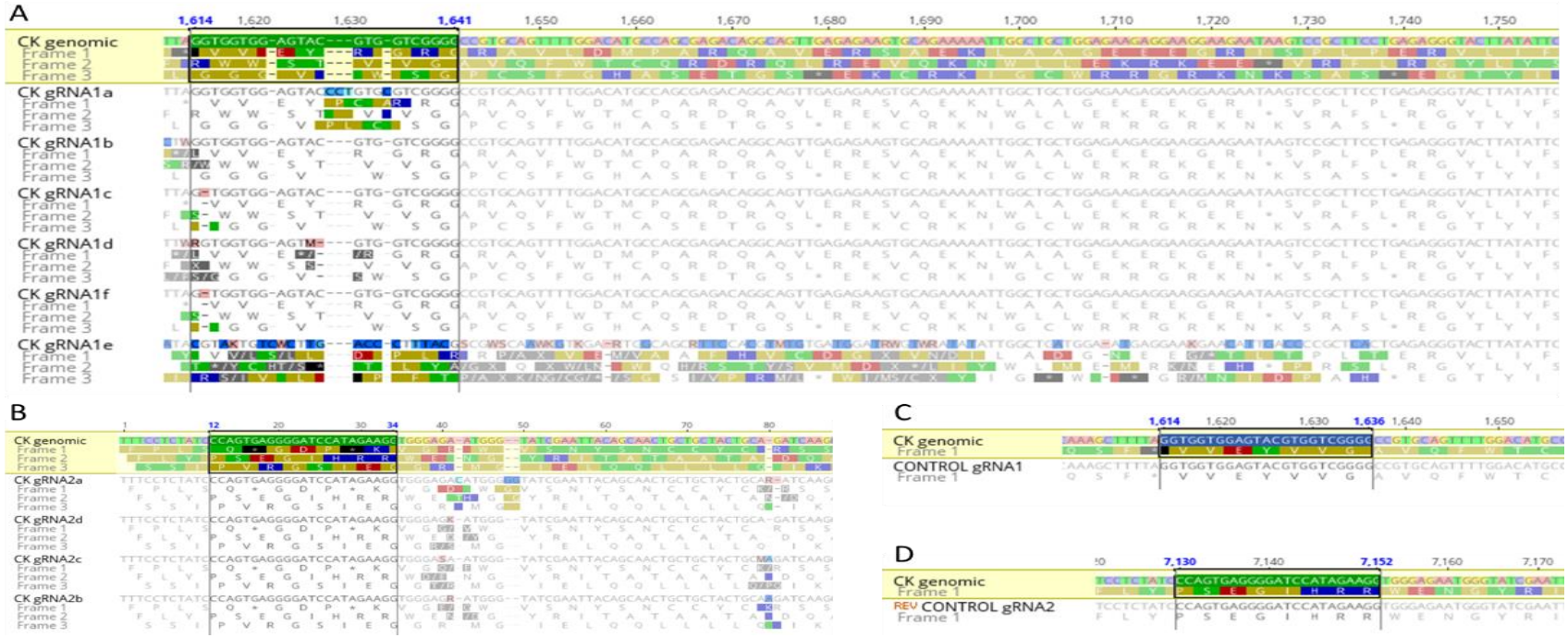


Figure 4.14 *Casein kinase, CKL1* (CK) transformed and untransformed protoplast DNA samples relative to Manes.16G117000 genomic sequence (reference) with sgRNA regions segregated as gRNA1 (A), gRNA2 (B). Targeted gRNA region highlighted in reference sequence (CK genomic) relative to control sequence A. gRNA1 transformed B. gRNA2 transformed C. gRNA2 untransformed (Control) D. gRNA1 untransformed (Control). Mutations in nucleotide sequences highlighted accordingly, Insertions shown in blue, deletions shown as a dash and substitutions in pink. All disagreements highlighted in translated frames relative to the reference CKL1 sequence (CK genomic).

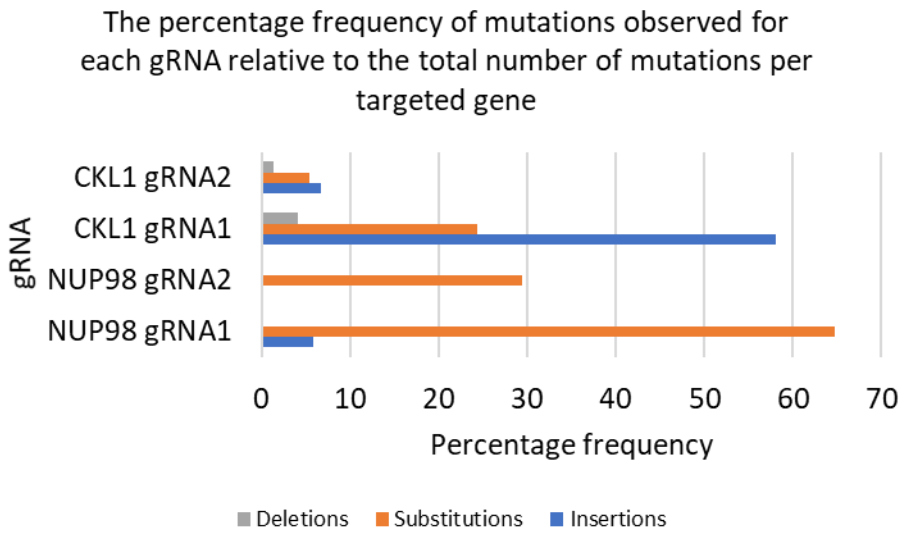


Figure 4.15 Mutation frequencies yielded by each gRNA region targeted. Mutations noted colour coded as follows: Deletions-grey, Substitutions-orange and Insertions-blue.

4.3.5 DETERMINATION OF pDeSp ACTIVITY

pDeSp+sgRNA (targeting *CKL1/NUP98/SDG16*) PCR amplification from cDNA was confirmed by very faint bands within the 1000 bp position **(Figure 4.16)**.

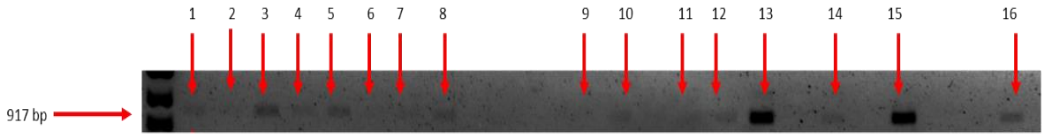
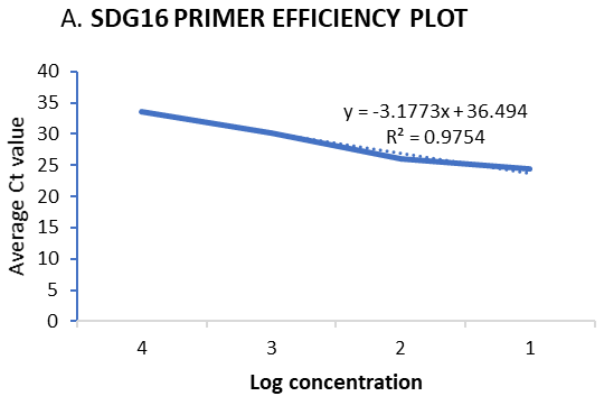
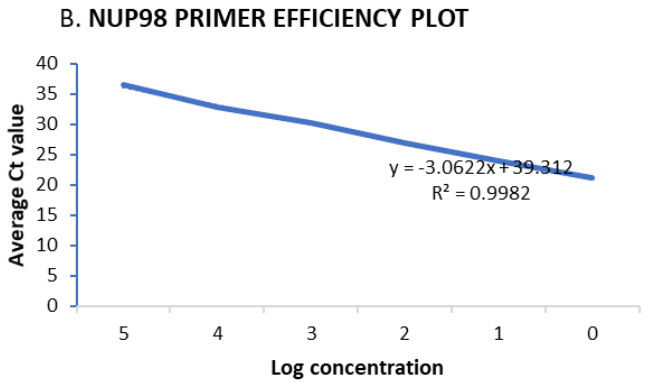


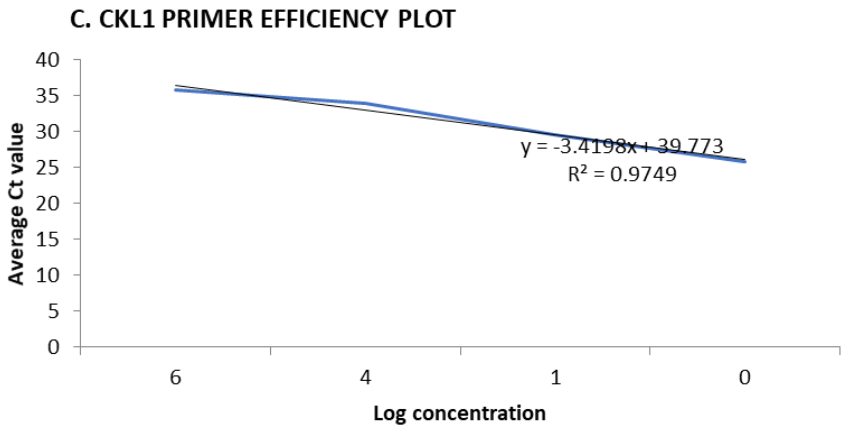
Figure 4.16 PCR amplification of protoplast cDNA to ascertain the transcription of pDeSp+sgRNA in transfected protoplast. Amplicon at 917 bp position. *CKL1* (lanes 1-5, 16-17), *NUP98* (lanes 6-9), *SDG16* (lanes 10-15). GeneRulerTM 1kb Plus DNA ladder (1).

4.3.6 PRIMER EFFICIENCY

Primer efficiencies ranged between 96.07 % and 112.11 % and correlation coefficients (R^2) were ranged between 0.98 and 1.00 **(Figure 4.17)**. Amplification factors were 2.06, 1.96 and 2.12 for *histone-lysine n-methyltransferase atx4-related (SDG16), casein kinase 1-like protein HD16 (CKL1) and nucleoporin autopeptidase (NUP98)*, respectively.







D. PRIMER EFFICIENCY

PRIMER PAIR	EFFICIENCY %	AMPLIFICATION FACTOR
SDG16	106.41	2.06
CKL1	96.07	1.96
NUP98	112.11	2.12

Figure 4.17 Primer efficiency plots for amplification of cassava TME3 target regions for quantitative real time PCR (qPCR) (A-C). Primer efficiency percentages and amplification factors for each gene target (D).

4.3.7 GENE EXPRESSION OF TARGET GENES IN PROTOPLAST cDNA

All Kruskal Wallis Test **(Appendix 4.1)** and Mann Whitney U **(Appendix 4.2, Appendix 4.3, Appendix 4.4, Appendix 4.5)** were calculated to evaluate expression data. Gene expression of target genes in treatments, SACMV transfected (S), SACMV+pDeSp+sgRNA (targeting *NUP98/SDG16/CKL1*) and pDeSp+sgRNA (targeting *NUP98/SDG16/CKL1*) were evaluated relative to the untransformed pDeSp vector only (pD1/pD2/pD3) **(Figure 4.18)**.

Histone-lysine n-methyltransferase atx4-related (SDG16) expression was enhanced in S2 and HS by 92 % and 13 % respectively. Silencing of *SDG16* expression showed a repression by 29 %. Kruskal Wallis H Test proved that there were statistically significant differences among samples for the expression of *SDG16* ($H_{calculated}$, 14.120 > $H_{tabulated}$, 7.240). The Mann Whitney U test showed that there were significant differences for all pairwise comparisons for *SDG16* (p-calculated, 0.002 < 0.05).

Nucleoporin autopeptidase (NUP98) did not show a great perturbation in gene expression. Expression was enhanced in SACMV (S1) by 17 %. Silencing of *NUP98* showed an enhanced 4 % expression. Kruskal Wallis H Test proved that there were statistically significant differences among samples for the expression of *NUP98* ($H_{calculated}$, 13.059 > $H_{tabulated}$,7.235). The Mann Whitney U test showed that there was no significant difference between CRISPR targeted *NUP98* and CRISPR targeted *NUP98* with SACMV (p-calculated, 0.12 > 0.05). All other pairwise comparisons were significantly different (p-calculated, 0.002 < 0.05).

Casein kinase 1-like protein HD16 (CKL1) expression was enhanced in S3 by 41 %. Kruskal Wallis H Test proved that there were statistically significant differences among samples for the expression of CKL1 ($H_{calculated}$, 16.714 > $H_{tabulated}$, 7.377). The Mann Whitney U test showed that there was no significant difference between CRISPR targeted *CKL1* and CRISPR targeted *CKL1* with SACMV (p-calculated, 0.12 > 0.05) albeit all other pairwise comparisons were significantly different (p-calculated, 0.0004 < 0.05).

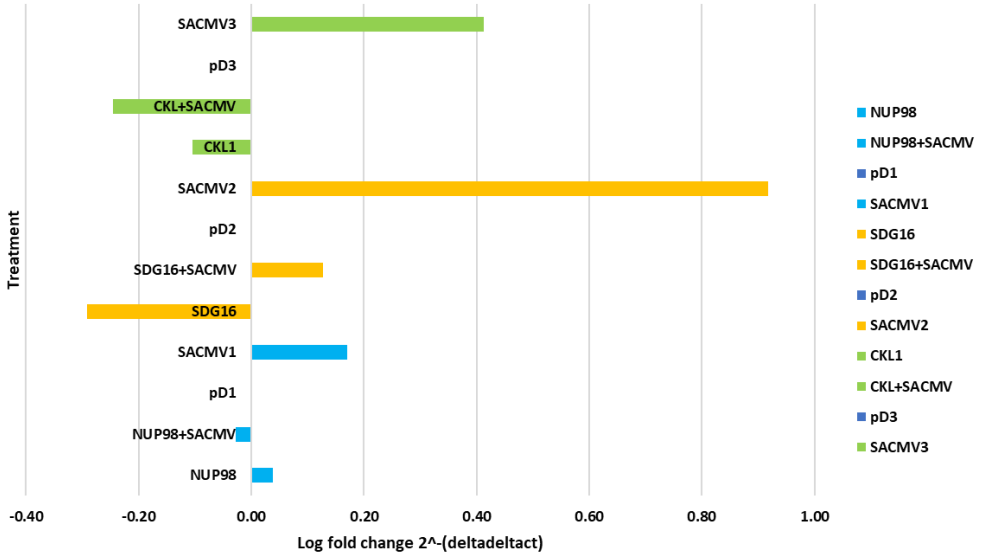


Figure 4.18 Real time PCR was carried out on cDNA from a total RNA extract for evaluation of gene expression relative to 18S rRNA as the reference gene. Gene expression was assessed in SACMV transfected protoplasts (S1, S2, S3), untransformed pDeSp (pD1, pD2, pD3), SACMV+pDeSp+sgRNA gene (NS, HS, CS denoting NUP98, SDG16, CKL1 with SACMV respectively), pDeSp+target gene (N, H, C denoting NUP98, SDG16, CKL1 respectively). The gene expression was visualised in a bar graph standardised against untransformed protoplasts.

4.4 DISCUSSION

4.4.1 SUCCESSFUL ISOLATION AND TRANSFORMATION OF CASSAVA TME3 PROTOPLASTS

Protoplast morphology was characterised by spherical protoplasts and was comparable to previous work on cassava (Chatukuta & Rey, 2020; Sofiari, 1996; Wu et al., 2017). The approximate diameter range of 10-35 µm was comparable to previous work in the Rey laboratory (Chatukuta & Rey, 2020). Protoplast viability within cassava with either Evans Blue Dye or Fluorescein diacetate (FDA) from leaf material or somatic embryos ranged from 40 % to 92.6 % in prior studies (Chatukuta & Rey, 2020; Sofiari, 1996; Wu et al., 2017). A similar extraction protocol in maize yielded a viability of 75 % with fluorescein diacetate (FDA) (Sant'ana et al., 2020). Increase in cellulase and macerozyme percentage composition is reduce protoplast viability. Percentages of 2.4 % cellulase and 1.2 % macerozyme in cassava yielded a protoplast viability of 83 % whilst 1.6 % cellulase and 0.8 % macerozyme yielded a viability of 85 % and 92.6 % (Wu et al., 2017; Chatukuta & Rey, 2020). Increased enzyme percentages result in increased protoplast yield and a reduction in protoplast viability due to toxicity of enzymes altering protoplast regulation and by compromising protoplast membrane integrity (Wu et al., 2017; Zhang et al., 2011). It is concluded that the lower percentages used within this study 1.2 % cellulase and 0.6 % macerozyme would have exceeded viability shown in prior work albeit the viability average showed a variable reduction (Wu et al., 2017; Chatukuta & Rey, 2020) by comparison. Further the results herein show that protoplast yield and viability is multifaceted, and is affected by parameters of enzyme concentration, pH, enzymatic incubation temperature and duration and mannitol concentration (Zhang et al., 2011). The FSC vs SSC gating of isolated protoplasts indicated the presence of cell debris and this may have been rectified by use of a smaller mesh size of at least 40 µm to 50 µm as opposed to the 75 µm used in this objective (Hsu et al., 2021; Zhang et al., 2011). Overall the range of protoplast yield (per gram fresh weight (FW) was higher 2.34 to 31.4 x 10^5 protoplasts/g FW (Sofiari, 1996), 4.90 to 6.36 × 10^6 protoplasts/g FW (Chatukuta & Rey, 2020) and lower, 4.4 × 10^7 protoplasts/g FW (Wu et al., 2017). However, all protoplast yields reported are sufficient for cassava protoplast transfection.

4.4.2 CRISPR-CAS9 VECTOR (pDeSp) EFFICIENT FOR CASSAVA TME3 PROTOPLAST TRANSFORMATION

The CRISPR/Cas expression vector pDeSp in this study and its derivatives proved to be effective, rapid and simplistic to use *in planta* by *Agrobacterium* mediated transformation for heritable mutations in *Arabidopsis thaliana* (Fauser et al., 2014; Schiml et al., 2014; Steinert et al., 2017), *Nicotiana benthamiana* (Schmitz et al., 2020) and grape (Nakajima et al., 2017). To date there has been no study that has utilised these expression plasmids

in protoplast studies and thus utilisation to this regard in cassava protoplasts may be considered as the first. To confirm the active transcription of pDeSp in transfected protoplasts, the PCR amplification of cDNA from RNA that had undergone DNAse treatment was a simple and effective way to do so. An improvement to this would have been the inclusion of western blotting with a Cas9 antibody for detection of actual Cas9 protein within transfected protoplasts (Nakajima et al., 2017). The expression vector utilised did not contain a fluorescent transformation marker such as GFP which is useful in detecting the stability and efficiency of transformation at least 24 h post transfection, and has been useful in protoplast mutagenesis studies (Chatukuta & Rey, 2020; Wu et al., 2017; Brandt et al., 2020). The pDeSp used in this study did not contain a fluorescent marker and functions on the basis of antibiotic resistance of transformants (Fauser et al., 2014). The regeneration of transformed protoplasts was not pertinent to this study and therefore antibiotic screening post transformation was not required. Transformation efficiency and actual mutagenesis was limited to confirmation by Sanger sequencing of the targeted gRNA regions. The inclusion of fluorescent markers aids not only as a preliminary marker for evaluation of transformation efficiency but also allow researchers to optimise experiments prior to downstream analysis and costly sequencing . Khan and colleagues modified the pEn-chimera and pDe-Cas9 by a simpler In-Fusion® HD cloning system to include a YFP marker to screen for the most promising gRNAs (Khan et al., 2017).

4.4.3 pDeSp IS ACTIVELY TRANSCRIBED IN CASSAVA TME3 PROTOPLASTS

In detection of Cas9 activity, there are three possible options for high throughput evaluation i.e., SURVEYOR nuclease assays or Plasmid- or ssODN-mediated HDR or by NGS sequencing (Ran et al., 2013). NGS sequencing is utilised for high throughput, however Sanger sequencing is a more accurate and ideal option for smaller studies (Slatko et al., 2018). Prior work and design of the pDeSp expression vector was used for the generation of heritable mutations in *Arabidopsis* and detection of Cas9 activity was by Sanger sequencing of randomly selected mutant plants conferred with antibiotic resistance from the pDeSp expression vector (Fauser et al., 2014; Schiml et al., 2016). For this objective with regards to protoplasts, two methods were utilised to ascertain pDeSp expression and consequently, representative of Cas9 activity. Cas9 was detected by amplification from cDNA and sequencing analysis of resulting mutations within target regions by Sanger sequencing was sufficient proof of Cas9 activity.

4.4.4 SACMV ACCUMULATION VERIFIED IN CASSAVA TME3 PROTOPLASTS

It has been long established that protoplasts are capable of actively replicating viral DNA (Kikkawa et al., 1982; Lee et al., 2001; Rottier, 1980). Geminivirus studies have utilised this replicative ability (Ugaki et al., 1991; Méndez-Lozano et al., 2003; García-Neriaand & Rivera-Bustamante, 2010). Detection and quantification of SACMV replication was

achieved in TME3 protoplasts in this study. Replication in differentiated cells was demonstrated in transfection with wheat dwarf virus (WDV) in endosperm derived protoplasts (Ugaki et al., 1991). Protoplasts from leaves have further been proved as representative tools of *in planta* interactions (Méndez-Lozano et al., 2003). Interactions between pepper huastecovirus (PHV) and pepper golden mosaic virus (PepGMV) were identical to those shown *in planta* in which the replication rate of both viruses was doubled in a non-host dependent manner and PHV complements PepGMV (Méndez-Lozano et al., 2003). Tolerance and susceptibility to these geminiviruses was also sufficiently demonstrated in protoplasts (García-Neriaand & Rivera-Bustamante, 2010). PepGMV accumulation in protoplasts derived from resistant pepper plants was 70 % less than that of protoplast derived from susceptible pepper plants (García-Neriaand & Rivera-Bustamante, 2010). The characterisation of geminivirus interactions with the host and among different geminiviruses has been effectively shown by studies in protoplasts however viral movement can only be assessed *in planta* (Ugaki et al., 1991; Méndez-Lozano et al., 2003; García-Neriaand & Rivera-Bustamante, 2010). Therefore, future *in planta* studies would be applicable in advancing work with SACMV MP systemic infection and MP movement. Furthermore, cassava TME3 is a tolerant genotype and therefore it would be ideal to carry out *in planta* studies to visually assess the impact of gene knockdown on plant morphology. Some transgenic tolerant cassava genotypes have been shown to accumulate cassava geminivirus in the absence of symptoms (Moralo, 2015). Therefore, SACMV accumulation alone may not be sufficient to ascertain the impact on susceptibility. However, the comparison of viral accumulation in tolerant cassava TME3 to susceptible T200 revealed a trend equivalent to that shown *in planta* studies and therefore with respect to this objective is reliable in evaluation of susceptibility (Chatukuta & Rey, 2020).

DpnI is a bacterial restriction endonuclease that specifically targets the methylated adenine (^{m}A) at every '5-GATC-3' sequence from replicated bacterial DNA (Lacks et al., 1986). With respect to this objective, it was critical to quantify only replicated and not input DNA in transfected protoplasts. Dpn1 was able to target the replicated pDeSp, pBIN19-SACMV-DNA-A and pBIN19-SACMV-DNA-B vectors in transfected protoplasts.

4.4.5 TARGET GENE EXPRESSION CORRELATES TO SACMV ACCUMULATION

Targeted silencing and transfection with SACMV resulted in reduced expression of SACMV relative to the control sample that was transfected with SACMV only for all three genes *SDG16, NUP98* and *CKL1*. This was indicative that these genes play a role in the accumulation and replication of SACMV. Statistically the viral accumulation noted in *SDG16, CKL1* and *NUP98* targeted samples was equivalent. In terms of fold change, the

targeting of *SDG16* resulted in the highest reduction in SACMV DNA whilst *NUP98* and *CKL1* were relatively equal with *NUP98* showing the lowest perturbation.

Previously in Chapter 2, conclusions had been made on the impact of the interaction of the proteins encoded by these three genes with SACMV MP protein. These were (i) CKL1-MP interaction enables the phosphorylation of MP and anchorage of MP to microtubules for movement, (ii) NUP98-MP interaction facilitates movement of MP anchored SACMV DNA and MP independently and (iii) NUP98-MP and SDG16-MP interactions activate viral DNA transcription. SDG16 and NUP98 were postulated to be crucial to enhancing host gene expression and may play a role in elevating host gene expression. Furthermore, the interaction with SACMV MP was also considered in its potential to play a limiting role in host defence (Lu et al., 2003; Huang et al., 2014), thereby influencing viral DNA replication. Putting into perspective the results yielded from protoplast transfection in this study, there appears to be an interplay in which expression and translation of the SACMV genome may be favoured by NUP98 and SDG16 expression. Although histone methyltransferases such as SDG16 may be generally involved in gene repression, much like histone acetyltransferases, both play a critical role in the post translational modifications of histones to induce gene expression (Hubert et al., 2009; Lee et al., 2020). Consequently NUP98 and SDG16 proteins may contribute to facilitating viral replication and transcription (Liu et al., 2004). The proposal for geminiviruses being packed as mini-chromosomes with host histones (Pilartz & Jeske, 1992; Jeske, 2009) to facilitate movement (Zhou et al., 2011) further adds credence to the requisite of SDG16 in SACMV infection. It has been well established that viral NSP anchors viral DNA for movement across the nuclear membrane (Carvalho & Lazarowitz, 2004). In this current work, the likelihood of an additional mechanism with NUP98 as a binary interaction with MP or a more complex interaction involving NSP, MP and the viral genome is plausible. Prior studies have elucidated the utilisation of host NUP98 for viral replication and systemic movement in mammalian viruses (De Jesús-González et al., 2021; Lau & Weber, 2020; Ebina et al., 2004). It may be postulated that SACMV also actively manipulates NUP98 for this purpose since SACMV transfection revealed an upregulation of *NUP98* gene expression. There were no significant differences between the CRISPR/Cas9 (targeted knockdown) and the CRISPR/Cas9+SACMV transfected samples in *NUP98* gene expression and these samples showed the lowest fold change relative to the untransformed control (< 0.04). The gene expression noted in the silenced and, silenced and SACMV infected samples may constitute a feedback loop mechanism or alternative splicing events or multiple transcriptional regulators for *NUP98*. In the human genome, NUP98 comprises a single gene that undergoes two alternative splicing events resulting in four mature mRNA isoforms: two *NUP98*, one *NUP96* and an 8 kDa polypeptide that is removed from *NUP98* by auto-cleavage (Tamura et al., 2010). More applicable to cassava, which is still not well

annotated, is the *Arabidopsis* plant genome. In *Arabidopsis*, *NUP96*, *NUP98A* and *NUP98B* are located in different genome locations (Cheng et al., 2020). Currently in cassava, cassava4.1_033148m and cassava4.1_000699m have been identified as homologous to all *NUP98/96* transcripts in *Arabidopsis* (Szklarczyk et al., 2017; Gough et al., 2001). This implies that in addition to a probable feedback loop mechanism, gene expression may have amplified transcripts homologous to the untargeted isoform.

Gene expression analysis of *SDG16* correlated with an increase in SACMV expression. The directed upregulation of *SDG16* by SACMV was noted in the increased expression of SDG16 transcripts in the presence of CRISPR/Cas9 mediating expression vectors for sgRNA1 and sgRNA2. The p12 protein of plant RNA virus, chrysanthemum virus B, acts as a transcription factor upregulating the expression of a gene designated *upp-L* that transcribes for a Zinc finger protein (bHLH-TF) involved in stimulating cell proliferation and growth (Lukhovitskaya et al., 2013). In Chapter 2, SACMV MP was shown to be associated with two zinc finger domain proteins, SDG16 and NAT (acyl-CoA N-acyltransferase). Zinc finger motifs in proteins and the association of virus proteins with zinc finger motifs is a feature of a transcriptional role (Sheppard & Werner, 2017; Lukhovitskaya et al., 2013). Therefore, MP may interact with Zn finger motifs in proteins such as SDG16 for transcriptional activation of the viral genome.

It was speculated in Chapter 2, that CKL1, acts as an essential post translational modifier, which when hindered by SACMV MP may result in the accumulation of susceptibility related hormone signalling pathways of GA and ABA (Allie et al., 2014; Sun et al., 2020). Thus, it was proposed that targeting the *CKL1* gene would be linked to an increase in viral load. However, considering the suppression of *CKL1* being directly proportional to SACMV viral accumulation, it is more plausible that CKL1 plays a critical role in viral protein phosphorylation, a critical aspect for functionality of movement proteins in geminiviruses (Kleinow et al., 2009; Bisaro, 1996) and other viruses (Matsushita et al., 2000; Ivanov et al., 2003; Nemes et al., 2017; Citovsky et al., 1993). The phosphorylation of PVA coat protein by a casein kinase protein (Ivanov et al., 2003) further adds credence to CKL1 impacting both SACMV MP and other geminivirus phosphorylation dependant viral proteins C4 (Kumar, 2019; Florentino et al., 2006; Mei et al., 2018; Mills-Lujan et al., 2015). AC2 protein of geminiviruses is a transcriptional regulator possessing a zinc finger-like domain that allows interaction with host proteins to activate virus genome transcription (Guerrero et al., 2020). CKL family protein, CKL6, has also been shown to localise to and bind cortical microtubules via a C-terminal signalling domain (Ben-Nissan et al., 2008) suggesting that CKL1 may deliberately be harnessed for mobility and microtubular rearrangements.

CKL1 transcripts in SACMV transfected protoplasts were upregulated although downregulated in samples transfected with CRISPR/Cas and SACMV+CRISPR/Cas expression vectors. Additionally, the downregulated expression of *CKL1* was two-fold that observed in protoplast targeted with CRISPR/Cas only. SACMV induces mutations in E3 ligase (Chatukuta & Rey, 2020). It may be feasibly considered that modifications in the *CKL1* sequence induced by geminivirus infection may result in an additive effect when combined with CRISPR/Cas mediated editing. It would therefore be recommended that in future work entire gene sequences (as opposed to regions ≤ 200 bp) in SACMV infected and non-SACMV infected samples be carried out to either verify or rule out the occurrence of such a mechanism.

4.4.6 MUTATION FREQUENCY IS MULTIFACETED

Targeted gene silencing of *NUP98* showed a low percentage of mutations (10% and 20 %) compared to *CKL1* (40 % and 60 %) per sgRNA. In *CKL1*, 70 % of mutations were indels whilst in *NUP98* only 6 % were attributed to insertions and none to deletions. The remaining percentages were substitutions (30 % and 94 %). A high throughput evaluation of the rate of substitutions arising from CRISPR gene editing in human cells revealed an average frequency of 0.8 % (Hwang et al., 2020). In *Arabidopsis* 2.2 % of all recorded mutations were from base substitution (Shkryl et al., 2021). Prior work by Fauser and colleagues with the same pDeSp derived expression vector showed an abundance of indels across all targeted genomes resulting in frameshifts and substitutions were also noted (Fauser et al., 2014). Consequently, substitutions were also anticipated in this study. Most studies tend toward the identification of indels and not substitutions as plausible mutations due to the presence of potential artefacts and errors that arise from sequencing experiments. Therefore, actual substitution frequencies are under reported. However, recently, the targeting of DNA ligase IV in the NHEJ pathway resulted in a reduction of substitution frequencies at target sites indicating that NHEJ is a primary enforcer of nucleotide substitutions (Hwang et al., 2020). In *Brassica* species, the induction of substitutions across fifteen gRNAs was absent in thirteen gRNAs and biased towards two gRNAs depicting 14.3 % and 21.4 % substitution frequencies respectively (Yang et al., 2017). The occurrence of substitutions may be sgRNA sequence dependent. NHEJ mediated frequency of mutations vary relative to species and target regions and all mutations recorded previously by Chatukuta and Rey were substitution mutations (Chatukuta & Rey, 2020), and may also be a characteristic of CRISPR/Cas mediated gene editing in cassava protoplasts. In this study, observed substitutions ranged from 24 % to 65 %.

Target sgRNAs generally result in variable mutational efficiencies due to experimental conditions (Doench et al., 2016) and inherent factors such as GC content and the presence of nucleosomes may impact Cas9 binding and editing (Horlbeck et al., 2016). Regions

comprising of poly(dA:dT) and poly(dG:dC) inhibit nucleosome formation, whilst non-homopolymeric GC-rich regions favour nucleosome formation (Struhl & Segal, 2013). *CKL1* sgRNAs, gRNA1 and gRNA2 both contained 4(dG:dC) polymeric regions whilst NUP98 sgRNAs, gRNA1 and 2 contained non-homopolymeric (dG:dC) albeit NUP98 gRNA1contained a 4(dA:dT) polymeric region. A higher GC content has proof of correlation to higher mutation frequencies (Ren et al., 2020; Chatukuta & Rey, 2020). NUP98 gRNA1 and 2 had 48 % and 43 % GC content, respectively. *CKL1* gRNA1 and 2 had 70 % and 57 % GC content, respectively. A GC content percentage minimum of 65 % is result in higher mutation efficiencies exceeding 50 % whilst gRNAs of 30 % GC content resulted in mutation efficiencies below 30 % (Ren et al., 2020). From this objective it was noted that a GC content of 70 % accounted for 86 % of total observed mutations affiliated to *CKL1* and 70 % of total observations in *CKL1* and *NUP98*. With regards to *NUP98* the lower GC content of gRNA1 and gRNA2 jointly only accounted for 19 % of total observed mutations. However, other work has shown that gRNAs of GC content of 35 % to 70 % were comparable (Sentmanat et al., 2018). Therefore, gRNA efficiency cannot be fully ascertained by GC content alone.

4.5 SUMMARY

From the experimental outcomes of gene knockout in cassava TME3 protoplasts, a model of MP functionality in relation to the functions of *NUP98*, *CKL1* and *SDG1*6 is proposed (**Figure 4.19**). MP-interactors discussed here contribute to SACMV replication *in vivo* (protoplasts). Though not shown in this study it would influence symptoms and morphological changes *in planta* in correlation with favouring viral load

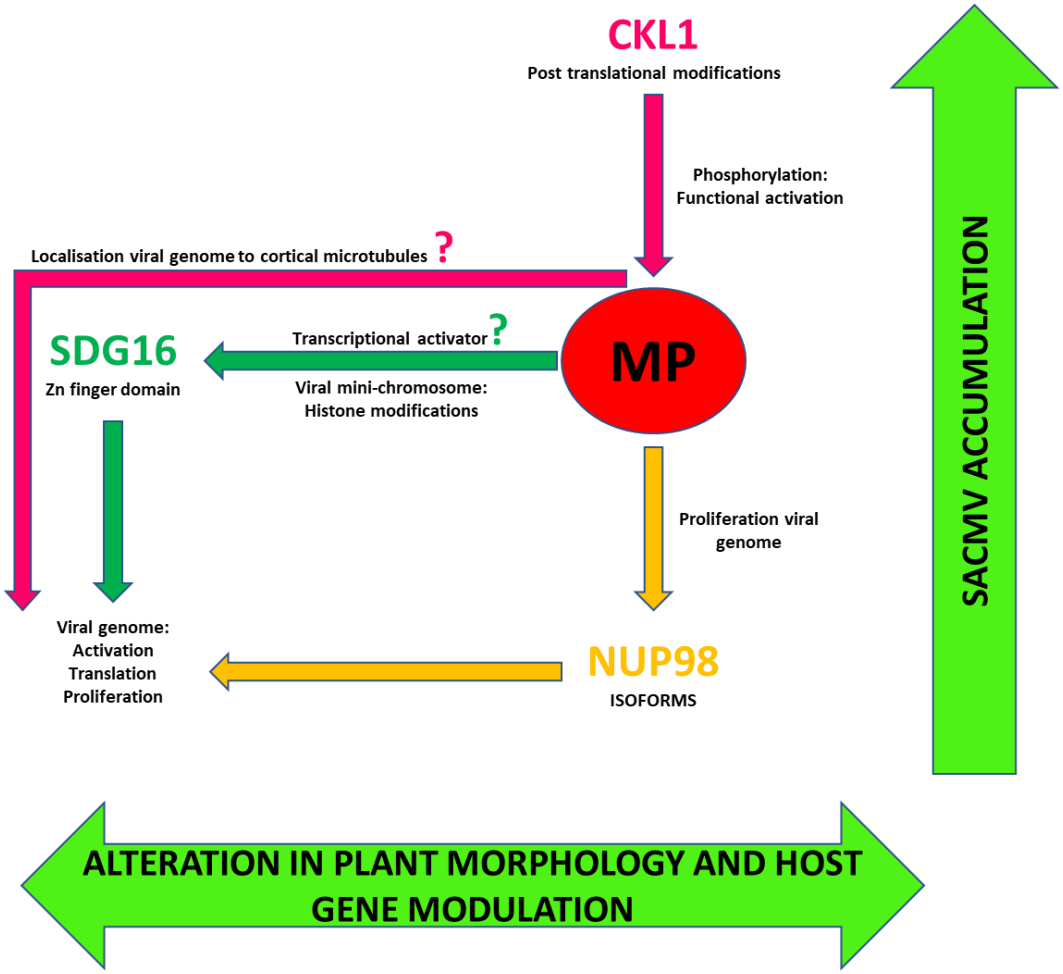


Figure 4.19 An illustration depicting the functional activation of SACMV MP post translational phosphorylation mediated by CKL1 (pink) and interaction with SDG16 (green) and NUP98 (orange) to mediate SACMV replication, translation and proliferation resulting in alterations of plant morphology and host gene modulation (lime green). Postulated mechanisms that would require further investigations followed by question marks. All three genes favour the accumulation of SACMV and interaction with MP would result in the alteration of host morphology and gene modulation. CKL1 activates MP by phosphorylation. SDG16 contains a zinc finger domain that may play a role in transcriptional activation. NUP98 and SDG16 play a key role in the activation, translation, and proliferation of viral SACMV. SDG16 facilitates histone modifications and the packaging of viral DNA in mini-chromosomes. NUP98 mediates viral DNA and MP movement across the nuclear membrane whilst SDG16 mediates viral DNA movement by translocation to cortical microtubules.

Chapter 5

SUMMARIES AND CONCLUSIONS

5.1 GENERAL OVERVIEW

The main aim of this research was to study and identify the specific pathways of interaction of the SACMV MP. Using Y2H and Co-IP assays and CRISPR/Cas9 mediated gene knockdown, it was possible to construct an elaborate model of the cassava- SACMV MP protein interactions and identify possible molecular processes that can be targeted to develop strategic disease resistance that can be potentially harnessed to minimise crop yield losses (Miles et al., 2016; Czosnek et al., 2013). Plant immunity involves multiple signalling proteins and complex molecular networks that require an integrated study approach that considers different interacting elements. Geminiviruses such as SACMV can devastate plant yield globally (Czosnek & Laterrot, 1997; Wasswa et al., 2011; Srivastava et al., 2017; Akinbade et al., 2010; Mandal et al., 2001; Qazi et al., 2007). Limited understanding of host-geminivirus protein interactions has confounded the development of strategies that confer resistance/tolerance against geminiviruses to increase plant yield for economic output. Identifying proteins of the geminiviral-host protein interactome provides information that is useful to understanding the mechanisms of geminiviral disease progression (Gergerich & Dolja, 2006). Within the host, the movement of the geminiviral genome is essential to successful systemic spread within the host (Frischmuth et al., 2007; Aberle et al., 2002; Krapp et al., 2017). With respect to this research, SACMV MP was explored both *in vivo* and *in vitro* using Y2H and Co-IP assays to identify binary interactions and complexes of interactors respectively. Both Y2H and Co-IP identified proteins involved in possible resistance associated pathways that suggest mechanisms involving various metabolic and replication associated pathways. This was followed by *in vivo* CRISPR-mediated gene editing for evaluation of a select number of identified host-MP interactors to bring into perspective the impact of knocking down host genes on the accumulation of SACMV. The accumulation of SACMV viral DNA quantified by quantitative PCR (qPCR) was significantly reduced in CRISPR mediated gene edited samples for genes *CKL1*, *SDG16* and *NUP98A* and therefore these genes were identified as essential to the transcription of proteins targeted by viral protein MP and therefore favourable for SACMV viral accumulation. Overall, our results show that SACMV MP interacts with proteins involved in defence mechanisms, alters host metabolism and the host cell cycle, and manipulates the host for viral DNA movement and replication.

5.2 NOVELTY OF FINDINGS

Histones and heat shock proteins and PUP1 have been previously identified by other researchers as interactors of geminiviral MP and nuclear shuttle protein (NSP) (Krenz et al., 2010; Krapp et al., 2017; Gouveia-mageste et al., 2021). This current research specifically identified Histone H3, HSP90 and PUP1 and therefore provided further evidence of the interaction of these proteins and their associated proteins in facilitating geminivirus systemic movement. Gene ontology categorisations identified ontologies unique to this current research. These were, for Cellular Component (CC) the "nuclear envelope" linked to interacting protein NUP98A; for Molecular Function (MF): "chromatin binding", "structural molecule activity" and "kinase activity" linked to PTEN2A, NAT, NUP98A, HSK and CKL1.; and for Biological Process (BP): "translation" and "cell-cell signalling" associated with KRS1 and CKL1 respectively. Host proteins NUP98, CKL1 and SDG16 were confirmed *in vivo* as MP interacting proteins and by gene editing with CRISPR-Cas9 further confirmed as susceptibility enhancing proteins that play roles in enhancing SACMV geminiviral replication. Therefore, further findings in this current research highlighted the association of MP to proteins associated with potential roles in direct movement of MP and the viral genome across the nuclear membrane (CC), MP post translational modifications and SACMV genome interactions (MF) and SACMV protein translation (BP) and replication.

5.3 ROLES ASSOCIATED TO MP GEMINIVIRAL PROTEIN

In general, geminiviral movement proteins are established in roles of movement of viral DNA and symptom progression conferred by post translational modifications for functionality (Kleinow et al., 2009; Wege & Jeske, 1998). In this current study, both Y2H and Co-IP assays, identified nineteen interactors as novel to SACMV MP in cassava and eighteen were novel to geminivirus MP that have not been previously shown. The multifunctionality of MP was appreciated in the diversity of the interacting proteins identified in their potential role to facilitate post-translational phosphorylation events (PTEN2A, CKL1) which would confer functionality for viral systemic intra- and inter-cellular movement (CKL1, LSH, PUP1, PTEN2A, NUP98, KRS1, SDG16, HSP90, CDC48), suppression of host immune signalling and other immune related responses (PTEN2A, SIP2, NAT, ACLB-2, SDG16, KRS1, GLDP2, NUP98, HSK), accumulation of susceptibility associated molecules (ABR1, CKL1, PUP1), dysregulation of the cell cycle (CDC48, LSH), viral replication and accumulation (LSH, KRS1, SDG16, NUP98, CDC48, H4) and symptom accumulation and pathogenesis (PsaD, mmDH1, SUS4). The multiple binding partners identified were further evidence of the hub protein nature of MP characteristic of its disordered secondary structure (Patil & Nakamura, 2006; Nankoo, 2018).

5.4 SCHEMATIC SUMMARY PROPOSED FOR MP FUNCTIONALITY WITHIN THE HOST

Figure 5.1 provides a schematic summary of the proposed impact of MP interaction with cassava TME3 host proteins identified in **Chapter 2** and **Chapter 3** and *in vivo* evaluations of interactions in **Chapter 4**.

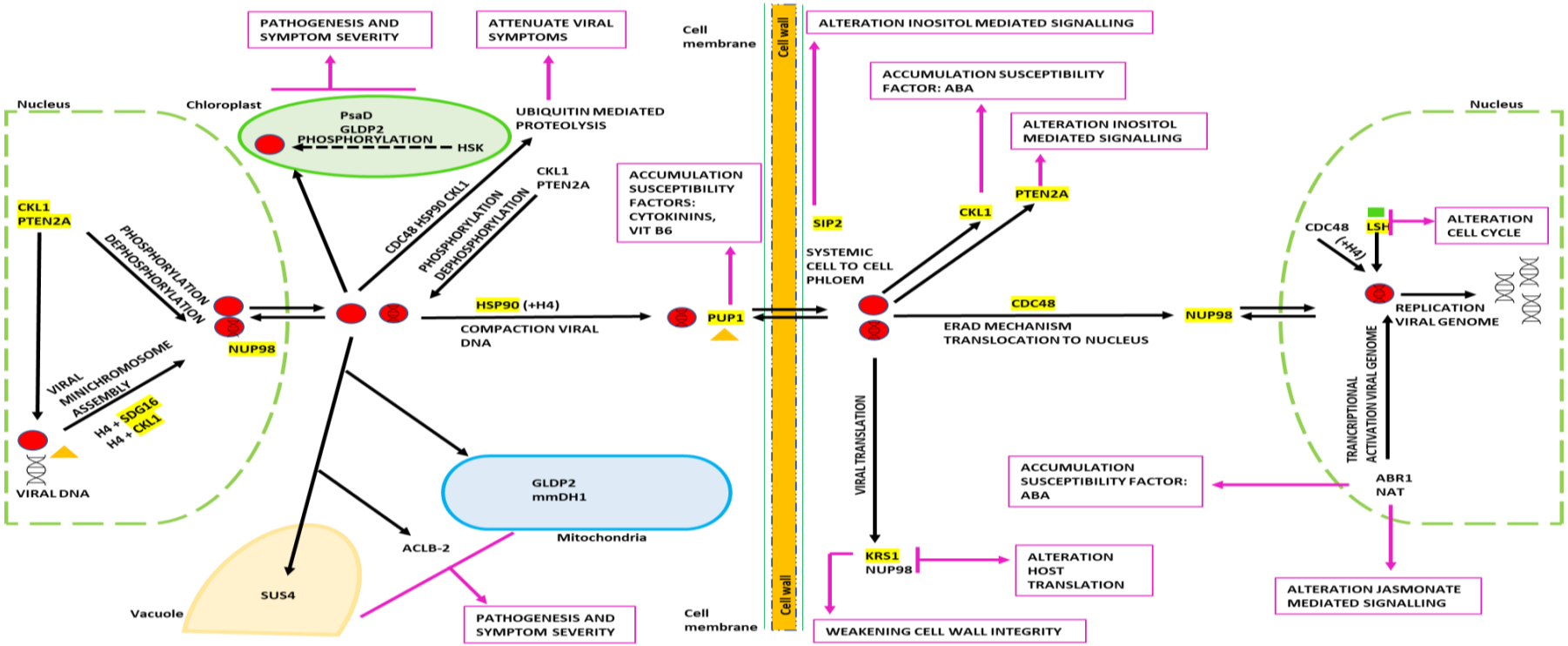


Figure 5.1 Schematic summary of MP activity. Interactions with *MP* are shown with black arrows. Interaction with uncertain inference is represented by a dashed black

arrow. *MP* (red circle), *MP* with SACMV DNA (red circle with DNA molecule). Viral proteins shown as, NSP (orange triangle), Rep (green square). Interactions are shown with black arrows. Movement of *MP* is shown via interactors highlighted in yellow. From the left: within the nucleus (and cytoplasm) the functionality of *MP* may be determined by dephosphorylation and phosphorylation events by CKL1 and PTEN2A. These events within the nucleus would allow *MP* to mediate SACMV viral minichromosome assembly by interaction with H4 and SDG16 to package viral DNA for movement in cooperation with viral NSP. Movement out of the nucleus into the cytoplasm may then be facilitated by NUP98. Within the cytoplasm interaction of *MP* with HSP90 would allow HSP90 to interact with viral H4 and further compact viral DNA for movement through cell membranes by interactions with PUP1 and viral NSP. The movement of *MP* systemically across cell walls is evidenced either via or by interaction with SIP2 which is localised to plasmodesma and involved in phloem trafficking. Additional interactions that occur within the cytoplasm may further facilitate movement by localisation to microtubules (CKL1, PTEN2A), a weakening of the cell wall (KRS1) and ERAD-mediated localisation to the nucleus (CDC48). Movement across the nuclear membrane may be facilitated by NUP98A. Additional interaction within the nucleus with LSH contribute to localisation of *MP* to microtubules. The proposed impact of interaction with host proteins is shown by pink arrows and text in pink rectangles. Within the cytoplasm, *MP* interacts with proteins located in cytoplasmic organelles (chloroplast, vacuole, mitochondria) potentially contributing to geminivirus pathogenesis and symptom severity by interference with various metabolic processes. These being, SUS4 (vacuole), mmDH1, GLDP2 (mitochondria, blue) and PsaD, GLDP2, HSK (chloroplast, green). HSK is specific to phosphorylation of homoserine but may interact with *MP* to this regard and is represented by dashed lines. CDC48, HSP90 and CKL1 facilitate ubiquitin mediated proteolysis which when based on the identified multifunctionality of *MP* may reduce both the viral DNA accumulation amelioration of symptoms. The reduction of *MP* would then reduce debilitating symptom severity. The interaction with PUP, CKL1, and ABR1 may result in the accumulation of susceptibility associated molecules of cytokinin, vitamin B6 and ABA. Interaction with PTEN2A and SIP2 would reduce inositol mediated defence signalling. Interaction with NAT would alter jasmonate mediated defence signalling. Viral translation may be promoted by interaction with KRS1, SDG16 and NUP98A. Interaction with KRS1 may lead to weakening of the cell wall whilst interaction with SDG16 and NUP98A may alter host translation. Activation of viral transcription within the nucleus may be facilitated by recruitment of host proteins ABR1 and NAT and viral genome accessibility by interaction with CDC48 and H4 to the host replisome. Interaction of *MP* with LSH may lead to dysregulation of the cell cycle cooperatively with viral Rep protein to enable viral genome replication.

5.5 LIMITATIONS AND FUTURE WORK

The MP of SACMV, remains under-explored in protein structure and protein interaction mechanisms. It would be ideal to identify phosphorylation and ATP binding sites present in MP to understand its functionality for targeted mutations to limit viral movement. Furthermore, this study has shown that MP function is closely linked to NSP (Gouveia-mageste et al., 2021) and replication (Rep) (Rey lab, unpublished data) proteins via identified MP interacting proteins (PUP1 and LSH) identified from Y2H screening assays. Therefore, it would be ideal to further investigate the mechanisms and MP regions involved concerning the co-operative roles of these geminiviral proteins with host proteins LSH and PUP1. Work in protoplasts is rapid and informative, however, movement cannot be assessed. *In planta* gene editing for overexpression and silencing in a heritable manner in either *Nicotiana benthamiana* or cassava would be ideal for each of the nineteen interactor proteins identified in this study. These plants could be assessed for both viral movement and replication. Furthermore, the assessment of their metabolite and proteomics profiles within silenced backgrounds could contribute to a more informative approach. All identified interactors within this study were identified from tolerant cassava TME3 during its

symptomatic infection phase by SACMV. It would be relevant to extend this work in the whole plant to study movement in susceptible and resistant cassava hosts and ascertain the impact on viral movement. Due to the relatedness of geminivirus bipartite species, the interactors identified in this research may also be extended to other bipartite geminivirus species in different plant host species of economic importance.

REFERENCES

Abd-El-Haliem, A.M. & Joosten, M.H.A.J., (2017). Plant phosphatidylinositol-specific phospholipase C at the center of plant innate immunity. *Journal of Integrative Plant Biology*, 59(3), pp.164–179.

Abell, B.M. et al., (2007). Post-translational integration of tail-anchored proteins is facilitated by defined molecular chaperones. *Journal of Cell Science*, 120(10), pp.1743–1751.

Aberle, H.J. et al., (2002). Localizing the movement proteins of Abutilon mosaic geminivirus in yeast by subcellular fractionation and freeze-fracture immuno-labelling. *Archives of Virology*, 147(1), pp.103–117.

Adriko, J. et al., (2011). Response of improved cassava varieties in Uganda to cassava mosaic disease (CMD) and their inherent resistance mechanisms. *African Journal of Agricultural Research*, 6(3), pp.521–531. Available at: http://www.academicjournals.org/AJAR.

Afshar-Mohammadian, M., Rezadoost, M.H. & Fallah, S.F., (2018). Comparative analysis and innovation of a simple and rapid method for high-quality RNA and DNA extraction of kiwifruit. *MethodsX*, 5(April), pp.352–361. Available at: https://doi.org/10.1016/j.mex.2018.03.008.

Akinbade, S.A. et al., (2010). First report of the East African cassava mosaic virus-Uganda (EACMV-UG) infecting cassava (Manihot esculenta) in Cameroon. *New Disease Reports*, 21(22).

Alazem, M., Lin, K.-Y. & Lin, N.-S., (2014). The Abscisic Acid Pathway Has Multifaceted Effects on the Accumulation of *Bamboo mosaic virus*. *Molecular Plant-Microbe Interactions*, 27(2), pp.177–189. Available at: http://apsjournals.apsnet.org/doi/10.1094/MPMI-08-13-0216-R.

Albers, S. & Czech, A., (2016). Exploiting tRNAs to boost virulence. *Life*, 6(1).

Alcântara, A. et al., (2019). Systematic Y2H Screening Reveals Extensive Effector-Complex Formation. *Frontiers in Plant Science*, 10(November).

Alene, A.D. et al., (2013). Economic impacts of cassava research and extension in Malawi and Zambia. *J. Dev. Agric. Econ.*, 5(11), pp.457–469. Available at: http://www.academicjournals.org/JDAE.

Ali, M.S. & Baek, K.H., (2020). Jasmonic acid signaling pathway in response to abiotic stresses in plants. *International Journal of Molecular Sciences*, 21(2).

Ali, Z. et al., (2016). CRISPR/Cas9-Mediated Immunity to Geminiviruses: Differential Interference and Evasion. *Scientific Reports*, 6(May). Available at: http://dx.doi.org/10.1038/srep26912.

Ali, Z. et al., (2015). CRISPR/Cas9-mediated viral interference in plants. *Genome Biology*, 16(238).

Alicai, T. et al., (2007). Re-emergence of Cassava Brown Streak Disease in Uganda. *Plant Disease*, 91(1), pp.24–29. Available at: http://apsjournals.apsnet.org/doi/10.1094/PD-91-0024.

Allem, A.C., (2001). The Origins and Taxonomy of Cassava. In J. M. Hillocks, R. J., Thresh, ed. CAB Internantional, pp. 1–16.

Allie, F., Pierce, Erica J, et al., (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, p.1006.

Allie, F., Pierce, Erica J., et al., (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(1).

Allie, F. & Rey, M.E.C., (2013). Transcriptional alterations in model host, Nicotiana benthamiana, in response to infection by South African cassava mosaic virus. *European Journal of Plant Pathology*, 137(4), pp.765–785.

Alonso-Peral, M.M., Sun, C. & Millar, A.A., (2012). MicroRNA159 Can Act as a Switch or Tuning MicroRNA Independently of Its Abundance in Arabidopsis.

Alvarez, J.M. et al., (2020). Transient genome-wide interactions of the master transcription factor NLP7 initiate a rapid nitrogen-response cascade. *Nature Communications*, 11(1), pp.1–13.

Alvarez, M.E. et al., (1998). Reactive Oxygen Intermediates Mediate a Systemic Signal Network in the Establishment of Plant Immunity. *Cell*, 92(6), pp.773–784.

Amuge, T. et al., (2017). A time series transcriptome analysis of cassava (Manihot esculenta Crantz) varieties challenged with Ugandan cassava brown streak virus. *Scientific Reports*, 7(1), p.9747. Available at: https://doi.org/10.1038/s41598-017-09617-z.

Andersson, M. et al., (2017). Efficient targeted multiallelic mutagenesis in tetraploid potato (Solanum tuberosum) by transient CRISPR-Cas9 expression in protoplasts. *Plant Cell Reports*, 36(1), pp.117–128.

Anjanappa, R.B. et al., (2017). Early transcriptome response to brown streak virus infection in susceptible and resistant cassava varieties. *bioRxiv*, pp.1–22.

Aranda, P.S., LaJoie, D.M. & Jorcyk, C.L., (2012). NIH Public Access: Bleach Gel. *NAtional Institutes Health*, 33(2), pp.366–369.

Artigas-Jerónimo, S. et al., (2020). A novel combined scientific and artistic approach for the advanced characterization of interactomes: The akirin/subolesin model. *Vaccines*, 8(1).

Ascencio-Ibanez, J.T. et al., (2008). Global Analysis of Arabidopsis Gene Expression Uncovers a Complex Array of Changes Impacting Pathogen Response and Cell Cycle during Geminivirus Infection. *Plant Physiology*, 148(1), pp.436–454. Available at: http://www.plantphysiol.org/cgi/doi/10.1104/pp.108.121038.

Ayres, F.M., (2018). Characterisation of the replication-associated protein of South African cassava mosaic virus and elucidation of protein binding partners in cassava.

Azali, H.A. et al., (2017). Occurrence of cassava brown streak disease and associated *Cassava brown streak virus* and *Ugandan cassava brown streak virus* in the Comoros Islands. *New Disease Reports*, 36, p.19. Available at: http://www.ndrs.org.uk/article.php?id=036019.

Bach-Pages, M. et al., (2020). Discovering the RNA-binding proteome of plant leaves with an improved RNA interactome capture method. *Biomolecules*, 10(4), pp.1–21.

Baliji, S., Lacatus, G. & Sunter, G., (2010). The Interaction between geminivirus pathogenicity proteins and adenosine kinase leads to increased expression of primary cytokinin-responsive genes. *Virology*, 402(2), pp.238–247. Available at: http://dx.doi.org/10.1016/j.virol.2010.03.023.

Barreiro-Alonso, A. et al., (2019). Characterization of HMGB1/2 interactome in prostate cancer by yeast two hybrid approach: Potential pathobiological implications. *Cancers*, 11(11), pp.1–21.

Bart, R. et al., (2012). Correction for " High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance, " by. *Proc Natl Acad Sci Plant Mol Biol*, 28109(18). Available at: www.pnas.org/cgi/doi/10.1073/pnas.1211014109.

Bartel, P. et al., Using the two-hybrid system to detect protein-protein interactions. In H. DA, ed. *In: Cellular Interactions in Development: A Practical Approach*. Oxford: Oxford University Press, pp. 153–179.

Bates, D.C. & Chant, S.R., (1975). The effects of virus infection on oxygen uptake and respiratory quotient of the leaves of French bean and cowpea. *Physiological Plant Pathology*, 6(2), pp.199–209.

Bäumler, J. et al., (2019). AtERF#111/ABR1 is a transcriptional activator involved in the wounding response. *Plant Journal*, 100(5), pp.969–990.

Baxter, C.J. et al., (2007). The Metabolic Response of Heterotrophic Arabidopsis Cells to Oxidative Stress 1 [W]. , 143(January), pp.312–325.

Bègue, H. et al., (2019). Toward the understanding of the role of CDC48, a major component of the protein quality control, in plant immunity. *Plant Science*, 279(November 2018), pp.34–44. Available at: https://doi.org/10.1016/j.plantsci.2018.10.029.

Behnam, B. et al., (2019). An optimized isolation protocol yields high-quality RNA from cassava tissues (Manihot esculenta Crantz). *FEBS Open Bio*, 9(4), pp.814–825.

Belkhadir, Y. et al., (2004). Arabidopsis RIN4 Negatively Regulates Disease Resistance Mediated by RPS2 and RPM1 Downstream or Independent of the NDR1 Signal Modulator and Is Not Required for the Virulence Functions of Bacterial Type III Effectors AvrRpt2 or AvrRpm1. *The Plant Cell*, 16(10), pp.2822–2835.

Bellin, D. et al., (2013). Nitric Oxide as a Mediator for Defense Responses. , 26(3), pp.271–277. Available at: http://dx.doi.org/10.1094/MPMI-09-12-0214-CR.

Ben-Nissan, G. et al., (2008). Arabidopsis casein kinase 1-like 6 contains a microtubule-binding domain and affects the organization of cortical microtubules. *Plant Physiology*, 148(4), pp.1897–1907.

Bengyella, L. et al., (2015). Virus tolerance and recovery from viral induced-symptoms in plants are associated with transcriptome reprograming. *Plant Molecular Biology*, 89(3), pp.243–252.

Bernoux, M., Ellis, J.G. & Dodds, P.N., (2011). New insights in plant immunity signaling activation. *Current Opinion in Plant Biology*.

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. *Journal of General Virology*.

Berry, S. & Rey, M.E.C., (2001). Molecular evidence for diverse populations of cassava-infecting begomoviruses in southern Africa. *Archives of Virology*.

Beyene, G., Chauhan, R.D. & Taylor, N.J., (2017). A rapid virus-induced gene silencing (VIGS) method for assessing resistance and susceptibility to cassava mosaic disease. *Virology Journal*, 14(1), pp.1–11.

Beyene Getu, Chauhan Raj Deepika, Wagaba Henry, Moll Theodore, Alicai Titus, Mian Doouglas , Carrington James C., T.N.J., (2016). Loss of CMD2-mediated resistance to cassava mosaic disease in plants regenerated through somatic embryogenesis. *Molecular Plant Pathology*, 17(7), pp.1095–110.

Bhattacharyya, D. et al., (2015). A geminivirus betasatellite damages the structural and

functional integrity of chloroplasts leading to symptom formation and inhibition of photosynthesis. *Journal of Experimental Botany*, 66(19), pp.5881–5895.

Bhor, S.A. et al., (2017). Inducible transgenic tobacco system to study the mechanisms underlying chlorosis mediated by the silencing of chloroplast heat shock protein 90. *VirusDisease*, 28(1), pp.81–92.

Bieniawska, Z. et al., (2007). Analysis of the sucrose synthase gene family in Arabidopsis. *Plant Journal*, 49(5), pp.810–828.

Bioinformatics, B. et al., (2006). Effect of data normalization on fuzzy clustering of DNA microarray data. *BMC Bioinformatics2006 BMC Bioinformatics*, 7(7). Available at: http://www.biomedcentral.com/1471-2105/7/134.

Bisaro, D.M., (1996). Geminivirus DNA Replication. In M. L. DePamphilis, ed. *DNA Replication in Eukaryofic Cells*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 833–854.

Bizabani, C., Rogans, S.J. & Rey, M.E.C., (2021). *Differential miRNA profiles in South African cassava mosaic virus-infected cassava landraces reveal clues to susceptibility and tolerance to cassava mosaic disease*, Elsevier B.V. Available at: https://doi.org/10.1016/j.virusres.2021.198400.

Black, J.C., Van Rechem, C. & Whetstine, J.R., (2012). Histone Lysine Methylation Dynamics: Establishment, Regulation, and Biological Impact. *Molecular Cell*, 48(4), pp.491–507. Available at: http://dx.doi.org/10.1016/j.molcel.2012.11.006.

Blevins, T. et al., (2006). Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic Acids Research*, 34(21), pp.6233–6246.

Bokanga M, (1995). Cassava : Opportunities for the food, feed, and other industries in Africa. In S. T. T. Agbor Egbe, A. Brauman, D. Griffon, ed. *Transformation Alimentaire du Manioc*. pp. 557–569.

Bologna, N.G. & Voinnet, O., (2014). The Diversity, Biogenesis, and Activities of Endogenous Silencing Small RNAs in Arabidopsis. *Annu. Rev. Plant Biol*, 65, pp.473–503.

Bonizec, M. et al., (2014). The ubiquitin-selective chaperone Cdc48/p97 associates with Ubx3 to modulate monoubiquitylation of histone H2B. *Nucleic Acids Research*, 42(17), pp.10975–10986.

Bosque, G. et al., (2014). Topology analysis and visualization of Potyvirus protein-protein interaction network. *BMC systems biology*, 8, p.129.

Bossing, T. et al., (2012). Disruption of Microtubule Integrity Initiates Mitosis during CNS Repair. *Developmental Cell*, 23(2), pp.433–440.

Böttcher, B. et al., (2004). Geminate Structures of African Cassava Mosaic Virus. *JOURNAL OF VIROLOGY*, 78(13), pp.6758–6765.

Boyko, V. et al., (2000). Function of microtubules in intercellular transport of plant virus RNA. *Nature Cell Biology*, 2(11), pp.826–832.

Brandt, K.M. et al., (2020). A Streamlined Protocol for Wheat (Triticum aestivum) Protoplast Isolation and Transformation With CRISPR-Cas Ribonucleoprotein Complexes. *Frontiers in Plant Science*, 11(June), pp.1–14.

Braun, P. et al., (2009). An experimentally derived confidence score for binary protein-protein interactions. *Nature Methods*, 6(1), pp.91–97.

Bredeson, J. V. et al., (2016). Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. *Nature*

Biotechnology, 34(5), pp.562–570.

Brohée Sylvain and van Helden Jacques, (2006). Evaluation of clustering algorithms for protein-protein interaction networks. *BMC Bioinformatics*, 7(488), pp.2463–2468. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12351582.

Broquist, H.P., (1991). Lysine-pipecolic acid metabolic relationships in microbes and mammals. *Annual Review of Nutrition*, 11, pp.435–448.

Brown, J.K. et al., (2015). Revision of Begomovirus taxonomy based on pairwise sequence comparisons. *Archives of virology*.

Bruce, G. et al., (2011). Influence of retinoblastoma-related gene silencing on the initiation of DNA replication by African cassava mosaic virus Rep in cells of mature leaves in Nicotiana benthamiana plants. *Virology Journal*, 8(1), pp.1–6. Available at: http://dx.doi.org/10.1155/2013/783253.

Brückner, A. et al., (2009). Yeast two-hybrid, a powerful tool for systems biology. *International Journal of Molecular Sciences*.

Brunetti, a et al., (1997). High expression of truncated viral rep protein confers resistance to tomato yellow leaf curl virus in transgenic tomato plants. , 10(5), pp.571–579.

Buchmann, R.C. et al., (2009). Geminivirus AL2 and L2 Proteins Suppress Transcriptional Gene Silencing and Cause Genome-Wide Reductions in Cytosine Methylation. *Journal of Virology*, 83(10), pp.5005–5013.

Burgess, J., Motoyoshi, F. & Fleming, E.N., (1973). The mechanism of infection of plant protoplasts by viruses. *Planta*, 112(4), pp.323–332.

CABI and EPPO, (2016). Data Sheets on Quarantine Pests. Available at: http://www.cabi.org/isc/datasheet/8927 [Accessed September 20, 2016].

Cagney, G. & Uetz, P., (2001). High-throughput screening for protein-protein interactions using yeast two-hybrid arrays. *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]*, Chapter 19, pp.1–12.

Cai, X.T. et al., (2014). Arabidopsis ERF109 mediates cross-talk between jasmonic acid and auxin biosynthesis during lateral root formation. *Nature Communications*, 5(May), pp.1–13.

Caisová, L. & Jobe, T.O., (2019). Regeneration and transient gene expression in protoplasts of Draparnaldia (chlorophytes), an emerging model for comparative analyses with basal streptophytes. *Plant Methods*, 15(1), pp.1–14. Available at: https://doi.org/10.1186/s13007-019-0460-6.

Cantoni, G.L., (1952). The nature of the active methyl donor formed enzymatically from l-methionine and adenosinetriphosphate. *Journal of the American Chemical Society*, 74(11), pp.2942–2943.

Cao, Y. et al., (2020). Control of Plant Viruses by CRISPR/Cas System-Mediated Adaptive Immunity. *Frontiers in Microbiology*, 11(October), pp.1–9.

Carmo, L.S.T. et al., (2013). Identification of host proteins modulated by the virulence factor AC2 of Tomato chlorotic mottle virus in Nicotiana benthamiana. *Proteomics*, 13(12–13), pp.1947–1960.

Carmo, L.S.T. et al., (2017). Plant responses to tomato chlorotic mottle virus: Proteomic view of the resistance mechanisms to a bipartite begomovirus in tomato. *Journal of Proteomics*, 151, pp.284–292. Available at: http://dx.doi.org/10.1016/j.jprot.2016.07.018.

Carvalho, C.M. et al., (2008). A novel nucleocytoplasmic traffic GTPase identified as a

functional target of the bipartite geminivirus nuclear shuttle protein. *Plant Journal*, 55(5), pp.869–880.

Carvalho, M.F. & Lazarowitz, S.G., (2004). Interaction of the Movement Protein NSP and the Arabidopsis Acetyltransferase AtNSI Is Necessary for Cabbage Leaf Curl Geminivirus Infection and Pathogenicity. *JOURNAL OF VIROLOGY*, 78(20), pp.11161–11171.

De Carvalho, R. & Guerra, M., (2002). Cytogenetics of Manihot esculenta Crantz (cassava) and eight related species. *Hereditas*, 136, pp.159–168.

Castillo-González, C. et al., (2015). Geminivirus-encoded TrAP suppressor inhibits the histone methyltransferase SUVH4/KYP to counter host defense. *eLife*, 4(September), pp.1–31.

Castillo, a G. et al., (2004). Interaction between a geminivirus replication protein and the plant sumoylation system. *Journal of virology*, 78(6), pp.2758–69. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=353736&tool=pmcentrez&rendertype=abstract.

Castillo, A.G. et al., (2003). Dual interaction of plant PCNA with geminivirus replication accessory protein (REn) and viral replication protein (Rep). *Virology*, 312(2), pp.381–394.

Cauchy, P. et al., (2019). 2HybridTools, a handy software to facilitate clone identification and mutation mapping from yeast two-hybrid screening. *PeerJ*, 2019(7).

Ceccaldi, R., Rondinelli, B. & D 'andrea, A.D., (2016). Repair Pathway Choices and Consequences at the Double- Strand Break Mechanisms of DNA DSB Repair. *Trends in Biochemical Sciences*, 26(1), pp.52–64. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4862604/pdf/nihms719094.pdf.

Ceniceros-Ojeda, E.A., Rodríguez-Negrete, E.A. & Rivera-Bustamante, R.F., (2016). Two Populations of Viral Minichromosomes Are Present in a Geminivirus-Infected Plant Showing Symptom Remission (Recovery). *Journal of Virology*, 90(8), pp.3828–3838.

Chanda, B. et al., (2011). Glycerol-3-phosphate is a critical mobile inducer of systemic immunity in plants. *Nat Genet*, 43(5), pp.421–427. Available at: http://dx.doi.org/10.1038/ng.798.

Chandra-Shekara, A.C. et al., (2006). Light-dependent hypersensitive response and resistance signaling against Turnip Crinkle Virus in Arabidopsis. *The Plant Journal*, 45, pp.320–334.

Chandrasekhar, T., Thangavel, K. & Elayaraja, E., (2011). Effective Clustering Algorithms for Gene Expression Data. *International Journal of Computer Applications*, 32(4), pp.975–8887.

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

Chaozu, R. et al., (2018). Membrane-binding and activation mechanism of PTEN. *Scientific Reports*, 8(1), pp.1–17. Available at: http://journal.frontiersin.org/article/10.3389/fpls.2013.00252/full.

Chatterji, A., Beachy, R.N. & Fauquet, C.M., (2001). Expression of the Oligomerization Domain of the Replication-associated Protein (Rep) of Tomato Leaf Curl New Delhi Virus Interferes with DNA Accumulation of Heterologous Geminiviruses. *Journal of Biological Chemistry*, 276(27), pp.25631–25638.

Chatukuta, P. & Rey, M.E.C., (2020). A cassava protoplast system for screening genes

associated with the response to South African cassava mosaic virus. *Virology Journal*, 17(1), pp.1–15. Available at: https://doi.org/10.1186/s12985-020-01453-4.

Chellappan, P., Vanitharani, R. & Fauquet, C.M., (2005). MicroRNA-binding viral protein interferes with Arabidopsis development. *Proceedings of the National Academy of Sciences of the United States of America* , 102(29), pp.10381–10386. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1177406/.

Chen, H.H. et al., (2018). EL1-like Casein Kinases Suppress ABA Signaling and Responses by Phosphorylating and Destabilizing the ABA Receptors PYR/PYLs in Arabidopsis. *Molecular Plant*, 11(5), pp.706–719.

Chen, L.Q. et al., (2017). ATX3, ATX4, and ATX5 encode putative H3K4 methyltransferases and are critical for plant development. *Plant Physiology*, 174(3), pp.1795–1806.

Chen, T. et al., (2013). Comparative transcriptome profiling of a resistant vs. susceptible tomato (Solanum lycopersicum) cultivar in response to infection by tomato yellow leaf curl virus. *PLoS ONE*, 8(11).

Chen, Y. & Xu, D., (2003). Computational analyses of high-throughput protein-protein interaction data. *Current protein & peptide science*, 4, pp.159–181.

Cheng, Z. et al., (2020). NUP96 and HOS1 are mutually stabilized and gate constans protein level, conferring long-day photoperiodic flowering regulation in Arabidopsis. *Plant Cell*, 32(2), pp.374–391.

Cheng, Z. et al., (2011). The bHLH transcription factor MYC3 interacts with the jasmonate ZIM-domain proteins to mediate jasmonate response in arabidopsis. *Molecular Plant*, 4(2), pp.279–288. Available at: http://dx.doi.org/10.1093/mp/ssq073.

Choe, S. et al., (2021). Tolerance to tomato yellow leaf curl virus in transgenic tomato overexpressing a cellulose synthase-like gene. *Plant Biotechnology Journal*, 19(4), pp.657–659.

Choi, Y. et al., (2016). Isolation and characterization of pepper genes interacting with the CMV-P1 helicase domain. *PLoS ONE*, 11(1), pp.1–14.

Christensen, N. et al., (2009). The 5′ cap of tobacco mosaic virus (TMV) is required for virion attachment to the actin/endoplasmic reticulum network during early infection. *Traffic*, 10(5), pp.536–551.

Chu, V.T. et al., (2015). Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nature Biotechnology*, 33(5), pp.543–548.

Citovsky, V. et al., (1993). Phosphorylation of tobacco mosaic virus cell-to-cell movement protein by a developmentally regulated plant cell wall-associated protein kinase. *Genes and Development*, 7(5), pp.904–910.

Coello, P. et al., (2012). Evidence that abscisic acid promotes degradation of SNF1-related protein kinase (SnRK) 1 in wheat and activation of a putative calcium-dependent SnRK2. *Journal of Experimental Botany*, 63(2), pp.913–924.

Coleman, H.D., Yan, J. & Mansfield, S.D., (2009). Sucrose synthase affects carbon partitioning to increase cellulose production and altered cell wall ultrastructure. *Proceedings of the National Academy of Sciences of the United States of America*, 106(31), pp.13118–13123.

Cong, L. et al., (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*, 339(6121), pp.819 LP – 823. Available at: http://science.sciencemag.org/content/339/6121/819.abstract.

Cookson, S.J. et al., (2014). Heterografting with nonself rootstocks induces genes involved in stress responses at the graft interface when compared with autografted controls. *Journal of Experimental Botany*.

Courdavault, V. & Besseau, S., (2020). *Virus-Induced Gene Silencing in Plants Methods and Protocols (Methods in Molecular Biology)*, Humana Press. Available at: http://www.springer.com/series/7651.

Cruz, C. et al., (2018). Tri-methylation of histone h3 lysine 4 facilitates gene expression in ageing cells. *eLife*, 7, pp.1–24.

Culver, J.N. & Padmanabhan, M.S., (2007). Virus-induced disease: altering host physiology one interaction at a time. *Annual review of phytopathology*, 45, pp.221–243.

Curnutte, L.B., Simmons, A.M. & Abd-Rabou, S., (2014). Climate Change and Bemisia tabaci (Hemiptera: Aleyrodidae): Impacts of Temperature and Carbon Dioxide on Life History. *Ann. Entomol. Soc. Am*, 107(5), pp.933–943. Available at: http://dx.doi.org/10.1603/AN13143.

Cutler, S.R. et al., (2010). *Abscisic acid: Emergence of a core signaling network*,

Czosnek, H. et al., (2013). Discovering host genes involved in the infection by the Tomato yellow leaf curl virus complex and in the establishment of resistance to the virus using Tobacco rattle virus-based post transcriptional gene silencing. *Viruses*, 5(3), pp.998–1022.

Czosnek, H. & Laterrot, H., (1997). A worldwide survey of tomato yellow leaf curl viruses. *Arch Virol*, 142(7), pp.1391–1406.

Day, B., Dahlbeck, D. & Staskawicz, B.J., (2006). NDR1 Interaction with RIN4 Mediates the Differential Activation of Multiple Disease Resistance Pathways in Arabidopsis. *The Plant cell*, 18(10), pp.2782–2791.

Deblasio, S.L. et al., (2015). Insights into the Polerovirus-plant interactome revealed by coimmunoprecipitation and mass spectrometry. *Molecular Plant-Microbe Interactions*, 28(4), pp.467–481.

DeBlasio, S.L. et al., (2018). Disruption of chloroplast function through downregulation of phytoene desaturase enhances the systemic accumulation of an aphid-borne, phloem-restricted virus. *Molecular Plant-Microbe Interactions*, 31(10), pp.1095–1110.

DeBlasio, S.L. et al., (2017). Evaluation of a Bead-Free Coimmunoprecipitation Technique for Identification of Virus–Host Protein Interactions Using High-Resolution Mass Spectrometry. *Journal of Biomolecular Techniques : JBT*, (August), p.jbt.17-2803-002. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5524270/.

DeJesus, R. et al., (2016). Functional CRISPR screening identifies the ufmylation pathway as a regulator of SQSTM1/p62. *eLife*.

Dharmaraj, K. et al., (2019). Construction of a kiwifruit yeast two-hybrid cDNA library to identify host targets of the Pseudomonas syringae pv. actinidiae effector AvrPto5. *BMC Research Notes*, 12(1), pp.1–6. Available at: https://doi.org/10.1186/s13104-019-4102-x.

Ding, B. et al., (1995). Cucumber Mosaic Virus 3a Protein Potentiates Cell-to-Cell Trafficking of CMV RNA in Tobacco Plants. *Virology*, 207(2), pp.345–353.

Ding, Z. et al., (2016). Transcriptome response of cassava leaves under natural shade. *Scientific Reports*.

Doench, J.G. et al., (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nature Biotechnology*, 34(2), pp.184–191. Available at: http://dx.doi.org/10.1038/nbt.3437.

Dreher, T.W., (2009). Role of tRNA-like structures in controlling plant virus replication. *Virus Research*, 139(2), pp.217–229.

Dubiella, U. et al., (2013). Calcium-dependent protein kinase/NADPH oxidase activation circuit is required for rapid defense signal propagation. *Proceedings of the National Academy of Sciences*, 110(21), pp.8744–8749.

Durfee, T. et al., (1993). The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. *Genes and Development*, 7(4), pp.555–569.

Dutta, S., Teresinski, H.J. & Smith, M.D., (2014). A split-ubiquitin yeast two-hybrid screen to examine the substrate specificity of atToc159 and atToc132, two arabidopsis chloroplast preprotein import receptors. *PLoS ONE*, 9(4).

Earley, K.W. & Poethig, R.S., (2011). Binding of the cyclophilin 40 ortholog SQUINT to Hsp90 protein is required for SQUINT function in Arabidopsis. *Journal of Biological Chemistry*.

Ebina, H. et al., (2004). Role of Nup98 in nuclear entry of human immunodeficiency virus type 1 cDNA. *Microbes and Infection*, 6(8), pp.715–724.

Elena, S.F. & Rodrigo, G., (2012). Towards an integrated molecular model of plant – virus interactions. *Current Opinion in Virology*, 2(6), pp.719–724. Available at: http://dx.doi.org/10.1016/j.coviro.2012.09.004.

Emily, A.M. et al., (2016). Occurrence and estimated losses caused by cassava viruses in Migori County, Kenya. *African Journal of Agricultural Research*, 11(24), pp.2064–2074.

Engel, N. et al., (2007). Deletion of glycine decarboxylase in arabidopsis is lethal under nonphotorespiratory conditions. *Plant Physiology*, 144(3), pp.1328–1335.

Evans, J.R. & Caemmerer, V., (1987). Would C4 rice produce more biomass than C3 rice ? *Ratio*, (Hatch).

FAOSTAT, (2020). FAOSTAT. Available at: http://www.fao.org/ag/save-and-grow/cassava/en/1/index.html [Accessed May 1, 2020].

Faria, M. & Wraight, S.P., (2001). Biological control of Bemisia tabaci with fungi. *Crop Protection*, 20, pp.767–778.

Fashena, S.J., Serebriiskii, I.G. & Golemis, E.A., (2000). LexA-based two-hybrid systems. *Methods in Enzymology*, 328(1991), pp.14–26.

Fatland, B.L. et al., (2002). Molecular characterization of a heteromeric ATP-citrate lyase that generates cytosolic acetyl-coenzyme A in arabidopsis. *Plant Physiology*, 130(2), pp.740–756.

Fauser, F., Schiml, S. & Puchta, H., (2014). Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. *Plant Journal*, 79(2), pp.348–359.

Fernández-Calvino, L. et al., (2014). Virus-induced alterations in primary metabolism modulate susceptibility to Tobacco rattle virus in arabidopsis. *Plant Physiology*, 166(4), pp.1821–1838.

Fields, S. & Song, O.K., (1989). A novel genetic system to detect protein-protein interactions. *Nature*, 340(6230), pp.245–246.

Filomeni, G., De Zio, D. & Cecconi, F., (2015). Oxidative stress and autophagy: The clash between damage and metabolic needs. *Cell Death and Differentiation*, 22(3), pp.377–388.

Florentino, L.H. et al., (2006). A PERK-Like Receptor Kinase Interacts with the Geminivirus Nuclear Shuttle Protein and Potentiates Viral Infection. *Journal of Virology*, 80(13), pp.6648–6656. Available at: http://jvi.asm.org/cgi/doi/10.1128/JVI.00173-06.

Fondong, V.N. et al., (2000). Evidence of synergism between African cassava mosaic virus and a new double-recombinant geminivirus infecting cassava in Cameroon. *Journal of General Virology*.

Fondong, V.N., (2013). Geminivirus protein structure and function. *Molecular Plant Pathology*, 14(6), pp.635–649.

Fontes, E.P.B. et al., (2004a). The geminivirus nuclear shuttle protein is a virulence factor that suppresses transmembrane receptor kinase activity. *Genes and Development*, 18(20), pp.2545–2556. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC529541/.

Fontes, E.P.B. et al., (2004b). The geminivirus nuclear shuttle protein is a virulence factor that suppresses transmembrane receptor kinase activity. *Genes and Development*, 18(20), pp.2545–2556.

Formstecher, E. et al., (2005). Protein interaction mapping: A Drosophila case study. *Genome Research*, 15(3), pp.376–384.

Franceschini, A. et al., (2013). STRING v9.1: Protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Research*.

Franks, T.M. et al., (2017). Nup98 recruits the Wdr82-Set1A/COMPASS complex to promoters to regulate H3K4 trimethylation in hematopoietic progenitor cells. *Genes and Development*, 31(22), pp.2222–2234.

Fregene, M. et al., (2004). Serial analysis of gene expression (SAGE) of host-plant resistance to the cassava mosaic disease (CMD). *Plant Molecular Biology*, 56(4), pp.563–571.

Frischmuth, S. et al., (2007). The movement protein BC1 promotes redirection of the nuclear shuttle protein BV1 of Abutilon mosaic geminivirus to the plasma membrane in fission yeast. *Protoplasma*, 230(1–2), pp.117–123.

Frischmuth, S. et al., (2004). Yeast two-hybrid systems confirm the membrane-association and oligomerization of BC1 but do not detect an interaction of the movement proteins BC1 and BV1 of Abutilon mosaic geminivirus. *Archives of Virology*, 149(12), pp.2349–2364.

Fromm, M., Taylor, L.P. & Walbot, V., (1985). Expression of genes transferred into monocot and dicot plant cells by electroporation. *Proceedings of the National Academy of Sciences of the United States of America*, 82(17), pp.5824–5828.

Fromont-racine, M., Rain, J. & Legrain, P., (1997). Through Exhaustive Two-Hybrid Screens. , 16(july), pp.25–28.

Frow, E. & Yearley, S., (2010). People, Policy and Plant Genomics. In A. A. G. Kole Chittaranjan, ed. *Principles and Practices of Plant Genomics, Volume 3: Advanced Genomics*. Crc press, pp. 528–570.

Fu, L. et al., (2016). Physiological Investigation and Transcriptome Analysis of Polyethylene Glycol (PEG)-Induced Dehydration Stress in Cassava. *International Journal of Molecular Sciences*.

Fu, Z.Q. & Dong, X., (2013). Systemic Acquired Resistance: Turning Local Infection into Global Defense. *Annu. Rev. Plant Biol*, 64, pp.839–63.

Gallemí, M. et al., (2016). DRACULA2 is a dynamic nucleoporin with a role in regulating the shade avoidance syndrome in Arabidopsis. *Development (Cambridge)*, 143(9), pp.1623–1631.

Ganguly, A. et al., (2020). FRA1 kinesin modulates the lateral stability of cortical microtubules through cellulose synthase–Microtubule uncoupling proteins. *Plant Cell*, 32(8), pp.2508–2524.

Gao, J.X. et al., (2015). Construction of a high-quality yeast two-hybrid library and its application in identification of interacting proteins with Brn1 in curvularia lunata. *Plant Pathology Journal*, 31(2), pp.108–114.

García-Neriaand, M.A. & Rivera-Bustamante, R.F., (2010). Characterization of Geminivirus Resistance in an Accession of Capsicum chinense Jacq. *The American Phytopathological Society*, 24(2), pp.72–182.

Garcia, D., Garcia, S. & Voinnet, O., (2014). Nonsense-mediated decay serves as a general viral restriction mechanism in plants. *Cell Host and Microbe*, 16(3), pp.391–402. Available at: http://dx.doi.org/10.1016/j.chom.2014.08.001.

Gaudet, P. et al., (2011). Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. *Briefings in Bioinformatics*, 12(5), pp.449–462.

Ge, S.X. et al., (2020). ShinyGO: A graphical gene-set enrichment tool for animals and plants. *Bioinformatics*, 36(8), pp.2628–2629.

Geisler-Lee, J. et al., (2007). A Predicted Interactome for Arabidopsis. *Plant Physiology*, 145(2), pp.317–329.

Gergerich, R.C. & Dolja, V. V., (2006). Introduction to plant viruses, the invisible foe. *Plant Health Instructor*.

Ghandi, A. et al., (2016). Tomato yellow leaf curl virus infection mitigates the heat stress response of plants grown at high temperatures. *Scientific Reports*, 6(January), pp.1–13.

Ghosh, I. et al., (2000). Antiparallel Leucine Zipper-Directed Protein Reassembly : Application to the Green Fluorescent Protein Department of Molecular Biophysics and Biochemistry The dissection and subsequent reassembly of a protein from peptidic fragments provides an avenue for. , (11), pp.5658–5659.

Gilioli, G. et al., (2014). Modelling the potential distribution of Bemisia tabaci in Europe in light of the climate change scenario. *Pest Management Science*.

Gnanasekaran, P., Ponnusamy, K. & Chakraborty, S., (2019). A geminivirus betasatellite encoded βC1 protein interacts with PsbP and subverts PsbP-mediated antiviral defence in plants. *Molecular Plant Pathology*, 20(7), pp.943–960.

Golemis, E.A. & Brent, R., (1992). Fused protein domains inhibit DNA binding by LexA. *Molecular and Cellular Biology*, 12(7), pp.3006–3014.

Gomez, M.A. et al., (2019). Simultaneous CRISPR/Cas9-mediated editing of cassava eIF4E isoforms nCBP-1 and nCBP-2 reduces cassava brown streak disease symptom severity and incidence. *Plant Biotechnology Journal*, 17(2), pp.421–434.

Góngora-Castillo, E. et al., (2012). Transcriptome analysis of symptomatic and recovered leaves of geminivirus-infected pepper (Capsicum annuum). *Virology Journal*, 9, pp.1–16.

Gorovits, R. et al., (2013). Recruitment of the Host Plant Heat Shock Protein 70 by

Tomato Yellow Leaf Curl Virus Coat Protein Is Required for Virus Infection. *PLoS ONE*, 8(7).

Gorovits, R. & Czosnek, H., (2017). The involvement of heat shock proteins in the establishment of Tomato Yellow Leaf Curl Virus infection. *Frontiers in Plant Science*, 8(March), pp.1–12.

Gough, J. et al., (2001). Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *Journal of Molecular Biology*, 313(4), pp.903–919.

Gouveia-mageste, B.C. et al., (2021). A plant-specific syntaxin-6 protein contributes to the intracytoplasmic route for begomoviruses. *Plant Physiol.*, 187(1), pp.158–173.

Grant, M. et al., (2006). Modifications to the Arabidopsis Defense Proteome Occur Prior to Significant Transcriptional Change in Response to Inoculation with Pseudomonas syringae. *Plant Physiology*, 142(4), pp.1603–1620.

Griffis, E.R. et al., (2002). Nup98 is a mobile nucleoporin with transcription-dependent dynamics. *Molecular Biology of the Cell*, 13(4), pp.1282–1297.

Guan, L. et al., (2019). An optimized method to obtain high-quality RNA from cassava storage root. *3 Biotech*, 9(4), pp.1–7. Available at: http://dx.doi.org/10.1007/s13205-019-1608-9.

Guerrero, J. et al., (2020). Manipulation of the Plant Host by the Geminivirus AC2/C2 Protein, a Central Player in the Infection Cycle. *Frontiers in Plant Science*, 11(May), pp.1–18.

Guo, D. et al., (2014). Cassava (Manihot esculenta Krantz) genome harbors KNOX genes differentially expressed during storage root development. *Genetics and Molecular Research*.

Guo, L. et al., (2019). Amino Acid Utilization May Explain Why Bemisia tabaci Q and B Differ in Their Performance on Plants Infected by the Tomato yellow leaf curl virus. *Frontiers in Physiology*, 10(MAY), pp.1–8.

Guo, W. et al., (2005). Protoplast transformation and regeneration of transgenic Valencia sweet orange plants containing a juice quality-related pectin methylesterase gene. *Plant Cell Reports*, 24(8), pp.482–486.

Gupta, N. et al., (2021). Plant responses to geminivirus infection: guardians of the plant immunity. *Virology Journal*, 18(1), pp.1–25. Available at: https://doi.org/10.1186/s12985-021-01612-1.

Gururani, M.A. et al., (2012). Plant disease resistance genes: Current status and future directions. *Physiological and Molecular Plant Pathology*.

Gutierrez, C. et al., (2004). Geminivirus DNA replication and cell cycle interactions. *Veterinary Microbiology*, 98(2), pp.111–119. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0378113503003262 [Accessed March 10, 2017].

Gutierrez, C., (2000). Geminiviruses and the plant cell cycle. *Plant Molecular Biology*, 43, pp.763–772.

Hahn, S.K., Terry, E.R. & Leuschner, K., (1980). Breeding cassava for resistance to cassava mosaic disease. *Euphytica*.

Haldrup, A., Lunde, C. & Scheller, H.V., (2003). Arabidopsis thaliana plants lacking the PSI-D subunit of photosystem I suffer severe photoinhibition, have unstable photosystem I complexes, and altered redox homeostasis in the chloroplast stroma. *Journal of Biological Chemistry*, 278(35), pp.33276–33283. Available at:

http://dx.doi.org/10.1074/jbc.M305106200.

Hamilton A, Voinnet O, Chappell L, B.D., (2002). Two classes of short interfering RNA in RNA silencing. *The EMBO Journal* , 21(17), pp.4671–4679.

Handford, M.G. & Carr, J.P., (2007). A defect in carbohydrate metabolism ameliorates symptom severity in virus-infected Arabidopsis thaliana. *Journal of General Virology*, 88(1), pp.337–341.

Hanke, G. & Mulo, P., (2013). Plant type ferredoxins and ferredoxin-dependent metabolism. *Plant, Cell and Environment*, 36(6), pp.1071–1084.

Hanley-Bowdoin, L. et al., (2013). Geminiviruses: masters at redirecting and reprogramming plant processes. *Nature Publishing Group*, 11.

Hao, L. et al., (2003). Geminivirus AL2 and L2 Proteins Interact with and Inactivate SNF1 Kinase. *The Plant Cell*, 15, pp.1034–1048.

Hao, T. et al., (2016). Reconstruction and Application of Protein–Protein Interaction Network. *International Journal of Molecular Sciences*.

Hara-Nishimura, I. & Hatsugai, N., (2011). The role of vacuole in plant cell death. *Cell Death and Differentiation*, 18(8), pp.1298–1304.

Havaux, M. et al., (2009). Vitamin B6 deficient plants display increased sensitivity to high light and photo-oxidative stress. *BMC Plant Biology*, 9, pp.1–22.

Haxim, Y. et al., (2017). Autophagy functions as an antiviral mechanism against geminiviruses in plants. *eLife*, 6, pp.1–17.

Haywood, V., Kragler, F. & Lucas, W.J., (2002). Plasmodesmata: Pathways for protein and ribonucleoprotein signaling. *Plant Cell*, 14(SUPPL.), pp.303–325.

He, Y. et al., (2019). Yeast two-hybrid screening for proteins that interact with PFT in wheat. *Scientific Reports*, 9(1), pp.1–10.

Heazlewood, J.L. et al., (2004). Experimental Analysis of the Arabidopsis Mitochondrial Proteome Highlights Signaling and Regulatory Components, Provides Assessment of Targeting Prediction Programs, and Indicates Plant-Specific Mitochondrial Proteins. *Plant Cell*, 16(1), pp.241–256.

Hehnle, S., Wege, C. & Jeske, H., (2004). Interaction of DNA with the movement proteins of geminiviruses revisited. *Journal of virology*, 78(14), pp.7698–7706.

Heinlein, M. et al., (1998). Changing patterns of localization of the tobacco mosaic virus movement protein and replicase to the endoplasmic reticulum and microtubules during infection. *Plant Cell*, 10(7), pp.1107–1120.

Henry, R. et al., (2019). The impact of cDNA normalization on long-read sequencing of a complex transcriptome. *Frontiers in Genetics*, 10(JUN), pp.1–17.

Hillocks, R.J.H. et al., (2001). 2001_E€ ects of Brown Streak Virus Disease on Yield and Quality of Cassava in Tanzania_Hillocks RJ.pdf. *Journal of Phytopathology*, 394(147), pp.389–394.

Hipp, K. et al., (2016). Properties of African cassava mosaic virus capsid protein expressed in fission yeast. *Viruses*, 8(7).

Hipp, K. et al., (2014). The RXL motif of the African cassava mosaic virus Rep protein is necessary for rereplication of yeast DNA and viral infection in plants. *Virology*, 462, pp.189–198.

Ho, E.S., Kuchie, J. & Duffy, S., (2014). Bioinformatic analysis reveals genome size reduction and the emergence of tyrosine phosphorylation site in the movement

protein of new world bipartite begomoviruses. *PLoS ONE*, 9(11).

Hong, Y., Saunders, K. & Stanley, J., (1997). Transactivation of Dianthin Transgene Expression by African Cassava Mosaic Virus AC2. *VIROLOGY*, 228, pp.383–387.

Horlbeck, M.A. et al., (2016). Nucleosomes impede cas9 access to DNA in vivo and in vitro. *eLife*, 5(MARCH2016), pp.1–21.

Hou, Y.M. et al., (2000). Transgenic plants expressing geminivirus movement proteins: Abnormal phenotypes and delayed infection by Tomato mottle virus in transgenic tomatoes expressing the Bean dwarf mosaic virus BV1 or BC1 proteins. *Molecular Plant-Microbe Interactions*, 13(3), pp.297–308.

Hoyle, A., (2018). *Identification of TNFAIP8L1 Binding Partners Through Co-Immunoprecipitation and Mass Spectrometry*. University of Maine. Available at: https://digitalcommons.library.umaine.edu/honors/335.

Hsu, C.-T. et al., (2021). Genome Editing and Protoplast Regeneration to Study Plant–Pathogen Interactions in the Model Plant Nicotiana benthamiana. *Frontiers in Genome Editing*, 2(January), pp.1–9.

Hu, J. et al., (2017). Nitric Oxide Regulates Protein Methylation during Stress Responses in Plants. *Molecular Cell*, 67(4), pp.702-710.e4. Available at: http://dx.doi.org/10.1016/j.molcel.2017.06.031.

Hu, T. et al., (2019). βc1 protein encoded in geminivirus satellite concertedly targets MKK2 and MPK4 to counter host defense. *PLoS Pathogens*, 15(4), pp.1–23.

Hu, Y.-X., Tao, Y.-B. & Xu, Z.-F., (2017). Overexpression of Jatropha Gibberellin 2-oxidase 6 (JcGA2ox6) Induces Dwarfism and Smaller Leaves, Flowers and Fruits in Arabidopsis and Jatropha. *Frontiers in Plant Science*, 8(December), pp.1–12. Available at: http://journal.frontiersin.org/article/10.3389/fpls.2017.02103/full.

Huang, B.X. & Kim, H.Y., (2013). Effective Identification of Akt Interacting Proteins by Two-Step Chemical Crosslinking, Co-Immunoprecipitation and Mass Spectrometry. *PLoS ONE*, 8(4).

Huang, H. & Bader, J.S., (2009). Precision and recall estimates for two-hybrid screens. *Bioinformatics*, 25(3), pp.372–378. Available at: http://bioinformatics.oxfordjournals.org/content/25/3/372.full#content-block.

Huang, S. et al., (2014). HSP90s are required for NLR immune receptor accumulation in Arabidopsis. *Plant Journal*, 79(3), pp.427–439.

Huang, X. et al., (2020). Friend or Enemy: A Dual Role of Autophagy in Plant Virus Infection. *Frontiers in Microbiology*, 11(April), pp.1–9.

Huang, Y.W. et al., (2012). Hsp90 interacts specifically with viral RNA and differentially regulates replication initiation of Bamboo mosaic virus and associated satellite RNA. *PLoS Pathogens*, 8(5).

Hubert, D.A. et al., (2009). Specific Arabidopsis HSP90.2 alleles recapitulate RAR1 cochaperone function in plant NB-LRR disease resistance protein regulation. *Proceedings of the National Academy of Sciences of the United States of America*, 106(24), pp.9556–9563.

Hüdig, M. et al., (2014). Plants Possess a Cyclic Mitochondrial Metabolic Pathway similar to the Mammalian Metabolic Repair Mechanism Involving Malate Dehydrogenase and l-2-Hydroxyglutarate Dehydrogenase. *Plant and Cell Physiology*, 56(9), pp.1820–1830.

Hull, R., (2009). Introduction to plant viruses. In *Comparative Plant Virology*. Academic Press, pp. 3–21. Available at:

https://booksite.elsevier.com/samplechapters/9780123741547/Sample Chapters/02~Chapter_1.pdf.

Hutvágner György, Z.P.D., (2002). Hutva´gner and Zamore, 2002. *Science*, 297(5589), pp.2056–2060. Available at: http://science.sciencemag.org/content/297/5589/2056.full.pdf+html.

Hwang, G.H. et al., (2020). CRISPR-sub: Analysis of DNA substitution mutations caused by CRISPR-Cas9 in human cells. *Computational and Structural Biotechnology Journal*, 18, pp.1686–1694. Available at: https://doi.org/10.1016/j.csbj.2020.06.026.

Hye, R.W. et al., (2007). VIM1, a methylcytosine-binding protein required for centromeric heterochromatinization. *Genes and Development*, 21(3), pp.267–277.

Hyodo, K. et al., (2015). Phosphatidic Acid Produced by Phospholipase D Promotes RNA Replication of a Plant RNA Virus. *PLoS Pathogens*, 11(5), pp.1–28.

Ihnatowicz, A. et al., (2004). Mutants for photosystem I subunit D of Arabidopsis thaliana: Effects on photosynthesis, photosystem I stability and expression of nuclear genes for chloroplast functions. *Plant Journal*, 37(6), pp.839–852.

Imai, J. et al., (2003). The molecular chaperone Hsp90 plays a role in the assembly and maintenance of the 26S proteasome. *EMBO Journal*, 22(14), pp.3557–3567.

Ingham, D.J., Pascal, E. & Lazarowitz, S.G., (1995). Both bipartite geminivirus movement proteins define viral host range, but only BL1 determines viral pathogenicity. *Virology*, 207(1), pp.191–204.

Iqbal, H., Akins, D.R. & Kenedy, M.R., (1960). Co-immunoprecipitation for Identifying Protein-Protein Interactions in Borrelia burgdorferi. *Methods Mol Biol*, 1690, pp.47–55.

Irigoyen, M.L. et al., (2020). Genome-wide analyses of cassava pathogenesis-related (pr) gene families reveal core transcriptome responses to whitefly infestation, salicylic acid and jasmonic acid. *BMC Genomics*, 21(1), pp.1–18.

Ismayil, A. et al., (2020). Cotton leaf curl Multan virus βC1 protein induces autophagy by disrupting the interaction of autophagy-related protein 3 with glyceraldehyde-3-phosphate dehydrogenases. *Plant Cell*, 32(4), pp.1124–1135.

Ivanov, K.I. et al., (2003). Phosphorylation of the potyvirus capsid protein by protein kinase CK2 and its relevance for virus infection. *Plant Cell*, 15(9), pp.2124–2139.

Jacobs, C. & Segrè, D., (2012). Organization Principles in Genetic Interaction Networks. In S. O. Soyer, ed. *Evolutionary Systems Biology*. New York, NY: Springer New York, pp. 53–78. Available at: http://dx.doi.org/10.1007/978-1-4614-3567-9_3.

Jarosch, E. et al., (2002). Protein dislocation from the ER requires polyubiquitination and the AAA-ATPase Cdc48. *Nature Cell Biology*, 4(2), pp.134–139.

Javanbakht, H. et al., (2003). The interaction between HIV-1 Gag and human lysyl-tRNA synthetase during viral assembly. *Journal of Biological Chemistry*, 278(30), pp.27644–27651. Available at: http://dx.doi.org/10.1074/jbc.M301840200.

Jeevalatha, A. et al., (2017). An insight into differentially regulated genes in resistant and susceptible genotypes of potato in response to tomato leaf curl New Delhi virus-[potato] infection. *Virus Research*, 232, pp.22–33.

Jennifer L. M. Soosaar, T.M.B.-S.& S.P.D.-K., (2005). MECHANISMS OF PLANT RESISTANCE TO VIRUSES. *NATURE REVIEWS | MICROBIOLOGY*, 3, pp.789–798. Available at: http://www.nature.com/nrmicro/journal/v3/n10/full/nrmicro1239.html.

Jennings, D.L. & Iglesias, C., (2001). Breeding for Crop Improvement. In Breeding for Crop Improvement & Jennings, eds. *Breeding for Crop Improvement Jennings*. New York: CAB eBooks, pp. 149–166. Available at: http://www.cabi.org/cabebooks/ebook/20073012661.

Jeong, Y.Y. et al., (2021). Optimization of protoplast regeneration in the model plant Arabidopsis thaliana. *Plant Methods*, 17(1), pp.1–16. Available at: https://doi.org/10.1186/s13007-021-00720-x.

Jeske, H., (2009). Geminiviruses. *Current Topics in Microbiology and Immunology*, 331, pp.185–226.

De Jesús-González, L.A. et al., (2021). The nuclear pore complex is a key target of viral proteases to promote viral replication. *Viruses*, 13(4).

Jia, Q. et al., (2016). CLCuMuB βC1 Subverts Ubiquitination by Interacting with NbSKP1s to Enhance Geminivirus Infection in Nicotiana benthamiana. *PLoS Pathogens*, 12(6), pp.1–30.

Jiang, F. & Doudna, J.A., (2017). CRISPR – Cas9 Structures and Mechanisms. , pp.505–531.

Jinek, M. et al., (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, 337(6096), pp.816–821.

John Odipio, Emmanuel Ogwok, Nigel J Taylor, Mark Halsey, Anton Bua, C.M.F.& T.A., (2014). RNAi-derived field resistance to Cassava brown streak disease persists across the vegetative cropping cycle. *GM Crops & Food*, 5(1), pp.16–19. Available at: http://www.tandfonline.com/action/journalInformation?journalCode=kgmc20.

Jones, C.P. et al., (2013). Molecular mimicry of human tRNALys anti-codon domain by HIV-1 RNA genome facilitates tRNA primer annealing. *Rna*, 19(2), pp.219–229.

Jones, H., Ooms, G. & Jones, M.G.K., (1989). Transient gene expression in electroporated Solanum protoplasts. *Plant Molecular Biology*, 13(5), pp.503–511.

Jones, J.D.G. & Dangl, J.L., (2006). The plant immune system. *Nature*, 444, pp.323–329.

Jones, S. & Thornton, J.M., (1996). Principles of protein-protein interactions. *Proceedings of the National Academy of Sciences of the United States of America*, 93(1), pp.13–20.

Joshi, P., Maidji, E. & Stoddart, C.A., (2016). Inhibition of heat shock protein 90 prevents HIV rebound. *Journal of Biological Chemistry*, 291(19), pp.10332–10346. Available at: http://dx.doi.org/10.1074/jbc.M116.717538.

Just, K. et al., (2017). Monitoring infection of tomato fruit by Tomato yellow leaf curl virus. *Plant Pathology*, 66(4), pp.522–528.

Kang, H.H., Naing, A.H. & Kim, C.K., (2020). Protoplast isolation and shoot regeneration from protoplast-derived callus of petunia hybrida cv. Mirage rose. *Biology*, 9(8), pp.1–13.

Karner, U. et al., (2004). myo-Inositol and sucrose concentrations affect the accumulation of raffinose family oligosaccharides in seeds. *Journal of Experimental Botany*, 55(405), pp.1981–1987.

Kashina, B.D. et al., (2013). Molecular identification of a new begomovirus associated with mosaic disease of Jatropha curcas L. in Nigeria. *Archives of Virology*, 158(2), pp.511–514.

Katagiri, T. et al., (2005). An important role of phosphatidic acid in ABA signaling during germination in Arabidopsis thaliana. *Plant Journal*, 43(1), pp.107–117.

Kawakami, S. et al., (2003). Defective Tobamovirus Movement Protein Lacking Wild-Type Phosphorylation Sites Can Be Complemented by Substitutions Found in Revertants. *Journal of Virology*, 77(2), pp.1452–1461.

Kawakami, S., Watanabe, Y. & Beachy, R.N., (2004). Tobacco mosaic virus infection spreads cell to cell as intact replication complexes. *Proceedings of the National Academy of Sciences of the United States of America*, 101(16), pp.6291–6296.

Kawamura, Y. & Uemura, M., (2003). Mass spectrometric approach for identifying putative plasma membrane proteins of Arabidopsis leaves associated with cold acclimation. *Plant Journal*, 36(2), pp.141–154.

Khan, A.A. et al., (2017). A highly efficient ligation-independent cloning system for CRISPR/Cas9 based genome editing in plants. *Plant Methods*, 13(1), pp.1–13.

Kikkawa, H. et al., (1982). Infection of Protoplasts from Tobacco Suspension Cultures by Tobacco Mosaic Virus. *Journal of General Virology*, 63(2), pp.451–456.

Kim, T.H. et al., (2010). Guard cell signal transduction network: Advances in understanding abscisic acid, CO2, and Ca2+ signaling. *Annual Review of Plant Biology*, 61, pp.561–591.

Kleinow, T., Tanwir, F., et al., (2009). Expression dynamics and ultrastructural localization of epitope-tagged Abutilon mosaic virus nuclear shuttle and movement proteins in Nicotiana benthamiana cells. *Virology*, 391(2), pp.212–220. Available at: http://dx.doi.org/10.1016/j.virol.2009.06.042.

Kleinow, T. et al., (2020). Phosphorylations of the Abutilon Mosaic Virus Movement Protein Affect Its Self-Interaction, Symptom Development, Viral DNA Accumulation, and Host Range. *Frontiers in Plant Science*, 11(July), pp.1–20.

Kleinow, T. et al., (2008). Post-translational modifications of Abutilon mosaic virus movement protein (BC1) in fission yeast. *Virus Research*, 131(1), pp.86–94.

Kleinow, T., Nischang, M., et al., (2009). Three C-terminal phosphorylation sites in the Abutilon mosaic virus movement protein affect symptom development and viral DNA accumulation. *Virology*, 390(1), pp.89–101. Available at: http://dx.doi.org/10.1016/j.virol.2009.04.018.

Klie, S. & Nikoloski, Z., (2012). The choice between MapMan and Gene ontology for automated gene function prediction in plant science. *Frontiers in Genetics*.

Von Kobbe, C. et al., (2000). Vesicular stomatitis virus matrix protein inhibits host cell gene expression by targeting the nucleoporin Nup98. *Molecular Cell*, 6(5), pp.1243–1252.

Kogovšek, P. et al., (2016). Primary metabolism, phenylpropanoids and antioxidant pathways are regulated in potato as a response to potato virus y infection. *PLoS ONE*, 11(1), pp.1–20.

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. *The Plant cell*, 14(8), pp.1817–32. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=151467&tool=pmcentrez&rendertype=abstract.

Koonin EV, G.M., (2003). Genome Annotation and Analysis. In *Sequence - Evolution - Function: Computational Approaches in Comparative Genomics.* Boston: Kluwer Academic. Available at: https://www.ncbi.nlm.nih.gov/books/NBK20253/#!po=1.25000.

Kourmpetis, Y.A.I. et al., (2010). Bayesian markov random field analysis for protein function prediction based on network data. *PLoS ONE*.

Kranthi, O., Mandadi, K. & Scholthof, K.-B.G., (2013). Plant Immune Responses Against Viruses: How Does a Virus Cause Disease? *The Plant cell*, 25(5), pp.1489–1505.

Krapp, S. et al., (2017). Begomoviral movement protein effects in human and plant cells: Towards new potential interaction partners. *Viruses*, 9(11).

Krens, F.A. et al., (1982). In vitro transformation of plant protoplasts with Ti-plasmid DNA. *Nature*, 296(5852), pp.72–74.

Krenz, B. et al., (2010). A plastid-targeted heat shock cognate 70kDa protein interacts with the Abutilon mosaic virus movement protein. *Virology*, 401(1), pp.6–17. Available at: http://dx.doi.org/10.1016/j.virol.2010.02.011.

Kruijt M, DE Kock MJ, de W.P., (2005). Receptor-like proteins involved in plant disease resistance. *MOLECULAR PLANT PATHOLOGY*, 6(1), pp.85–97.

Kumakura, N. et al., (2009). SGS3 and RDR6 interact and colocalize in cytoplasmic SGS3/RDR6-bodies. *FEBS Letters*.

Kumar, R.V., (2019). Plant antiviral immunity against geminiviruses and viral counter-defense for survival. *Frontiers in Microbiology*, 10(JUN), pp.1–17.

Kumar, S. et al., (2020). Interaction of potyvirus helper component-proteinase (HcPro) with RuBisCO and nucleosome in viral infections of plants. *Plant Physiology and Biochemistry*, 151(March), pp.313–322. Available at: https://doi.org/10.1016/j.plaphy.2020.03.036.

Kumar, V. et al., (2015). Mungbean yellow mosaic Indian virus encoded AC2 protein suppresses RNA silencing by inhibiting Arabidopsis RDR6 and AGO1 activities. *Virology*, 486, pp.158–172. Available at: http://dx.doi.org/10.1016/j.virol.2015.08.015.

Kuon, J.E. et al., (2019). Haplotype-resolved genomes of geminivirus-resistant and geminivirus-susceptible African cassava cultivars. *BMC Biology*, 17(1), pp.1–15.

Kushwaha, N. et al., (2015). Chilli leaf curl virus infection highlights the differential expression of genes involved in protein homeostasis and defense in resistant chilli plants. *Applied Microbiology and Biotechnology*, 99(11), pp.4757–4770.

Kushwaha, N.K., Bhardwaj, M. & Chakraborty, S., (2017). *The replication initiator protein of a geminivirus interacts with host monoubiquitination machinery and stimulates transcription of the viral genome*,

Lacatus, G. & Sunter, G., (2009). The Arabidopsis PEAPOD2 transcription factor interacts with geminivirus AL2 protein and the coat protein promoter. *Virology*, 392(2), pp.196–202. Available at: http://dx.doi.org/10.1016/j.virol.2009.07.004.

Lacks, S.A. et al., (1986). Genetic basis of the complementary Dpnl and Dpnll restriction systems of S. pneumoniae: An intercellular cassette mechanism. *Cell*, 46(7), pp.993–1000.

Lamesch, P. et al., (2012). The Arabidopsis Information Resource (TAIR): Improved gene annotation and new tools. *Nucleic Acids Research*.

Lario, L.D. et al., (2013). ANTI-SILENCING FUNCTION1 proteins are involved in ultraviolet-induced DNA damage repair and are cell cycle regulated by E2F transcription factors in arabidopsis. *Plant Physiology*, 162(2), pp.1164–1177.

Lau, S. & Weber, F., (2020). Nuclear pore protein Nup98 is involved in replication of Rift Valley fever virus and nuclear import of virulence factor NSs. *Journal of General*

Virology, 101(7), pp.712–716.

Laxalt, A.M. & Munnik, T., (2002). Phospholipid signalling in plant defence. *Current Opinion in Plant Biology*, 5(4), pp.332–338.

Laya, A. & Koubala, B.B., (2020). Polyphenols in cassava leaves (Manihot esculenta Crantz) and their stability in antioxidant potential after in vitro gastrointestinal digestion. *Heliyon*, 6(3), p.e03567. Available at: https://doi.org/10.1016/j.heliyon.2020.e03567.

Lee, D.Y. et al., (2005). Role of protein methylation in regulation of transcription. *Endocrine Reviews*, 26(2), pp.147–170.

Lee, H.T. et al., (2020). The key role of dna methylation and histone acetylation in epigenetics of atherosclerosis. *Journal of Lipid and Atherosclerosis*, 9(3), pp.419–434.

Lee, K.C. et al., (2001). Cucurbit protoplast isolation for the study of plant virus replication. *Journal of Virological Methods*, 91(1), pp.21–27.

Lee, M. et al., (2005). Methionine and threonine synthesis are limited by homoserine availability and not the activity of homoserine kinase in Arabidopsis thaliana. *Plant Journal*, 41(5), pp.685–696.

Lee, M. & Leustek, T., (1999). Identification of the gene encoding homoserine kinase from Arabidopsis thaliana and characterization of the recombinant enzyme derived from the gene. *Archives of Biochemistry and Biophysics*, 372(1), pp.135–142.

Lee, S. et al., (2016). Global regulation of plant immunity by histone lysine methyl transferases. *Plant Cell*, 28(7), pp.1640–1661.

Lefeuvre, P. et al., (2007). Avoidance of protein fold disruption in natural virus recombinants. *PLoS Pathogens*.

Legg, J. et al., (2014). A global alliance declaring war on cassava viruses in Africa. *Food Security*.

Legg, J.P. et al., (2015). Cassava virus diseases: Biology, epidemiology, and management. *Advances in Virus Research*.

Lellis, A.D. et al., (2002). Loss-of-susceptibility mutants of Arabidopsis thaliana reveal an essential role for eIF(iso)4E during potyvirus infection. *Current Biology*, 12(12), pp.1046–1051.

Lentz, E.M. et al., (2018). Cassava geminivirus agroclones for virus-induced gene silencing in cassava leaves and roots. *Plant Methods*, 14(1), pp.1–9. Available at: https://doi.org/10.1186/s13007-018-0340-5.

Lewis, J.D. & Lazarowitz, S.G., (2010). Arabidopsis synaptotagmin SYTA regulates endocytosis and virus movement protein cell-to-cell transport. *Proceedings of the National Academy of Sciences of the United States of America*, 107(6), pp.2491–2496.

Li, C. et al., (2020). Human coronavirus dependency on host heat shock protein 90 reveals an antiviral target. *Emerging Microbes and Infections*, 9(1), pp.2663–2672.

Li, F. et al., (2017). A calmodulin-like protein suppresses RNA silencing and promotes geminivirus infection by degrading SGS3 via the autophagy pathway in Nicotiana benthamiana. *PLoS Pathogens*, 13(2), pp.1–25.

Li, F. et al., (2020). Nuclear autophagy degrades a geminivirus nuclear protein to restrict viral infection in solanaceous plants. *New Phytologist*, 225(4), pp.1746–1761.

Li, H. et al., (2018). S-acylation of a geminivirus C4 protein is essential for regulating the CLAVATA pathway in symptom determination. *Journal of Experimental Botany*, 69(18), pp.4459–4468.

Li, J.F. et al., (2013). Multiplex and homologous recombination-mediated plant genome editing via guide RNA/Cas9. *Nat Biotechnol*, 31(8), pp.688–691.

Li, P. et al., (2011). AtPID: The overall hierarchical functional protein interaction network interface and analytic platform for arabidopsis. *Nucleic Acids Research*.

Li, R. et al., (2014). Virulence Factors of Geminivirus Interact with MYC2 to Subvert Plant Resistance and Promote Vector Performance. *The Plant Cell Online*, 26(12), pp.4991–5008. Available at: http://www.plantcell.org/lookup/doi/10.1105/tpc.114.133181.

Li, Y., Pearl, S.A. & Jackson, S.A., (2015). Gene Networks in Plant Biology: Approaches in Reconstruction and Analysis. *Trends in Plant Science*.

Li, Y.Z. et al., (2016). Characters related to higher starch accumulation in cassava storage roots. *Scientific Reports*, 6(February 2015), pp.1–17.

Liang, P. et al., (2015). Identification and characterization of a novel geminivirus with a monopartite genome infecting apple trees. *Journal of General Virology*, 96(8), pp.2411–2420.

Lim, G.-H. et al., (2017). Fatty Acid– and Lipid-Mediated Signaling in Plant Defense. *Annual Review of Phytopathology*, 55(1), pp.505–536.

Lim, G.H. et al., (2017). Fatty Acid- and Lipid-Mediated Signaling in Plant Defense. *Annual Review of Phytopathology*, 55, pp.505–536.

Lin, C.S. et al., (2018). Application of protoplast technology to CRISPR/Cas9 mutagenesis: from single-cell mutation detection to mutant plant regeneration. *Plant Biotechnology Journal*, 16(7), pp.1295–1310.

Lindén, P. et al., (2016). Reduced mitochondrial malate dehydrogenase activity has a strong effect on photorespiratory metabolism as revealed by 13C labelling. *Journal of Experimental Botany*, 67(10), pp.3123–3135.

Ling, M.H. et al., (2013). Extending MapMan Ontology to Tobacco for Visualization of Gene Expression. *Dataset Papers in Biology*, 2013(706465.).

Liu, J. et al., (2014). Why mosaic? Gene expression profiling of African cassava mosaic virus-infected cassava reveals the effect of chlorophyll degradation on symptom development. *Journal of Integrative Plant Biology*, 56(2).

Liu, L. et al., (2014). Altered expression of Arabidopsis genes in response to a multifunctional geminivirus pathogenicity protein.

Liu, S. et al., (1997). Efficient whitefly transmission of African cassava mosaic geminivirus requires sequences from both genomic components. *Journal of General Virology*.

Liu, S., Gao, Y. & Vakser, I.A., (2008). Dockground protein-protein docking decoy set. *Bioinformatics*.

Liu, Y. et al., (2004). Molecular Chaperone Hsp90 Associates with Resistance Protein N and Its Signaling Proteins SGT1 and Rar1 to Modulate an Innate Immune Response in Plants. *Journal of Biological Chemistry*, 279(3), pp.2101–2108. Available at: http://dx.doi.org/10.1074/jbc.M310029200.

Liu, Y. et al., (2002). Role of SCF Ubiquitin-Ligase and the COP9 Signalosome in the N Gene–Mediated Resistance Response to Tobacco mosaic virus. *The Plant Cell*, 14, pp.1483–1496.

Liu, Z. et al., (2020). Application of different types of CRISPR/Cas-based systems in bacteria. *Microbial Cell Factories*, 19(1), pp.1–14. Available at: https://doi.org/10.1186/s12934-020-01431-z.

Loconsole, G. et al., (2012). Identification of a single-stranded DNA virus associated with citrus chlorotic dwarf disease, a new member in the family Geminiviridae. *Virology*, 432(1), pp.162–172. Available at: http://dx.doi.org/10.1016/j.virol.2012.06.005.

Loebenstein, G., (1959). Effect of a virus infection in sweet potatoes upon the host plant respiration. *Virology*, 9(1), pp.62–71.

Lorenzo, O. et al., (2004). JASMONATE-INSENSITIVE1 encodes a MYC transcription factor essential to discriminate between different jasmonate-regulated defense responses in arabidopsis. *Plant Cell*, 16(7), pp.1938–1950.

Lorković, Z.J., (2009). Role of plant RNA-binding proteins in development, stress response and genome organization. *Trends in Plant Science*, 14(4), pp.229–236.

Louis, B. & Rey, C., (2015). Resistance gene analogs involved in tolerant cassava–geminivirus interaction that shows a recovery phenotype. *Virus Genes*.

Lozano-Durán, R. et al., (2012). Geminivirus C2 protein represses genes involved in sulphur assimilation and this effect can be counteracted by jasmonate treatment. *European Journal of Plant Pathology*, 134(1), pp.49–59.

Lozano-Durán, R. et al., (2011). Geminiviruses subvert ubiquitination by altering CSN-mediated derubylation of SCF E3 ligase complexes and inhibit jasmonate signaling in Arabidopsis thaliana. *Plant Cell*, 23(3), pp.1014–1032.

Lozano, R. et al., (2011). Identification and distribution of the NBS-LRR gene family in the Cassava genome.

Lu, H. et al., (2013). Overexpression of a citrus NDR1 ortholog increases disease resistance in Arabidopsis. , 31(11), pp.35–1.

Lu, H. et al., (2015). Signal regulators of systemic acquired resistance.

Lu, R. et al., (2003). High throughput virus-induced gene silencing implicates heat shock protein 90 in plant disease resistance. *The EMBO journal*, 22(21), pp.5690–9. Available at: http://emboj.embopress.org/content/22/21/5690.abstract.

Lü, S. et al., (2009). Arabidopsis CER8 encodes LONG-CHAIN ACYL-COA SYNTHETASE 1 (LACS1) that has overlapping functions with LACS2 in plant wax and cutin synthesis. *Plant Journal*, 59(4), pp.553–564.

Ludvigsen, M. & Honoré, B., (2018). Transcriptomics and Proteomics: Integration? *eLS*, pp.1–7.

Lukhovitskaya, N.I. et al., (2013). An RNA virus-encoded zinc-finger protein acts as a plant transcription factor and induces a regulator of cell size and proliferation in two tobacco Species. *Plant Cell*, 25(3), pp.960–973.

Luque, A. et al., (2002). Interaction of Geminivirus Rep Protein with Replication Factor C and Its Potential Role during Geminivirus DNA Replication. *Virology*, 302(1), pp.83–94.

Ma, J. & Ptashne, M., (1988). Converting a eukaryotic transcriptional inhibitor into an activator. *Cell*, 55(3), pp.443–446.

Maas, C. et al., (1995). Preparation and Transformation of Monocot Protoplasts. *Methods in Cell Biology*, 50(C), pp.383–399.

Maghuly, F., Ramkat, R.C. & Laimer, M., (2014). Virus versus host plant microRNAs:

Who determines the outcome of the interaction? *PLoS ONE*.

Maier, F. et al., (2011). Nonexpressor Of Pathogenesis-Related Proteins1 (NPR1) and some NPR1-related proteins are sensitive to salicylic acid. *Molecular Plant Pathology*.

Maio, F. et al., (2020). Identification of Tomato Proteins That Interact With Replication Initiator Protein (Rep) of the Geminivirus TYLCV. *Frontiers in Plant Science*, 11(July), pp.1–15.

Mäkinen, K. & Swarnalok, D., (2019). The significance of methionine cycle enzymes in plant virus infections. *Current Opinion in Plant Biology*, 50, pp.67–75.

Mandal, B. et al., (2001). First Report of Cabbage leaf curl virus (Family Geminiviridae) in Georgia. *Plant Disease*, 85(5), p.561. Available at: https://doi.org/10.1094/PDIS.2001.85.5.561A.

Maredza, A.T. et al., (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, pp.1467–1485.

Maruthi, M.N. et al., (2014). Transcriptional response of virus-infected cassava and identification of putative sources of resistance for cassava brown streak disease. *PLoS ONE*, 9(5).

Marwal, A., Sahu, A.K. & Gaur, R.K., (2014). Transmission and host interaction of Geminivirus in weeds. In *Plant Virus–Host Interaction*. Elsevier, pp. 143–161. Available at: http://linkinghub.elsevier.com/retrieve/pii/B978012411584200007X [Accessed May 14, 2018].

Más, P. & Beachy, R.N., (2000). Role of microtubules in the intracellular distribution of tobacco mosaic virus movement protein. *Proceedings of the National Academy of Sciences of the United States of America*, 97(22), pp.12345–12349.

Mascarin, G.M. et al., (2013). The virulence of entomopathogenic fungi against Bemisia tabaci biotype B (Hemiptera: Aleyrodidae) and their conidial production using solid substrate fermentation. *Biological Control*.

Mathur, J. & Hülskamp, M., (2002). Microtubules and microfilaments in cell morphogenesis in higher plants. *Current Biology*, 12(19), pp.669–676.

Matiolli, C.C. & Melotto, M., (2018). A comprehensive arabidopsis yeast two-hybrid library for protein-protein interaction studies: A resource to the plant research community. *Molecular Plant-Microbe Interactions*, 31(9), pp.899–902.

Matsushita, Y. et al., (2000). In vitro phosphorylation of the movement protein of tomato mosaic tobamovirus by a cellular kinase. *Journal of General Virology*, 81(8), pp.2095–2102.

Mayer, K.A. et al., (2019). Hijacking the supplies: Metabolism as a novel facet of virus-host interaction. *Frontiers in Immunology*, 10(JULY), pp.1–12.

McGarry, R.C. et al., (2003). A novel Arabidopsis acetyltransferase interacts with the geminivirus movement protein NSP. *The Plant cell*, 15(7), pp.1605–1618. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=165404&tool=pmcentrez&rendertype=abstract.

McGivern, D.R. et al., (2005). An intact RBR-binding motif is not required for infectivity of Maize streak virus in cereals, but is required for invasion of mesophyll cells. *Journal of General Virology*.

McLoughlin, F. & Testerink, C., (2013). Phosphatidic acid, a versatile water-stress signal

in roots. *Frontiers in Plant Science*, 4(DEC), pp.1–8.

Mcwhite, C.D. et al., (2021). novel assemblies. , 181(2), pp.460–474.

Mehla, J., Caufield, J.H. & Uetz, P., (2015). Mapping protein-protein interactions using yeast two-hybrid assays. *Cold Spring Harbor Protocols*, 2015(5), pp.442–452.

Mehta, D. et al., (2018). CRISPR-Cas9 interference in cassava linked to the evolution of editing-resistant geminiviruses. *bioRxiv*, p.314542.

Mei, Y. et al., (2020). Geminivirus C4 antagonizes the HIR1-mediated hypersensitive response by inhibiting the HIR1 self-interaction and promoting degradation of the protein. *New Phytologist*, 225(3), pp.1311–1326.

Mei, Y. et al., (2018). Nucleocytoplasmic Shuttling of Geminivirus C4 Protein Mediated by Phosphorylation and Myristoylation Is Critical for Viral Pathogenicity. *Molecular Plant*, 11(12), pp.1466–1481. Available at: https://doi.org/10.1016/j.molp.2018.10.004.

Méndez-Lozano, J. et al., (2003). Interactions between geminiviruses in a naturally occurring mixture: Pepper huasteco virus and Pepper golden mosaic virus. *Phytopathology*, 93(3), pp.270–277.

Mérai, Z. et al., (2014). The AAA-ATPase molecular chaperone Cdc48/p97 disassembles sumoylated centromeres, decondenses heterochromatin, and activates ribosomal RNA genes. *Proceedings of the National Academy of Sciences of the United States of America*, 111(45), pp.16166–16171.

Meyer, M. & Schomburg, D., (2008). Protein Interactions. *Biotechnology*, 5, pp.87–108.

Mil', E.M. & Biniukov, V.. I., (1988). Local homology of amino acid sequences of histones H3, H4 and the protein repressor lambda Cro. Possible conformation of homologous histone sites in DNP. *Biokhimiia*, 53(12), pp.1980–1986.

Miles, L.A. et al., (2016). Design, execution, and analysis of pooled in vitro CRISPR/Cas9 screens. *The FEBS Journal*, 283(17), pp.3170–3180.

Mills-Lujan, K. et al., (2015). The roles of phosphorylation and SHAGGY-like protein kinases in geminivirus C4 protein induced hyperplasia. *PLoS ONE*, 10(3), pp.1–26.

Miozzi, L. et al., (2014). Transcriptomics of the interaction between the monopartite phloem-limited geminivirus tomato yellow leaf curl sardinia virus and solanum lycopersicum highlights a role for plant hormones, autophagy and plant immune system fine tuning during infection. *PLoS ONE*, 9(2).

Missich, R., Ramirez-Parra, E. & Gutierrez, C., (2000). Relationship of oligomerization to DNA binding of Wheat dwarf virus RepA and Rep proteins. *Virology*, 273(1), pp.178–188.

Mistry, J. et al., (2021). Pfam: The protein families database in 2021. *Nucleic Acids Research*, 49(D1), pp.D412–D419.

Möckli, N. & Auerbach, D., (2004). Quantitative β-galactosidase assay suitable for high-throughput applications in the yeast two-hybrid system. *BioTechniques*, 36(5), pp.872–876.

Mohr, P.G. & Cahill, D.M., (2007). Suppression by ABA of salicylic acid and lignin accumulation and the expression of multiple genes, in Arabidopsis infected with Pseudomonas syringae pv. tomato. *Functional & Integrative Genomics*, 7(3), pp.181–191. Available at: https://doi.org/10.1007/s10142-006-0041-4.

Monaghan, J. & Zipfel, C., (2012). Plant pattern recognition receptor complexes at the plasma membrane. *Current Opinion in Plant Biology*.

Montero-Rojas, M., Correa, A.M. & Siritunga, D., (2011). Molecular differentiation and diversity of cassava (manihot esculenta) taken from 162 locations across Puerto rico and assessed with microsatellite markers. *AoB PLANTS*.

Mooney, S. et al., (2009). Vitamin B6: A long known compound of surprising complexity. *Molecules*, 14(1), pp.329–351.

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. , p.241.

More, P., Agarwal, Parinita & Agarwal, Pradeep, (2021). The Jatropha leaf curl Gujarat virus on infection regulates the sugar and tricarboxylic acid cycle metabolic pathways in Jatropha. *Research Square*.

Moreno, D.F. et al., (2019). Competition in the chaperone-client network subordinates cell-cycle entry to growth and stress. *Life Science Alliance*, 2(2), pp.1–16.

Moshe, A. et al., (2012). Stress Responses to Tomato Yellow Leaf Curl Virus (TYLCV) Infection of Resistant and Susceptible Tomato Plants are Different. *Journal of Postgenomics Drug & Biomarker Development*, s1(01), pp.1–13. Available at: https://www.omicsonline.org/stress-responses-to-tomato-yellow-leaf-curl-virus-(tylcv)-infection-of resistant-and-susceptible-tomato-plants-are-different-2153-0769.S1-006.php?aid=6028.

Moshe, A. et al., (2016). Tomato plant cell death induced by inhibition of HSP90 is alleviated by Tomato yellow leaf curl virus infection. *Molecular Plant Pathology*, 17(2), pp.247–260.

Mourrain, P. et al., (2000). Arabidopsis SGS2 and SGS3 Genes Are Required for Posttranscriptional Gene Silencing and Natural Virus Resistance gene required for PTGS, a possible mechanistic link. *Cell*, 101, pp.533–542.

Mubarik, M.S. et al., (2021). Engineering broad-spectrum resistance to cotton leaf curl disease by CRISPR-Cas9 based multiplex editing in plants. *GM Crops and Food*, 00(00), pp.1–12. Available at: https://doi.org/10.1080/21645698.2021.1938488.

Mukhtar, M.S. et al., (2011). Plant Immune System Network. *Science*, 333(July), pp.596–601.

Mulimbi, W. et al., (2012). First report of Ugandan cassava brown streak virus on cassava in Democratic Republic of Congo. *New Disease Reports*, 26, p.11. Available at: http://www.ndrs.org.uk/article.php?id=026011.

Munganyinka, E. et al., (2018). Cassava brown streak disease in Rwanda, the associated viruses and disease phenotypes. *Plant Pathology*, 67(2), pp.377–387.

Murashige, T. & Skoog, F., (1962). A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum*, 15(3), pp.473–497.

Murovec, J. et al., (2018). DNA-free genome editing of brassica oleracea and B. Rapa protoplasts using CRISPR-cas9 ribonucleoprotein complexes. *Frontiers in Plant Science*, 871(November), pp.1–9.

Mwaba, I. & Rey, M.E.C., (2017). Nitric oxide associated protein 1 is associated with chloroplast perturbation and disease symptoms in Nicotiana benthamiana infected with South African cassava mosaic virus. *Virus Research*, 238(March), pp.75–83. Available at: http://dx.doi.org/10.1016/j.virusres.2017.05.022.

Mwaba, I.I., (2011). *Development of a VIGS vector based on SACMV for genomic studies in tobacco and cassava*. University of the Witwatersrand. Available at: http://wiredspace.wits.ac.za/handle/10539/9265?show=full.

Nadakuduti, S.S. et al., (2019). Genome Editing in Potato with CRISPR/Cas9. Plant Genome Editing with CRISPR Systems. In Q. Yiping, ed. *Plant Genome Editing with CRISPR Systems: Methods and Protocols, Methods in Molecular Biology*. © Springer Science+Business Media, LLC, part of Springer Nature, pp. 183–201.

Nagar, S., Hanley-Bowdoin, L. & Robertson, D., (2002). Host DNA replication is induced by geminivirus infection of differentiated plant cells. *The Plant cell*, 14(12), pp.2995–3007.

Naguib, A. et al., (2015). PTEN functions by recruitment to cytoplasmic vesicles. *Molecular Cell*, 58(2), pp.255–268. Available at: http://dx.doi.org/10.1016/j.molcel.2015.03.011.

Nakajima, I. et al., (2017). CRISPR/Cas9-mediated targeted mutagenesis in grape. *PLoS ONE*, 12(5), pp.1–16.

Nakano, K. et al., (2013). Viral interference with host mRNA surveillance, the nonsense-mediated mRNA decay (NMD) pathway, through a new function of HTLV-1 Rex: Implications for retroviral replication. *Microbes and Infection*, 15(6–7), pp.491–505. Available at: http://dx.doi.org/10.1016/j.micinf.2013.03.006.

Nankoo, N., (2018). *The recombinant expression and structural characterization of movement protein BC1 from South African cassava mosaic virus Nikita Nankoo Supervised by : Prof M . E . C .* The University of the Witwatersrand. Available at: https://hdl.handle.net/10539/27066.

Naseem, S. & Winter, S., (2016). Quantification of African cassava mosaic virus (ACMV) and East African cassava mosaic virus (EACMV-UG) in single and mixed infected Cassava (Manihot esculenta Crantz) using quantitative PCR. *Journal of Virological Methods*, 227, pp.23–32.

Nawaz-ul-Rehman, M.S. & Fauquet, C.M., (2009). Evolution of geminiviruses and their satellites. *FEBS Letters*.

Ndunguru, J. et al., (2016). Two Novel DNAs that Enhance Symptoms and Overcome CMD2 Resistance to Cassava Mosaic Disease. *Journal of Virology*, 90(8), pp.4160–4173.

Nelson, D.C. et al., (2010). Karrikins enhance light responses during germination and seedling development in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America*, 107(15), pp.7095–7100.

Nemes, K. et al., (2017). Phosphorylation regulates the subcellular localization of Cucumber Mosaic Virus 2b protein. *Scientific Reports*, 7(1), pp.1–12.

Nichols, R.F.W., (1950). The Brown Streak Disease of Cassava. *The East African Agricultural Journal*, 15(3), pp.154–160. Available at: https://doi.org/10.1080/03670074.1950.11664727.

Niehl, A. et al., (2012). Control of Tobacco mosaic virus movement protein fate by CELL-DIVISION-CYCLE protein48. *Plant Physiology*, 160(4), pp.2093–2108.

Nishizawa-Yokoi, A., Yabuta, Y. & Shigeoka, S., (2008). The contribution of carbohydrates including raffinose family oligosaccharides and sugar alcohols to protection of plant cells from oxidative damage. *Plant Signaling and Behavior*, 3(11), pp.1016–1018.

Niu, L. et al., (2018). Modified TCA/acetone precipitation of plant proteins for proteomic analysis. *PLoS ONE*, 13(12), pp.1–13.

Nkaa, F.A. et al., (2013). Elimination of African Cassava Mosaic Virus (ACMV) and East African Cassava Mosaic Virus (EACMV) from cassava (Manihot esculenta Crantz)

cv. “Nwugo” via somatic embryogenesis. *American Journal Of Biotechnology And Molecular Sciences*, 3(2), pp.33–40. Available at: http://www.scihub.org/AJBMS.

Noueiry, A.O., Lucas, W.J. & Gilbertson, R.L., (1994). Two proteins of a plant DNA virus coordinate nuclear and plasmodesmal transport. *Cell*, 76(5), pp.925–932.

Ntui, V.O. et al., (2015). Resistance to Sri Lankan cassava mosaic virus (SLCMV) in genetically engineered cassava cv. KU50 through RNA silencing. *PLoS ONE*.

Oblessuc, P.R., Matiolli, C.C. & Melotto, M., (2020). Novel molecular components involved in callose-mediated Arabidopsis defense against Salmonella enterica and Escherichia coli O157:H7. *BMC Plant Biology*, 20(1), pp.1–13.

Ochsenbein, C. et al., (2006). The role of EDS1 (enhanced disease susceptibility) during singlet oxygen-mediated stress responses of Arabidopsis. *Plant Journal*.

Ogawa, D., Yamaguchi, K. & Nishiuchi, T., (2007). High-level overexpression of the Arabidopsis HsfA2 gene confers not only increased themotolerance but also salt/osmotic stress tolerance and enhanced callus growth. *Journal of Experimental Botany*, 58(12), pp.3373–3383.

Okogbenin, E. et al., (2007). Marker-assisted introgression of resistance to cassava mosaic disease into latin American germplasm for the genetic improvement of cassava in Africa. *Crop Science*.

Okogbenin, E. et al., (2012). Molecular marker analysis and validation of resistance to cassava mosaic disease in elite cassava genotypes in Nigeria. *Crop Science*.

Oliveira, C. et al., (2017). RNA-binding proteins and their role in the regulation of gene expression in trypanosoma cruzi and saccharomyces cerevisiae. *Genetics and Molecular Biology*, 40(1), pp.22–30.

Oliveira Gilmara Alvarenga Fachardo, de O.E.J. do C.D.C. da S.M.S., (2015). Molecular-assisted selection for resistance to cassava mosaic disease in Manihot. *Scientia Agricola*, 72(6), pp.520–527.

Opara, E.U., (2015). An overview on the role of virus vectors in host plant Intraction-transmission strategy. *World Journal of Agricultural Sciences*, 3(2), pp.9–12. Available at: http://wsrjournals.org/journal/wjas.

Osakabe, K. et al., (2020). Genome editing in plants using CRISPR type I-D nuclease. *Communications Biology*, 3(1), pp.1–10. Available at: http://dx.doi.org/10.1038/s42003-020-01366-6.

Osmond, C.B. et al., (1998). Chlorophyll fluorescence quenching during photosynthetic induction in leaves of Abutilon striatum dicks. Infected with Abutilon mosaic virus, observed with a field-portable imaging system. *Botanica Acta*, 111(5), pp.390–397.

Padidam, M., Sawyer, S. & Fauquet, C.M., (1999). Possible Emergence of New Geminiviruses by Frequent Recombination. *Virology*, 265(2), pp.218–25. Available at: 10.1006/viro.1999.0056.

Pan, J. et al., (2021). Computational prediction of protein - protein interactions in plants using only sequence information. *Research square*, Preprint(cited July 9,2021).

Pandey, G.K. et al., (2005). ABR1, an APETALA2-domain transcription factor that functions as a repressor of ABA response in Arabidopsis. *Plant Physiology*, 139(3), pp.1185–1193.

Parisi, E. et al., (2018). Cdc48/p97 segregase is modulated by cyclin-dependent kinase to determine cyclin fate during G1 progression. *The EMBO Journal*, 37(16), pp.1–14.

Parkes, Elizabeth et al., (2015). Developing Cassava Mosaic Disease resistant cassava

varieties in Ghana using a marker assisted selection approach. *Euphytica*, 203, pp.549–556.

Parry, M.A.J. et al., (1987). Variations in properties of ribulose-1,5-bisphosphate carboxylase from various species related to differences in amino acid sequences. *Journal of Experimental Botany*, 38(8), pp.1260–1271.

Pascal, E. et al., (1994). The geminivirus BR1 movement protein binds single-stranded DNA and localizes the cell nucleus. *Plant Cell*, 6(7), pp.995–1006.

Pascal, E et al., (1993). Transgenic tobacco plants expressing the geminivirus BL1 protein exhibit symptoms of viral disease. *The Plant cell*, 5(7), pp.795–807. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=160317&tool=pmcentrez&rendertype=abstract.

Pascal, Erica et al., (1993). Transgenic tobacco plants expressing the geminivirus BL1 protein exhibit symptoms of viral disease. *Plant Cell*, 5(7), pp.795–807.

Pashkova, N. et al., (2016). DEEPN as an Approach for Batch Processing of Yeast 2-Hybrid Interactions. *Cell Reports*, 17(1), pp.303–315. Available at: http://dx.doi.org/10.1016/j.celrep.2016.08.095.

Patanun, O. et al., (2012). Computational Identification of MicroRNAs and Their Targets in Cassava (Manihot esculenta Crantz.). *Molecular Biotechnology*.

Patarroyo, C., Laliberté, J.F. & Zheng, H., (2013). Hijack it, change it: How do plant viruses utilize the host secretory pathway for efficient viral replication and spread? *Frontiers in Plant Science*, 3(JAN), pp.1–8.

Patil, A. & Nakamura, H., (2006). Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks. *FEBS Letters*, 580(8), pp.2041–2045.

Patil, B.L. et al., (2015). Cassava brown streak disease: A threat to food security in Africa. *Journal of General Virology*.

Patil, B.L. & Fauquet, C.M., (2009). Cassava mosaic geminiviruses: Actual knowledge and perspectives. *Molecular Plant Pathology*, 10(5), pp.685–701.

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(3).

Pauwels, L. et al., (2020). NINJA connects the co-repressor TOPLESS to jasmonate signalling. , 464(7289), pp.788–791.

Pelgrom, A.J.E. et al., (2020). Host interactors of effector proteins of the lettuce downy mildew Bremia lactucae obtained by yeast two-hybrid screening. *PLoS ONE*, 15(5), pp.1–26. Available at: http://dx.doi.org/10.1371/journal.pone.0226540.

Peters, S. et al., (2010). Functional identification of Arabidopsis ATSIP2 (At3g57520) as an alkaline α-galactosidase with a substrate specificity for raffinose and an apparent sink-specific expression pattern. *Plant and Cell Physiology*, 51(10), pp.1815–1819.

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*.

Pilartz, M. & Jeske, H., (1992). Abutilon mosaic geminivirus double-stranded DNA is packed into minichromosomes. *Virology*, 189(2), pp.800–802.

Pinkert, S., Schultz, J. & Reichardt, J., (2010). Protein interaction networks-more than

mere modules. *PLoS Computational Biology*.

Plant, E.P. et al., (2007). Differentiating between Near- and Non-Cognate Codons in Saccharomyces cerevisiae. *PLoS ONE*, 2(6).

Plant village, (2015). Cassava (Manioc). Available at: https://www.plantvillage.org/en/topics/cassava-manioc/diseases_and_pests_description_uses_propagation [Accessed October 1, 2015].

Pooggin, M.M., (2013). How can plant DNA viruses evade siRNA-directed DNA methylation and silencing? *International Journal of Molecular Sciences*.

Poole, N. et al., (2010). Constraints to Smallholder Participation in Cassava Value Chain Development in Zambia. Available at: http://www.fao.org/fileadmin/templates/est/AAACP/eastafrica/FAO_AAACP_Paper_Series_No_15_Constraints_to_smallholder_participaÃ__1_.pdf.

Potrykus, I. et al., (1985). Direct gene transfer State of the Art and Future Potential. *Plant Molecular Biology Reporter*, 3(3), pp.117–128.

Poudel, A.N. et al., (2019). 12-Hydroxy-Jasmonoyl-l-Isoleucine Is an Active Jasmonate That Signals through CORONATINE INSENSITIVE 1 and Contributes to the Wound Response in Arabidopsis. *Plant and Cell Physiology*, 60(10), pp.2152–2166.

Powers, M.A. et al., (1997). The vertebrate GLFG nucleoporin, Nup98, is an essential component of multiple RNA export pathways. *Journal of Cell Biology*, 136(2), pp.241–250.

Pribat, A. et al., (2012). A novel class of PTEN protein in *Arabidopsis* displays unusual phosphoinositide phosphatase activity and efficiently binds phosphatidic acid. *Biochemical Journal*, 441(1), pp.161–171. Available at: http://biochemj.org/lookup/doi/10.1042/BJ20110776.

Prochnik, S. et al., (2012). The Cassava Genome: Current Progress, Future Directions. *Tropical Plant Biology*.

Punna Ramu, J.-L.J.M.W.I.Y.R.C.E.M.H.R.K.P.K.R.L.D.P.D.C., (2016). Genome-wide association and prediction reveals the genetic architecture of cassava mosaic disease resistance and prospects for rapid genetic improvement. *The Plant Genome*, pp.1–248. Available at: http://biorxiv.org/content/early/2015/11/11/031179.article-info.

Pyott, D.E., Sheehan, E. & Molnar, A., (2016). Engineering of CRISPR/Cas9-mediated potyvirus resistance in transgene-free Arabidopsis plants. *Molecular plant pathology*, 17(8), pp.1276–1288.

Qazi, J. et al., (2007). Legume yellow mosaic viruses: Genetically isolated begomoviruses. *Molecular Plant Pathology*, 8(4), pp.343–348.

Qiu, K. et al., (2015). EIN3 and ORE1 Accelerate Degreening during Ethylene-Mediated Leaf Senescence by Directly Activating Chlorophyll Catabolic Genes in Arabidopsis. *PLoS Genetics*, 11(7), pp.1–20.

Rabbi, I.Y. et al., (2014). High-resolution mapping of resistance to cassava mosaic geminiviruses in cassava using genotyping-by-sequencing and its implications for breeding. *Virus Research*.

Rain, J.C. et al., (2001). The protein-protein interaction map of Helicobacter pylori. *Nature*, 409(6817), pp.211–215.

Raja, P., (2010). Methylation of Geminivirus Genomes: Investigating its role as a host defense and evaluating its efficacy as a model to study chromatin methylation in

plants . *Molecular, Cellular, and Developmental Biology*, Ph.D., p.232. Available at: http://etd.ohiolink.edu/view.cgi?acc_num=osu1274728228.

Raja, P. et al., (2008). Viral Genome Methylation as an Epigenetic Defense against Geminiviruses. *Journal of Virology*, 82(18), pp.8997–9007.

Ran, A.F. et al., (2013). Genome engineering using the CRISPR-Cas9 system. *Nature protocols*, 8(11), pp.2281–2308. Available at: http://dx.doi.org/10.1038/nprot.2013.143%5Cnhttp://www.nature.com/nprot/journal/v8/n11/abs/nprot.2013.143.html.

Rancour, D.M. et al., (2002). Characterization of AtCDC48. Evidence for multiple membrane fusion mechanisms at the plane of cell division in plants. *Plant Physiology*, 130(3), pp.1241–1253.

Reagan, B.C. & Burch-Smith, T.M., (2020). Viruses reveal the secrets of plasmodesmal cell biology. *Molecular Plant-Microbe Interactions*, 33(1), pp.26–39.

Reichel, C. & Beachy, R.N., (2000). Degradation of Tobacco Mosaic Virus Movement Protein by the 26S Proteasome. *Journal of Virology*, 74(7), pp.3330–3337.

Ren, R. et al., (2020). Highly efficient protoplast isolation and transient expression system for functional characterization of flowering related genes in Cymbidium orchids. *International Journal of Molecular Sciences*, 21(7).

Rhoades, M.W. et al., (2002). Prediction of plant microRNA targets. *Cell*.

Richter, K.S. et al., (2016). The contribution of translesion synthesis polymerases on geminiviral replication. *Virology*, 488, pp.137–148.

Robaglia, C. & Caranta, C., (2006). Translation initiation factors: A weak link in plant RNA virus infection. *Trends in Plant Science*, 11(1), pp.40–45.

Rodríguez-Negrete, E.A., Carrillo-Tripp, J. & Rivera-Bustamante, R.F., (2009). RNA Silencing against Geminivirus: Complementary Action of Posttranscriptional Gene Silencing and Transcriptional Gene Silencing in Host Recovery. *JOURNAL OF VIROLOGY*, 83(3), pp.1332–1340.

Rogans, S.J. & Rey, C., (2016). Unveiling the micronome of cassava (manihot esculenta crantz). *PLoS ONE*, 11(1), pp.1–30.

Rojas, M.R. et al., (1998). Bean dwarf mosaic geminivirus movement proteins recognize DNA in a form- and size-specific manner. *Cell*, 95(1), pp.105–113.

Román, V. et al., (2014). Construction of a cassava PR protein-interacting network during Xanthomonas axonopodis pv. manihotis infection. *Plant Pathology*.

Rosas-Diaz, T. et al., (2018). A virus-targeted plant receptor-like kinase promotes cell-to-cell spread of RNAi. *Proceedings of the National Academy of Sciences*.

Rosas-Díaz, T. et al., (2016). The C2 Protein from the Geminivirus Tomato Yellow Leaf Curl Sardinia Virus Decreases Sensitivity to Jasmonates and Suppresses Jasmonate-Mediated Defences. *Plants*, 5(1), p.8.

Rose, A. et al., (2004). Genome-wide identification of Arabidopsis coiled-coil proteins and establishment of the ARABI-COIL database. *Plant physiology*, 134(3), pp.927–939.

Rotter, A. et al., (2007). Adaptation of the MapMan ontology to biotic stress responses: application in solanaceous species. *Plant Methods*, 3(10). Available at: http://www.plantmethods.com/content/3/1/10.

Rottier, P., (1980). Viral protein synthesis in cowpea mosaic virus infceted protoplasts. , 3, p.97. Available at: https://edepot.wur.nl/211280.

Rowland, O. et al., (2006). CER4 Encodes an Alcohol-Forming Fatty Acyl-Coenzyme A Reductase Involved in Cuticular Wax Production in Arabidopsis. *Plant Physiology*, 142(3), pp.866–877. Available at: http://www.plantphysiol.org/cgi/doi/10.1104/pp.106.086785.

Roy, A. et al., (2019). Multiplexed editing of a begomovirus genome restricts escape mutant formation and disease development. *PLoS ONE*, 14(10), pp.1–18.

Roy, B. et al., (2015). Nonsense suppression by near-cognate tRNAs employs alternative base pairing at codon positions 1 and 3. *Proceedings of the National Academy of Sciences of the United States of America*, 112(10), pp.3038–3043.

Ruan, J. et al., (2019). Jasmonic acid signaling pathway in plants. *International Journal of Molecular Sciences*, 20(10).

Sable, A. et al., (2018). Inhibition of Heat Shock proteins HSP90 and HSP70 induce oxidative stress, suppressing cotton fiber development. *Scientific Reports*, 8(1), pp.1–17. Available at: http://dx.doi.org/10.1038/s41598-018-21866-0.

Sah, S.K., Reddy, K.R. & Li, J., (2016). Abscisic acid and abiotic stress tolerance in crop plants. *Frontiers in Plant Science*, 7(MAY2016), pp.1–26.

Sahu, P.P. et al., (2016). Tomato 26S Proteasome subunit RPT4a regulates ToLCNDV transcription and activates hypersensitive response in tomato. *Scientific Reports*, 6(June), pp.1–12. Available at: http://dx.doi.org/10.1038/srep27078.

Samsatly, J., Bayen, S. & Jabaji, S.H., (2020). Vitamin B6 Is Under a Tight Balance During Disease Development by Rhizoctonia solani on Different Cultivars of Potato and on Arabidopsis thaliana Mutants. *Frontiers in Plant Science*, 11(June), pp.1–18.

Sanderfoot, A.A. & Lazarowitz, S.G., (1995). Cooperation in viral movement: The geminivirus BL1 movement protein interacts with BR1 and redirects it from the nucleus to the cell periphery. *Plant Cell*, 7(8), pp.1185–1194.

Sano, S. et al., (2014). Light-dependent expression of flg22-induced defense genes in Arabidopsis.

Sant'ana, R.R.A. et al., (2020). PEG-delivered CRISPR-Cas9 ribonucleoproteins system for gene-editing screening of maize protoplasts. *Genes*, 11(9), pp.1–14.

Santoro, M.G., Amici, C. & Rossi, A., (2009). *Role of Heat Shock Proteins in Viral Infection*,

Sarry, J.E. et al., (2006). The early responses of Arabidopsis thaliana cells to cadmium exposure explored by protein and metabolite profiling analyses. *Proteomics*, 6(7), pp.2180–2198.

Satterly, N. et al., (2007). Influenza virus targets the mRNA export machinery and the nuclear pore complex. *Proceedings of the National Academy of Sciences of the United States of America*, 104(6), pp.1853–1858.

Saunders, K. et al., (2003). The earliest recorded plant virus disease. *Nature*, 422, p.831. Available at: http://dx.doi.org/10.1038/422831a.

Saunders, K., Lucy, A. & Stanley, J., (1991). DNA forms of the geminivirus African cassava mosaic virus consistent with a rolling circle mechanism of replication. *Nucleic Acids Research*, 19(9), pp.2325–2330.

Schiml, S., Fauser, F. & Puchta, H., (2016). CRISPR/Cas-Mediated Site-Specifi c Mutagenesis in Arabidopsis thaliana Using Cas9 Nucleases and Paired Nickases. In M. Murata, ed. *Methods in Molecular Biology*. Humana Press, pp. 111–122.

Schiml, S., Fauser, F. & Puchta, H., (2014). The CRISPR/Cas system can be used as

nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in Arabidopsis resulting in heritable progeny. *Plant Journal*, 80(6), pp.1139–1150.

Schindele, P. & Puchta, H., (2020). Engineering CRISPR/LbCas12a for highly efficient, temperature-tolerant plant gene editing. *Plant Biotechnology Journal*, 18(5), pp.1118–1120.

Schmitz, D.J. et al., (2020). CRISPR/Cas9 Mutagenesis by Translocation of Cas9 Protein Into Plant Cells via the Agrobacterium Type IV Secretion System. *Frontiers in Genome Editing*, 2(July), pp.1–11.

Scholthof, K.B.G. et al., (2011). Top 10 plant viruses in molecular plant pathology. *Molecular Plant Pathology*.

Schroeder, A. et al., (2006). The RIN: An RNA integrity number for assigning integrity values to RNA measurements. *BMC Molecular Biology*, 7, pp.1–14.

Schuberth, C. & Buchberger, A., (2005). Membrane-bound Ubx2 recruits Cdc48 to ubiquitin ligases and their substrates to ensure efficient ER-associated protein degradation. *Nature Cell Biology*, 7(10), pp.999–1006.

Sciuto, M.R. et al., (2018). Two-Step Coimmunoprecipitation (TIP) enables efficient and highly selective isolation of native protein complexes. *Molecular and Cellular Proteomics*, 17(5), pp.993–1009.

Seal S. E., vandenBosch F.& J.M.J., (2006). Factors Influencing Begomovirus Evolution and Their Increasing Global Significance: Implications for Sustainable Control. *Critical Reviews in Plant Sciences*, 25(1), pp.23–46.

Segrè, D. et al., (2004). Modular epistasis in yeast metabolism. *Nature Genetics*, 37, pp.77–83.

Sentmanat, M.F. et al., (2018). A Survey of Validation Strategies for CRISPR-Cas9 Editing. *Scientific Reports*, 8(1), pp.1–8. Available at: http://dx.doi.org/10.1038/s41598-018-19441-8.

Seo, J.K. et al., (2018). Molecular dissection of distinct symptoms induced by tomato chlorosis virus and tomato yellow leaf curl virus based on comparative transcriptome analysis. *Virology*, 516(January), pp.1–20. Available at: https://doi.org/10.1016/j.virol.2018.01.001.

Serebriiskii, I.G. & Golemis, E.A., (2000). Uses of lacZ to study gene function: Evaluation of β-galactosidase assays employed in the yeast two-hybrid system. *Analytical Biochemistry*, 285(1), pp.1–15.

Serebriiskii, I.G., Toby, G.G. & Golemis, E.A., (2000). Streamlined yeast colorimetric reporter activity assays using scanners and plate readers. *BioTechniques*, 29(2), pp.278–288.

Serre, N.B.C. et al., (2020). Protein lysine methylation contributes to modulating the response of sensitive and tolerant Arabidopsis species to cadmium stress. *Plant Cell and Environment*, 43(3), pp.760–774.

Shadwick, F.S., (2007). Viral Infection and Propagation in Plant Tissue Culture. , (January).

Sharan, R., Ulitsky, I. & Shamir, R., (2007). Network-based prediction of protein function. *Molecular Systems Biology*, 13.

Shen, J. et al., (2014). Isolation, culture, and transienttransformation of plant protoplasts. *Current Protocols in Cell Biology*, (SUPPL.63), pp.1–17.

Shen, Q. et al., (2016). Tobacco RING E3 Ligase NtRFP1 Mediates Ubiquitination and Proteasomal Degradation of a Geminivirus-Encoded βC1. *Molecular Plant*, 9(6), pp.911–925. Available at: http://dx.doi.org/10.1016/j.molp.2016.03.008.

Shen, Q. et al., (2011). Tomato S1SnRK1 protein interacts with and phosphorylates βC1, a Pathogenesis protein encoded by a geminivirus β-satellite. *Plant Physiology*, 157(3), pp.1394–1406.

Shen, Q., Bao, M. & Zhou, X., (2012). A plant kinase plays roles in defense response against geminivirus by phosphorylation of a viral pathogenesis protein. , (July), pp.888–892.

Shen, W. et al., (2014). SnRK1 Phosphorylation of AL2 Delays Cabbage Leaf Curl Virus Infection in Arabidopsis. *Journal of Virology*, 88(18), pp.10598–10612.

Shen, W. et al., (2018). Sucrose Non-Fermenting 1-Related Protein Kinase 1 Phosphorylates a Geminivirus Rep Protein to Impair Viral Replication and Infection. *Plant Physiology*, (Mcb 1052218), p.pp.00268.2018. Available at: http://www.plantphysiol.org/lookup/doi/10.1104/pp.18.00268.

Shen, W., Reyes, M.I. & Hanley-Bowdoin, L., (2009). Arabidopsis Protein Kinases GRIK1 and GRIK2 Specifically Activate SnRK1 by Phosphorylating Its Activation Loop. *Plant Physiology*, 150(2), pp.996–1005. Available at: http://www.plantphysiol.org/cgi/doi/10.1104/pp.108.132787.

Sheppard, C. & Werner, F., (2017). Structure and mechanisms of viral transcription factors in archaea. *Extremophiles*, 21(5), pp.829–838.

Shih, Y.K. & Parthasarathy, S., (2012). Identifying functional modules in interaction networks through overlapping Markov clustering. *Bioinformatics*.

Shivhare, D. et al., (2021). Removing auto-activators from yeast-two-hybrid assays by conditional negative selection. *Scientific Reports*, 11(1), pp.1–9. Available at: https://doi.org/10.1038/s41598-021-84608-9.

Shkryl, Y. et al., (2021). Crispr/cas9-mediated knockout of hos1 reveals its role in the regulation of secondary metabolism in arabidopsis thaliana. *Plants*, 10(1), pp.1–18.

Silva, F.N. et al., (2014). Recombination and pseudorecombination driving the evolution of the begomoviruses Tomato severe rugose virus (ToSRV) and Tomato rugose mosaic virus (ToRMV): two recombinant DNA-A components sharing the same DNA-B. *Virology Journal*, 11. Available at: http://www.virologyj.com/content/11/1/66.

Singh, B.K. & Shaner, D.L., (1995). Biosynthesis of branched chain amino acids: From test tube to field. *Plant Cell*, 7(7), pp.935–944.

Slatko, B.E., Gardner, A.F. & Ausubel, F.M., (2018). Overview of Next-Generation Sequencing Technologies. *Current Protocols in Molecular Biology*, 122(1), pp.1–15.

Smakowska-Luzan, E. et al., (2018). An extracellular network of Arabidopsis leucine-rich repeat receptor kinases. *Nature*, 553(7688), pp.342–346. Available at: http://dx.doi.org/10.1038/nature25184.

Smith, D.R. et al., (2010). Inhibition of heat-shock protein 90 reduces Ebola virus replication. *Antiviral Research*, 87(2), pp.187–194.

Sofiari, E., (1996). *Regeneration and Transformation of Cassava (Manihot Esculenta Crantz.)*, Available at: http://books.google.com.co/books?id=9LT8tgAACAAJ.

Soitamo, A.J., Jada, B. & Lehto, K., (2012). Expression of geminiviral AC2 RNA silencing suppressor changes sugar and jasmonate responsive gene expression in transgenic tobacco plants. *BMC Plant Biology*, 12. Available at: http://www.biomedcentral.com/1471-2229/12/204.

Souza, P.F.N., Garcia-Ruiz, H. & Carvalho, F.E.L., (2019). What proteomics can reveal about plant–virus interactions? Photosynthesis-related proteins on the spotlight. *Theoretical and Experimental Plant Physiology*, 31(1), pp.227–248.

Sprenger, N. & Keller, F., (2000). Allocation of raffinose family oligosaccharides to transport and storage pools in Ajuga reptans: The roles of two distinct galactinol synthases. *Plant Journal*, 21(3), pp.249–258.

Srivastava, A. et al., (2015). Characterization of a novel begomovirus associated with yellow mosaic disease of three ornamental species of Jatropha grown in India. *Virus Research*, 201, pp.41–49. Available at: http://dx.doi.org/10.1016/j.virusres.2015.02.015.

Srivastava, A. et al., (2017). Screening of chilli pepper (Capsicum spp.) lines for resistance to the begomoviruses causing chilli leaf curl disease in India. *Crop Protection*, 100, pp.177–185. Available at: http://dx.doi.org/10.1016/j.cropro.2017.06.015.

Staack, L. et al., (2019). Cassava cell wall characterization and degradation by a multicomponent NSP-targeting enzyme (NSPase). *Scientific Reports*, 9(1), pp.1–11.

Stael, S. et al., (2015). Plant innate immunity - sunny side up? *Trends in Plant Science*.

Staley, J.P. & Woolford, J.L., (2009). Assembly of ribosomes and spliceosomes: complex ribonucleoprotein machines. *Current Opinion in Cell Biology*, 21(1), pp.109–118.

Stanley, J. & Albrecht, von A., (1992). Determinants of Tomato Golden Located on DNA B. *Virology*, 293, pp.286–293.

Stark, C. et al., (2006). BioGRID: a general repository for interaction datasets. *Nucleic acids research*, 34(Database issue), pp.535–539.

Stein, O. & Granot, D., (2018). Plant fructokinases: Evolutionary, developmental, and metabolic aspects in sink tissues. *Frontiers in Plant Science*, 9(March), pp.1–12.

Steinert, J., Schmidt, C. & Puchta, H., (2017). Use of the Cas9 Orthologs from Streptococcus thermophilus and Staphylococcus aureus for Non-Homologous End-Joining Mediated Site-Specific Mutagenesis in Arabidopsis thaliana. Plant Germline Development. *Plant Germline Development*, 1669, pp.365–376.

Stellberger, T. et al., (2010). Improving the yeast two-hybrid system with permutated fusions proteins: The Varicella Zoster Virus interactome. *Proteome Science*, 8, pp.1–9.

Stemmer, M. et al., (2015). CCTop: An intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. *PLoS ONE*, 10(4), pp.1–11.

Storey, H.H. & Nichols, R.F.W., (1938). Studies of the mosaic disease of cassava. *Annals of Applied Biology*, 25(1906), pp.790–806. Available at: http://dx.doi.org/10.1111/j.1744-7348.1938.tb02354.x.

Struhl, K. & Segal, E., (2013). Determinants of nucleosome positioning. *Nature Structural and Molecular Biology*, 20(3), pp.267–273.

Su, S. et al., (2010). Cucumber Mosaic Virus movement protein severs actin filaments to increase the plasmodesmal size exclusion limit in tobacco. *Plant Cell*, 22(4), pp.1373–1387.

Sump, B. & Brickner, J.H., (2017). Nup98 regulation of histone methylation promotes normal gene expression and may drive leukemogenesis. *Genes and Development*, 31(22), pp.2201–2203.

Sun, R. et al., (2020). The ac2 protein of a bipartite geminivirus stimulates the

transcription of the bv1 gene through abscisic acid responsive promoter elements. *Viruses*, 12(12), pp.1–13.

Sun, X. et al., (2015). Targeted mutagenesis in soybean using the CRISPR-Cas9 system. *Scientific Reports*, 5, pp.1–10.

Sun, Y.W. et al., (2015). Attenuation of Histone Methyltransferase KRYPTONITE-mediated transcriptional gene silencing by Geminivirus. *Scientific Reports*, 5(November), pp.1–12.

Sunter, G. & Bisaro, D.M., (1992). Transactivation of geminivirus AR1 and BR1 gene expression by the viral AL2 gene product occurs at the level of transcription. *Plant Cell*, 4(10), pp.1321–1331.

Szklarczyk, D. et al., (2017). The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Research*, 45(D1), pp.D362–D368. Available at: http://dx.doi.org/10.1093/nar/gkw937.

Tahir, M. et al., (2015). Ageratum enation virus—A Begomovirus of weeds with the potential to infect crops. *Viruses*, 7(2), pp.647–665.

Takeda, S. et al., (2011). CUP-SHAPED COTYLEDON1 transcription factor activates the expression of LSH4 and LSH3, two members of the ALOG gene family, in shoot organ boundary cells. *Plant Journal*, 66(6), pp.1066–1077.

Tamura, K. et al., (2010). Identification and characterization of nuclear pore complex components in Arabidopsis thaliana. *Plant Cell*, 22(12), pp.4084–4097.

Tan, Y.F. et al., (2010). Divalent metal ions in plant mitochondria and their role in interactions with proteins and oxidative stress-induced damage to respiratory function. *Plant Physiology*, 152(2), pp.747–761.

Teixeira, R.M. et al., (2021). Geminiviral triggers and suppressors of plant antiviral immunity. *Microorganisms*, 9(4).

Testerink, C. et al., (2004). Isolation and identification of phosphatidic acid targets from plants. *Plant Journal*, 39(4), pp.527–536.

Thaker, S.K., Ch'ng, J. & Christofk, H.R., (2019). Viral hijacking of cellular metabolism. *BMC biology*, 17(1), p.59.

Theodoris, G., Inada, N. & Freeling, M., (2003). Conservation and molecular dissection of ROUGH SHEATH2 and ASYMMETRIC LEAVES1 function in leaf development. *Proceedings of the National Academy of Sciences of the United States of America*, 100(11), pp.6837–6842.

Thimm, O. et al., (2004). MAPMAN: A user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant Journal*.

Tomaz, T. et al., (2010). Mitochondrial malate dehydrogenase lowers leaf respiration and alters photorespiration and plant growth in Arabidopsis. *Plant Physiology*, 154(3), pp.1143–1157.

Toualy, M.N.Y. et al., (2014). Incidence and distribution of cassava mosaic begomoviruses in Côte d'Ivoire.

Trapet, P. et al., (2015). NO signaling in plant immunity: A tale of messengers. *Phytochemistry*.

Trigg, S.A. et al., (2017). CrY2H-seq: A massively multiplexed assay for deep-coverage interactome mapping. *Nature Methods*, 14(8), pp.819–825.

Truman, W. et al., (2007). Arabidopsis systemic immunity uses conserved defense

signaling pathways and is mediated by jasmonates. *Proceedings of the National Academy of Sciences*, 104(3), pp.1075–1080.

Tsai, M.Y. et al., (2000). Release of kinesin from vesicles by hsc70 and regulation of fast axonal transport. *Molecular Biology of the Cell*, 11(6), pp.2161–2173.

Uchiyama, A. et al., (2014). The Arabidopsis synaptotagmin SYTA regulates the cell-to-cell movement of diverse plant viruses. *Frontiers in Plant Science*, 5(November), pp.1–17.

Ugaki, M. et al., (1991). Replication of a geminivirus derived shuttle vector in maize endosperm cells. *Nucleic Acids Research*, 19(2), pp.371–377.

Uhrig, J.F. & MacFarlane, S.A., (2008). Protein-Protein Interactions in Plant Virus Movement and Pathogenicity. In C. S. Nautiyal & P. Dion, eds. *Molecular Mechanisms of Plant and Microbe Coexistence*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 319–338. Available at: http://dx.doi.org/10.1007/978-3-540-75575-3_13.

Ujino, S. et al., (2009). Heat-shock protein 90 is essential for stabilization of the hepatitis C virus nonstructural protein NS3. *Journal of Biological Chemistry*, 284(11), pp.6841–6846. Available at: http://dx.doi.org/10.1074/jbc.M806452200.

Unseld, S. et al., (2001). Subcellular Targeting of the Coat Protein of African Cassava Mosaic Geminivirus. *Virology*, 286(2), pp.373–383.

Usadel, B. et al., (2009). A guide to using MapMan to visualize and compare Omics data in plants: A case study in the crop species, Maize. *Plant, Cell and Environment*.

Valluru, R. & Van den Ende, W., (2011). Myo-inositol and beyond - Emerging networks under stress. *Plant Science*, 181(4), pp.387–400. Available at: http://dx.doi.org/10.1016/j.plantsci.2011.07.009.

Vanderschuren, H. et al., (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*.

Vanderschuren, H. et al., (2013). Strategies for vitamin B6 biofortification of plants: A dual role as a micronutrient and a stress protectant. *Frontiers in Plant Science*, 4(MAY), pp.1–7.

Vanitharani, R. et al., (2004). Differential Roles of AC2 and AC4 of Cassava Geminiviruses in Mediating Synergism and Suppression of Posttranscriptional Gene Silencing. *JOURNAL OF VIROLOGY*, 78(17), pp.9487–9498.

Vargason, J.M. et al., (2003). Size Selective Recognition of siRNA by an RNA Silencing Suppressor. *Cell*, 115(7), pp.799–811.

Varma, A. & Malathi, V.G., (2003). Emerging geminivirus problems: A serious threat to crop production. *Annals of Applied Biology*, 142(2), pp.145–164.

Vaucheret, H., (2006). Post-transcriptional small RNA pathways in plants: mechanisms and regulations. *Genes and Development*, 20, pp.759–771. Available at: http://genesdev.cshlp.org/content/20/7/759.long.

Vega, S. et al., (2013). NS3 protease from hepatitis C virus: Biophysical studies on an intrinsically disordered protein domain. *International Journal of Molecular Sciences*, 14(7), pp.13282–13306.

Velásquez-Zapata, V. et al., (2021). Next-generation yeast-two-hybrid analysis with Y2H-SCORES identifies novel interactors of the MLA immune receptor. *PLoS Computational Biology*, 17(4), pp.1–30.

Velásquez-Zapata, V. et al., (2020). Y2H-SCORES: A statistical framework to infer protein-protein interactions from next-generation yeast-two-hybrid sequence data. *bioRxiv*, pp.1–44.

Vojtek, A.B. & Hollenberg, S.M., (1995). Ras-Raf Interaction: Two-Hybrid Analysis. *Methods in Enzymology*, 255(C), pp.331–342.

Wada, M. et al., (2018). Interplay between coronavirus, a cytoplasmic RNA virus, and nonsense-mediated mRNA decay pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 115(43), pp.E10157–E10166.

Wang, A. & Krishnaswamy, S., (2012). Eukaryotic translation initiation factor 4E-mediated recessive resistance to plant viruses and its utility in crop improvement. *Molecular Plant Pathology*, 13(7), pp.795–803.

Wang, B. et al., (2014). V2 of tomato yellow leaf curl virus can suppress methylation-mediated transcriptional gene silencing in plants. *Journal of General Virology*, 95(PART 1), pp.225–230.

Wang, C. et al., (2018). Pipecolic acid confers systemic immunity by regulating free radicals. *Science Advances*, 4(5), pp.1–11.

Wang, D. & Maule, A.J., (1995). Inhibition of host gene expression associated with plant virus replication. *Science*, 267(5195), pp.229–231.

Wang, H. et al., (2003). Adenosine Kinase Is Inactivated by Geminivirus AL2 and L2 Proteins. *Plant Cell*, 15(12), pp.3020–3032.

Wang, L. et al., (2017). Inference of a geminivirus−host protein−protein interaction network through affinity purification and mass spectrometry analysis. *Viruses*, 9(10).

Wang, M.-B. et al., (2012). RNA Silencing and Plant Viral Diseases. , 25(10), pp.1275–1285. Available at: http://dx.doi.org/10.1094.

Wang, Q. et al., (2017). Identification of host cellular proteins that interact with the M protein of a highly pathogenic porcine reproductive and respiratory syndrome virus vaccine strain. *Virology Journal*, 14(1), pp.1–11.

Wang, W. et al., (2014). ARTICLE Cassava genome from a wild ancestor to cultivated varieties. *Nature Communications*, 5.

Wang, W. et al., (2011). Timing of plant immune responses by a central circadian regulator. *Nature*, 470(7332), pp.110–115.

Wang, X. & Chapman, K.D., (2013). Lipid signaling in plants. *Frontiers in Plant Science*, 4(JUN), pp.1–2.

Wang, Y.Q. et al., (2013). Proteomic analysis of chromoplasts from six crop species reveals insights into chromoplast function and development. *Journal of Experimental Botany*.

Ward, B.M. et al., (1997). The geminivirus BL1 movement protein is associated with endoplasmic reticulum-derived tubules in developing phloem cells. *Journal of virology*, 71(5), pp.3726–3733. Available at: http://jvi.asm.org/content/71/5/3726.abstract.

Wasswa, P. et al., (2011). First identification of a sweet potato begomovirus (sweepovirus) in Uganda: Characterization, detection and distribution. *Plant Pathology*, 60(6), pp.1030–1039.

Wasternack, C., (2007). Jasmonates: An update on biosynthesis, signal transduction and action in plant stress response, growth and development. *Annals of Botany*, 100(4), pp.681–697.

Weber, H., (2002). Fatty acid-derived signals in plants. *Trends in Plant Science*, 7(5), pp.217–224.

Wege, C. & Jeske, H., (1998). Abutilon mosaic geminivirus proteins expressed and phosphorylated in Escherichia coli. *Journal of Phytopathology*, 146(11–12), pp.613–621.

Wei, T. et al., (2010). Formation of complexes at plasmodesmata for potyvirus intercellular movement is mediated by the viral protein P3N-PIPO. *PLoS Pathogens*, 6(6).

Wen, F. et al., (2020). Plant regeneration via protoplast electrofusion in cassava. *Journal of Integrative Agriculture*, 19(3), pp.632–642. Available at: http://dx.doi.org/10.1016/S2095-3119(19)62711-5.

Weßling, R. et al., (2014). Convergent targeting of a common host protein-network by pathogen effectors from three kingdoms of life. *Cell Host and Microbe*, 16(3), pp.364–375.

Westhof, E., Yusupov, M. & Yusupova, G., (2014). Recognition of Watson-Crick base pairs: Constraints and limits due to geometric selection and tautomerism. *F1000Prime Reports*, 6(April).

Willets, H, J. & Wong, J.A.L., (1980). Metabolic phenomena associated with virus infection in plants. *The Botanical Review*, 46(4), pp.101–165.

Winterbach, W. et al., (2013). Topology of molecular interaction networks. *BMC Systems Biology*, 7(1), pp.1–15. Available at: http://dx.doi.org/10.1186/1752-0509-7-90.

Wojcik, J., Boneca, I.G. & Legrain, P., (2002). Prediction, assessment and validation of protein interaction maps in bacteria. *Journal of Molecular Biology*, 323(4), pp.763–770.

Wu, J.Z. et al., (2017). Highly efficient mesophyll protoplast isolation and PEG-mediated transient gene expression for rapid and large-scale gene characterization in cassava (Manihot esculenta Crantz). *BMC Biotechnology*, 17(1), pp.1–8.

Wu, M. et al., (2021). Plant DNA polymerases α and δ mediate replication of geminiviruses. *Nature Communications*, 12(1). Available at: http://dx.doi.org/10.1038/s41467-021-23013-2.

Wu, M. et al., (2019). Transcriptional reprogramming caused by the geminivirus Tomato yellow leaf curl virus in local or systemic infections in Nicotiana benthamiana. *BMC Genomics*, 20(1), pp.1–17.

Wu, X., Gong, F. & Wang, W., (2014). Protein extraction from plant tissues for 2DE and its application in proteomic analysis. *Proteomics*, 14(6), pp.645–658.

Wu, X.R. et al., (2007). Altered expression of plant lysyl tRNA synthetase promotes tRNA misacylation and translational recoding of lysine. *Plant Journal*, 50(4), pp.627–636.

Wylie, S.J. et al., (2018). ICTV Virus Taxonomy Profile: Potyviridae. , (May), pp.352–354.

Xia, Z. et al., (2014). AFSM sequencing approach: a simple and rapid method for genome-wide SNP and methylation site discovery and genetic mapping. *Scientific Reports* , 4(7300). Available at: http://www.nature.com/articles/srep07300.

Xie, Q. et al., (1999). GRAB proteins, novel members of the NAC domain family, isolated by their interaction with a geminivirus protein. *Plant Molecular Biology*, 39(4), pp.647–656.

Xing, J. et al., (2021). Coordination of Phospholipid-Based Signaling and Membrane Trafficking in Plant Immunity. *Trends in Plant Science*, 26(4), pp.407–420. Available

at: https://doi.org/10.1016/j.tplants.2020.11.010.

Xiong, R. et al., (2009). Characterization and subcellular localization of an RNA silencing suppressor encoded by Rice stripe tenuivirus. *Virology*.

Xu, P. et al., (2008). Virus infection improves drought tolerance. *New Phytologist*, 180(4), pp.911–921.

Xu, Y. et al., (2020). Construction and characterization of a high-quality cDNA library of Cymbidium faberi suitable for yeast one- And two-hybrid assays. *BMC Biotechnology*, 20(1), pp.1–9.

Xu, Y. et al., (2021). The metabolic origins of non-photorespiratory CO2 release during photosynthesis: A metabolic flux analysis. *Plant Physiology*, 186(1), pp.297–314.

Yanfang, F. et al., (2013). High frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol.*, 31(9), pp.822–826.

Yang, H. et al., (2017). CRISPR/Cas9-mediated genome editing efficiently creates specific mutations at multiple loci using one sgRNA in Brassica napus. *Scientific Reports*, 7(1), pp.1–13. Available at: http://dx.doi.org/10.1038/s41598-017-07871-9.

Yang, X. et al., (2007). Functional Modulation of the Geminivirus AL2 Transcription Factor and Silencing Suppressor by Self-Interaction. *Journal of Virology*, 81(21), pp.11972–11981.

Yang, X. et al., (2020). Involvement of the chloroplast gene ferredoxin 1 in multiple responses of Nicotiana benthamiana to Potato virus X infection. *Journal of Experimental Botany*, 71(6), pp.2142–2156.

Yang, X. et al., (2011). Suppression of methylation-mediated transcriptional gene silencing by βc1-sahh protein interaction during geminivirus-betasatellite infection. *PLoS Pathogens*, 7(10).

Yao, D., Gonzales-Vigil, E. & Mansfield, S.D., (2020). Arabidopsis sucrose synthase localization indicates a primary role in sucrose translocation in phloem. *Journal of Experimental Botany*, 71(6), pp.1858–1869.

Ye, B. Bin et al., (2020). AP2/ERF Transcription Factors Integrate Age and Wound Signals for Root Regeneration. *The Plant cell*, 32(1), pp.226–241.

Yong Chung, H., Lacatus, G. & Sunter, G., (2014). Geminivirus AL2 protein induces expression of, and interacts with, a calmodulin-like gene, an endogenous regulator of gene silencing. *Virology*, 460–461(1), pp.108–118. Available at: http://dx.doi.org/10.1016/j.virol.2014.04.034.

Yoshida, K. & Hisabori, T., (2016). Adenine nucleotide-dependent and redox-independent control of mitochondrial malate dehydrogenase activity in Arabidopsis thaliana. *Biochimica et Biophysica Acta - Bioenergetics*, 1857(6), pp.810–818. Available at: http://dx.doi.org/10.1016/j.bbabio.2016.03.001.

Yunus, I.S. et al., (2015). Phosphatidic acid is a major phospholipid class in reproductive organs of arabidopsis thaliana. *Plant Signaling and Behavior*, 10(8), pp.3–6.

Zaidi, S.S. e. A. et al., (2016). Engineering Plants for Geminivirus Resistance with CRISPR/Cas9 System. *Trends in Plant Science*, 21(4), pp.279–281. Available at: http://dx.doi.org/10.1016/j.tplants.2016.01.023.

Zanini, A.A. et al., (2021). Cassava common mosaic virus infection causes alterations in chloroplast ultrastructure, function, and carbohydrate metabolism of cassava plants. *Plant Pathology*, 70(1), pp.195–205.

Zechmann, B., Müller, M. & Zellnig, G., (2003). Cytological modifications in zucchini

yellow mosaic virus (ZYMV)-infected Styrian pumpkin plants. *Archives of Virology*, 148(6), pp.1119–1133.

Zeier, J., (2013). New insights into the regulation of plant immunity by amino acid metabolic pathways. *Plant, Cell and Environment*, 36(12), pp.2085–2103.

Zerbini, F.M. et al., (2017). ICTV ICTV Virus Taxonomy Profile : Geminiviridae. *Journal of General Virology*, 98(2), pp.131–133. Available at: http://www.microbiologyresearch.org/docserver/fulltext/jgv/98/2/131_vir000738.pdf?expires=1519415641&id=id&accname=guest&checksum=E96F5FAEF6B1775E85500DFC4E621EAA%0Ahttp://www.ncbi.nlm.nih.gov/pubmed/28284245%0Ahttp://www.microbiologyresearch.org/content/.

Zhang, C., Hicks, G.R. & Raikhel, N. V., (2014). Plant vacuole morphology and vacuolar trafficking. *Frontiers in Plant Science*, 5(SEP), pp.1–9.

Zhang, J. et al., (2011). Factors that influence the yield and viability of protoplasts from Carica papaya L. *African Journal of Biotechnology*, 10(26), pp.5137–5142.

Zhang, L. et al., (2011). An AP2 domain-containing gene, ESE1, targeted by the ethylene signaling component EIN3 is important for the salt response in Arabidopsis. *Plant Physiology*, 157(2), pp.854–865.

Zhang, P. et al., (2005). Resistance to cassava mosaic disease in transgenic cassava expressing antisense RNAs targeting virus replication genes. *Plant Biotechnology Journal*.

Zhang, S.C., Ghosh, R. & Jeske, H., (2002). Subcellular targeting domains of Abutilon mosaic geminivirus movement protein BC1. *Archives of Virology*, 147(12), pp.2349–2363.

Zhang, Y. et al., (2020). Mechanistic insights into chromatin targeting by leukemic NUP98-PHF23 fusion. *Nature Communications*, 11(1), pp.1–10. Available at: http://dx.doi.org/10.1038/s41467-020-17098-4.

Zhang, Y. et al., (2015). Transcriptional response to petiole heat girdling in cassava. *Scientific Reports*, 5(8414).

Zhao, J. et al., (2013). Arabidopsis thaliana AHL family modulates hypocotyl growth redundantly by interacting with each other via the PPC/DUF296 domain. *Proceedings of the National Academy of Sciences of the United States of America*, 110(48), pp.1–10.

Zhao, J. et al., (2016). Chloroplast in plant-virus interaction. *Frontiers in Microbiology*, 7(OCT), pp.1–20.

Zhao, J. et al., (2019). Genome editing in large animals: Current status and future prospects. *National Science Review*, 6(3), pp.402–420.

Zhao, J., (2015). Phospholipase D and phosphatidic acid in plant defence response: From protein-protein and lipid-protein interactions to hormone signalling. *Journal of Experimental Botany*, 66(7), pp.1721–1736.

Zhao, W. et al., (2016). Different pathogenicities of Rice stripe virus from the insect vector and from viruliferous plants. *New Phytologist*, 210(1), pp.196–207.

Zheng, H., (2005). Disruptions of the Arabidopsis Enoyl-CoA Reductase Gene Reveal an Essential Role for Very-Long-Chain Fatty Acid Synthesis in Cell Expansion during Plant Morphogenesis. *the Plant Cell Online*, 17(5), pp.1467–1481. Available at: http://www.plantcell.org/cgi/doi/10.1105/tpc.104.030155.

Zheng, H. et al., (2018). MLK1 and MLK2 coordinate RGA and CCA1 activity to regulate hypocotyl elongation in Arabidopsis thaliana. *Plant Cell*, 30(1), pp.67–82.

Zheng, Q. et al., (2016). Nuclear distribution of eIF3g and its interacting nuclear proteins in breast cancer cells. *Molecular Medicine Reports*, 13(4), pp.2973–2980.

Zhou, X. et al., (1997). Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. *Journal of General Virology*, 78(8), pp.2101–2111.

Zhou, X., Liu, Y., et al., (1998). Four DNA-A variants among Pakistani isolates of cotton leaf curl virus and their affinities to DNA-A of geminivirus isolates from okra. *Journal of General Virology*, 79(4), pp.915–923.

Zhou, X., Robinson, D.J. & Harrison, B.D., (1998). Types of variation in DNA-A among isolates of East African cassava mosaic virus from Kenya, Malawi and Tanzania. *Journal of General Virology*, 79(11), pp.2835–2840.

Zhou, Y. et al., (2011a). Histone H3 Interacts and Colocalizes with the Nuclear Shuttle Protein and the Movement Protein of a Geminivirus. *Journal of Virology*, 85(22), pp.11821–11832. Available at: http://jvi.asm.org/cgi/doi/10.1128/JVI.00082-11.

Zhou, Y. et al., (2011b). Histone H3 Interacts and Colocalizes with the Nuclear Shuttle Protein and the Movement Protein of a Geminivirus. *Journal of Virology*, 85(22), pp.11821–11832.

Zhu, X., Gerstein, M. & Snyder, M., (2007). Getting connected: Analysis and principles of biological networks. *Genes and Development*.

Zhu, X., Zelmer, A. & Wellmann, S., (2017). Visualization of protein-protein interaction in nuclear and cytoplasmic fractions by co-immunoprecipitation and in situ proximity ligation assay. *Journal of Visualized Experiments*, 2017(119), pp.1–5.

Zhu, Y. et al., (2005). Characterization of a novel developmentally retarded mutant (drm1) associated with the autonomous flowering pathway in Arabidopsis. *Cell Research*, 15(2), pp.133–140.

Zhu, Z. & Lee, B., (2015). Friends or foes: New insights in jasmonate and ethylene co-actions. *Plant and Cell Physiology*.

Zinga, I. et al., (2013). Epidemiological assessment of cassava mosaic disease in Central African Republic reveals the importance of mixed viral infection and poor health of plant cuttings. *Crop Protection*.

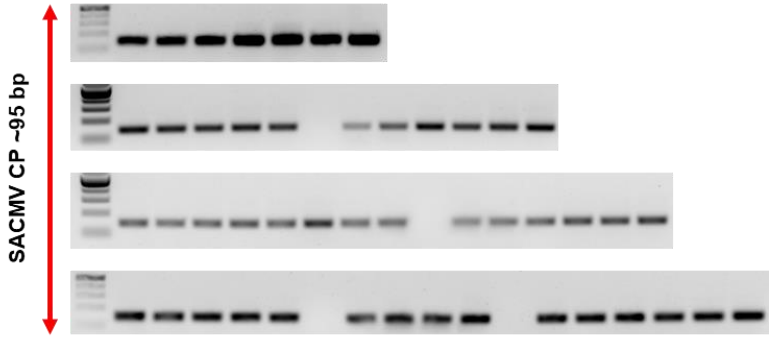
Zürcher, E. et al., (2016). Plant development regulated by cytokinin sinks. *Science*, 353(6303), pp.1027–1030.

APPENDICES

Appendix 2.1 Plant height (cm) and symptom score (0-3) of cassava TME3 plants infected with SACMV infectious clones. Measurements made from 0-32 days post-inoculation (dpi). Symptom scores, 0 (no symptoms/healthy), 1 (mild leaf distortion or mild chlorosis), 2 (moderate leaf distortion and chlorosis) and 3 (severe leaf distortion and chlorosis with reduction in leaf area)

Plant number	TME3-Treatment	0 dpi		32 dpi	
		Height (cm)	Symptom score	Height (cm)	Symptom score
1	TME3-SACMV	3.7	0	5	1
2	TME3-SACMV	4.8	0	9	1
3	TME3-SACMV	8	0	9	1
4	TME3-SACMV	3.5	0	9	2
5	TME3-SACMV	4	0	7	2
6	TME3-SACMV	3	0	11	2
7	TME3-SACMV	4.5	0	10	2
8	TME3-SACMV	6	0	8	2
9	TME3-SACMV	5	0	9	2
10	TME3-SACMV	4	0	6	2
11	TME3-SACMV	3.2	0	5	2
12	TME3-SACMV	4.8	0	12	2
13	TME3-SACMV	2.9	0	6	2
14	TME3-SACMV	4.6	0	6	2
15	TME3-SACMV	2.9	0	6	2
16	TME3-SACMV	11.5	0	16	2
17	TME3-SACMV	8	0	9	2
18	TME3-SACMV	4	0	5	2
19	TME3-SACMV	2.7	0	3	3
20	TME3-SACMV	5.5	0	10	3
21	TME3-SACMV	4	0	10	3
22	TME3-SACMV	5	0	7	3
23	TME3-SACMV	2.6	0	6	3
24	TME3-SACMV	5	0	6	3
25	TME3-SACMV	5.6	0	6	3
26	TME3-SACMV	6.5	0	8	3
27	TME3-SACMV	7.5	0	12	3
28	TME3-SACMV	6	0	12	3
29	TME3-SACMV	4.5	0	5	3
30	TME3-SACMV	4	0	5	3

Appendix 2.2 SACMV CP amplification for confirmation of SACMV mediated symptoms in infected cassava TME3 plants to be used for Y2H assays and Co-IP experiments.



Appendix 2.3 A summary of functional roles of identified BC1 interactors and the implications in BC1 mechanisms and SACMV infection.

Abbreviation	Functional role	Proposed potential role of interaction with BC1 during SACMV infection
LSH	Suppression of organ differentiation (Takeda et al., 2011). Contributes to lateral stability of cortical microtubules by binding to AT3G27960 (BioGRID:7240), a kinesin that contributes to lateral stability of cortical microtubules (Ganguly et al., 2020). Shown by Y2H to interact with viral protein Rep of SACMV (Rey, unpublished).	Hinder role of KRS1 in cell differentiation, increase mitosis and weaken cell wall to favour viral replication (Nagar et al., 2002). The binding of BC1 with LSH may interfere with LSH binding to AT3G27960 leading to an integrity, both factors that would favour viral replication and proliferation (Kong & Hanley-Bowdoin, 2002).
HSK	Phosphorylation of L-homoserine in threonine and methionine biosynthesis (Lee et al., 2005; Lee & Leustek, 1999) .	Interfere with amino acid biosynthesis pathways that are crucial to modulation of plant defence responses (Cantoni, 1952; Poudel et al., 2019).
ABR1	Regulation JA, ABA, and ethylene signalling (Cai et al., 2014; Ye et al., 2020; Zhang et al., 2011).	Interference of hormone signalling pathway regulation essential to plant function and defence (Cai et al., 2014; Ye et al., 2020; Zhang et al., 2011).
NAT	JAZ protein domains that regulate repression of jasmonate biosynthesis pathway (Ali & Baek, 2020; Ruan et al., 2019).	Interference of JA signalling pathway regulation. Geminiviruses interfere with jasmonate responses (Allie et al., 2014; Pierce & Rey, 2013; Rosas-Díaz et al., 2016)
PTEN2A	Localises to the nucleus and cytoplasm, anchored to microtubular vesicles via phosphatidylinositol 3 phosphate (PI3P) (Naguib et al., 2015). An electrostatic cell membrane binding mechanism facilitates its translocation (Chaozu et al., 2018). Key role in dephosphorylation (Pribat et al., 2012)	Facilitate microtubular movement of viral DNA, disrupts cell signalling and regulation of cellular processes and hinder immune signalling and dephosphorylation of BC1. Geminivirus movement proteins localise to microtubules (Frischmuth et al., 2007; Aberle et al., 2002; Krapp et al., 2017). Phosphorylation events regulate geminivirus MP functionality and abundance (Wege & Jeske, 1998; Kleinow et al., 2009).
SDG16	Post translational methylation of H3 and H4 histones (Chen et al., 2017). Gene activation and consequent transcription and translation (Black et al., 2012).	Inhibition of the methyltransferase role in gene activation and relegation of transcripts essential to immune responses. Activation of viral DNA transcription. Other geminiviruses proteins implicated in trimethylation of H3K4 of viral mini-chromosomes (Kushwaha et al., 2017) and interference with host antiviral methylation mechanisms favouring viral replication (Castillo-González et al., 2015; Sun et al., 2015; Raja et al., 2008).
SIP2	Transport carbohydrate sugars across the phloem (Sprenger & Keller, 2000). Localises to the plasmodesma (Haywood et al., 2002).	Interfere with inositol mediated signalling by altering allocation of raffinose which would impact the accumulation of inositol and its derivatives (Karner et al., 2004). Additional evidence of systemic localisation of SACMV/BC1 across the cell walls.
CKL1	Regulation ABA and GA hormone signalling pathways (Zheng et al., 2018; Chen et al., 2018). Suppression ABA signalling (Chen et al., 2018). Localises to and binds cortical microtubules (Ben-Nissan et al., 2008). Histone phosphorylation and protein phosphorylation (Zheng et al., 2018; Gaudet et al., 2011; Chen et al., 2018).	ABA and GA hormone dysregulation altering plant growth and development. Accumulation ABA linked to enhancing plant susceptibility (Alazem et al., 2014; Mohr & Cahill, 2007; Qiu et al., 2015) by activation of sucrose non-fermenting 1-related subfamily 2 protein kinases (SnRK2s) antagonistic (Sun et al., 2020; Chen et al., 2018) to plant geminivirus defence protein SnRK1 (Shen et al., 2009; Shen et al., 2011; Testerink et al., 2004). Indirect role in facilitating NSP transcription by enabling the ABA pathway (Sun et al., 2020). Facilitate viral DNA movement via microtubules . Condense viral DNA by phosphorylation histones of viral minichromosomes. Phosphorylation BC1.
SUS4	Catalyses the cleavage of sucrose to fructose and UDP-glucose (Bieniawska et al., 2007) and these products then feed into cell metabolic pathways (Stein & Granot, 2018). High sucrose synthase activity in leaf tissue increases the source to sink accumulation of starch in cassava tubers (Li et al., 2016). Efficient starch storage efficiency in cassava plants with a strong xylem architecture (Li et al., 2016) may be attributed to sucrose synthase role contributing to cellulose biosynthesis in carbon partitioning (Coleman et al., 2009).	Poor accumulation of starch in cassava tubers during geminivirus infection (Varma & Malathi, 2003). An interference of cellulose biosynthesis pathways and silencing of SUS4 contributed to features of plant stunting and chlorosis respectively (Choe et al., 2021; Yao et al., 2020) features of geminivirus infection.

NUP98	Export of mRNA and RNA across the nuclear membrane into the cytoplasm, import of proteins into the nucleus (Gaudet et al., 2011; Gallemí et al., 2016; Powers et al., 1997). gene activation and by inducing trimethylation of H3K4 methyltransferase thus resulting in increased gene expression (Zhang et al., 2020; Sump & Brickner, 2017; Cruz et al., 2018; Franks et al., 2017).	Movement viral DNA and BC1. Interfere recruitment of effective immune signalling by the plant for defence response. Promotes viral replication and translocation of viral genomes and proteins (De Jesús-González et al., 2021; Lau & Weber, 2020; Ebina et al., 2004).
ACLB2	Cytosolic enzyme involved in the primary synthesis of acetyl-CoA, a primary precursor for the synthesis of C18-C24 fatty acid biosynthesis (Fatland et al., 2002).	Interfere with the initial onset of plant defence dependent on fatty acid biosynthesis, cell architecture, energy utilisation and metabolic signalling (Weber, 2002; Lim et al., 2017; Laxalt & Munnik, 2002).
GLDP2	Essential Methionine cycle (MTC) enzyme involved in photorespiration and non-photorespiratory pathways (TCA cycle reactions, the G6P/OPP shunt, fatty acid biosynthesis and starch and sucrose synthesis (Engel et al., 2007 Xu et al., 2021; Mäkinen & Swarnalok, 2019).	Interference of plant metabolites relevant to antiviral defence. Methionine cycle (MTC) enzymes are targeted by viral proteins in the negation of host antiviral responses during geminivirus and potyvirus infection (Mäkinen & Swarnalok, 2019).
PUP1	Transmembrane protein that functions in proton coupled transport of purine nucleobases and their derivatives (adenine, trans-zeatin, pyridoxine (vitamin B6), caffeine, adenosine, cytokinin) via the xylem and phloem to and from shoot tissues (Bürkle et al., 2003; Szydlowski et al., 2013; Gillissen et al., 2000).	Interference of PUP1 activity resulting in an accumulation of bioactive cytokinins. An upregulation of cytokinin responsive genes is induced by pathogenicity proteins (AL2/C2) of begomo- and curto-viruses via interaction with adenosine kinase (ADK), impeding its ability to phosphorylate cytokinin and thus increase bioavailable cytokinin (Baliji et al., 2010). PUP1 identified as an NSP and MP interacting protein in geminivirus cabbage leaf curl virus (CabLCV) (Gouveia-Mageste et al., 2020).
PsAD	Photosynthetic protein that forms complexes with ferredoxin a key electron acceptor in photosynthesis and as a key component in the biosynthesis of pigments and assimilation of sulphur and nitrogen (Hanke & Mulo, 2013). PsaD is both a protein and mRNA binding protein (Dutta et al., 2014; Bach-Pages et al., 2020).	Phenotypic deviations (stunting, size reduction and leaf abnormalities) manifested during SACMV infection due to alterations in carbon metabolism. (Yang et al., 2020)
KRS1	Nucleotide binding ligase involved mainly in the ATP dependent transfer of a cognate and non-cognate transfer of amino acids for translation of mRNA for protein synthesis (Wu et al., 2007; Gaudet et al., 2011).	May interfere with efficiency and accuracy of host protein synthesis. Viral protein translation. The non-cognate nature of other tRNA ligases is a common occurrence (Plant et al., 2007; Wu et al., 2007) and KRS1 may attribute to the occurrence of mutations in host proteins arising from SACMV infection (Chatukuta & Rey, 2020). Geminiviruses require host machinery for replication and translation (Hanley-Bowdoin et al., 2013; Ramesh et al., 2017) and BC1 may manipulate host tRNA ligase for viral protein translation.
mMDH1	Operates in the TCA cycle and converts malate to oxaloacetate in a reversible manner, thus providing precursors for acetyl CoA, carbon dioxide and pyruvate and regulates carbon partitioning (Yoshida & Hisabori, 2016; Tomaz et al., 2010; Hüdig et al., 2014; Lindén et al., 2016; Tan et al., 2010; Kawamura & Uemura, 2003; Sarry et al., 2006; Grant et al., 2006).	Competitively interfere with mMDH1 function and contribute to symptom progression (Hou et al., 2000; Kleinow et al., 2009) and a dysregulation of carbon assimilation noted in virus compatible studies (Handford & Carr, 2007). It is proposed that mMDH1 interaction with BC1 could potentially alter the methionine cycle and alter amino acid biosynthesis, secondary metabolites and methylation products involved in regulation of gene expression (Mäkinen & Swarnalok, 2019).
H4	Geminiviruses form minichromosomes with host histones (Pilartz & Jeske, 1992; Jeske, 2009)	For minichromosome assembly of the viral genome and point of post translational histone modifications that impact viral genome packaging and activation (refer CKL1, SDG16, HSP90 and CDC48). Recruitment of histone H3 and complex interaction with NSP and MP facilitate compaction of geminiviral DNA to facilitate movement through plasmodesmata and nuclear pores (Zhou et al., 2011).
HSP90	HSP90-mediated ubiquitin/proteasome plays a specific role in degradation of viral protein (Moshe et al., 2016). Molecular chaperones in facilitating ubiquitin mediated proteolysis (Imai et al., 2003; Schuberth & Buchberger, 2005) and play key roles in the accumulation of cyclins for initiation of the cell cycle (Moreno et al., 2019; Parisi et al., 2018).	Role in contributing to symptoms by competitively binding to and thus inhibiting HSP90 functions. BC1 is a predominantly disordered protein (Nankoo, 2018). The molecular chaperone role essential for viral protein folding, stability and nucleocapsid assembly as noted in other plant and animal virus studies (Santoro et al., 2009; Ujino et al., 2009; Li et al., 2020).

CDC48	Molecular chaperones in facilitating ubiquitin mediated proteolysis (Imai et al., 2003; Schuberth & Buchberger, 2005) and play key roles in the accumulation of cyclins for initiation of the cell cycle (Moreno et al., 2019; Parisi et al., 2018). CDC48 complex releases and prevents degradation of ubiquitinated Cln3 and G1 and D1 cyclins promoting their nuclear localisation and accumulation to induce the start of the cell cycle (Parisi et al., 2018). Euchromatisation of DNA thus enhancing transcription and protein synthesis (Mérai et al., 2014; Parisi et al., 2018).	Decondense viral DNA. Alter cell cycle to facilitate conditions for geminiviral replication, transcription, and protein synthesis (Bruce et al., 2011; Hipp et al., 2014). ERAD mediated movement of viral proteins and viral DNA. Interference with the ubiquination/proteasome pathways contributing to susceptibility/infection.

Appendix 3.1 Plant height (cm) and symptom score (0-3) of cassava TME3 plants infected with SACMV infectious clones. Measurements made from 0-32 days post-inoculation (dpi). Symptom scores, 0 (no symptoms/healthy), 2 (moderate leaf distortion and chlorosis) and 3 (severe leaf distortion and chlorosis with reduction in leaf area). Scores for mock infected samples remained at 0 (not shown).

Plant number	TME3-Treatment	Score: 0 dpi	Score: 32 dpi
1	TME3-SACMV	0	3
2	TME3-SACMV	0	3
3	TME3-SACMV	0	3
4	TME3-SACMV	0	3
5	TME3-SACMV	0	3
6	TME3-SACMV	0	3
7	TME3-SACMV	0	3
8	TME3-SACMV	0	3
9	TME3-SACMV	0	3
10	TME3-SACMV	0	3
11	TME3-SACMV	0	3
12	TME3-SACMV	0	3
13	TME3-SACMV	0	2
14	TME3-SACMV	0	2
15	TME3-SACMV	0	2
16	TME3-SACMV	0	2
17	TME3-SACMV	0	2
18	TME3-SACMV	0	2
19	TME3-SACMV	0	2
20	TME3-SACMV	0	2
21	TME3-SACMV	0	2
22	TME3-SACMV	0	2
23	TME3-SACMV	0	2
24	TME3-SACMV	0	2
25	TME3-SACMV	0	2
26	TME3-SACMV	0	2

27	TME3-SACMV	0	2
28	TME3-SACMV	0	2
29	TME3-SACMV	0	2
30	TME3-SACMV	0	2

Appendix 3.2 KEGG pathway enrichment analysis of all proteins identified from SDS-PAGE in-gel digest and mass spectrometry of proteins extracted from mock infected leaf tissue samples.

Enrichment FDR	Genes in list	Total genes	Functional Category
3.97E-100	403	1793	Metabolic pathways
5.61E-74	252	962	Biosynthesis of secondary metabolites
1.05E-62	111	227	Carbon metabolism
7.55E-57	133	362	Microbial metabolism in diverse environments
6.90E-44	95	239	Biosynthesis of amino acids
4.02E-27	38	60	Carbon fixation in photosynthetic organisms
7.75E-22	33	57	Citrate cycle (TCA cycle)
8.32E-22	38	77	Pyruvate metabolism
1.01E-20	32	57	Aminoacyl-tRNA biosynthesis
1.03E-20	43	105	Glycolysis / Gluconeogenesis
4.07E-19	72	292	Ribosome
6.54E-18	29	55	Proteasome
8.83E-17	31	68	Glycine, serine, and threonine metabolism
1.75E-15	27	56	Glyoxylate and dicarboxylate metabolism
2.17E-12	27	71	2-Oxocarboxylic acid metabolism
2.50E-11	22	52	Pentose phosphate pathway
3.29E-11	21	48	Porphyrin and chlorophyll metabolism
1.15E-10	20	46	Alanine, aspartate, and glutamate metabolism
1.27E-09	28	97	Cysteine and methionine metabolism
1.30E-08	23	75	Photosynthesis
1.79E-08	18	48	Fructose and mannose metabolism
4.98E-08	25	93	Glutathione metabolism
6.94E-08	29	122	Amino sugar and nucleotide sugar metabolism
5.34E-07	10	18	Selenocompound metabolism
5.93E-07	17	53	Phenylalanine, tyrosine, and tryptophan biosynthesis
6.49E-07	12	27	Propanoate metabolism
1.62E-06	12	29	Fatty acid biosynthesis
3.46E-06	9	17	Butanoate metabolism
3.69E-06	19	73	Arginine and proline metabolism
7.90E-06	10	23	Valine, leucine, and isoleucine biosynthesis
1.06E-05	16	58	Terpenoid backbone biosynthesis
1.16E-05	11	29	Pantothenate and CoA biosynthesis
1.43E-05	18	73	Fatty acid metabolism

Enrichment FDR	Genes in list	Total genes	Functional Category
1.60E-05	9	20	One carbon pool by folate
4.99E-05	27	151	Purine metabolism
8.00E-05	12	41	Tyrosine metabolism
9.52E-05	22	115	Pyrimidine metabolism
0.000146437	14	57	Cyanoamino acid metabolism
0.000177514	11	38	Sulphur metabolism
0.000193555	12	45	Valine, leucine, and isoleucine degradation
0.000193555	7	16	Lysine biosynthesis
0.000275819	11	40	Nitrogen metabolism
0.000362983	12	48	Galactose metabolism
0.000508027	5	9	C5-Branched dibasic acid metabolism
0.000671649	6	14	Biotin metabolism
0.000892402	7	20	Photosynthesis - antenna proteins
0.000934356	10	39	Ascorbate and aldarate metabolism
0.001122771	10	40	Fatty acid degradation
0.001122771	30	213	Protein processing in endoplasmic reticulum
0.001168338	27	185	Starch and sucrose metabolism
0.001267476	15	79	Peroxisome
0.001299373	10	41	beta-Alanine metabolism
0.002084331	24	164	RNA transport
0.002194277	22	146	Oxidative phosphorylation
0.002581146	22	148	Phenylpropanoid biosynthesis
0.002787789	8	31	alpha-Linolenic acid metabolism
0.003111975	5	13	Nicotinate and nicotinamide metabolism
0.011448318	5	17	Histidine metabolism
0.019326866	7	34	Tropane, piperidine and pyridine alkaloid biosynthesis
0.021308016	4	13	Vitamin B6 metabolism
0.026830741	15	110	Phenylalanine metabolism
0.027480825	5	21	Flavonoid biosynthesis
0.039538451	5	23	Isoquinoline alkaloid biosynthesis
0.047335192	2	4	Synthesis and degradation of ketone bodies
0.049108853	11	79	Phagosome
0.049768283	8	51	Glycerolipid metabolism
0.049768283	3	10	Thiamine metabolism
0.049768283	3	10	Degradation of aromatic compounds
0.002194277	22	146	Oxidative phosphorylation
0.002581146	22	148	Phenylpropanoid biosynthesis
0.002787789	8	31	alpha-Linolenic acid metabolism
0.003111975	5	13	Nicotinate and nicotinamide metabolism
0.011448318	5	17	Histidine metabolism
0.019326866	7	34	Tropane, piperidine and pyridine alkaloid biosynthesis
0.021308016	4	13	Vitamin B6 metabolism

Enrichment FDR	Genes in list	Total genes	Functional Category
0.026830741	15	110	Phenylalanine metabolism
0.027480825	5	21	Flavonoid biosynthesis
0.039538451	5	23	Isoquinoline alkaloid biosynthesis
0.047335192	2	4	Synthesis and degradation of ketone bodies
0.049108853	11	79	Phagosome
0.049768283	8	51	Glycerolipid metabolism
0.049768283	3	10	Thiamine metabolism
0.049768283	3	10	Degradation of aromatic compounds

Appendix 3.3 KEGG pathway enrichment analysis of all proteins identified from SDS-PAGE in-gel digest and mass spectrometry of proteins extracted from SACMV infected leaf tissue samples.

Enrichment FDR	Genes in list	Total genes	Functional Category
3.59E-92	404	1793	Metabolic pathways
5.07E-62	242	962	Biosynthesis of secondary metabolites
6.59E-57	108	227	Carbon metabolism
1.25E-49	128	362	Microbial metabolism in diverse environments
7.36E-41	94	239	Biosynthesis of amino acids
5.44E-26	85	292	Ribosome
9.97E-24	36	60	Carbon fixation in photosynthetic organisms
6.16E-20	32	57	Citrate cycle (TCA cycle)
6.16E-20	37	77	Pyruvate metabolism
9.07E-20	43	105	Glycolysis / Gluconeogenesis
2.37E-18	30	55	Proteasome
1.09E-16	29	57	Aminoacyl-tRNA biosynthesis
6.59E-16	28	56	Glyoxylate and dicarboxylate metabolism
3.93E-15	30	68	Glycine, serine, and threonine metabolism
1.33E-13	29	71	2-Oxocarboxylic acid metabolism
3.49E-12	22	46	Alanine, aspartate, and glutamate metabolism
7.10E-11	22	52	Pentose phosphate pathway
8.99E-11	21	48	Porphyrin and chlorophyll metabolism
7.18E-09	24	75	Photosynthesis
7.46E-09	27	93	Glutathione metabolism
1.91E-08	27	97	Cysteine and methionine metabolism
1.91E-08	23	73	Arginine and proline metabolism
2.36E-07	29	122	Amino sugar and nucleotide sugar metabolism
2.48E-07	11	20	One carbon pool by folate
2.52E-07	17	48	Fructose and mannose metabolism
5.54E-07	16	45	Valine, leucine, and isoleucine degradation
1.14E-06	12	27	Propanoate metabolism
1.37E-06	11	23	Valine, leucine, and isoleucine biosynthesis

1.90E-06	20	73	Fatty acid metabolism
2.58E-06	9	16	Lysine biosynthesis
4.94E-06	9	17	Butanoate metabolism
1.53E-05	13	39	Ascorbate and aldarate metabolism
2.04E-05	13	40	Fatty acid degradation
9.03E-05	8	18	Selenocompound metabolism
0.000107585	12	40	Nitrogen metabolism
0.000122023	10	29	Fatty acid biosynthesis
0.000125003	27	151	Purine metabolism
0.000156746	34	213	Protein processing in endoplasmic reticulum
0.000174453	26	146	Oxidative phosphorylation
0.00019158	8	20	Photosynthesis - antenna proteins
0.000192484	28	164	RNA transport
0.000240008	14	57	Cyanoamino acid metabolism
0.000286859	14	58	Terpenoid backbone biosynthesis
0.000543137	11	41	Tyrosine metabolism
0.000588837	9	29	Pantothenate and CoA biosynthesis
0.000598839	29	185	Starch and sucrose metabolism
0.000610127	5	9	C5-Branched dibasic acid metabolism
0.000776253	16	79	Peroxisome
0.001100431	10	38	Sulphur metabolism
0.001381532	12	53	Phenylalanine, tyrosine, and tryptophan biosynthesis
0.002018602	10	41	beta-Alanine metabolism
0.004228331	8	31	alpha-Linolenic acid metabolism
0.006039573	14	79	Phagosome
0.006041216	5	14	Biotin metabolism
0.0068182	18	115	Pyrimidine metabolism
0.013115731	6	23	Isoquinoline alkaloid biosynthesis
0.01450322	5	17	Histidine metabolism
0.019773819	9	48	Galactose metabolism
0.025377805	7	34	Tropane, piperidine and pyridine alkaloid biosynthesis
0.028079829	9	51	Glycerolipid metabolism
0.032684584	3	8	Riboflavin metabolism
0.032684584	7	36	Biosynthesis of unsaturated fatty acids
0.037070649	19	148	Phenylpropanoid biosynthesis
0.047418372	5	23	Lysine degradation

Appendix 3.4 MP interacting cassava TME3 proteins identified by mass spectrometry and analysed by SAINT software.

Treatment	AvgSpec	SaintScore	Fold Change	BFDR	Interactor	Description
BC1-Ab	6.00	0.00	1.12	0.32	A4QJU0	ATP synthase subunit beta
BC1-Ab	20.00	0.00	0.39	0.33	A4QKT9	Ribulose bisphosphate carboxylase large chain (RuBisCO large subunit)
BC1-Ab	4.00	0.00	1.20	0.29	O48549	40S ribosomal protein S6-1
BC1-Ab	2.00	0.00	0.75	0.33	O49506	Peroxisomal (S)-2-hydroxy-acid oxidase GLO5
BC1-Ab	4.00	0.00	0.71	0.33	P10871	Ribulose bisphosphate carboxylase/oxygenase activase
BC1-Ab	11.00	0.00	0.70	0.33	P10896	Ribulose bisphosphate carboxylase/oxygenase activase
BC1-Ab	10.00	1.00	100.00	0.00	P14981	Movement protein BC1 (33.7 kDa protein) (Movement protein BL1)
BC1-Ab	3.00	0.72	3.00	0.03	P21238	Chaperonin 60 subunit alpha 1
BC1-Ab	4.00	0.01	1.20	0.27	P21240	Chaperonin 60 subunit beta 1
BC1-Ab	2.00	0.00	0.75	0.33	P25851	Fructose-1,6-bisphosphatase 1
BC1-Ab	4.00	1.00	40.00	0.00	P25856	Glyceraldehyde-3-phosphate dehydrogenase GAPA1
BC1-Ab	3.00	0.30	1.80	0.16	P25873	50S ribosomal protein L15
BC1-Ab	3.00	0.00	1.00	0.33	P34788	40S ribosomal protein S18
BC1-Ab	1.00	0.00	1.00	0.33	P41127	60S ribosomal protein L13-1 (Protein BBC1 homolog)
BC1-Ab	6.00	0.00	0.82	0.33	P50318	Phosphoglycerate kinase 2
BC1-Ab	2.00	0.47	2.00	0.13	P53496	Actin-11
BC1-Ab	2.00	0.67	3.00	0.05	P53497	Actin-12
BC1-Ab	4.00	0.93	6.00	0.01	P61841	30S ribosomal protein S7
BC1-Ab	2.00	0.67	3.00	0.05	Q8LCL3	60S ribosomal protein L27-2
BC1-Ab	3.00	1.00	30.00	0.00	Q8LD46	60S ribosomal protein L23a-1 (AtRPL23A-1)
BC1-Ab	2.00	0.67	3.00	0.05	Q8RXX5	50S ribosomal protein L19-2
BC1-Ab	2.00	0.00	0.33	0.33	Q8W4H7	Elongation factor 1-alpha 2 (EF-1-alpha 2)
BC1-Ab	2.00	0.67	3.00	0.05	Q9FIF3	40S ribosomal protein S8-2
BC1-Ab	2.00	0.67	3.00	0.05	Q9FJX2	60S ribosomal protein L26-2
BC1-Ab	2.00	0.16	1.50	0.21	Q9FZ47	ACT domain-containing protein ACR11
BC1-Ab	2.00	0.00	0.46	0.33	Q9LRS0	(S)-2-hydroxy-acid oxidase GLO2
BC1-Ab	2.00	0.99	20.00	0.00	Q9LXG1	40S ribosomal protein S9-1
BC1-Ab	2.00	0.99	20.00	0.00	Q9M385	50S ribosomal protein L17
BC1-Ab	3.00	1.00	30.00	0.00	Q9SF40	60S ribosomal protein L4-1 (L1)
BC1-Ab	2.00	0.00	0.75	0.33	Q9SIH0	40S ribosomal protein S14-1
BC1-Ab	3.00	0.53	2.25	0.12	Q9SZJ5	Serine hydroxymethyltransferase 1
BC1-Ab	2.00	0.67	3.00	0.05	Q9T029	40S ribosomal protein S25-4
SACMV-Ab	51.00	0.00	0.99	0.33	A4QKT9	Ribulose bisphosphate carboxylase large chain (RuBisCO large subunit)
SACMV-Ab	11.00	1.00	110.00	0.00	C0Z361	Chaperonin 60 subunit beta 3
SACMV-Ab	14.00	1.00	140.00	0.00	O03042	Ribulose bisphosphate carboxylase large chain (RuBisCO large subunit)
SACMV-Ab	8.00	1.00	80.00	0.00	O80988	Glycine dehydrogenase (decarboxylating) 2
SACMV-Ab	1.00	0.00	0.18	0.33	P10871	Ribulose bisphosphate carboxylase/oxygenase activase
SACMV-Ab	2.00	0.00	0.13	0.33	P10896	Ribulose bisphosphate carboxylase/oxygenase activase

Treatment	AvgSpec	SaintScore	Fold Change	BFDR	Interactor	Description
SACMV-Ab	6.00	1.00	60.00	0.00	P14981	Movement protein BC1 (33.7 kDa protein) (Movement protein BL1)
SACMV-Ab	12.00	1.00	12.00	0.00	P21238	Chaperonin 60 subunit alpha 1
SACMV-Ab	4.00	0.01	1.20	0.27	P21240	Chaperonin 60 subunit beta 1
SACMV-Ab	5.00	0.12	1.88	0.22	P25851	Fructose-1,6-bisphosphatase 1
SACMV-Ab	8.00	1.00	80.00	0.00	P25856	Glyceraldehyde-3-phosphate dehydrogenase GAPA1
SACMV-Ab	7.00	0.00	0.78	0.33	P25857	Glyceraldehyde-3-phosphate dehydrogenase GAPB
SACMV-Ab	9.00	0.00	1.59	0.30	P25858	Glyceraldehyde-3-phosphate dehydrogenase GAPC1
SACMV-Ab	2.00	0.10	1.20	0.23	P25873	50S ribosomal protein L15
SACMV-Ab	9.00	0.89	3.86	0.01	P29197	Chaperonin CPN60
SACMV-Ab	2.00	0.00	0.67	0.33	P34788	40S ribosomal protein S18
SACMV-Ab	1.00	0.00	1.00	0.33	P41127	60S ribosomal protein L13-
SACMV-Ab	2.00	0.00	0.75	0.33	P42737	Beta carbonic anhydrase 2
SACMV-Ab	5.00	0.00	0.88	0.33	P53492	Actin-7 (Actin-2)
SACMV-Ab	2.00	0.47	2.00	0.13	P53496	Actin-11
SACMV-Ab	2.00	0.67	3.00	0.05	P55228	Glucose-1-phosphate adenylyltransferase small subunit
SACMV-Ab	6.00	1.00	60.00	0.00	P55737	Heat shock protein 90-2 (AtHSP90.2)
SACMV-Ab	2.00	0.67	3.00	0.05	P61841	30S ribosomal protein S7
SACMV-Ab	3.00	0.72	3.00	0.03	Q56WH1	Tubulin alpha-3 chain
SACMV-Ab	2.00	0.47	2.00	0.13	Q56YA5	Serine-glyoxylate aminotransferase
SACMV-Ab	5.00	0.78	3.00	0.01	Q8RWV0	Transketolase-1
SACMV-Ab	2.00	0.67	3.00	0.05	Q8RXX5	50S ribosomal protein L19-2
SACMV-Ab	9.00	0.00	1.50	0.31	Q8W4H7	Elongation factor 1-alpha 2 (EF-1-alpha 2)
SACMV-Ab	1.00	0.00	10.00	0.33	Q93VG5	40S ribosomal protein S8-1
SACMV-Ab	2.00	0.24	1.50	0.17	Q944G9	Fructose-bisphosphate aldolase 2
SACMV-Ab	4.00	0.93	6.00	0.01	Q9ASR1	Elongation factor 2 (EF-2)
SACMV-Ab	3.00	1.00	30.00	0.00	Q9FGX1	ATP-citrate synthase beta chain protein 2
SACMV-Ab	6.00	0.99	9.00	0.00	Q9FJX2	60S ribosomal protein L26-2
SACMV-Ab	10.00	1.00	15.00	0.00	Q9LD57	Phosphoglycerate kinase 1
SACMV-Ab	7.00	0.20	2.10	0.20	Q9LJE4	Chaperonin 60 subunit beta 2
SACMV-Ab	3.00	0.00	0.26	0.33	Q9LPW0	Glyceraldehyde-3-phosphate dehydrogenase GAPA2
SACMV-Ab	4.00	0.00	0.92	0.33	Q9LRS0	(S)-2-hydroxy-acid oxidase GLO2
SACMV-Ab	3.00	1.00	30.00	0.00	Q9LXL5	Sucrose synthase 4 (AtSUS4)
SACMV-Ab	4.00	1.00	40.00	0.00	Q9LZF6	Cell division control protein 48 homolog E (AtCDC48e)
SACMV-Ab	3.00	1.00	30.00	0.00	Q9SF40	60S ribosomal protein L4-1 (L1)
SACMV-Ab	4.00	0.93	6.00	0.01	Q9SGC1	Probable phosphoglucomutase
SACMV-Ab	3.00	0.08	1.50	0.24	Q9SXJ7	Chaperone protein ClpC2
SACMV-Ab	4.00	0.74	3.00	0.02	Q9SZJ5	Serine hydroxymethyltransferase 1
SACMV-Ab	3.00	1.00	30.00	0.00	Q9ZP06	Malate dehydrogenase 1
nBC1-Ab	27.00	0.00	0.52	0.33	A4QKT9	Ribulose bisphosphate carboxylase large chain (RuBisCO large subunit)

Treatment	AvgSpec	SaintScore	Fold Change	BFDR	Interactor	Description
nBC1-Ab	2.00	0.99	20.00	0.00	C0Z361	Chaperonin 60 subunit beta 3
nBC1-Ab	8.00	1.00	80.00	0.00	O03042	Ribulose bisphosphate carboxylase large chain (RuBisCO large subunit)
nBC1-Ab	3.00	0.01	1.12	0.26	O49506	Peroxisomal (S)-2-hydroxy-acid oxidase GLO5
nBC1-Ab	2.00	0.00	1.00	0.33	P0CJ47	Actin-3
nBC1-Ab	5.00	0.00	0.88	0.33	P10871	Ribulose bisphosphate carboxylase/oxygenase activase
nBC1-Ab	3.00	0.00	0.19	0.33	P10896	Ribulose bisphosphate carboxylase/oxygenase activase
nBC1-Ab	2.00	0.99	20.00	0.00	P14981	Movement protein BC1 (33.7 kDa protein) (Movement protein BL1)
nBC1-Ab	2.00	0.67	3.00	0.05	P19366	ATP synthase subunit beta
nBC1-Ab	2.00	0.47	2.00	0.13	P21238	Chaperonin 60 subunit alpha 1
nBC1-Ab	2.00	0.67	3.00	0.05	P25819	Catalase-2 (EC 1.11.1.6)
nBC1-Ab	3.00	0.00	0.33	0.33	P25857	Glyceraldehyde-3-phosphate dehydrogenase GAPB
nBC1-Ab	3.00	0.72	3.00	0.03	P53496	Actin-11
nBC1-Ab	2.00	0.67	3.00	0.05	P53497	Actin-12
nBC1-Ab	2.00	0.99	20.00	0.00	P56757	ATP synthase subunit alpha
nBC1-Ab	2.00	0.99	20.00	0.00	P59259	Histone H4
nBC1-Ab	6.00	0.00	1.00	0.33	Q8W4H7	Elongation factor 1-alpha 2 (EF-1-alpha 2)
nBC1-Ab	2.00	0.99	20.00	0.00	Q96528	Catalase-1 (EC 1.11.1.6)
nBC1-Ab	3.00	0.00	0.69	0.33	Q9LRS0	(S)-2-hydroxy-acid oxidase GLO2
nBC1-Ab	2.00	0.24	1.50	0.17	Q9SGA6	40S ribosomal protein S19-1
nBC1-Ab	2.00	0.24	1.50	0.17	Q9SZJ5	Serine hydroxymethyltransferase 1

Appendix 3.5 MP interacting cassava TME3 proteins identified by mass spectrometry and analysed by Y2H Scores software.

Treatment	Enrichment score	Specificity score	Sum scores	Borda scores	Borda quantile	Interactor	Description
BC1-Ab	0.88	0.34	1.22	4.96	0.17	A4QJU0	ATP synthase subunit beta
SACMV-Ab	0.00	0.06	0.06	1.20	0.90	A4QJU0	ATP synthase subunit beta
nSACMV-Ab	0.91	0.62	1.54	15.56	0.04	A4QJU0	ATP synthase subunit beta
BC1-Ab	0.94	0.00	0.94	2.00	0.47	A4QKT9	Ribulose bisphosphate carboxylase large chain (RuBisCO large subunit)
SACMV-Ab	0.00	0.73	0.73	2.43	0.35	A4QKT9	Ribulose bisphosphate carboxylase large chain (RuBisCO large subunit)
nSACMV-Ab	0.00	0.49	0.49	2.03	0.44	A4QKT9	Ribulose bisphosphate carboxylase large chain (RuBisCO large subunit)
nBC1-Ab	0.00	0.23	0.23	1.67	0.63	A4QKT9	Ribulose bisphosphate carboxylase large chain (RuBisCO large subunit)
BC1-Ab	0.22	0.00	0.22	1.26	0.85	C0Z361	Chaperonin 60 subunit beta 3
SACMV-Ab	0.91	0.68	1.59	19.31	0.03	C0Z361	Chaperonin 60 subunit beta 3
nSACMV-Ab	0.42	0.02	0.43	1.42	0.74	C0Z361	Chaperonin 60 subunit beta 3
nBC1-Ab	0.79	0.33	1.13	4.18	0.19	C0Z361	Chaperonin 60 subunit beta 3
BC1-Ab	0.20	0.00	0.20	1.25	0.86	O03042	Ribulose bisphosphate carboxylase large chain (RuBisCO large subunit)
SACMV-Ab	0.93	0.67	1.60	21.54	0.02	O03042	Ribulose bisphosphate carboxylase large chain (RuBisCO large subunit)
nSACMV-Ab	0.40	0.02	0.42	1.40	0.75	O03042	Ribulose bisphosphate carboxylase large chain (RuBisCO large subunit)
nBC1-Ab	0.93	0.42	1.35	6.51	0.11	O03042	Ribulose bisphosphate carboxylase large chain (RuBisCO large subunit)
BC1-Ab	0.85	0.34	1.19	4.75	0.17	O48549	40S ribosomal protein S6-1
SACMV-Ab	0.04	0.06	0.10	1.21	0.89	O48549	40S ribosomal protein S6-1
nSACMV-Ab	0.00	0.32	0.32	1.76	0.59	O48549	40S ribosomal protein S6-1
BC1-Ab	0.71	0.16	0.87	2.95	0.30	O49506	Peroxisomal (S)-2-hydroxy-acid oxidase GLO5
SACMV-Ab	0.05	0.00	0.05	1.16	0.94	O49506	Peroxisomal (S)-2-hydroxy-acid oxidase GLO5
nSACMV-Ab	0.02	0.62	0.64	2.34	0.37	O49506	Peroxisomal (S)-2-hydroxy-acid oxidase GLO5
nBC1-Ab	0.00	0.40	0.40	1.83	0.55	O49506	Peroxisomal (S)-2-hydroxy-acid oxidase GLO5
BC1-Ab	0.27	0.10	0.37	1.60	0.66	O80988	Glycine dehydrogenase (decarboxylating) 2
SACMV-Ab	0.89	0.58	1.47	10.57	0.08	O80988	Glycine dehydrogenase (decarboxylating) 2
nSACMV-Ab	0.46	0.12	0.58	1.94	0.50	O80988	Glycine dehydrogenase (decarboxylating) 2
BC1-Ab	0.26	0.12	0.39	1.71	0.61	P0CJ46	Actin-1
SACMV-Ab	0.00	0.06	0.06	1.20	0.90	P0CJ46	Actin-1
nSACMV-Ab	0.59	0.38	0.97	3.29	0.26	P0CJ46	Actin-1
BC1-Ab	0.18	0.03	0.21	1.26	0.85	P0CJ47	Actin-3
nSACMV-Ab	0.83	0.61	1.44	9.49	0.08	P0CJ47	Actin-3
nBC1-Ab	0.80	0.32	1.12	4.03	0.21	P0CJ47	Actin-3

BC1-Ab	0.85	0.24	1.09	4.06	0.20	P10871	Ribulose bisphosphate carboxylase/oxygenase activase
nSACMV-Ab	0.96	0.71	1.66	40.00	0.01	P10871	Ribulose bisphosphate carboxylase/oxygenase activase
nBC1-Ab	0.00	0.42	0.42	1.85	0.54	P10871	Ribulose bisphosphate carboxylase/oxygenase activase
BC1-Ab	0.92	0.49	1.41	8.75	0.09	P10896	Ribulose bisphosphate carboxylase/oxygenase activase
nSACMV-Ab	1.00	0.75	1.74	280.00	0.00	P10896	Ribulose bisphosphate carboxylase/oxygenase activase
nBC1-Ab	0.00	0.22	0.22	1.64	0.65	P10896	Ribulose bisphosphate carboxylase/oxygenase activase
BC1-Ab	0.90	0.69	1.60	19.31	0.03	P14981	Movement protein BC1 (33.7 kDa protein) (Movement protein BL1)
SACMV-Ab	0.87	0.45	1.32	6.02	0.12	P14981	Movement protein BC1 (33.7 kDa protein) (Movement protein BL1)
nSACMV-Ab	0.39	0.00	0.39	1.38	0.77	P14981	Movement protein BC1 (33.7 kDa protein) (Movement protein BL1)
nBC1-Ab	0.80	0.17	0.97	3.54	0.25	P14981	Movement protein BC1 (33.7 kDa protein) (Movement protein BL1)
BC1-Ab	0.19	0.05	0.24	1.29	0.83	P19366	ATP synthase subunit beta
SACMV-Ab	0.05	0.00	0.05	1.15	0.95	P19366	ATP synthase subunit beta
nSACMV-Ab	0.00	0.03	0.03	1.13	0.98	P19366	ATP synthase subunit beta
nBC1-Ab	0.00	0.48	0.48	1.96	0.49	P19366	ATP synthase subunit beta
BC1-Ab	0.77	0.42	1.18	4.48	0.18	P21238	Chaperonin 60 subunit alpha 1
SACMV-Ab	0.99	0.71	1.70	112.00	0.01	P21238	Chaperonin 60 subunit alpha 1
nSACMV-Ab	0.39	0.00	0.39	1.37	0.78	P21238	Chaperonin 60 subunit alpha 1
nBC1-Ab	0.81	0.17	0.98	3.64	0.24	P21238	Chaperonin 60 subunit alpha 1
BC1-Ab	0.84	0.16	1.00	3.50	0.25	P21240	Chaperonin 60 subunit beta 1
SACMV-Ab	0.82	0.13	0.95	2.98	0.29	P21240	Chaperonin 60 subunit beta 1
nSACMV-Ab	0.90	0.64	1.54	16.47	0.04	P21240	Chaperonin 60 subunit beta 1
BC1-Ab	0.24	0.06	0.30	1.36	0.79	P25819	Catalase-2 (EC 1.11.1.6)
nSACMV-Ab	0.44	0.03	0.48	1.46	0.73	P25819	Catalase-2 (EC 1.11.1.6)
nBC1-Ab	0.79	0.48	1.27	5.71	0.13	P25819	Catalase-2 (EC 1.11.1.6)
BC1-Ab	0.70	0.13	0.83	2.53	0.33	P25851	Fructose-1,6-bisphosphatase 1
SACMV-Ab	0.96	0.38	1.34	6.51	0.11	P25851	Fructose-1,6-bisphosphatase 1
nSACMV-Ab	0.86	0.40	1.26	5.33	0.15	P25851	Fructose-1,6-bisphosphatase 1
BC1-Ab	0.83	0.33	1.17	4.52	0.18	P25856	Glyceraldehyde-3-phosphate dehydrogenase GAPA1
SACMV-Ab	0.89	0.61	1.50	12.44	0.06	P25856	Glyceraldehyde-3-phosphate dehydrogenase GAPA1
nSACMV-Ab	0.41	0.08	0.50	1.70	0.62	P25856	Glyceraldehyde-3-phosphate dehydrogenase GAPA1
BC1-Ab	0.14	0.00	0.14	1.20	0.90	P25857	Glyceraldehyde-3-phosphate dehydrogenase GAPB
SACMV-Ab	0.00	0.45	0.45	1.85	0.54	P25857	Glyceraldehyde-3-phosphate dehydrogenase GAPB

nSACMV-Ab	0.99	0.70	1.69	62.22	0.01	P25857	Glyceraldehyde-3-phosphate dehydrogenase GAPB
nBC1-Ab	0.00	0.19	0.19	1.61	0.66	P25857	Glyceraldehyde-3-phosphate dehydrogenase GAPB
BC1-Ab	0.15	0.07	0.22	1.39	0.76	P25858	Glyceraldehyde-3-phosphate dehydrogenase GAPC1
SACMV-Ab	0.95	0.60	1.55	16.47	0.04	P25858	Glyceraldehyde-3-phosphate dehydrogenase GAPC1
nSACMV-Ab	0.98	0.37	1.35	6.83	0.11	P25858	Glyceraldehyde-3-phosphate dehydrogenase GAPC1
BC1-Ab	0.76	0.35	1.11	4.15	0.20	P25873	50S ribosomal protein L15
SACMV-Ab	0.60	0.12	0.72	2.20	0.38	P25873	50S ribosomal protein L15
nSACMV-Ab	0.86	0.60	1.46	10.37	0.08	P25873	50S ribosomal protein L15
BC1-Ab	0.19	0.11	0.30	1.59	0.68	P27140	Beta carbonic anhydrase 1
SACMV-Ab	0.00	0.06	0.06	1.16	0.94	P27140	Beta carbonic anhydrase 1
nSACMV-Ab	0.00	0.11	0.11	1.35	0.79	P27140	Beta carbonic anhydrase 1
BC1-Ab	0.27	0.14	0.41	1.84	0.54	P29185	Chaperonin CPN60-1
SACMV-Ab	0.00	0.06	0.06	1.18	0.91	P29185	Chaperonin CPN60-1
nSACMV-Ab	0.46	0.13	0.58	2.02	0.45	P29185	Chaperonin CPN60-1
BC1-Ab	0.21	0.10	0.31	1.54	0.69	P29197	Chaperonin CPN60
SACMV-Ab	0.94	0.59	1.53	14.36	0.05	P29197	Chaperonin CPN60
nSACMV-Ab	0.40	0.07	0.48	1.64	0.64	P29197	Chaperonin CPN60
BC1-Ab	0.77	0.35	1.13	4.21	0.19	P34788	40S ribosomal protein S18
SACMV-Ab	0.62	0.12	0.73	2.24	0.38	P34788	40S ribosomal protein S18
nSACMV-Ab	0.00	0.57	0.57	2.06	0.43	P34788	40S ribosomal protein S18
BC1-Ab	0.59	0.14	0.73	2.37	0.36	P41127	60S ribosomal protein L13-1 (Protein BBC1 homolog)
SACMV-Ab	0.58	0.11	0.69	2.12	0.40	P41127	60S ribosomal protein L13-1 (Protein BBC1 homolog)
nSACMV-Ab	0.75	0.59	1.34	6.91	0.10	P41127	60S ribosomal protein L13-1 (Protein BBC1 homolog)
BC1-Ab	0.35	0.15	0.49	2.07	0.42	P42643	14-3-3-like protein GF14 chi (General regulatory factor 1)
SACMV-Ab	0.00	0.07	0.07	1.22	0.87	P42643	14-3-3-like protein GF14 chi (General regulatory factor 1)
nSACMV-Ab	0.54	0.16	0.70	2.60	0.31	P42643	14-3-3-like protein GF14 chi (General regulatory factor 1)
BC1-Ab	0.16	0.07	0.23	1.40	0.75	P42737	Beta carbonic anhydrase 2
SACMV-Ab	0.61	0.29	0.90	3.01	0.29	P42737	Beta carbonic anhydrase 2
nSACMV-Ab	0.02	0.57	0.59	2.22	0.38	P42737	Beta carbonic anhydrase 2
BC1-Ab	0.88	0.34	1.22	5.05	0.16	P50318	Phosphoglycerate kinase 2
SACMV-Ab	0.00	0.06	0.06	1.14	0.96	P50318	Phosphoglycerate kinase 2
nSACMV-Ab	0.92	0.62	1.54	16.47	0.04	P50318	Phosphoglycerate kinase 2
BC1-Ab	0.37	0.14	0.51	1.96	0.48	P51419	60S ribosomal protein L27-3
SACMV-Ab	0.12	0.06	0.18	1.32	0.81	P51419	60S ribosomal protein L27-3
nSACMV-Ab	0.62	0.45	1.07	3.86	0.22	P51419	60S ribosomal protein L27-3
BC1-Ab	0.16	0.10	0.26	1.51	0.73	P51430	40S ribosomal protein S6-2 (Protein EMBRYO DEFECTIVE 3010)

SACMV-Ab	0.00	0.05	0.05	1.14	0.97	P51430	40S ribosomal protein S6-2 (Protein EMBRYO DEFECTIVE 3010)
nSACMV-Ab	0.98	0.54	1.52	14.00	0.05	P51430	40S ribosomal protein S6-2 (Protein EMBRYO DEFECTIVE 3010)
BC1-Ab	0.15	0.07	0.22	1.39	0.76	P53492	Actin-7 (Actin-2)
SACMV-Ab	0.00	0.59	0.59	2.06	0.43	P53492	Actin-7 (Actin-2)
nSACMV-Ab	0.00	0.26	0.26	1.60	0.67	P53492	Actin-7 (Actin-2)
BC1-Ab	0.70	0.21	0.91	3.20	0.28	P53496	Actin-11
SACMV-Ab	0.61	0.18	0.79	2.83	0.30	P53496	Actin-11
nBC1-Ab	0.00	0.63	0.63	2.13	0.40	P53496	Actin-11
BC1-Ab	0.69	0.36	1.05	3.76	0.23	P53497	Actin-12
SACMV-Ab	0.06	0.00	0.06	1.15	0.94	P53497	Actin-12
nSACMV-Ab	0.00	0.03	0.03	1.10	0.99	P53497	Actin-12
nBC1-Ab	0.00	0.32	0.32	1.64	0.65	P53497	Actin-12
BC1-Ab	0.25	0.08	0.34	1.54	0.70	P55228	Glucose-1-phosphate adenylyltransferase small subunit
SACMV-Ab	0.01	0.46	0.47	1.98	0.48	P55228	Glucose-1-phosphate adenylyltransferase small subunit
nSACMV-Ab	0.45	0.08	0.54	1.74	0.60	P55228	Glucose-1-phosphate adenylyltransferase small subunit
BC1-Ab	0.28	0.09	0.37	1.59	0.67	P55737	Heat shock protein 90-2 (AtHSP90.2) (AtHsp90-2)
SACMV-Ab	0.87	0.56	1.42	8.89	0.09	P55737	Heat shock protein 90-2 (AtHSP90.2)
nSACMV-Ab	0.47	0.10	0.57	1.88	0.52	P55737	Heat shock protein 90-2 (AtHSP90.2)
BC1-Ab	0.30	0.05	0.35	1.38	0.77	P56757	ATP synthase subunit alpha
SACMV-Ab	0.10	0.00	0.10	1.17	0.92	P56757	ATP synthase subunit alpha
nSACMV-Ab	0.50	0.04	0.54	1.53	0.71	P56757	ATP synthase subunit alpha
nBC1-Ab	0.78	0.47	1.25	5.60	0.13	P56757	ATP synthase subunit alpha
BC1-Ab	0.20	0.12	0.33	1.65	0.63	P59224	40S ribosomal protein S13-2
SACMV-Ab	0.07	0.06	0.13	1.24	0.86	P59224	40S ribosomal protein S13-2
nSACMV-Ab	0.00	0.11	0.11	1.32	0.81	P59224	40S ribosomal protein S13-2
BC1-Ab	0.30	0.05	0.35	1.37	0.78	P59259	Histone H4
SACMV-Ab	0.10	0.00	0.10	1.17	0.93	P59259	Histone H4
nSACMV-Ab	0.50	0.04	0.54	1.53	0.71	P59259	Histone H4
nBC1-Ab	0.78	0.47	1.25	5.49	0.13	P59259	Histone H4
BC1-Ab	0.84	0.57	1.41	8.36	0.09	P61841	30S ribosomal protein S7
SACMV-Ab	0.60	0.29	0.89	2.98	0.29	P61841	30S ribosomal protein S7
nSACMV-Ab	0.00	0.07	0.07	1.22	0.88	P61841	30S ribosomal protein S7
BC1-Ab	0.35	0.15	0.49	2.06	0.43	Q42112	60S acidic ribosomal protein P0-2
SACMV-Ab	0.00	0.07	0.07	1.19	0.91	Q42112	60S acidic ribosomal protein P0-2
nSACMV-Ab	0.54	0.16	0.70	2.58	0.32	Q42112	60S acidic ribosomal protein P0-2
BC1-Ab	0.35	0.15	0.49	2.04	0.44	Q43127	Glutamine synthetase
SACMV-Ab	0.00	0.07	0.07	1.19	0.91	Q43127	Glutamine synthetase
nSACMV-Ab	0.54	0.16	0.70	2.56	0.32	Q43127	Glutamine synthetase

BC1-Ab	0.22	0.08	0.31	1.51	0.73	Q56WH1	Tubulin alpha-3 chain
SACMV-Ab	0.01	0.51	0.52	2.11	0.40	Q56WH1	Tubulin alpha-3 chain
nSACMV-Ab	0.43	0.09	0.52	1.77	0.58	Q56WH1	Tubulin alpha-3 chain
BC1-Ab	0.24	0.09	0.33	1.54	0.70	Q56YA5	Serine--glyoxylate aminotransferase
SACMV-Ab	0.00	0.46	0.46	1.79	0.56	Q56YA5	Serine--glyoxylate aminotransferase
nSACMV-Ab	0.45	0.10	0.55	1.85	0.53	Q56YA5	Serine--glyoxylate aminotransferase
BC1-Ab	0.28	0.13	0.41	1.79	0.56	Q8H156	GTP-binding nuclear protein Ran-3 (Ras-related nuclear protein 3)
SACMV-Ab	0.00	0.06	0.06	1.15	0.95	Q8H156	GTP-binding nuclear protein Ran-3 (Ras-related nuclear protein 3)
nSACMV-Ab	0.47	0.13	0.59	2.07	0.42	Q8H156	GTP-binding nuclear protein Ran-3 (Ras-related nuclear protein 3)
BC1-Ab	0.35	0.15	0.49	2.03	0.44	Q8L719	THO complex subunit 4B (ALYREF homolog 2) (AtALY2)
SACMV-Ab	0.00	0.07	0.07	1.18	0.92	Q8L719	THO complex subunit 4B (ALYREF homolog 2) (AtALY2)
nSACMV-Ab	0.54	0.16	0.70	2.53	0.33	Q8L719	THO complex subunit 4B (ALYREF homolog 2) (AtALY2)
BC1-Ab	0.67	0.29	0.96	3.29	0.26	Q8LCL3	60S ribosomal protein L27-2
SACMV-Ab	0.08	0.06	0.14	1.27	0.83	Q8LCL3	60S ribosomal protein L27-2
nSACMV-Ab	0.64	0.27	0.91	3.06	0.28	Q8LCL3	60S ribosomal protein L27-2
BC1-Ab	0.75	0.50	1.25	5.71	0.13	Q8LD46	60S ribosomal protein L23a-1 (AtRPL23A-1)
SACMV-Ab	0.12	0.07	0.19	1.33	0.80	Q8LD46	60S ribosomal protein L23a-1 (AtRPL23A-1)
nSACMV-Ab	0.51	0.13	0.64	2.16	0.39	Q8LD46	60S ribosomal protein L23a-1 (AtRPL23A-1)
BC1-Ab	0.16	0.07	0.23	1.41	0.75	Q8RWV0	Transketolase-1
SACMV-Ab	0.98	0.54	1.52	14.00	0.05	Q8RWV0	Transketolase-1
nSACMV-Ab	0.00	0.08	0.08	1.21	0.89	Q8RWV0	Transketolase-1
BC1-Ab	0.67	0.15	0.82	2.68	0.30	Q8RXX5	50S ribosomal protein L19-2
SACMV-Ab	0.60	0.12	0.71	2.17	0.39	Q8RXX5	50S ribosomal protein L19-2
nSACMV-Ab	0.65	0.15	0.79	2.59	0.31	Q8RXX5	50S ribosomal protein L19-2
BC1-Ab	0.72	0.00	0.72	1.65	0.64	Q8W4H7	Elongation factor 1-alpha 2 (EF-1-alpha 2)
SACMV-Ab	0.95	0.71	1.66	35.00	0.02	Q8W4H7	Elongation factor 1-alpha 2 (EF-1-alpha 2)
nSACMV-Ab	0.00	0.23	0.23	1.53	0.71	Q8W4H7	Elongation factor 1-alpha 2 (EF-1-alpha 2)
nBC1-Ab	0.00	0.46	0.46	1.78	0.58	Q8W4H7	Elongation factor 1-alpha 2 (EF-1-alpha 2)
BC1-Ab	0.38	0.11	0.49	1.80	0.55	Q93VG5	40S ribosomal protein S8-1
SACMV-Ab	0.58	0.39	0.97	3.27	0.27	Q93VG5	40S ribosomal protein S8-1
nSACMV-Ab	0.57	0.11	0.68	2.11	0.41	Q93VG5	40S ribosomal protein S8-1
BC1-Ab	0.23	0.10	0.33	1.58	0.68	Q944G9	Fructose-bisphosphate aldolase 2
SACMV-Ab	0.00	0.46	0.46	1.77	0.58	Q944G9	Fructose-bisphosphate aldolase 2
nSACMV-Ab	0.42	0.08	0.50	1.70	0.62	Q944G9	Fructose-bisphosphate aldolase 2
BC1-Ab	0.32	0.15	0.48	2.01	0.46	Q94C74	Serine hydroxymethyltransferase 2
SACMV-Ab	0.11	0.07	0.18	1.35	0.79	Q94C74	Serine hydroxymethyltransferase 2

nSACMV-Ab	0.74	0.50	1.24	5.38	0.15	Q94C74	Serine hydroxymethyltransferase 2
BC1-Ab	0.35	0.15	0.49	2.01	0.46	Q96291	2-Cys peroxiredoxin BAS1
SACMV-Ab	0.00	0.07	0.07	1.16	0.93	Q96291	2-Cys peroxiredoxin BAS1
nSACMV-Ab	0.54	0.16	0.70	2.51	0.33	Q96291	2-Cys peroxiredoxin BAS1
BC1-Ab	0.29	0.14	0.43	1.86	0.53	Q96300	14-3-3-like protein GF14 nu (General regulatory factor 7)
SACMV-Ab	0.00	0.06	0.06	1.14	0.97	Q96300	14-3-3-like protein GF14 nu (General regulatory factor 7)
nSACMV-Ab	0.48	0.13	0.61	2.09	0.41	Q96300	14-3-3-like protein GF14 nu (General regulatory factor 7)
BC1-Ab	0.30	0.05	0.35	1.36	0.78	Q96528	Catalase-1 (EC 1.11.1.6)
SACMV-Ab	0.10	0.00	0.10	1.16	0.94	Q96528	Catalase-1 (EC 1.11.1.6)
nSACMV-Ab	0.50	0.04	0.54	1.52	0.72	Q96528	Catalase-1 (EC 1.11.1.6)
nBC1-Ab	0.78	0.47	1.25	5.38	0.15	Q96528	Catalase-1 (EC 1.11.1.6)
BC1-Ab	0.23	0.09	0.32	1.54	0.69	Q9ASR1	Elongation factor 2 (EF-2)
SACMV-Ab	0.97	0.53	1.50	12.73	0.06	Q9ASR1	Elongation factor 2 (EF-2)
nSACMV-Ab	0.44	0.09	0.52	1.75	0.59	Q9ASR1	Elongation factor 2 (EF-2)
BC1-Ab	0.31	0.12	0.44	1.81	0.55	Q9FGX1	ATP-citrate synthase beta chain protein 2 (ATP-citrate synthase B-2)
SACMV-Ab	0.73	0.50	1.23	5.44	0.14	Q9FGX1	ATP-citrate synthase beta chain protein 2 (ATP-citrate synthase B-2)
nSACMV-Ab	0.49	0.10	0.59	1.90	0.51	Q9FGX1	ATP-citrate synthase beta chain protein 2 (ATP-citrate synthase B-2)
BC1-Ab	0.67	0.29	0.96	3.26	0.27	Q9FIF3	40S ribosomal protein S8-2
SACMV-Ab	0.08	0.06	0.14	1.26	0.85	Q9FIF3	40S ribosomal protein S8-2
nSACMV-Ab	0.64	0.27	0.91	3.03	0.28	Q9FIF3	40S ribosomal protein S8-2
BC1-Ab	0.68	0.12	0.81	2.47	0.34	Q9FJX2	60S ribosomal protein L26-2
SACMV-Ab	0.87	0.62	1.49	11.91	0.07	Q9FJX2	60S ribosomal protein L26-2
nSACMV-Ab	0.65	0.14	0.80	2.59	0.31	Q9FJX2	60S ribosomal protein L26-2
BC1-Ab	0.69	0.28	0.97	3.27	0.27	Q9FZ47	ACT domain-containing protein ACR11 (Protein ACT DOMAIN REPEATS 11)
SACMV-Ab	0.05	0.05	0.11	1.22	0.88	Q9FZ47	ACT domain-containing protein ACR11 (Protein ACT DOMAIN REPEATS 11)
nSACMV-Ab	0.03	0.30	0.33	1.79	0.57	Q9FZ47	ACT domain-containing protein ACR11 (Protein ACT DOMAIN REPEATS 11)
BC1-Ab	0.17	0.08	0.25	1.43	0.74	Q9LD57	Phosphoglycerate kinase 1
SACMV-Ab	0.90	0.59	1.49	11.67	0.07	Q9LD57	Phosphoglycerate kinase 1
nSACMV-Ab	0.00	0.08	0.08	1.19	0.91	Q9LD57	Phosphoglycerate kinase 1
BC1-Ab	0.21	0.08	0.29	1.48	0.73	Q9LJE4	Chaperonin 60 subunit beta 2
SACMV-Ab	0.00	0.58	0.58	1.90	0.51	Q9LJE4	Chaperonin 60 subunit beta 2
nSACMV-Ab	0.41	0.10	0.50	1.75	0.59	Q9LJE4	Chaperonin 60 subunit beta 2
BC1-Ab	0.35	0.15	0.49	2.00	0.47	Q9LKA3	Malate dehydrogenase 2
SACMV-Ab	0.00	0.07	0.07	1.15	0.95	Q9LKA3	Malate dehydrogenase 2

nSACMV-Ab	0.54	0.16	0.70	2.49	0.34	Q9LKA3	Malate dehydrogenase 2
BC1-Ab	0.14	0.07	0.21	1.38	0.77	Q9LPW0	Glyceraldehyde-3-phosphate dehydrogenase GAPA2
SACMV-Ab	0.00	0.32	0.32	1.57	0.69	Q9LPW0	Glyceraldehyde-3-phosphate dehydrogenase GAPA2
nSACMV-Ab	0.57	0.64	1.21	5.14	0.16	Q9LPW0	Glyceraldehyde-3-phosphate dehydrogenase GAPA2
BC1-Ab	0.35	0.15	0.49	1.99	0.47	Q9LQC8	ADP-ribosylation factor 2-A (AtARF2)
SACMV-Ab	0.00	0.07	0.07	1.15	0.96	Q9LQC8	ADP-ribosylation factor 2-A (AtARF2)
nSACMV-Ab	0.54	0.16	0.70	2.47	0.34	Q9LQC8	ADP-ribosylation factor 2-A (AtARF2)
BC1-Ab	0.71	0.00	0.71	1.63	0.65	Q9LRS0	(S)-2-hydroxy-acid oxidase GLO2
SACMV-Ab	0.00	0.66	0.66	2.00	0.47	Q9LRS0	(S)-2-hydroxy-acid oxidase GLO2
nSACMV-Ab	0.00	0.02	0.02	1.05	1.00	Q9LRS0	(S)-2-hydroxy-acid oxidase GLO2
nBC1-Ab	0.00	0.44	0.44	1.65	0.64	Q9LRS0	(S)-2-hydroxy-acid oxidase GLO2
BC1-Ab	0.35	0.15	0.49	1.97	0.48	Q9LVI8	Glutamine synthetase cytosolic isozyme 1-3 (GS1)
SACMV-Ab	0.00	0.07	0.07	1.13	0.98	Q9LVI8	Glutamine synthetase cytosolic isozyme 1-3 (GS1)
nSACMV-Ab	0.54	0.16	0.70	2.45	0.35	Q9LVI8	Glutamine synthetase cytosolic isozyme 1-3 (GS1)
BC1-Ab	0.37	0.14	0.51	1.95	0.49	Q9LX88	40S ribosomal protein S15a-4
SACMV-Ab	0.12	0.06	0.18	1.31	0.82	Q9LX88	40S ribosomal protein S15a-4
nSACMV-Ab	0.62	0.45	1.07	3.81	0.23	Q9LX88	40S ribosomal protein S15a-4
BC1-Ab	0.66	0.44	1.10	3.92	0.21	Q9LXG1	40S ribosomal protein S9-1
SACMV-Ab	0.13	0.06	0.20	1.33	0.80	Q9LXG1	40S ribosomal protein S9-1
nSACMV-Ab	0.56	0.14	0.70	2.35	0.36	Q9LXG1	40S ribosomal protein S9-1
BC1-Ab	0.31	0.12	0.44	1.79	0.56	Q9LXL5	Sucrose synthase 4 (AtSUS4)
SACMV-Ab	0.73	0.50	1.23	5.33	0.15	Q9LXL5	Sucrose synthase 4 (AtSUS4)
nSACMV-Ab	0.49	0.10	0.59	1.89	0.51	Q9LXL5	Sucrose synthase 4 (AtSUS4)
BC1-Ab	0.28	0.09	0.38	1.60	0.66	Q9LZF6	Cell division control protein 48 homolog E (AtCDC48e)
SACMV-Ab	0.82	0.53	1.34	6.91	0.10	Q9LZF6	Cell division control protein 48 homolog E (AtCDC48e)
nSACMV-Ab	0.47	0.10	0.58	1.90	0.51	Q9LZF6	Cell division control protein 48 homolog E (AtCDC48e)
BC1-Ab	0.66	0.44	1.10	3.86	0.22	Q9M385	50S ribosomal protein L17
SACMV-Ab	0.13	0.06	0.20	1.33	0.81	Q9M385	50S ribosomal protein L17
nSACMV-Ab	0.56	0.14	0.70	2.33	0.37	Q9M385	50S ribosomal protein L17
BC1-Ab	0.75	0.35	1.10	4.06	0.20	Q9SF40	60S ribosomal protein L4-1 (L1)
SACMV-Ab	0.74	0.32	1.05	3.68	0.24	Q9SF40	60S ribosomal protein L4-1 (L1)
nSACMV-Ab	0.43	0.08	0.51	1.70	0.62	Q9SF40	60S ribosomal protein L4-1 (L1)
BC1-Ab	0.29	0.14	0.43	1.85	0.54	Q9SFH9	Delta-aminolevulinic acid dehydratase 1
SACMV-Ab	0.00	0.06	0.06	1.11	0.99	Q9SFH9	Delta-aminolevulinic acid dehydratase 1
nSACMV-Ab	0.48	0.13	0.61	2.07	0.42	Q9SFH9	Delta-aminolevulinic acid dehydratase 1
BC1-Ab	0.17	0.05	0.22	1.27	0.83	Q9SGA6	40S ribosomal protein S19-1
SACMV-Ab	0.07	0.00	0.07	1.14	0.96	Q9SGA6	40S ribosomal protein S19-1

nSACMV-Ab	0.00	0.03	0.03	1.05	1.00	Q9SGA6	40S ribosomal protein S19-1
nBC1-Ab	0.00	0.49	0.49	1.73	0.60	Q9SGA6	40S ribosomal protein S19-1
BC1-Ab	0.23	0.09	0.32	1.53	0.71	Q9SGC1	Probable phosphoglucomutase
SACMV-Ab	0.97	0.53	1.50	12.17	0.06	Q9SGC1	Probable phosphoglucomutase
nSACMV-Ab	0.44	0.09	0.52	1.74	0.60	Q9SGC1	Probable phosphoglucomutase
BC1-Ab	0.70	0.28	0.98	3.33	0.25	Q9SIH0	40S ribosomal protein S14-1
SACMV-Ab	0.04	0.05	0.09	1.20	0.90	Q9SIH0	40S ribosomal protein S14-1
nSACMV-Ab	0.01	0.55	0.56	2.19	0.39	Q9SIH0	40S ribosomal protein S14-1
BC1-Ab	0.37	0.14	0.51	1.94	0.50	Q9SIP7	40S ribosomal protein S3-1
SACMV-Ab	0.12	0.06	0.18	1.31	0.82	Q9SIP7	40S ribosomal protein S3-1
nSACMV-Ab	0.62	0.45	1.07	3.76	0.23	Q9SIP7	40S ribosomal protein S3-1
BC1-Ab	0.35	0.15	0.49	1.96	0.49	Q9SJQ9	Fructose-bisphosphate aldolase 6
SACMV-Ab	0.00	0.07	0.07	1.12	0.98	Q9SJQ9	Fructose-bisphosphate aldolase 6
nSACMV-Ab	0.54	0.16	0.70	2.42	0.35	Q9SJQ9	Fructose-bisphosphate aldolase 6
BC1-Ab	0.25	0.12	0.37	1.69	0.63	Q9SJU4	Fructose-bisphosphate aldolase 1
SACMV-Ab	0.09	0.06	0.15	1.27	0.84	Q9SJU4	Fructose-bisphosphate aldolase 1
nSACMV-Ab	0.00	0.12	0.12	1.26	0.85	Q9SJU4	Fructose-bisphosphate aldolase 1
BC1-Ab	0.26	0.12	0.38	1.70	0.61	Q9SKX4	50S ribosomal protein L3-1
SACMV-Ab	0.00	0.06	0.06	1.09	0.99	Q9SKX4	50S ribosomal protein L3-1
nSACMV-Ab	0.64	0.45	1.10	3.97	0.21	Q9SKX4	50S ribosomal protein L3-1
BC1-Ab	0.20	0.11	0.30	1.58	0.68	Q9SQZ1	40S ribosomal protein S17-3
SACMV-Ab	0.06	0.06	0.12	1.23	0.87	Q9SQZ1	40S ribosomal protein S17-3
nSACMV-Ab	0.03	0.46	0.49	2.02	0.45	Q9SQZ1	40S ribosomal protein S17-3
BC1-Ab	0.18	0.08	0.26	1.44	0.74	Q9SXJ7	Chaperone protein ClpC2
SACMV-Ab	0.94	0.32	1.26	5.38	0.15	Q9SXJ7	Chaperone protein ClpC2
nSACMV-Ab	0.83	0.54	1.37	7.47	0.10	Q9SXJ7	Chaperone protein ClpC2
BC1-Ab	0.76	0.42	1.18	4.48	0.18	Q9SZJ5	Serine hydroxymethyltransferase 1
SACMV-Ab	0.00	0.63	0.63	2.34	0.37	Q9SZJ5	Serine hydroxymethyltransferase 1
nSACMV-Ab	0.39	0.00	0.39	1.34	0.80	Q9SZJ5	Serine hydroxymethyltransferase 1
nBC1-Ab	0.81	0.17	0.98	3.59	0.24	Q9SZJ5	Serine hydroxymethyltransferase 1
BC1-Ab	0.68	0.46	1.14	4.27	0.19	Q9T029	40S ribosomal protein S25-4
SACMV-Ab	0.08	0.06	0.13	1.25	0.85	Q9T029	40S ribosomal protein S25-4
nSACMV-Ab	0.00	0.11	0.11	1.23	0.86	Q9T029	40S ribosomal protein S25-4
BC1-Ab	0.31	0.12	0.44	1.78	0.57	Q9ZP06	Malate dehydrogenase 1
SACMV-Ab	0.73	0.50	1.23	5.23	0.16	Q9ZP06	Malate dehydrogenase 1
nSACMV-Ab	0.49	0.10	0.59	1.88	0.52	Q9ZP06	Malate dehydrogenase 1

Appendix 4.1 The Kruskal Wallis H Test based on the null hypothesis (Ho): All viral load/ gene expression median values across all samples are equal. H values sourced from www.dataaHalytics.org.uk/critical-values-for-the-kruskal-wallis-test/.

SACMV VIRAL LOAD	
H calculated	15.6
H value	7.501
SDG16 EXPRESSION	
H calculated	-22.57
H value	7.235
CKL1 EXPRESSION	
H calculated	-29.250
H value	7.377
NUP98 EXPRESSION	
H calculated	-30.81092464
H value	7.235

Appendix 4.2 Mann Whitney U Test for pairwise comparisons based on the null hypothesis (H_o): The two samples represented as treatments and with four expression values have the same expression value. Where p calculated > p tabulated (0.05) the H_o was rejected. HS, NS and CS denote SACMV+CRISPR/Cas transfected samples SDG16, NUP98 and CKL1, respectively.

Pairwise comparisons	**U statistic**	**z and p statistics**	
HS	31	z	-1.14
NS	18	p calculated	0.26
HS-NS	18		
HS	29	z	-0.79
CS	20	p calculated	0.43
HS-CS	20		
HS	49	z	-4.29
S	0	p calculated	0.00002
HS-S	0		
NS	24	z	-0.09
CS	25	p calculated	0.93
NS-CS	24		
NS	49	z	-4.29
S	0	p calculated	0.00002
NS-S	0		
CS	49	z	-4.29
S	0	p calculated	0.00002
CS-S	0		

Appendix 4.3 Mann Whitney U Test for pairwise comparisons based on the null hypothesis (H_o): The two samples represented as treatments and with four expression values have the same expression value. Where p calculated > p tabulated (0.05) the H_o was rejected. HS, H, S and pD denotes SACMV+ CRISPR/Cas (SDG16), CRISPR/Cas (SDG16), SACMV and untransformed samples, respectively.

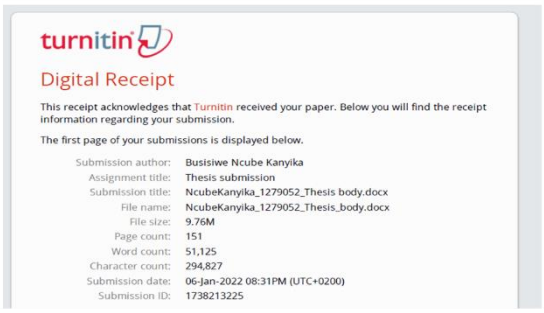
SDG16 pairwise comparisons			
H	16	z	-3.098386677
HS	0	p calculated	0.001945774
H-HS	0		
H	16	z	-3.098386677
pD	0	p calculated	0.001945774
H-pD	0		
H	16	z	-3.098386677
S	0	p calculated	0.001945774
H-S	0		
HS	0	z	-3.098386677
pD	16	p calculated	0.001945774
HS-pD	0		
HS	16	z	-3.098386677
S	0	p calculated	0.001945774
HS-S	0		
pD	16	z	-3.098386677
S	0	p calculated	0.001945774
pD-S	0		

Appendix 4.4 Mann Whitney U Test for pairwise comparisons based on the null hypothesis (H_o): The two samples represented as treatments and with four expression values have the same expression value. Where p calculated > p tabulated (0.05) the H_o was rejected. CS, C, S and pD denotes SACMV+ CRISPR/Cas (CKL1), CRISPR/Cas (CKL1), SACMV and untransformed samples, respectively.

CKL1pairwise comparisons			
C	7	z	-1.555634919
CS	20	p calculated	0.119794931
C-CS	7		
C	25	z	-3.535533906
pD	0	p calculated	0.000406952
C-pD	0		
C	25	z	-3.535533906
S	0	p calculated	0.000406952
C-S	0		
CS	25	z	-3.535533906
pD	0	p calculated	0.000406952
CS-pD	0		
CS	25	z	-3.535533906
S	0	p calculated	0.000406952
CS-S	0		
pD	25	z	-3.535533906
S	0	p calculated	0.000406952
pD-S	0		

Appendix 4.5 Mann Whitney U Test for pairwise comparisons based on the null hypothesis (H_o): The two samples represented as treatments and with four expression values have the same expression value. Where p calculated > p tabulated (0.05) the H_o was rejected. NS, N, S and pD denotes SACMV+ CRISPR/Cas (NUP98), CRISPR/Cas (NUP98), SACMV and untransformed samples, respectively.

NUP98 pairwise comparisons			
N	0	z	-3.098386677
NS	16	p calculated	0.001945774
N-NS	0		
N	0	z	-3.098386677
pD	16	p calculated	0.001945774
N-pD	0		
N	16	z	-3.098386677
S	0	p calculated	0.001945774
N-S	0		
NS	12	z	-1.549193338
pD	4	p calculated	0.121335225
NS-pD	4		
NS	16	z	-3.098386677
S	0	p calculated	0.001945774
NS-S	0		
pD	16	z	-3.098386677
S	0	p calculated	0.001945774
pD-S	0		



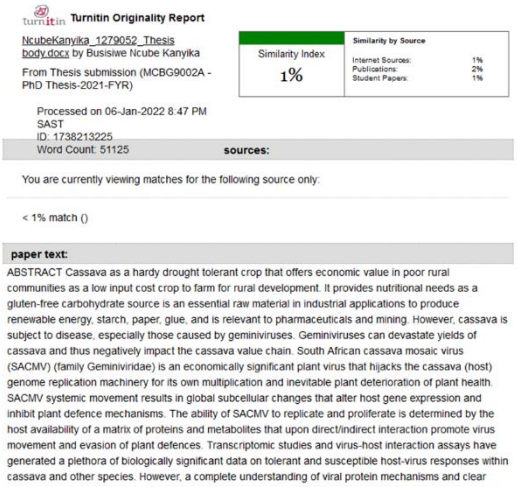
turnitin

Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Busisiwe Ncube Kanyika
Assignment title:	Thesis submission
Submission title:	NcubeKanyika_1279052_Thesis body.docx
File name:	NcubeKanyika_1279052_Thesis_body.docx
File size:	9.76M
Page count:	151
Word count:	51,125
Character count:	294,827
Submission date:	06-Jan-2022 08:31PM (UTC+0200)
Submission ID:	1738213225



turnitin Turnitin Originality Report

NcubeKanyika_1279052_Thesis body.docx by Busisiwe Ncube Kanyika

From Thesis submission (MCBG9002A - PhD Thesis-2021-FYR)

Processed on 06-Jan-2022 8:47 PM SAST
ID: 1738213225
Word Count: 51125

Similarity Index 1%

Similarity by Source
Internet Sources: 1%
Publications: 2%
Student Papers: 1%

sources:

You are currently viewing matches for the following source only:

< 1% match ()

paper text:

ABSTRACT Cassava as a hardy drought tolerant crop that offers economic value in poor rural communities as a low input cost crop to farm for rural development. It provides nutritional needs as a gluten-free carbohydrate source is an essential raw material in industrial applications to produce renewable energy, starch, paper, glue, and is relevant to pharmaceuticals and mining. However, cassava is subject to disease, especially those caused by geminiviruses. Geminiviruses can devastate yields of cassava and thus negatively impact the cassava value chain. South African cassava mosaic virus (SACMV) (family Geminiviridae) is an economically significant plant virus that hijacks the cassava (host) genome replication machinery for its own multiplication and inevitable plant deterioration of plant health. SACMV systemic movement results in global subcellular changes that alter host gene expression and inhibit plant defence mechanisms. The ability of SACMV to replicate and proliferate is determined by the host availability of a matrix of proteins and metabolites that upon direct/indirect interaction promote virus movement and evasion of plant defences. Transcriptomic studies and virus-host interaction assays have generated a plethora of biologically significant data on tolerant and susceptible host-virus responses within cassava and other species. However, a complete understanding of viral protein mechanisms and clear