THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

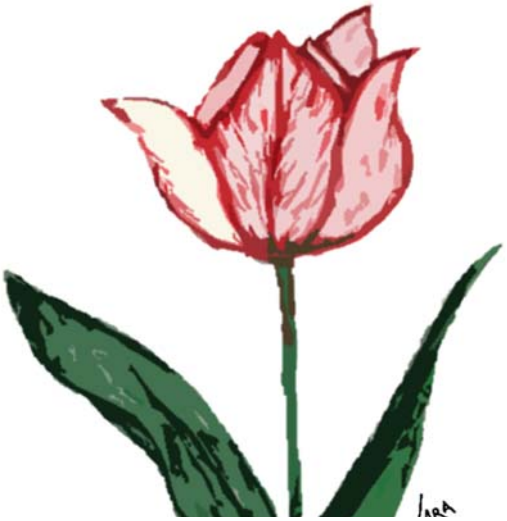
Investigation of virus-induced defence modulations during plant-virus interactions

Subtitle

The Plant-Virus Relationship Status:

It's Complicated

Lara-Simone Pretorius



Flower-breaking symptoms in a Tulip

A thesis submitted for the degree of Doctor of Philosophy at The University of Queensland in 2018

School of Agriculture & Food Sciences

Abstract

Turnip mosaic virus (TuMV) is a single stranded RNA (sRNAs) virus belonging to the genus Potyvirus .*Turnip mosaic virus* has a wide host range including several important crop plants, making it economically significant. This study investigates the relation between a strain of TuMV and two model plant systems; *Arabidopsis thaliana* and *Nicotiana benthamiana*, focusing on the viral sequence, host gene expression and defence mechanisms, and viral derived small RNAs (viRNAs).

Turnip mosaic virus has been well studied and characterised with 183 complete genome nucleotide sequences available in GenBank (5 June 2018). The TuMV isolate used in this study was sequenced and submitted to GenBank, as well as the original isolate collected in 1994, under the names TuMV-QLD1b and TuMV-QLD1a, respectively (accession numbers KX641465 and KX641466). The original TuMV isolate was PCR sequenced while the 2015 isolate was sequenced by deep RNA sequencing. A comparison between the two sequences showed minor variations with 18 single-nucleotide-polymorphisms (SNPs). Another aspect of the study involved the sequencing of a Cucumber mosaic virus (CMV) isolate (strain K) as well as an Australian Cauliflower mosaic virus (CaMV) isolate belonging to the genus Cucumovirus and Caulimovirus respectively. This strain of CaMV represents the first Australian isolate to be fully sequenced. Both the CMV and the CaMV isolates were used to study plant defence pathways in Chapter 4 of this thesis. The sequence of both isolates were published and make up Chapter 2.

Chapter 3 investigates the early defence response of *A. thaliana* 6, 24 and 48 hours after TuMV inoculation. Marker gene expression results suggest that the virus upregulated the jasmonic acid (JA) pathway 24 hours after infection. This is significant, as viruses are classified as biotrophic pathogens which usually upregulate the salicylic acid (SA) pathway leading to hyper-sensitive response and programmed cell death, preventing the virus from spreading. It is hypothesised that the upregulation of the JA pathway may favour the virus allowing it to establish more easily and systemically infect its host.

The JA and other defence pathways are further researched through studying the interactions between a mediator mutant and four different viruses; TuMV, CMV, CaMV and *Alternanthera mosaic virus* (AltMV). The mediator complex consists of several subunits and is highly conserved among eukaryotes as it regulates transcription. Previous studies have shown that *med18* plants are more resistant to *Fusarium oxysporum* which was also found to upregulate the JA pathway in WT plants. Results show a similar trend with viral infected *med18* plants having less viral RNA than WT plants 14 days after infection, though most were not significant due to large variations of viral load between individual plants. However, *med18* plants infected with CMV had significantly less viral load than WT plants.

Some of phenotypic symptoms caused by viral infection may be a secondary effect of RNA silencing. It is hypothesised that viRNAs can interfere with the plant's regulations and development causing phenotypic symptoms. Similar to previous studies, small RNA sequencing of virus infected *A. thaliana* and *N. benthamiana* suggested that there was a significant increase in the number of small RNAs (sRNAs), specifically those of 21 nucleotides (nt) in length. This length would also suggest that these are produced through a specific biogenesis pathway. Studies have reported that certain areas of a virus genome, called "hotspots", are more prone to being acted upon by the plant's sRNAs biogenesis machinery. Chapter 5 results suggest this to be true, with certain regions producing more sRNAs and causing certain viRNAs to be more abundant.

When comparing the most abundant viRNAs to host genes we identified possible targets based on complementarity. As a high degree of complementarity is believed to be required for endogenous small interfering RNAs (siRNAs) directed silencing it was hypothesized that highly abundant and complementary viRNAs had the potential to target and inhibit certain genes. The dual-luciferase report system was used to attempt to quickly validate possible viRNA targeting of host transcript sequences. The quantitative nature of this assay allowed us to determine whether viRNAs interacted with the target transcript based on the expression ratio. A mutated target sequence was included as a control to confirm viRNA interaction. Many targets from both *A. thaliana* and *N. benthamiana* were tested, though only one target sequence interaction was repeatedly confirmed using this assay system. The *N. benthamiana* gene 3160g02007 was confirmed to be targeted by TuMV as there was a clear decrease in the dual LUC expression ratio in the presence of the virus. The expression was restored when the target was mutated. Furthermore, no decrease was apparent when plants were not infected.

Declaration by Author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, financial support and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my higher degree by research candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis and have sought permission from co-authors for any jointly authored works included in the thesis.

Publications during Candidature

Peer-Reviewed-Papers

Pretorius, L., Hussein, N., Al-Amery, A., Moyle, R., Schenk, P. (2017). Natural and Engineered Defences Against Plant Viruses. Current Biotechnology 6, 339-348.

Pretorius, L., Moyle, R.L., Dalton-Morgan, J., Schwinghamer, M.W., Crew, K., Schenk, P.M., Geering, A.D.W. (2017). First fully sequenced genome of an Australian isolate of *Cauliflower mosaic virus*. Australasian Plant Pathology 46, 597-599.

Moyle, R.L., Carvalhais, L.C., Pretorius, L.-S., Nowak, E., Subramaniam, G., Dalton-Morgan, J., Schenk, P.M. (2017). An Optimized Transient Dual Luciferase Assay for Quantifying MicroRNA Directed Repression of Targeted Sequences. Frontiers in Plant Science 8.

Moyle, R.L., Pretorius, L., Carvalhais, L.C., Schenk, P.M. (2017). Complete Nucleotide Sequence of Australian *Tomato spotted wilt virus* Isolate TSWV-QLD2. Genome Announcements 5.

Moyle, R., Pretorius, L. S., Shuey, L. S., Nowak, E., & Schenk, P. M. (2018). Analysis of the Complete Genome Sequence of *Cucumber mosaic virus* Strain K. Genome announcements, 6(7), e00053-18.

Moyle, R., Pretorius, L. S., Dalton-Morgan, J., Persley, D., & Schenk, P. (2016). Analysis of the first complete genome sequence of an Australian *Tomato spotted wilt* virus isolate. Australasian Plant Pathology, 45(5), 509-512.

Pretorius L, Moyle RL, Dalton-Morgan J, Hussein N, Schenk PM. (2016). Complete Nucleotide Sequence of an Australian Isolate of *Turnip mosaic virus* before and after Seven Years of Serial Passaging. Genome Announcements, 4(6):e01269-16.

Conference Abstracts and Presentations

Pretorius, L.S., Kidd, B.N. and Schenk, P.M (2013). Chlorosis causing compounds in the *Arabidopsis – Fusarium* interaction. *24th International Conference on Arabidopsis Research*. Australia. 24-28 June. (Poster).

Pretorius, L.S., Geering, A. and Schenk, P.M (2014). *Turnip mosaic virus* targets host gene HVA22D: a new opportunity to develop resistant crops. *11th Australasian Plant Virology Workshop*. Australia. 13-15 August. (Oral).

Pretorius, L.S., Moyle, R.L., Geering, A. and Schenk, P.M (2015). MicroRNA-like fragments from Turnip mosaic virus targets host gene HVA22D – a new opportunity to develop resistant crops. *XVIII International Plant Protection Congress.* Germany. 24-27 August. (Oral).

Pretorius L.S, Moyle R.L., Nowak K, Carvalhais L, Dalton-Morgan J, Geering A and Schenk PM (2017) Viral small RNAs that target plant genes and artificial miRNA that target plant viruses. *Science Protecting Plant Health*. Brisbane. 26-28 September. (Oral).

Publications Included in this Thesis

Pretorius, L., Hussein, N., Al-Amery, A., Moyle, R., Schenk, P. (2017). Natural and Engineered Defences against Plant Viruses. Current Biotechnology 6, 339-348.

Contributor	Statement of contribution
Author Pretorius, L (Candidate)	Conception and design of the project: 80% Analysis and interpretation: N/A: Drafting significant parts (contributing to interpretation): 71%
Author Hussein, N	Conception and design of the project: 0% Analysis and interpretation: N/A Drafting significant parts (contributing to interpretation): 2%
Author Al-Amery, A	Conception and design of the project: 0% Analysis and interpretation: N/A Drafting significant parts (contributing to interpretation): 2%
Author Moyle, R	Conception and design of the project: 0% Analysis and interpretation: N/A Drafting significant parts (contributing to interpretation): 15%
Author Schenk, P	Conception and design of the project :20% Analysis and interpretation: N/A Drafting significant parts (contributing to interpretation): 10%

Pretorius, L., Moyle, R.L., Dalton-Morgan, J., Schwinghamer, M.W., Crew, K., Schenk, P.M., Geering, A.D.W. (2017). First fully sequenced genome of an Australian isolate of *Cauliflower mosaic virus*. Australasian Plant Pathology 46, 597-599.

Contributor	Statement of contribution
Author Pretorius L. (Candidate)	Conception and design of the project: 40% Analysis and interpretation: 70% Drafting significant parts (contributing to interpretation): 50%
Author Moyle, R. L.	Conception and design of the project: 5% Analysis and interpretation: 0% Drafting significant parts (contributing to interpretation): 5%
Author Dalton-Morgan, J.	Conception and design of the project: 0% Analysis and interpretation: 5% Drafting significant parts (contributing to interpretation): 0%
Author Schwinghamer, M.W.	Conception and design of the project: 5% Analysis and interpretation: 5% Drafting significant parts (contributing to interpretation): 0%
Author Crew, K.	Conception and design of the project: 0% Analysis and interpretation: 5% Drafting significant parts (contributing to interpretation): 0%
Author Schenk, P. M.	Conception and design of the project: 5% Analysis and interpretation: 0% Drafting significant parts (contributing to interpretation): 5%
Author Geering A.D.W.	Conception and design of the project: 45% Analysis and interpretation: 15% Drafting significant parts (contributing to interpretation): 40%

Moyle, R., Pretorius, L. S., Shuey, L. S., Nowak, E., & Schenk, P. M. (2018). Analysis of the Complete Genome Sequence of *Cucumber mosaic virus* Strain K. Genome announcements, 6(7), e00053-18.

Contributor	Statement of contribution
Author Moyle, R. L.	Conception and design of the project: 45% Analysis and interpretation: 65% Drafting significant parts (contributing to interpretation): 61%
Author Pretorius, L. (Candidate)	Conception and design of the project: 40% Analysis and interpretation: 20% Drafting significant parts (contributing to interpretation): 30%
Shuey, L.	Conception and design of the project: 5% Analysis and interpretation: 10% Drafting significant parts (contributing to interpretation): 2%
Nowak, E.	Conception and design of the project: 5% Analysis and interpretation: 5% Drafting significant parts (contributing to interpretation): 2%
Schenk, P.M.	Conception and design of the project: 5% Analysis and interpretation: 0% Drafting significant parts (contributing to interpretation): 5%

Pretorius L, Moyle RL, Dalton-Morgan J, Hussein N, Schenk PM. (2016). Complete Nucleotide Sequence of an Australian Isolate of *Turnip mosaic virus* before and after Seven Years of Serial Passaging. Genome Announcements, 4(6):e01269-16.

Contributor	Statement of contribution
Author Pretorius, L. (Candidate)	Conception and design of the project: 75% Analysis and interpretation: 65% Drafting significant parts (contributing to interpretation): 85%
Author Moyle, R. L.	Conception and design of the project: 20% Analysis and interpretation: 15% Drafting significant parts (contributing to interpretation): 10%
Author Dalton-Morgan, J.	Conception and design of the project: 0% Analysis and interpretation: 5% Drafting significant parts (contributing to interpretation): 0%
Author Hussein, N.	Conception and design of the project: 0% Analysis and interpretation: 15% Drafting significant parts (contributing to interpretation): 0%
Author Schenk, P. M.	Conception and design of the project: 5% Analysis and interpretation: 0% Drafting significant parts (contributing to interpretation): 5%

Contributions by Others to the Thesis

Professor Peer Schenk and Dr Andrew Geering assisted with experimental design, data analysis and manuscript editing. Dr Richard Moyle assisted with laboratory supervision, experimental design, data analysis and manuscript editing. Dr Jessica Morgan-Dalton helped with bioinformatics.

Statement of Parts of the Thesis Submitted to Qualify for the Award of another Degree

None.

Acknowledgements

Life's uncertain; eat dessert first and only do a PhD after some serious consideration. It is emotionally draining, and a strong support team is vital for success. Therefore, I'd like to thank all those who prevented me from falling apart during these past "few" years.

Firstly, I'd like to thank Professor Peer Schenk for giving me the opportunity to study in his lab, appreciating my "art" by allowing me to draw in lab meetings, understanding my sense of humour and most importantly purchasing the coffee machine for the office!

To Dr Andrew Geering, thank you for agreeing to be part of my supervisory team and for all your valuable viral knowledge and understanding what an uphill battle PhD's can be.

Huge thanks to Dr Richard Moyle who was a late edition to my supervisory team but who I feel single-handedly saved my PhD and gave me hope that I could and would finish with something to be proud of. He has always supported and believed in me. I have learnt so much from him, making me a better scientist and an even more avid listener of 97.3 (thank you 97.3 for supplying the soundtrack).

To incredible Lilia, even though you weren't part of my supervisory team you were part of my emotional support team. I cannot thank you enough for always being there for me and providing the tissues. I loved sharing an office with you - you were my favourite procrastination excuse.

Thank you, Katia for being my "Lilia-patch" allowing me to rant and procrastinate to my heart's content. I will miss the smell of that Thai fish curry which was a sensory delight in my opinion. Thank you for laughing at my not-so-politically-correct jokes.

I'd like to thank Dr Brendan Kidd, who is an exceptional human being and scientist, for never being too busy to help with lab work or writing. To all my peeps in the Schenk lab, you guys rock! You made the daily grind bearable but more importantly you all laughed at my jokes. Special thanks to Hooman (who understood my pain), Vivian (who got my office), Elvis (who always went the extra mile), Sharifa (who loves my not so loving cats and for being a ray of sunshine) and Ali (who promised me a million dollars - once he's a millionaire). I would love to thank you all individually but too many trees would need to die.

To my Liam, I love you so incredibly much and can't thank you enough for being my sanity and for always making me laugh. Even though this PhD hasn't been the smoothest, you were always there to give me the strengthen I needed to continue. This PhD and our relationship started at the same time, I'm just happy the former is finally ending, and our relationship survived! You're the best!

Thank you for all the love and support from my family both here and in South Africa. I specifically want to thank Lexington, Trev and my mum. You are all such inspirational people and I'm so honoured to be part of the Cuss family. To Lexington, thank you for being the greatest sister and the most loyal, loving, cruelty-free vegan/soon-to-be-lawyer. I will always cherish the "honours times"! To my dear, Trev, whose strength and determination helped me understand what really matters - tea first, everything else second.

Lastly, I'd like to give the most special thanks to my mom, who has always made me strive to do my best even though I couldn't tell the time or spell my surname until I was 9 - in my defence it is a stupidly long surname! Just as I conquered that battle (or as you like to say slayed that dragon), with you by side I have finally conquered the PhD in the same fashion! That is why I dedicate this thesis to you, because you have made me who I am today and I would never have achieved as much as I have if it wasn't for you!

Financial Support

This research was supported by an Australian Postgraduate Award (APA) Scholarship.

Keywords

Biogenesis pathway, defence pathways, dual-luciferase assay, gene expression, jasmonic acid, *med18*, small RNAs, *Turnip mosaic virus*, viral evolution.

Australian and New Zealand Standard Research Classifications (ANZSRC)

060702 Plant Cell and Molecular Biology 20%

060704 Plant Pathology 70%

060705 Plant Physiology 10%

Fields of Research (FoR) Classification

0607 Plant Biology 60%

0605 Microbiology 25%

0604 Genetics 15%

Preface

The plant-virus relationship is a product of evolution, evident in the intricacies of interactions. One of the first coincidental documented plant-virus interactions was during the seventeenth century when tulips were all the rage. Tulip mania was the term for this period in the Dutch Golden Age as tulip bulbs were a hot commodity and sold for unbelievable prices. Unbeknown to tulip enthusiasts, the flowers which were considered most valuable were in fact infected with a virus and the flares of assorted colours intermingled in the petals were symptoms making viruses economically significant.

Today our idea of why viruses are economically important is slightly different; the words plant virus are usually associated with millions of dollars of agricultural losses. It is true that plant viruses are extremely damaging to the agricultural economy; even so it is amazing to think that something as microscopic can have such a macroscopic effect.

The study of plant virology began just over a century ago. Since then we have made great progress, however we are still in the process of understanding exactly how the plant-virus relationship works, as it's complicated.

This study investigates virus-induced defence modulation during the plant-virus interaction, focusing on *Turnip mosaic virus* that belongs to the same genus as the viruses responsible for the flower-breaking symptoms of the valuable tulips during tulip mania as illustrated on the front cover. The only kind of flower-breaking that *Turnip mosaic virus* does is illustrated below.

Contents

List of Figures

List of Tables

List of Abbreviations

(+/-)ssRNAs (Positive/negative sensed) single-stranded RNA

6K1 6kDa protein 1

6K2 6kDa protein 2

ACMV African cassava mosaic virus

AGO Argonaute protein

AltMV *Alternanthera mosaic virus*

amiRNAs Artificial miRNAs

BAK1 BRI1-associated kinase 1

BBTV Banana bunchy top virus

BMV Brome mosaic virus

BSV Banana streak virus

BYDV Barley yellow dwarf virus

CaMV Cauliflower mosaic virus

CI Cylindrical or cytoplasmic inclusion protein

CMV Cucumber mosaic virus

Col-0 Columbia ecotype 0

CP Coat protein

DAF The Department of Agriculture and Fisheries

dai Days after inoculation

DCL Dicer-like protein

dsDNA Double-stranded DNA

dsRNAs Double-stranded RNA

eIF(iso)4E Eukaryotic Initiation Factor

Er Endoplasmic reticulum

ERF1 ETHYLENE RESPONSE FACTOR 1

ERF6 ETHYLENE RESPONSE FACTOR 6

ET Ethylene

GRSV Groundnut ringspot virus

hai Hours after inoculation

HC-Pro Helper component protease

HR Hypersensitive response

HYL1 HYPONASTIC LEAVES1

ISR Induced systemic resistance

JA Jasmonic acid

kb Kilobase

kDa Kilodalton

MDMV Maize dwarf mosaic virus

MED18 Mediator subunit 18

Minutes min

miRNA MicroRNA

mRNA messenger RNA

MSV Maize streak virus

natsiRNAs Natural antisense siRNAs

NB-LRR Nucleotide binding site, leucine-rich repeat

NIa/b Nuclear inclusion protein a/b

nt Nucleotides

P1 Protein1

P3 Protein3

PAMPs Pathogen-associated molecular patterns

PCD Programmed cell death

PDF1.2 PLANT DEFENSIN 1.2

PdRp RNA-dependent RNA polymerase

Pol Polymerase

PPV Plum pox virus

PR Pathogen related

PR1 PATHOGENESIS-RELATED 1

PR5 PATHOGENESIS-RELATED 5

pre-miRNA Precursor miRNA

PTI Pattern-triggered immunity

PVX Potato virus X

PVY Potato virus Y

qRT-PCR Real-time quantitative reverse transcription PCR

R genes Resistance genes

rasiRNAs Repeat-associated siRNAs

RISC RNA-induced silencing complex

RNAi RNA interference

RNA-Seq RNA sequencing

ROS Reactive oxygen species

RTBV Rice tungro bacilliform virus

RYMV Rice yellow mottle virus

SA Salicylic acid

SAR Systemic acquired resistance

SE SERRATE

Seconds s

siRNAs Small interfering RNA

SPFMV Sweet potato feathery mottle virus

sRNAs Small RNAs

ssDNA Double-stranded DNA

tasiRNAs Trans-acting siRNAs

TCSV Tomato chlorotic spot virus

TMV Tobacco mosaic virus

TSWV Tomato spotted wilt virus

TuMV Turnip mosaic virus

TYLCV Tomato yellow leaf curl virus

UV Ultraviolet

VPgs Viral protein-linked genomes

VSP2 VEGETATIVE STORAGE PROTEIN

VSRs Viral suppressors of RNA silencing

WT Wild type

Research Questions and Aims

Viruses are successfully able to infect most living organisms, and as such have been classified as one of the most important disease-causing agents (Wang *et al*. 2012). Viruses are known to cause significant economic damage to the agricultural industry as all crops are potential hosts, reducing productivity and yields throughout the world (Oerke *et al.* 2012).

This project focuses on virus-induced defence modulation during plant-virus interactions; identifying essential interactions enabling disease or infection. As the plant-virus relationship has evolved in a gene-for-gene manner, interactions have become increasingly important in the search for durable resistance. However, these interactions are not straightforward as environmental factors have a significant influence (Lannou 2012). Nevertheless, studying these interactions is invaluable as once the pathogen's "attack strategy" is known, a counter-attack could be implicated and perhaps prevent future outbreaks. Two model plant systems, *A. thaliana* and *N. benthamiana*, are used to study plant-virus interactions using viruses from different families. It is hypothesised that viruses can manipulate the plant's gene expression by various means creating a more favourable environment allowing for successful establishment and systemic spread. The following aims were developed to address this question:

Aim 1: Fully sequence the genomes of plant virus isolates used in the study, including sequencing TuMV before and after seven years of serial passage. To determine the rate of mutation and how this strain relates to other strains in Australia and throughout the world.

Aim 2: To investigate the role of defence pathways in viral infection.

Aim 3: To determine if the mediator complex subunit 18 has an effect in viral infection.

Aim 4: To determine whether viRNAs can target host gene transcripts based on abundance and homology.

This thesis has been structured as follows: the first section is a review paper focusing on the study of virology, viral vectors and replication, plant defences and viral suppressors of plant defence and finally engineered resistance. A methodology section is included which details the methods used throughout this study.

Chapter 1 presents the sequencing and phylogenetic analysis of the TuMV viral isolates used in this study. *Turnip mosaic virus* has a positive sRNAs genome which is prone to mutation and recombination (Roossinck 2003). Therefore, identifying its changes over the past 7 years is important to establish its mutation rate (Aim 1). Chapter 2 focuses on the sequencing of other viral isolates specifically CaMV and CMV (Aim 1). Chapter 3 introduces the early defence responses which

highlights the importance of the JA pathway (Aim 2). Chapter 4 focuses on the mediator complex through the study of *med18* plants infected with different viruses to determine the viral expression levels when certain defence pathways are not functioning correctly (Aim 3). Finally, chapter 5 identifies possible host genes in *A. thaliana* and *N. benthamiana* which could be targeted by TuMV viRNAs, confirming one through a quantitative transient expression assay (Aim 4). All results contribute to an increased understanding of the plant-virus relationship, possibly helping create robust management strategies and more durable engineered viral resistance for the future.

A literature review "Natural and Engineered Defence against Plant Viruses" published in Current Biotechnology is included as a PDF in appendix 2 as access to this article is by subscription.

Literature Review: Natural and Engineered Defenses against Plant Viruses

Lara-Simone Pretorius, Anahid Al-Amery, Richard L. Moyle, Peer M. Schenk*

Nexgen Plants Pty Ltd and Plant-Microbe Interactions Laboratory, School of Agriculture and Food Sciences, The University of Queensland, Brisbane, Queensland 4072, Australia

*Corresponding author

Abstract

Plants live in a dynamic environment subjected to frequent pathogenic and environmental stresses and have evolved diverse self-defense pathways for their survival. Many of these defense responses are energy-intensive and antagonistic against each other to enable prioritization of strong targeted responses. Viruses cause significant damage in the agricultural and natural world. They are obligate intracellular pathogens that depend on their host cell for replication and movement. While programmed cell death is an appropriate initial defense response to isolate viruses, RNA silencing or the RNAi pathway presents a highly effective defense mechanism that plants employ against virus attack. But RNA silencing is also essential for normal gene regulation within the host. RNA silencing largely dominates the literature on plant virus defense, but there are a number of other effective defense mechanisms that will be discussed. This review first covers different viral replication strategies based on genome organization, then presents an overview of known virus defense mechanisms and pathways, and finally highlights recent work and new strategies to engineer viral resistance in crop plants.

Keywords: Plant defense, plant virus; RNAi; small RNA; sustainable agriculture

Introduction

The study of plant viruses is more than a century old; however we are now aware that many disease symptoms described earlier than the discovery of the first virus were caused by viruses. In 1898 Martinus W. Beijerinck was the first to distinguish a virus, which he described as a contagious living fluid, from a bacterium which he noted was able to reproduce in living tissue [3]. According to many scientists, Beijerinck's discovery was the birth of plant virology as he coined the term virus to distinguish the infectious agent from bacteria.

From here, plant virology can be divided into five major ages [4]. First was the prehistoric age, before 1898 and prior to the recognition of the viruses as their own entity by Beijerinck. The biological age came between 1900 and 1935, during which time viral transmission was studied. The early 1930s saw the beginning of the biochemical and physical age where the first viruses were isolated, characterized and visualized. In this age it was confirmed that the genome of TMV was made up of nucleic acid and surrounded by coat proteins that encapsulate the genome in a rod shape. We are currently in the molecular age which began circa 1960, when the full sequence of the TMV coat protein was determined [5,6]. The sequencing of plant viruses is ongoing, increasing our knowledge of the relationship between viruses and their host. Initially, this required isolation of the virus and cloning of its nucleic acid, followed by Sanger sequencing. However, recently direct RNA sequencing (RNA-Seq) from infected tissue combined with powerful bioinformatics assembly platforms enabled the rapid portrayal of viral genomes in diseased plants.

As virtually every plant is a potential host, plant viruses are responsible for reduced productivity of crops throughout the world [1]. Their impact is often underestimated as viral outbreaks can be irregular and may only occur when environmental conditions are conducive to virus infection. Table 1 provides an overview of the plant viruses that are currently considered most important for both research and economic significance. Plants have developed numerous resistance strategies that may prevent an invading virus from replicating and spreading. Likewise, viruses have evolved to counter plant defense strategies, and when a virus successfully circumvents the plant's defense mechanisms, outbreaks of disease occur. This review provides an overview of various virus replication mechanisms, their transmission, early recognition and post-entry defense focusing on resistance genes and RNA silencing, and viral counter-defense strategies. Finally,- we discuss approaches to engineer resistance to plant viruses.

Table1: Examples of plant viruses considered most important for molecular plant pathology and food production

Viruses (shown in alphabetical order) were ranked by plant virologists based on their significance to the field of Molecular Plant

Virus	***Genome***	***Main vector***	***Affected crops***	***Countries***	***Estimated losses***	***Reference***
African cassava mosaic virus* (ACMV) *(Geminivirus)	ssDNA	Whiteflies	Cassava, tobacco (wide host range)	Worldwide	Yield losses from different countries estimated at 20-95%	[7-9]
Banana bunchy top virus* (BBTV) *(Nanovirus)	ssDNA	Aphids	Banana	Worldwide	In Australia losses have been reported to range between 90-95%	[10-12]
Banana streak virus* (BSV) *(Badnavirus)	dsDNA	Vegetative propagation, mealybugs, seed	Banana	Worldwide	Yield losses from different countries estimated at 6-90% depending on disease severity	[13,14]
Barley yellow dwarf virus* (BYDV) *(Luteoviruses)	+ssRNA	Aphids	Barley, oat, wheat, maize, rice (wide host range)	Worldwide	yield losses can be >35%	[15-18]
Brome mosaic virus (BMV) *(Bromovirus)*	+ssRNA	Nematodes and aphids	Wheat, oat, maize (wide host range)	Worldwide	Unknown	[19,20]
Cauliflower mosaic virus (CaMV) *(Caulimovirus)*	dsDNA	Aphids	Cauliflower, turnip, tobacco (wide host range)	Worldwide	Depending on crop losses range from 25-90%	[21,22]
Cucumber mosaic virus* (CMV) *(Cucumovirus)	+ssRNA	Aphids	Squash, melons, pepper, bean, tomato, lettuce, celery, spinach (wide host range)	Worldwide	On average crop losses are bewteen 10-20% however can be higher depending on aphid population	[23,24]
Maize dwarf mosaic virus/Sugarcane mosaic virus* (MDMV) *(Potyviruses)	+ssRNA	Aphids	Maize, sorghum, sugarcane	Worldwide	Depending on cultivar yield losses are estiamted between 15-79%	[25-31]
Maize streak virus* (MSV) *(Mastrevirus)	ssDNA	Leafhoppers	Wheat, millet, sugarcane oat, rye, (wide host range)	Africa	Crop losses range between 24-76%	[32,33]
Potato virusX (PVX) *(Potexvirus)*	+ssRNA	only mechanical	Potato	Worldwide	Crop losses range between 10-36% for different countries	[34,35]
Potato virus Y (PVY) (*Potyvirus)*	+ssRNA	Aphids	Potato, tobacco, tomato, pepper (wide host range)	Worldwide	Yield losses average at 57% however can increase to 71% when crops are also infected with PVX	[36-39]
Plum pox virus (PPV) (*Potyvirus)*	+ssRNA	Aphids	Stone fruit species (wide host range)	Worldwide	Suspetable cultivars can occur losses between 80-100%	[40,41]
Rice yellow mottle virus* (RYMV) *(Sobemovirus)	+ssRNA	Leafhopper, beetles	Rice	Africa	Yield losses reported from different countries range between 58-97% depending on the cultivar	[42-45]
Rice tungro bacilliform virus* (RTBV) (*Tungrovirus)	dsDNA	Leafhoppers	Rice	Asia	Up to 74% losses have been reported	[46,47]
Sweet potato feathery mottle virus* (SPFMV) *(Potyvirus)	+ssRNA	Aphids	Sweet potato	Worldwide	A yield loss across African countries was report torange from 57-90%	[48-50]
Tobacco mosaic virus(TMV) *(Tobamovirus)*	+ssRNA	Aphids	Tobacco, tomato, pepper, cucumber (wide host range)	Worldwide	When associated with other viruses yield losses have been as high as 90%.	[51]
Tomato spotted wilt virus (TSWV) *(Tospovirus)*	-ssRNA	Thrips	Tomato, watermelon, capsicum, zucchini (wide host range)	Worldwide	Crop yields have been report to be between 50-100% depending on the crop	[52,53]
Tomato yellow leaf curl virus (TYLCV) *(Luteovirus)*	ssDNA	Whiteflies	Tomato, pepper, bean, eggplant, potato, tobacco (wide host range)	Worldwide	Yield loss of 30-100% report for different countries	[54-56]

Pathology [57] or economic importance (shown in bold) [58]. Their genomes include double-stranded (ds), or positive (+) or negative (-) single-stranded (ss) RNA or DNA.

Viral Replication

Single-stranded RNA viral replication is thought to involve replication complexes formed from the host's resources, creating a safe environment protected from RNA nucleases. The complexes are derived from a number of various host organelles, including endoplasmic reticulum (ER), vacuole and mitochondrial membranes [67]. This method of replication is highly conserved amongst plant and animal positive-stranded RNA viruses where RNA acts as the template for both replication and translation [68]. However, as the RNA dependent RNA polymerase RdRp is not encapsulated along with the viral genome, the replication of the virus cannot proceed until the RNA genome is translated to produce its specialized polymerase. Host factors are thought to translate the viral genome as well as help in other important processes dictated by the virus [69]. Typically, once the RdRp is produced it is able to produce a negative template creating multiple copies of positive RNA genomes which will ultimately be encapsulated by the viral protein coat and spread through the plant via movement protein [70]. These replication complexes are actually thought to be the sites of translation as well as replication and transcription, as shown by Cotton *et al.* [71]. Vesicles were also shown to be mobile due to their interaction with microfilaments which are thought to traffic the vesicles from cell to cell. This form of RNA replication is specific to single-stranded positive RNA viruses. Replication of negative and ambisense viruses is slightly different due the orientation of the genome, though they also replicate in the cytoplasm forming viral replication complexes.

Plant *Rhabdoviruses* are single stranded negative sensed RNA virus which form part of the Rhabdoviridae family and Mononegavirales order. Rabdoviruses are divided into two groups based on where the virus replicates; as suggested by the names, Cytorhabdovirus replicate in the cytoplasm and Nucleorhabdovirus replicates in the nucleus. Both code for five structural proteins and follow similar replication styles. Once inside the host cell the viral genome will continue to be transcribed by the viral polymerase until the concentration of viral proteins is considered sufficient to switch to replication (Morphology, genome organization, transcription and replication of rhabdoviruses). The negative strand is then used to generate a positive strand template for replication. The negative strands are encapsidated by the N protein along with other viral proteins creating new viral particles (Plant rhabdoviruses: new insights and research needs in the interplay of negative-strand RNA viruses with plant and insect hosts).

Tospoviruses are described as spherical envelopes, containing a tripartite ssRNA genome as well as a few molecules of RdRp [88]. This attribute of having their RdRp already translated means that viral replication can begin without any help from the host. The benefit of having RdRp enzyme encapsulated in the virions means the virus is able to replicate in its vector as well as in its host. The translation of genome segments is performed by RdRp through a mechanism known as cap snatching,

where the host mRNA is cleaved off by 10-20 nucleotides at the 5' end which is then used by the RdRp as a capped leader to initiate viral transcription [91]. This mechanism of capping works as a negative-sense genome (L segment) which is transcribed from 5' termini and truncated at the 3' end in accordance with the template, while the ambisense genomes (S and M segments) are transcribed from the intergenic region located in the middle of the strand also in accordance with the 5'-3' direction [92]. Interestingly, due to their orientations, the segments are able to act as template for transcription and replication, though the switch from transcription to replication by RdRp is poorly understood. [93].

Double-stranded RNA (dsRNA) viruses are also thought to replicate in the cytoplasm in replication complexes comprised of viral proteins, forming a protective boundary allowing the virus to replicate without evoking the plant's viral defense. Once the viral transcriptase is activated, its replication is similar to that of single-stranded RNA viruses, however there is no need to create a template strand due to their double-stranded nature [72]. The positive strand is translated into viral proteins then encapsidated by coat protein making new viral particles. The negative strand is then synthesized inside the viral particles from the positive strand, completing the viral replication cycle.

Single stranded DNA virus such as geminiviruses are a large family of viruses consisting of many economically significant viruses possessing either monpartite or bipartite genomes. Geminiviruses have a unique replication method called the rolling circle. This is where a negative strand is synthesized from the positive strand creating a double stranded intermediate known as the replicative form (Geminiviruses: Models for Plant DNA Replication, Transcription, and Cell Cycle Regulation). Due to the small number of proteins their genomes transcribe they are highly dependent on the host machinery and as such have become masters at reprogramming the plant cells creating the perfect replicating environment.

Caulimoviruses, such as *Cauliflower mosaic virus* (CaMV) possesses a circular, dsDNA molecule, roughly 8 kb in length [74]. CaMV replicates via reverse transcription which results in discontinued sequence at the 5' ends creating triple-stranded overlaps; the position and number varies depending on the strain. Similar to potyviruses, CaMV replication takes place in replication complexes located in the cytoplasm [75,76].

Plant Defense Mechanisms against Viruses

Early Recognition

The plant's first line of defense against all pathogens is its preformed and inducible defenses. These work particularly well against viruses as they are unable to breach the cuticle and cell wall without wounding providing an entry point [94]. Wounding can be caused by mechanical damage via wind, herbivores, humans or insects. A small percentage of viruses are able to be vertically transmitted via seeds [95]; however the majority of plant viruses usually have a specific vector. Vectors are described as mobile organisms whose feeding behavior allows the virus to gain access to the cytoplasm, via wounding caused during feeding. According to Brault *et al.* [96], aphids are responsible for the transmission of approximately 30% of all plant viruses; this could be due to their specific feeding habits. Aphids have perfectly engineered mouth parts, comprising a thin stylus which probes and "taste-tests" the plant mesophyll cells without damaging them, prior to feeding on the phloem [97]. Viruses can be transmitted to the plant as well as contracted by the aphid at any point during probing or feeding. This feeding habit allows for viruses with tissue specificity to be transmitted to their desired location. Other insect vectors of virus transmission include whitefly, thrips, leafhoppers, beetles and mealybugs.

Nematodes are another example of a viral vector, however it was not until 1958 that it was proven that the nematode *Xiphinema index* was the natural vector of the soil-borne nepovirus causing fanleaf in grapevines. Similar to aphids, nematodes are able to puncture the plant's cell wall with their unique feeding equipment, sucking out its contents and infect the plant or be infected with the virus at any time during this process. There are only a few viruses which can be transmitted by nematodes, as there is a high degree of specificity between the vector and the virus due to the vector's ability to retain the virus. The nematodes which are able to retain virus particles are equipped with unique sites located at the stylets [98]. There is also a certain viral molecular determinant which is also related to the specificity between the virus and its vector which was found to be a vital part in successful transmission. For example, the capsid protein of nepoviruses was found to be the only transmission determinant [99,100], while other nematode-transmitted viruses, such as the tobraviruses, were found to have multiple transmission determinants depending on the strain or isolate [101].

Once past the plant's preformed defenses, a virus is not easily detected by the plant. Plants can usually identify pathogens by their pathogen-associated molecular patterns (PAMPs), followed by a cascade of inducible defenses aimed at preventing further invasion from the pathogen [102]. However, as there have been no identifiable PAMPs associated with viruses thus far, the plant's main lines of defense are inducible defense-related proteins and RNA silencing [103,104]. Even though viruses cannot be recognized by PAMPs, viruses are still able to induce pattern-triggered immunity (PTI)

[96]. A recent study by Kørner *et al.* [105] proved that BRI1-associated kinase 1 (BAK1), an important regulator of PTI, had a role in antiviral defense as the *Arabidopsis* mutants had increased susceptibility to three different compatible RNA viruses.

Plants are also able to activate resistance (*R*) genes from the nucleotide-binding site-leucine-rich repeat (NB-LLR) family as they are able to recognize specific non-viral effectors and viral avirulence (*avr*) proteins. *R* genes act intracellularly; recognizing pathogens, viral or non-viral, resulting in effector-triggered immunity [106]; similar to PTI but more intense, as this form of immunity triggers a cascade of signals leading to the production of reactive oxygen species (ROS), hypersensitive response (HR) and programmed cell death (PCD) [102]. This response is an attempt to locally counteract the virus by restricting its spread throughout the plant and requires activation of the salicylic acid (SA) pathway that can lead to systemic acquired resistance (SAR).

Post-Entry Events

Once the virus has gained access to the cell, the events that occur next are a series of complex host and pathogen interactions and whether the host is resistant to the pathogen will be determined by its response. Hosts that are resistant to the pathogen lead to a visible phenotype, usually necrosis, however hosts that are disease resistant have no symptoms but the pathogen may still invade it as if the host was susceptible. This can also be seen as the compatibility between the host and the virus. An incompatible interaction is when the host is able to recognize its attacker resulting in the induction of one of the host defense mechanisms. In contrast, a compatible interaction is when the host R gene corresponding to the pathogen avr factor is missing or vice versa, resulting in successful infection [107]. Interestingly, both these interactions have a specific HR which is induced; the incompatible reaction usually leads to HR induced by R genes, while a compatible response results in symptoms and may include a systemic necrosis response [108]. A systemic necrosis response is induced much later in the infection stage than HR and, unlike HR, it is usually lethal. The response was first described by Holmes [109] where it was observed in *Nicotiana rustica* plants infected with TMV. There is relatively little known about systemic necrosis response compared to HR, however recent studies have found their molecular and biochemical processes to be similar [110,111]. Similarly, there is also a contrast in knowledge on the global expression of genes as well as the host and viral factor interactions during the relatively well studied compatible reaction, compared to those expressed during an incompatible reaction. This is because compatible disease-causing plant-viral interactions often result in economic losses in crops. But it is the incompatible interactions that should be studied in more detail because it is from these interactions that we can learn about the plant's immune system that successfully defends the plant.

Certain incompatible reactions require specific pathogenic and host factors as they are essential for successful infection and colonization. One such example is the interaction between the Potyvirus avirulence factor, viral protein-linked genomes (VPgs) protein, and the Eukaryotic Initiation Factor eIF (iso)4E found in *Arabidopsis thaliana* [112-114]. Numerous studies have shown that certain potyviral strains are unable to infect or colonize a host when the eIF (iso)4E gene has certain point mutations and it is suspected that these mutations create natural recessive resistance genes as the mutations do not affect the cellular role of the protein but confer a broad-range resistance specific to potyviruses [115-117]. The eIF4E gene family is highly conserved in all eukaryotes [118] and four family members have been identified in *Arabidopsis*; eIF4F, eIF3, eIF4B and eIF2 [119]. The mechanism of how this interaction works or why it is essential is still unclear, however a recent paper by Contreras-Paredes *et al*. [120] tested three potential reasons why the interactions between these two proteins are necessary. First, as the eIF (iso)4E protein has a major role in the initiation of translation, the virus may require the protein for RNA translation, or secondly the interaction may serve to protect the viral mRNA from degradation while in the cytoplasm. Thirdly, it is thought that the interaction may be beneficial, as the virus would have access to long-distance transport via the microtubules while it is part of the ribonucleoprotein complex. Their results suggested that the eIF (iso)4E was needed for replication of *Turnip mosaic potyvirus* (TuMV), as no viral RNA could be found in the eIF (iso)4E knockout plants. However when infecting the plants with an isolate of *Tobacco etch potyvirus* with the ability to overcome this resistance, viral RNA was found. They also found that the TuMV VPgs interacted with the eIF4E as well as the eIF (iso)4E. However the coat protein of TuMV was found to only interact with the eIF (iso)4E and as the coat protein is a vital component of viral replication and spread, it was concluded that eIF (iso)4E is essential for the systemic spread of the virus.

_R_ Genes

The recognition between the host's *R* gene products and the pathogen's *avr* products causes a chain reaction of signals which may result in the inhibition of further growth of the pathogen. For the host to successfully prevent the spread of the pathogen, the corresponding *R* gene to the specific *avr* product needs to be present. If either is absent or the pathogen's *avr* product is not recognized by any of the host's *R* genes, the pathogen can successfully infect and cause disease in its host. It is assumed that *R* genes are needed to increase the response time to pathogen attack by efficiently and effectively activating defense mechanisms of both innate and adaptive immunity pathways [121].

An example of an *R* gene known to function in viral defense is the *Rx* gene conferring resistance against PVX. The *Rx* gene is part of the nucleotide binding site, leucine-rich repeat (NB-LRR) class of proteins, which is the largest class of R proteins, which are thought to act in the cytoplasm as well

as in the intracellular space. Unlike most *R* genes of this class, the activation of the *Rx* gene does not lead to HR. Instead this response was found to involve epistasis [122]. The NB-LRR *R* genes are highly evolved; functioning exclusively in plant resistance and defense, and most of the *R* genes which confer viral resistances are from this class of proteins.

Greater understanding is needed to clarify how these *R* gene products interact with the pathogen products to cause a cascade of signals resulting in the correct response, as there is still speculation surrounding many mechanisms and functions. It is clear that *R* genes are activated by *avr* products which are thought to act as signaling molecules. This has been demonstrated by excluding the pathogens and exposing the plant to just the *avr* product [124].

RNA Silencing

RNA silencing is one of the plant's main defenses against viruses as it acts on dsRNA that typically forms as an intermediate in viral replication of RNA viruses. Viruses have evolved to encode counter attacks which subsequently result in continuous co-evolution where attack and defense can reach several levels. RNA silencing begins with the production of dsRNA (viral or host) during the replication of RNA. Viral dsRNA can act as a signaling molecule, evoking the plant's gene silencing machinery in an attempt to destroy the foreign RNA and defend itself against the perceived viral attack. RNA silencing not only acts as the plant defense mechanism against viruses but, for example, also plays a role in the regulation of genes, chromatin condensation, and suppression of transposable elements within the plant's genome [103,126]. RNA silencing is mediated by dsRNA which has a high level of consensus to the target sequence. The RNase III-like Dicer (DCL) protein cleaves the dsRNA, forming predominately short 21-24 nucleotide small RNAs (sRNAs). The majority of sRNAs are classified as either short interfering RNAs (siRNAs) or microRNAs (miRNAs), depending on the fragments they were derived from [127]. There is much speculation around the difference between siRNAs and miRNAs, in function as well as where they are derived from. However, it is generally accepted that miRNAs are derived from an imperfect stem-loop precursor while siRNAs are derived from perfect RNA duplexes [128]. .

With regards to function, siRNAs are thought to have a crucial role in plant antiviral defense [2], while miRNAs are involved in development, signal transduction, the degradation of protein, as well as response to both biotic and abiotic stressors [129]. The basic processes for both sRNAs are similar; after the dsRNA or hairpin RNA is cleaved by dicer proteins, it is loaded into an enzyme which is part of the Argonaute (AGO) family, forming the RNA-induced silencing complex (RISC). The sRNA sequence is then used as a guide to direct and perform the different roles of RNA silencing; RNA degradation, gene regulation, or DNA methylation of sequences homologous to the target sequence [2,126-127].

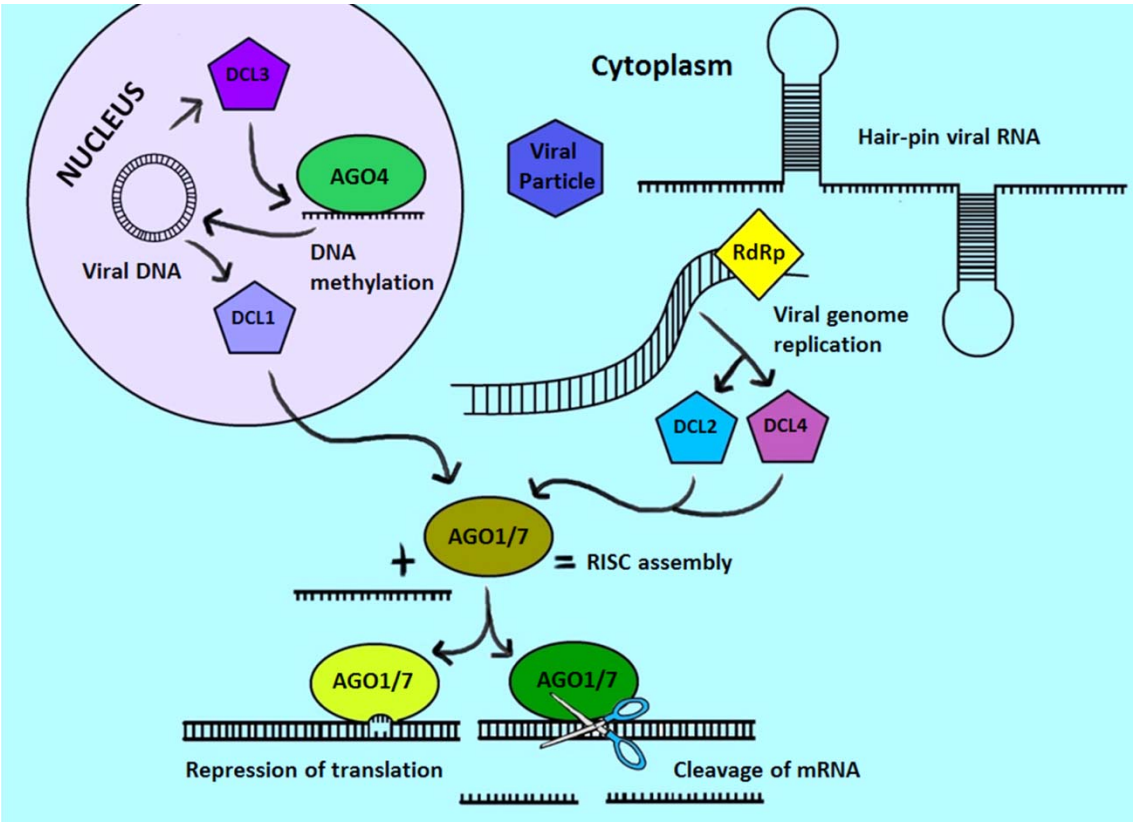


Figure 1. RNA silencing of viral products occurs in both the nucleus and cytoplasm depending on where the virus replicates. DNA viruses usually replicate in the nucleus while RNA viruses are believed to replicate in the cytoplasm. DCL3 and DCL1 were found to have a role in antiviral activity against DNA viruses, with DCL3 creating 24 nt siRNAs which methylate regions of the DNA virus genome. While DCL1 produces viral dsRNA which is presumed to be processed similarly to the endogenous miRNAs, exported to the cytoplasm, incorporated with an AGO to form RISC and then used to target viral mRNA. dsRNA from RNA viruses can be produced either through hairpins in the viral genome or viral replication. These are then acted upon by either DCL2 or DCL4 which produce 22 nt and 21 nt siRNAs, respectively. DCL4 is considered to be more dominant, as 21 nt siRNAs are more abundant. Similarly, these are incorporated into the RISC molecule which then targets viral mRNA, resulting in silencing either through cleavage or inhibition of translation.

Viral Counter Defense Mechanisms

As RNA silencing is the plant's main defense against viruses, it is not surprising that viruses have evolved their own counter defense strategies, specifically viral suppressors of RNA silencing. Counter defense mechanisms are intended to disrupt and inhibit the plant's RNA silencing pathways which ultimately prevent the plant from defending itself. Viral suppressors of RNA silencing have been shown to prevent RNA silencing in a number of ways, ranging from inhibiting crucial RNA silencing enzymes such as DCL [130] and AGO [131], to modification of its own as well as the plant's genome [132]. Through these actions as well as others, viruses are able to weaken the plant's RNA silencing pathway and use it to their own advantage as this interference will affect the host's mRNA production and ultimately its productivity.

Some of these viral counter defense mechanisms can be seen phenotypically in the disease symptoms, for example Smith *et al.* [133] reported that the yellowing symptoms from *Cucumber mosaic virus* (CMV) were the result of a siRNA derived from a viral satellite which silenced the chlorophyll biosynthesis gene, *CHLI*. Viral satellites and viroids appear to play a major role in the disease symptoms of some viruses, and it would seem they act through RNA silencing, expressing sRNAs with high similarity to target genes within the host genome, whereby inhibiting their expression [134]. Conversely, some plant viruses may purposely induce certain host miRNAs, for example the p19 protein of *Cymbidium ring spot virus* was reported as an inducer of the host miR168 which prevents the expression of *AGO-1*, required for antiviral defense and the correct functioning of RISC [135]. The p19 protein is also involved in preventing siRNAs from binding to an AGO protein. The study by [136].found that p19 binds to certain 21-nt ds siRNAs which could inhibit the signal transduction activate by siRNAs as well as systemic virus-induce posttranscriptional gene silencing. The binding also prevents the formation of an active RISC molecule and preventing the plants antiviral defense mechanism from functioning correctly [137].

Some viruses are also able to express their own sRNAs; viral miRNAs are common amongst animal DNA viruses and have been shown to actively regulate the host sRNA metabolism [137]. Some speculate that plant RNA viruses do not produce miRNA because their replication takes place in the cytoplasm and not in the nucleus, as with certain DNA viruses, therefore they do not have access to the correct cellular machinery.

Interestingly, viral suppressors of RNA silencing are reconditioned by the plant which elicits a counter-counter defense that is processed to work in a gene-for-gene manner according to Li *et al.* [142]. In their paper they demonstrate that the plant counteracts the 2b CMV viral suppressor of RNA silencing that otherwise prevents the plant from defending itself by initiating RNA silencing.

Engineered Defense against Plant Viruses

Over the past few decades a number of strategies have been used to confer virus resistance, specifically through plant genetic modifications. These include the expression of a viral coat protein sequence which was found to protect plants against the virus, from which the coat protein sequence was derived [144] and that this resistance may be lost when infected with another virus [143]. This RNA interference (RNAi)-based approach can use any part of the viral genome and works best when constructs are in the form of dsRNA [146], for example in the form of a hairpin-shaped RNA fitted with a spacer region. Another early strategy used to confer resistance was the introduction of a truncated replicase gene from the virus of interest into the plant genome which resulted in prolonged specific resistance [147]. An example of broad range resistance was demonstrated by Prins *et al.* [148] where viral nucleoprotein genes were stably expressed in transgenic tobacco plants for three tospoviruses, specifically *Tomato spotted wilt virus* (TSMV), *Tomato chlorotic spot virus* (TCSV) and *Groundnut ringspot virus* (GRSV). The tobacco line which stably expressed all three viral nucleoproteins was found to display high levels of resistance to all three viruses. As our understanding of the plant's natural viral defense mechanism has improved, a number of ingenious resistance strategies have been devised, based on RNA silencing (for a thorough review on this topic see Ramesh *et al.* [149]). Recently, a popular method for engineering resistance against viruses appears to be via the use of artificial miRNAs (amiRNAs). In this approach, the mature miRNA sequence of a plant miRNA is replaced with sequence targeting viral mRNA sequences. However, the high variability and genetic plasticity of RNA virus populations could potentially limit the success of engineered transgenic amiRNA and RNAi-based resistance [140,141].

Antiviral RNAi strategies used in animals are designed using the RNA polymerase promoters II and III to either drive the expression of specific siRNAs [150] or they use the pre-miRNA backbone to generate amiRNAs directed against the virus [151]. Similarly, plant antiviral strategies are also reliant on siRNAs as well as amiRNAs; specifically long-pathogen-derived hairpin RNAs have been introduced into the plant that are then processed into siRNAs inducing post-transcriptional gene silencing of viral sequences. According to Smith *et al.* [146] this method has been demonstrated to confer 100% resistance when directed against viral pathogens. As mentioned above, viral resistance has also been engineered through the modification of miRNAs to target specific sequences using the pre-miRNA backbone. This amiRNA technique was found to be more specific than the hairpin RNAs [152]; by engineering the amiRNA to target a 21 nt sequence of the mature miRNA of a specific miRNA family, one could silence all sequence-related viral strains [153]. siRNAs are less specific, as a diverse range of siRNAs are usually produced from hairpin RNA, resulting in the potential for off-target effects, whereas amiRNAs, if carefully designed, target specific viral gene sequences thereby avoiding any off-target effects [154].

Another emerging antiviral defense strategy may come from the modulation of defense pathways. SA defense signaling is required for HR and PCD but parts of this pathway are inhibited by jasmonic acid (JA) signaling, a pathway required to prevent cell death to defend against necrotrophic pathogens. Both, SA and JA signaling are generally antagonistic, as SA upregulates, but JA downregulated ROS in plants. Hence, by upregulating the SA pathway or by downregulating the JA pathway, a more pronounced SA-mediated antiviral defense response is anticipated when viruses attack. A drawback of this promising antiviral approach could be that plants may become more susceptible to necrotrophic pathogens.

Conclusion

Plant viruses continue to pose major threats to agriculture. Their irregular emergence makes it hard to incorporate multiple, broad-spectrum virus resistance in breeding programs. The most promising approaches to control plant viruses include the use of resistance genes in breeding programs, as well as RNAi constructs (currently the most widely applied method), artificial miRNA, and the modulation of defense signaling pathways for engineered resistance. Many viruses co-exist in plants without causing real damage to plants. Therefore, further research in plant virology may focus on better identifying the mechanisms behind symptom development in plant cells, as these are the real issues causing yield losses. In many cases these could be caused by the plant mounting ineffective defense pathways that weaken, rather than strengthen plant cells when defending against viruses. Hence, there are great opportunities by engineering multiple, broad defense strategies against viruses to protect plants against multiple viral strains and species simultaneously. These may include RNAi against conserved virus sequences, several amiRNAs that target multiple regions and adequate, carefully-measured modulation of plant defense pathways.

References

[1] Oerke E-C, Dehne H-W, Schönbeck F, Weber A. Crop production and crop protection: estimated losses in major food and cash crops. Elsevier 2012.

[2] Wang MB, Masuta C, Smith NA, Shimura H. RNA silencing and plant viral diseases. Mol Plant-Microb Interact 2012; 25: 1275-85.

[3] Beijerinck M. Concerning a contagium *vivum fluidum* as cause of the spot disease of tobacco leaves. 1898

[4] Hull R. Comparative plant virology. 2009; Elsevier.

[31] Smith GR, Van de Velde R. Detection of sugarcane mosaic virus and Fiji disease virus in diseased sugarcane using the polymerase chain reaction. Plant Disease 1994; 78: 557-61.

[5] Anderer FA, Uhlig H, Weber E, Schramm G. Primary structure of the protein of Tobacco mosaic virus. Nature 1960; 186: 922-25.

[6] Tsugita A, Gish D, Young J, Fraenkel-Conrat H, Knight C, Stanley W. The complete amino acid sequence of the protein of tobacco mosaic virus. Proc Natl Acad Sci USA 1960; 46:1463.

[7] Hanley-Bowdoin L, Settlage SB, Orozco BM, Nagar S, Robertson D. Geminiviruses: models for plant DNA replication, transcription, and cell cycle regulation. Crit Rev Plant Sci 1999; 18: 71-106.

[8] Legg J, Owor B , Sseruwagi P, Ndunguru J. Cassava mosaic virus disease in East and Central Africa: epidemiology and management of a regional pandemic. Adv Virus Res 2006; 67: 355-418.

[9] Legg JP, Fauquet CM. Cassava mosaic geminiviruses in Africa. Plant Mol Biol 2004; 56: 585-99.

[10] Burns TM, Harding RM, Dale JL. The genome organization of banana bunchy top virus: analysis of six ssDNA components. J Gen Virol 1995; 76: 1471-82.

[11] Horser CL, Harding RM, Dale JL. Banana bunchy top nanovirus DNA-1 encodes the 'master'replication initiation protein. J Gen Virol 2001; 82: 459-64.

[12] Hu J-M, Fu H-C, Lin C-H, Su H-J, Yeh H-H. Reassortment and ncerted evolution in Banana bunchy top virus genomes. J Virol 2007; 81: 1746-61.

[13] Iskra-Caruana M-l, Chabannes M, Duroy P-O, Muller E. A possible scenario for the evolution of Banana streak virus in banana. Virus Res 2014; 186: 155-62.

[14] Dangl JL, Jones JD. Plant pathogens and integrated defence responses to infection. Nature 2001; 411: 826-33.

[14] Lockhart B. Banana streak badnavirus infection in Musa: epidemiology, diagnosis and control. ASPAC Food & Fertilizer Technology Center 1995.

[15] Ad, WAM, Rasochová L. Barley Yellow Dwarf Viruses. Ann Rev Phytopathol 1997; 35: 167-190.

[16] Lister RM, Ranieri R. Distribution and economic importance of barley yellow dwarf. Barley yellow dwarf 1995; 40: 29-53.

[17] Miller WA, Liu S, Beckett R. Barley yellow dwarf virus: Luteoviridae or Tombusviridae? Mol Plant Pathol 2002; 3: 177-83.

[18] Sadeghi S, Bjur J, Ingwell L, Unger L, Bosque-Pérez N, Eigenbrode S. Interactions between *Metopolophium festucae cerealium* (Hemiptera: Aphididae) and Barley yellow dwarf virus (BYDV-PAV). J Insect Sci 2016; 16: 21.

[19] French R, Ahlquist P. Characterization and engineering of sequences controlling in vivo synthesis of Brome mosaic virus subgenomic RNA. Journal of Virology 1988; 62: 2411-20.

[20] Shih DS, Kaesberg P. Translation of brome mosaic viral ribonucleic acid in a cell-free system derived from wheat embryo. Proc Natl Acad Sci USA 1973; 70: 1799-1803.

[21] Hoh F, Uzest M, Drucker M, Plisson-Chastang C, Bron P, Blanc S et al. Structural insights into the molecular mechanisms of cauliflower mosaic virus transmission by its insect vector. J Virol 2010; 84: 4706-13.

[22] Menissier J, Murcia G De, Lebeurier G , Hirth L. Electron microscopic studies of the different topological forms of the cauliflower mosaic virus DNA: knotted encapsidated DNA and nuclear minichromosome. EMBO J 1983; 2: 1067.

[23] Gallitelli D. The ecology of Cucumber mosaic virus and sustainable agriculture. Virus Res 2000; 71: 9-21.

[24] Palukaitis P, García-Arenal F. Cucumoviruses Advances in Virus Research 2003; 62: 241-323.

[25] Tosic M, Ford RE, Shukla DD, Jilka J. Differentiation of sugarcane, maize dwarf, johnsongrass, and sorghum mosaic viruses based on reactions of oat and some sorghum cultivars. Plant Disease 1990; 74: 549-52.

[26] Cronin S, Verchot J, Haldeman-Cahill R, Schaad MC, Carrington JC. Long-distance movement factor: a transport function of the potyvirus helper component proteinase. Plant Cell 1995; 7: 549-59.

[27] Wu J-Y, Ding J-Q, Du Y-X, Xu Y-B, Zhang X-C. Genetic analysis and molecular mapping of two dominant complementary genes determining resistance to sugarcane mosaic virus in maize. Euphytica 2007; 156: 355-64.

[28] Zhang Z-Y, Fu F-L, Gou L, Wang H-G, Li WC. RNA interference-based transgenic maize resistant to maize dwarf mosaic virus. J Plant Biol 2010; 53: 297-305.

[29] Luo Q, Ahmad K, Fu HY, Wang JD, Chen RK, Gao SJ. Genetic diversity and population structure of Sorghum mosaic virus infecting Saccharum spp hybrids. Annal Appl Biol 2016; 169: 398-407.

[30] Mollov D, Tahir M, Wei C, Kaye C, Lockhart B, Comstock J et al. First Report of Sugarcane mosaic virus Infecting Columbus Grass (*Sorghum almum*) in the United States. Plant Disease 2016; 100: 1510.

[32] Peterschmitt M, Quiot JB, Reynaud B, Baudin P. Detection of maize streak virus antigens over time in different parts of maize plants of a sensitive and a so-called tolerant cultivar by ELISA. Annal Appl Biol 1992; 121: 641-53.

[33] Shepherd, DN, Martin DP, van der Walt E, Dent K, Varsani A, Rybicki EP. Maize streak virus: an old and complex 'emerging'pathogen. Mol Plant Pathol 2010; 11: 1-12.

[34] Adams M, Antoniw J, Bar-Joseph M, Brunt A, Candresse T, Foster G et al. Virology Division News: The new plant virus family Flexiviridae and assessment of molecular criteria for species demarcation. Arch Virol 2004; 149: 1045-1060.

[35] KutnjakD, Silvestre R, Cuellar W, Perez W, Müller G, Ravnikar M. et al. Complete genome sequences of new divergent potato virus X isolates and discrimination between strains in a mixed infection using small RNAs sequencing approach. Virus Res 2014; 191: 45-50.

[36] Anfoka G, Abu-Obaida M, Altaleb M. First report of Potato virus Y strain N-Wilga affecting potato in Jordan. Plant Disease 2016; 100: 2176.

[37] Boonham N, Hims M, Barker I, Spence N. Potato virus Y from petunia can cause symptoms of potato tuber necrotic ringspot disease (PTNRD). Eur J Plant Pathol 1999; 105: 617-21.

[38] Chung BY-W, Miller WA, Atkins JF, Firth AE. An overlapping essential gene in the Potyviridae. Proc Natl Acad Sci USA 2008; 105: 5897-902.

[39] Radcliffe EB, Ragsdale DW. Aphid-transmitted potato viruses: the importance of understanding vector biology. Am J Potato Res 2002; 79: 353-86.

[40] Cambra M, Capote N, Myrta A, Llácer G. Plum pox virus and the estimated costs associated with sharka disease. EPPO Bull 2006; 36: 202-4.

[41] Candresse T, Cambra M. Causal agent of sharka disease: historical perspective and current status of Plum pox virus strains. EPPO Bull 2006; 36: 239-46.

[42] Abo M, M, Alegbejo A, Misari Sy, Misari S. An overview of the mode of transmission, host plants and methods of detection of Rice yellow mottle virus. Journal of Sustainable Agriculture 2000; 17: 19-36.

[43] Matthews R. Plant Virology 1991 3rd edit Acad Press, New York.

[44] Odedara O, Ademolu K, Ayo-John E. Prevalence of Rice yellow mottle virus (RYMV) on rice plants grown in selected farms in Ogun state: preliminary results. Nigerian J Biotechnol 2016; 31: 96-102.

[45] Oludare A, Tossou HT, Kini K, Silué D. Diversity of Rice yellow mottle virus in Benin and Togo and screening for resistant accessions. J Phytopathol 2016. 164: 924-35.

[46] Bunawan H, Dusik L, Bunawan SN, Amin NM. Rice tungro disease: From identification to disease control. World Appl Sci J 2014; 31: 1221-26.

[47] Ling K. Rice virus diseases. International Rice Research Institute Los Banos, Philippines 1972.

[48] Moyer J, Cali B. Properties of sweet potato feathery mottle virus RNA and capsid protein. J Gen Virol 1985; 66: 1185-9.

[49] Rännäli M, Czekaj V, Jones R, Fletcher J, Davis R, Mu L et al. Molecular characterization of Sweet potato feathery mottle virus (SPFMV) isolates from Easter Island, French Polynesia, New Zealand, and southern Africa. Plant Disease 2009; 93: 933-9.

[50] Tairo F, Mukasa SB, Jones RA, Kullaya A, Rubaihayo PR, Valkonen J. Unravelling the genetic diversity of the three main viruses involved in sweet potato virus disease (SPVD), and its practical implications. Mol Plant Pathol 2005; 6: 199-211.

[51] Scholthof K-BG. Tobacco mosaic virus: a model system for plant biology. Annu Rev Phytopathol 2004; 42: 13-34.

[52] Whitfield AE, Ullman DE, German TL. Tospovirus-thrips interactions. Annu Rev Phytopathol 2005; 43: 459-89.

[53] Pappu H, Jones R, Jain R. Global status of tospovirus epidemics in diverse cropping systems: successes achieved and challenges ahead. Virus Res 2009; 141: 219-36.

[54] Anbinder I, Reuveni M, Azari R, Paran I, Nahon S, Shlomo H et al. Molecular dissection of Tomato leaf curl virus resistance in tomato line TY172 derived from *Solanum peruvianum.* Theoret Appl Genet 2009; 119: 519-30.

[55] Díaz-Pendón JA, Cañizares MC, Moriones E, Bejarano ER, Czosnek H, Navas--Castillo J. Tomato yellow leaf curl viruses: ménage à trois between the virus complex, the plant and the whitefly vector. Mol Plant Pathol 2010; 11: 441-50.

[56] Lefeuvre P, Martin DP, Harkins G, Lemey P, Gray AJ, Meredith S et al. The spread of Tomato yellow leaf curl virus from the Middle East to the world. PLoS Pathog 2010; 6: e1001164.

[57] Scholthof KBG, Adkins S, Czosnek H, Palukaitis P, Jacquot E, Hohn T et al. Top 10 plant viruses in molecular plant pathology. Mol Plant Pathol 2011; 12: 938-54.

[58] Rybicki EP. A Top Ten list for economically important plant viruses. Arch Virol 2015; 160: 17-20.

[59] Forterre P. The origin of viruses and their possible roles in major evolutionary transitions. Virus Res 2006; 117: 5-16.

[60] Koonin EV, Martin W. On the origin of genomes and cells within inorganic compartments. Trends Genet 2005; 21: 647-54.

[61] Andersson SG, Zomorodipour A, Andersson JO, Sicheritz-Ponten T, Alsmark UC, Podowski RM, Naslund AK, Eriksson AS, Winkler HH, Kurland CG. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature 1998; 396: 133-140.

[62] Claverie J –M. Viruses take center stage in cellular evolution. Genome Biol 2006; 7: 110.

[63] Gadelle D, Filée J, Buhler C, Forterre P. Phylogenomics of type II DNA topoisomerases. Bioessays 2003; 25: 232-42.

[64] Miller ES, Kutter E, Mosig G, Arisaka F, Kunisawa T, Rüger W. Bacteriophage T4 genome. Microbiol Mol Biol Rev 2003: 67: 86-156.

[65] Botstein D. A theory of modular evolution for bacteriophages. Annal New York Acad Sci 1980; 354: 484-91.

[66] Roossinck MJ. Mechanisms of plant virus evolution. Ann Rev Phytopathol 1997; 35: 191-209.

[67] Nagy PD, Pogany J. The dependence of viral RNA replication on co-opted host factors. Nature Rev Microbiol 2012; 10: 137-49.

[68] Ahlquist P, Noueiry AO, Lee W-M, Kushner DB, Dye BT. Host factors in positive-strand RNA virus genome replication. J Virol 2003; 77: 8181-86.

[69] Lai MMC. Cellular factors in the transcription and replication of viral RNA genomes: a parallel to DNA-dependent RNA transcription. Virology 1998; 244: 1-12.

[70] Novoa RR, Calderita G, Arranz R, Fontana J, Granzow H, Risco C. Virus factories: associations of cell organelles for viral replication and morphogenesis. Biol Cell 2005; 97: 147-72.

[71] Cotton S, Grangeon R, Thivierge K, Mathieu I, Ide C, Wei T, Wang A, Laliberté J-F. Turnip mosaic virus RNA replication complex vesicles are mobile, align with microfilaments, and are each derived from a single viral genome. J Virol 2009; 83: 10460-71.

[72] Piron M, Vende P, Cohen J, Poncet D. Rotavirus RNA-binding protein NSP3 interacts with eIF4GI and evicts the poly (A) binding protein from eIF4F. EMBO J 1998; 17: 5811-21.

[73] Kovalev N, Pogany J, Nagy PD. Template role of double-stranded RNA in tombusvirus replication. J Virol 2014; 88: 5638-51.

[74] Hull R. Caulimoviridae (Plant Pararetroviruses). 2001; eLS John Wiley & Sons.

[75] Covey SN, Hull R. Transcription of cauliflower mosaic virus DNA. Detection of transcripts, properties, and location of the gene encoding the virus inclusion body protein. Virology 1981; 111: 463-74.

[76] Bak A, Gargani D, Macia J-L, Malouvet E, Vernerey M-S, Blanc S, Drucker M. Virus factories of Cauliflower mosaic virus are virion reservoirs that engage actively in vector transmission. J Virol 2013; 87: 12207-15.

[77] Woolston CJ, Covey SN, Penswick JR, Davies JW. Aphid transmission and a polypeptide are specified by a defined region of the Cauliflower mosaic virus genome. Gene 1983; 23: 15-23.

[78] Tang W, Leisner S. Methylation of nonintegrated multiple copy DNA in plants. Biochem Biophys Res Comm 1998; 245: 403-6.

[79] Kiss-László Z, Blanc S, Hohn T. Splicing of Cauliflower mosaic virus 35S RNA is essential for viral infectivity. EMBO J 1995; 14: 3552-62.

[80] Bonneville J, Sanfaçon H, Fütterer J, Hohn T. Posttranscriptional trans-activation in Cauliflower mosaic virus. Cell 1989; 59: 1135-43.

[80] Haas M, Geldreich A, Bureau M, Dupuis L, Leh V, Vetter G, Kobayashi K, Hohn T, Ryabova L, Yot P, Keller M. The open reading frame VI product of Cauliflower mosaic virus is a nucleocytoplasmic protein: its N terminus mediates its nuclear export and formation of electron-dense viroplasms . Plant Cell 2005; 17: 927-43.

[81] Schoelz JE, Shepherd RJ. Host range control of Cauliflower mosaic virus. Virology 1988; 162: 30-7.

[82] Broglio E. Mutational analysis of Cauliflower mosaic virus gene VI: changes in host range, symptoms, and discovery of transactivation-positive, noninfectious mutants. Mol Plant-Microb Interact 1995; 8; 755-60.

[83] Perbal M-C, Thomas C, Maule A. Cauliflower mosaic virus gene I product (P1) forms tubular structures which extend from the surface of infected protoplasts. Virology 1993; 195: 281-5.

[84] Huang Z, Han Y, Howell SH. Effects of movement protein mutations on the formation of tubules in plant protoplasts expressing a fusion between the green fluorescent protein and Cauliflower mosaic virus movement protein. Mol Plant-Microb Interact 2001; 14: 1026-31.

[85] Leh V, Jacquot E, Geldreich A, Haas M, Blanc S, Keller M, Yot P. Interaction between the open reading frame III product and the coat protein is required for transmission of Cauliflower mosaic virus by aphids. J Virol 2001; 75: 100-6.

[86] Haas M, Bureau M, Geldreich A, Yot P, Keller M. Cauliflower mosaic virus: still in the news. Mol Plant Pathol 2002; 3: 419-29.

[87] Kobayashi K, Hohn T. Dissection of Cauliflower mosaic virus transactivator/viroplasmin reveals distinct essential functions in basic virus replication. J Virol 2003; 77: 8577-83.

[88] Luo M. Negative strand RNA virus. World Scientific 2011.

[89] Wijkamp I, van Lent J, Kormelink R, Goldbach R, Peters D. Multiplication of Tomato spotted wilt virus in its insect vector, *Frankliniella occidentalis*. J Gen Virol 1993; 74: 341-9.

[90] Kikkert M, Van Lent J, Storms M, Bodegom P, Kormelink R, Goldbach R. Tomato spotted wilt virus particle morphogenesis in plant cells. J Virol 1999; 73: 2288-97.

[91] van Knippenberg I, Lamine M, Goldbach R, Kormelink R. Tomato spotted wilt virus transcriptase in vitro displays a preference for cap donors with multiple base complementarity to the viral template. Virology 2005; 335: 122-30.

[92] Geerts-Dimitriadou C, Lu Y-Y, Geertsema C, Goldbach R, Kormelink R. Analysis of the Tomato spotted wilt virus ambisense S RNA-encoded hairpin structure in translation. PLoS ONE 2012; 7: e31013.

[93] Walter CT, Barr JN. Recent advances in the molecular and cellular biology of bunyaviruses. J Gen Virol 2011; 92: 2467-84.

[94] Spoel SH, Dong X. How do plants achieve immunity? Defence without specialized immune cells. Nature Rev Immunol 2012; 12: 89-100.

[95] Hull R. Plant virology. 2001; Gulf Professional Publishing.

[96] Brault V, Uzest M, Monsion B, Jacquot E, Blanc S. Aphids as transport devices for plant viruses. Comptes Rendus Biologies 2010; 333: 524-38.

[97] Pollard D. Plant penetration by feeding aphids (Hemiptera, Aphidoidea): a review. Bull Entomol Res 1973; 62: 631-714.

[98] Taylor CE, Robertson WM. Sites of virus retention in the alimentary tract of the nematode vectors, *Xiphinema diversicaudatum* (Micol) and *X index* (Thorne and Allen). Annal Appl Biol 1970; 66: 375-80.

[99] Macfarlane SA. Molecular determinants of the transmission of plant viruses by nematodes. Mol Plant Pathol 2003; 4: 211-5.

[100] Andret-Link P, Fuchs M. Transmission specificity of plant viruses by vectors. J Plant Pathol 2005, 153-165.

[101] Macfarlane SA, Wallis CV, Brown DJF. Multiple virus genes involved in the nematode transmission of Pea early browning virus. Virology 1996; 219: 417-22.

[102] Nürnberger T, Brunner F. Innate immunity in plants and animals: emerging parallels between the recognition of general elicitors and pathogen-associated molecular patterns. Curr Opin Plant Biol 2002; 5: 318-24.

[103] Zvereva AS, Pooggin MM. Silencing and innate immunity in plant defense against viral and non-viral pathogens. Viruses 2012; 4: 2578-97.

[104] Zipfel C. Plant pattern-recognition receptors. Trends Immunol 2014; 35: 345-51.

[105] Kørner CJ, Klauser D, Niehl A, Domínguez-Ferreras A, Chinchilla D, Boller T, Heinlein M, Hann DR. The immunity regulator BAK1 contributes to resistance against diverse RNA viruses. Mol Plant-Microb Interact 2013; 26: 1271-80.

[106] Petersen CP, Bordeleau M-E, Pelletier J, Sharp PA. Short RNAs repress translation after initiation in mammalian cells. Mol Cell 2006; 21: 533-42

[107] Dixon RA, Lamb CJ. Molecular communication in interactions between plants and microbial pathogens. Ann Rev Plant Biol 1990; 41: 339-67.

[108] Mandadi KK, Scholthof K-BG. Plant immune responses against viruses: how does a virus cause disease? Plant Cell 2013; 25:1489-1505.

[109] Holmes, F. Interspecific transfer of a gene governing type of response to tobacco-mosaic infection. Phytopathol 1936; 26: 1007-14.

[110] Xu P, Roossinck MJ. Cucumber mosaic virus D satellite RNA–induced programmed cell death in tomato. Plant Cell 2000; 12: 1079-92.

[111] Xu P, Wang H, Coker F, Ma J-y, Tang Y, Taylor M, Roossinck MJ. Genetic loci controlling lethal cell death in tomato caused by viral satellite RNA infection. Mol Plant-Microb Interact 2012; 25: 1034-44.

[112] Wittmann S, Chatel H, Fortin MG, Laliberté J-F. Interaction of the viral protein genome linked of *Turnip mosaic potyvirus* with the translational eukaryotic initiation factor (iso) 4e of *Arabidopsis thaliana* using the yeast two-hybrid system. Virology 1997; 234: 84-92.

[113] Kang BC, Yeam I, Frantz JD, Murphy JF, Jahn MM. The pvr1 locus in capsicum encodes a translation initiation factor eIF4E that interacts with Tobacco etch virus VPg. Plant J 2005; 42: 392-405.

[114] Piron F, Nicolaï M, Minoïa S, Piednoir E, Moretti A, Salgues A, Zamir D, Caranta C, Bendahmane A. An induced mutation in tomato eIF4E leads to immunity to two potyviruses. PLoS ONE 2010; 5: e11313.

[115] Ruffel S, Dussault MH, Palloix A, Moury B, Bendahmane A, Robaglia C, Caranta C. A natural recessive resistance gene against Potato virus Y in pepper corresponds to the eukaryotic initiation factor 4E (eIF4E). Plant J 2002; 32: 1067-75.

[116] Robaglia C, Caranta C. Translation initiation factors: a weak link in plant RNA virus infection. Trends Plant Sci 2006; 11:40-5.

[117] Maule AJ, Caranta C, Boulton MI.Sources of natural resistance to plant viruses: status and prospects. Mol Plant Pathol 2007; 8: 223-31.

[118] Joshi B, Cameron A, Jagus R. Characterization of mammalian eIF4E-family members. Eur J Biochem 2004; 271: 2189-203.

[119] Browning KS. Plant translation initiation factors: it is not easy to be green. Biochem Soc Transact 2004; 32: 589-91.

[120] Contreras-Paredes CA, Silva-Rosales L, Daròs J-A, Alejandri-Ramírez ND, Dinkova T D. The absence of eukaryotic initiation factor eIF (iso) 4E affects the systemic spread of a Tobacco etch virus isolate in *Arabidopsis thaliana.* Mol Plant-Microb Interact 2013; 26: 461-70.

[122] Bendahmane A, Kanyuka K, Baulcombe DC. The Rx gene from potato controls separate virus resistance and cell death responses. Plant Cell 1999; 11: 781-91.

[123] Jones DA, Jones JDG. The role of leucine-rich repeat proteins in plant defences In: J H Andrews, I. C. T., Callow, J. A. (Eds). Advances in Botanical Research Academic Press 1997; 89-167.

[124] Knogge W. Fungal infection of plants. Plant Cell 1996; 8: 1711.

[125] Ding S-W, Voinnet O. Antiviral immunity directed by small RNAs. Cell 2007; 130: 413-26.

[126] Baulcombe D. RNA silencing in plants. Nature 2004; 431: 356-63.

[127] Balmer D, Mauch-Mani B. Small yet mighty microRNAs in plant-microbe interactions. MicroRNA 2013; 2: 73-80.

[128] Axtell MJ. Classification and comparison of small RNAs from plants. Annual review of plant biology. 2013; 64: 137-59..

[129] Lu Y-D, Gan Q-H, Chi X-Y, Qin S. Roles of microRNA in plant defense and virus offense interaction. Plant Cell Reports 2008; 27: 1571-9.

[130] Mérai Z, Kerényi Z, Kertész S, Magna M, Lakatos L, Silhavy D. Double-stranded RNA binding may be a general plant RNA viral strategy to suppress RNA silencing. J Virol 2006; 80: 5747-56.

[131] Azevedo J, Garcia D, Pontier D, Ohnesorge S, Yu A, Garcia S, Braun L, Bergdoll M, Hakimi MA, Lagrange T. Argonaute quenching and global changes in Dicer homeostasis caused by a pathogen-encoded GW repeat protein. Genes Develop 2010; 24: 904-15.

[132] Kanazawa A, Inaba Ji, Shimura H, Otagaki S, Tsukahara S, Matsuzawa A, Kim BM, Goto K, Masuta C. Virus-mediated efficient induction of epigenetic modifications of endogenous genes with phenotypic changes in plants. Plant J 2011; 65: 156-68.

[133] Smith NA, Eamens AL, Wang M-B. Viral small interfering RNAs target host genes to mediate disease symptoms in plants. PLoS Pathog 2011; 7: e1002022.

[134] Wang M-B, Bian X-Y, Wu L-M, Liu L-X, Smith NA, Isenegger D, Wu R-M, Masuta C, Vance VB, Watson JM. On the role of RNA silencing in the pathogenicity and evolution of viroids and viral satellites. Proc Natl Acad Sci USA 2004; 101: 3275-80.

[135] Várallyay É, Válóczi A, Ágyi Á, Burgyán J, Havelda Z. Plant virus-mediated induction of miR168 is associated with repression of ARGONAUTE1 accumulation. EMBO J 2010; 29: 3507-19.

[136] Silhavy D, Molnár A, Lucioli A, Szittya G, Hornyik C, Tavazza M, Burgyán J. A viral protein suppresses RNA silencing and binds silencing-generated, 21-to 25-nucleotide double-stranded RNAs. The EMBO journal. 2002; 21 :3070-80.

[136] Sullivan CS, Ganem D. MicroRNAs and viral infection. Mol Cell 2005; 20: 3-7.

[137] Lakatos L, Szittya G, Silhavy D, Burgyán J. Molecular mechanism of RNA silencing suppression mediated by p19 protein of tombusviruses. The EMBO journal. 2004; 23: 876-84.

[138] Pfeffer S, Sewer A, Lagos-Quintana M, Sheridan R, Sander C, Grässer FA, van Dyk L F, Ho CK, Shuman S, Chien M. Identification of microRNAs of the herpesvirus family. Nature Meth 2005; 2: 269-76.

[139] Iram S. Role of viral and host microRNAs in plant-virus interactions. 2010; The University of Queensland PhD Thesis.

[140] Simón-Mateo C, García JA. MicroRNA-guided processing impairs Plum pox virus replication, but the virus readily evolves to escape this silencing mechanism. J Virol 2006; 80: 2429-36.

[141] Lin S-S, Wu H-W, Elena SF, Chen K-C, Niu Q-W, Yeh S-D, Chen C-C, Chua N-H. Molecular evolution of a viral non-coding sequence under the selective pressure of amiRNA-mediated silencing. PLoS Pathog 2009; 5: e1000312.

[142] Li H-W, Lucy AP, Guo H-S, Li W-X, Ji L-H, Wong S-M, Ding S-W. Strong host resistance targeted against a viral suppressor of the plant gene silencing defence mechanism. EMBO J 1999; 18: 2683-91.

[143] Sansregret R, Dufour V, Langlois M, Daayf F, Dunoyer P, Voinnet O, Bouarab K. Extreme resistance as a host counter-counter defense against viral suppression of RNA silencing. PLoS Pathog 2013; 9: e1003435.

[144] Abel PP, Nelson RS, De B, Hoffmann N, Rogers SG, Fraley RT, Beachy RN. Delay of disease development in transgenic plants that express the Tobacco mosaic virus coat protein gene. Science 1986; 232: 738-743.

[145] Savenkov E, Valkonen J. Coat protein gene-mediated resistance to Potato virus A in transgenic plants is suppressed following infection with another potyvirus. J Gen Virol 2001; 82: 2275-78.

[146] Smith NA, Singh SP, Wang M-B, Stoutjesdijk PA, Green AG, Waterhouse PM. Gene expression: total silencing by intron-spliced hairpin RNAs. Nature 2000; 407: 319-20.

[147] Anderson JM, Palukaitis P, Zaitlin M. A defective replicase gene induces resistance to cucumber mosaic virus in transgenic tobacco plants. Proc Natl Acad Sci USA 1992; 89: 8759-63.

[148] Prins M, De Haan P, Luyten R, Van Veller M, Van Grinsven M, Goldbach R. Broad resistance to tospoviruses in transgenic tobacco plants expressing three tospoviral nucleoprotein gene sequences. Mol Plant-Microb Interact 1995; 8: 85-91.

[149] Ramesh SV, Ratnaparkhe MB, Kumawat G, Gupta GK, Husain SM. Plant miRNAome and antiviral resistance: a retrospective view and prospective challenges. Virus Genes 2014; 48:1-14.

[150] Brummelkamp TR, Bernards R, Agami R. A system for stable expression of short interfering RNAs in mammalian cells. Science 2002; 296: 550-53.

[151] Boden D, Pusch O, Silbermann R, Lee F, Tucker L, Ramratnam B, Enhanced gene silencing of HIV 1 specific siRNA using microRNA designed hairpins. Nucl Acid Res 2004; 32: 1154-58.

[152] Schwab R, Ossowski S, Riester M, Warthmann N, Weigel D. Highly specific gene silencing by artificial microRNAs in Arabidopsis. Plant Cell 2006; 18: 1121-33.

[153] Eamens AL, Agius C, Smith NA, Waterhouse PM, Wang M-B. Efficient silencing of endogenous microRNAs using artificial microRNAs in *Arabidopsis thaliana.* Mol Plant 2011; 4: 157-70.

[154] Jackson AL, Bartz SR, Schelter J, Kobayashi SV, Burchard J, Mao M, Li B, Cavet G, Linsley PS. Expression profiling reveals off-target gene regulation by RNAi. Nature Biotechnol 2003; 21: 635-37.

Methodology Used in the Thesis

Sources of the Virus Isolates

The TuMV-QLD1b isolate used in this study was a serially passaged isolate of an original sample previously sourced from DAF by the Schenk Lab in 2007. Four Department of Agriculture and Fisheries (DAF) collection samples (VIR-0484, VIR-0745, VIR-1280 and VIR1281), kindly supplied by Dr Geering, were used to determine which was the original TuMV isolate (renamed TuMV–QLD1a). Similarly, CaMV- Dar78694 was also supplied by Dr Geering from the DAF collection. The CMV isolate K was part of former PhD project and given to the Schenk Lab by John Randle (personal communication) (ca 2004).

Virus Inoculation

The virus inoculation buffer is made up of sodium phosphate dibasic heptahydrate ($Na_2HPO_4.7H_2O$ 141.96g in 1L) and monosodium phosphate (NaH_2PO_4 156.01g in 1L). Additions of 77.4 mL of $Na_2HPO_4.7H_2O$ and 22.6 mL of NaH_2PO_4 dissolved in 1 L of distilled water gave a final concentration of 100 mM and a pH of 7.4 with a final addition of sodium sulphite (Na_2SO_3 1 g/L).

Infected fresh young leaf tissue (ca 10g) was ground with a mortar and pestle in 20 mL of inoculation buffer with ca 0.2 g celite. Plants were inoculated using a cotton swab, and rinsed with distilled water after inoculation.

TuMV was propagated in *N. benthamiana*. CaMV was propagated in *Brassica rapa* subspecies *Chinensis* from a freeze -dried sample. CMV was propagated from fresh tissue in *Solanum lycopersicum* Money Maker variety, however the virus inoculation buffer was diluted 10 fold. *A. thaliana* plants were four-five weeks old (five leaf stage) and *N. benthamiana, B. rapa and S. lycopersicum* plants were two-three weeks old prior to virus inoculation.

Fresh inoculum was used to perform all virus inoculations of *A. thaliana*. Mock inoculations were performed first to avoid cross-contamination; leaves were treated with same buffer and abrasives as virus inoculated plants

Symptom development in *N. benthamiana, B. rapa* and *S. lycopersicum* can be seen after two weeks however after one week the new leaves are often seen to curl being a good indication that the infection was successful. Symptom development is different in *A. thaliana* and will usually only be seen after four weeks but is not as distinct.

Plant Growth Conditions

Col-0 plants were sown on UC mix soil and placed at 4˚C for two days. Seedlings were then transferred to a growth chamber with the following conditions: 8 hours light at 24˚C and 16 hours dark at 21˚C. Seedlings were carefully transplanted into a 5x6 seedling tray after two-three weeks and then inoculated with TuMV at five weeks old. Seedlings were watered every second day with minimal water preventing algal growth. *N. benthamiana* and *B. rapa* seedlings were also sown on UC mix and transplanted into small square pots after two weeks. Seedlings were kept in a growth chamber: 16 hours of light at 26˚C and the dark cycle was set to 24˚C. *S. lycopersicum* seeds were individually sown in pots on UQ23 soil and kept in the same condition as above. As *N. benthamiana, B. rapa* and *S. lycopersicum* were used for viral propagation this was repeated every two months or once the plants were flowering. Plants were watered from the bottom with distilled water preventing algal and fungal growth.

Tissue Sampling and RNA Extraction

Young leaf tissue was collected from infected *N. benthamiana* plants for RNA extraction and PCR. Tissue was collected in a sterile 2 mL Eppendorf tube with an RNase-free ball bearing. Tissue was frozen in liquid nitrogen and ground using a TissueLyser (Qiagen). Similarly, young leaf tissue was also collected for RNA sequencing of TuMV, CMV and CaMV.

For the early defence response assay whole Col-0 plants were collected at three-time points: 6 hours, 24 hours and 48 hours. Plants were pooled (ca 20) with three biological replicates per time point. Plants were cut at the base, only using the foliar parts for analysis, and immediately frozen in liquid nitrogen.

Infected *med18* and Col-0 non-inoculated leaves were collected 14 days after inoculation (dai), in sterile 2 mL Eppendorf tubes with RNase-free ball bearings. Tissue was frozen in liquid nitrogen and ground using a TissueLyser (Qiagen).

The Maxwell ® RSC Plant RNA Kit (Promega) was used per the manufacture's instruction to extract RNA from all samples. RNase inhibitor (Thermo Fisher Scientific) was add to the eluted RNA to prevent degradation.

Deep RNA Sequencing

RNA from one replicate per treatment (mock 7 and 14 dai and virus 7 and 14 dai) were normalised to 1 µg with nuclease-free water. RNA libraries were prepared using the TruSeq RNA Library Preparation kit (Illumina) per manufacturer's instructions. Once both strands of the libraries were

synthesised unique adaptors were ligated to each library, used to identify the four samples; adapter 2 was used to identify mock 7 dai, adapter 4 mock 14 dai, adapter 5 virus 7 dai and adapter 6 14 dai.

The libraries were cleaned, and the DNA was enriched through a standard PCR followed by a second cleaning round. Concentrations were measured using the Qubit broad range DNA buffer and reagents according to the manufacturer's instruction. Samples were sent to the Australian Genome Research Facility (AGRF) for bioanalyser quantification.

Samples were diluted to 10 μM with nuclease-free water. The RNA Sample Preparation Kit (Illumina) was used per manufacturer's instructions to further prepared samples to be run on the Illumina MiSeq Gene and small Genome sequencer. Samples were loaded into separate cartridges and run per manufacturer's instructions.

TRIzol RNA Extraction and Small RNA Sequencing

As total RNA was needed for small RNA sequencing of TuMV, RNA was extracted using a phenol based method.

Ground plant tissue (ca 10 mg) was added to a sterile 2 mL tube with 600 μL of TRIsure reagent (Bioline) and vortexed for 15 s. Another 900 μL of TRIsure was added and vortexed for a further 30 s. Samples were incubated at room temperature for 5 min.

An aliquot of 300 μL of chloroform was added and mixed thoroughly by shaking. Samples were centrifuged for 15 min at 4˚C. The aqueous phase was transferred to a new tube containing 750 μL of isopropyl alcohol and incubated on ice for 15 min. Samples were centrifuged at 4˚C for 10 min. The supernatant was removed, and the RNA pellet was washed with 1 mL of 70% ethanol. Samples were centrifuged for 5 min at room temperature. Ethanol was pipetted off and samples were dried in a vacuum dome for 5 min. Samples were resuspended in 66μL of RNase-Free water.

Data Analysis

Data analysis was performed using Geneious R8 software and the free web-based RNA sequencing analysis platform Galaxy, installed and run from The University of Queensland's cloud. The raw data was uploaded into Galaxy as well as the unique adapter sequences for each sample.

The adapters were trimmed using the Cutadapt tool under the NGS: QC and manipulation heading. Once samples were trimmed they were exported into Geneious as FASTA files where a BLAST search was performed to remove all chloroplast RNA, ribosomal RNA and t-RNA. The remaining sequences were then mapped to the *Arabidopsis* genome using the Map to Reference function found under the Align/Assemble tab. This left a number of unused reads which were used to create the

TuMV genome using the De Novo Assembly tool under the Align/Assemble tab. A consensus sequence was extracted and named TuMV-QLD1b.

cDNA Synthesis and PCR

RNA was diluted to roughly 2 µg in a volume of 13 µL to synthesis cDNA using the Tetro cDNA Synthesis Kit (Bioline) per the manufacturer's instructions with minor adjustments. As certain PCRs were purely for amplifying a fragment of virus, only random hexamers were used in the synthesis and not oligo dT. This addition changes the thermocycling program to include an extra step at the beginning of the protocol. When synthesising cDNA for real-time quantitative reverse transcription PCR a combination of random hexamers and oligo dTs were added. The Thermocycling program was set with the following specifications: 10 min at 25 ˚C, 30 min at 45 ˚C. The reaction is terminated by incubating at 85˚C for 5 min.

Concentrate cDNA was used as the template in a 50µl PCR reaction. The high fidelity Phusion (New England Biolabs) enzyme was used per manufacturer's instructions with the following cycling conditions: 98˚C for 30 s followed by 30 cycles of 98˚C for 10 s, annealing temperature and time depending on primers and 72˚C extension with the time depending on amplicon length. Final extension 72˚C for 5 min and a 4˚C hold.

Real-Time Quantitative Reverse Transcription PCR

cDNA was prepared using the Tetro cDNA Synthesis Kit (Bioline) per manufacturer's instructions. RNA was normalised to the sample with the lowest concentration.

Real time quantitative reverse transcription PCR (qRT-PCRwas performed using the ViiA 7 Real-Time PCR system with the following thermocycling program: 10 min at 95˚C followed by 40 cycles of 15 s at 95˚C, 1 min at 60˚C. The melting curve conditions were 15 s at 95˚C, 1 min at 60˚C and 15 s at 95˚C.

SYBR green was used as the report for transcription levels with two technical replicates per gene. Primers were designed for qRT-PCR using Primer Express software and were used to quantify the gene expression of several genes all relative to the reference genes *b-ACTIN2*, *b-ACTIN7*, and *b-ACTIN8*. LinReg PCR software was used to calculate the change in Rn values. The ΔΔCT method was used to calculate the relative expression of all genes.

Table 1: Reference and marker genes primers used in chapter 3 and 4.

Primer name	Sequence
rt_Actin8_r	GAGGATAGCATGTGGAACTGAGAA
rt_Actin7_r	GAGGAAGAGCATTCCCCTCGTA
rt_Actin Uni_f	AGTGGTCGTACAACCGGTATTGT
rt_Actin2_r	GATGGCATGGAGGAAGAGAGAAAC
PDF1.2_F	AAGTTTGCTTCCATCATCACC C
PDF1.2_R	ATTGCCGGTGCGTCGAAAG
VSP2_F	GAAAACCATCTTTGGGAACG
VSP2_R	CGGTTTTGGAGTCGTATTGG
PR1_F	GTC TCC GCC GTG AAC ATG T
PR1_R	CGT GTT CGC AGC GTA GTT GT
PR5_F	AATGTCAAGCTGGGGA
PR5_R	AGGTGCTCGTTTCGTC
ERF1_F	AAAGCAGCTTGATCGTAGGC
ERF1_R	ATTCGACTAGAAACGGTATTAGGG
ERF6_F	ACGGTGGTTGAGAAAGTGCTAAAG
ERF6_R	CATGCTCAGAAACTCCGTCAAATC
RD22_F	ATTGTGCGACGTCTTTGGAGT
RD22_R	TGCGTTCTTCTTAGCCACCTC

Electrophoresis and Clean-up

Agarose gel (1.0%) was set up with ethidium bromide as the fluorescent tag. Loading dye (6X) was added to the sample and loaded onto the gel and run at approximately 110 volts for 45 min ensuring adequate band separation. The voltage and time was adjusted depending on the size of the fragment.

The gel was visualised on a UV gel dock. The Wizard® SV Gel and PCR Clean-Up System (Promega) was used per the manufacturer's instructions to extract DNA fragments from the gel.

Cloning and Sanger Sequencing

Cleaned PCR fragments were ligated into the PCR-Blunt vector using the Zero Blunt® PCR Cloning Kit (Invitrogen) per the manufacturer's instructions. The maximum volume of 5 µL of fresh PCR product was used in the reaction and only 1 µL of nuclease-free water. According to the protocol incubating the reaction for 5 min at room temperature was sufficient, however the reaction was incubated for an extra 25 min to increase the cloning efficiency.

The Alpha-Select Silver Efficiency (Bioline) commercial competent *E. coli* cells were used with 2-5 µL of the ligation reaction depending on the volume of cells used. Cells were left on ice for 10 min and heat shock transformation was performed at 42˚C for 35 s. After heat shock, cells were immediately placed on ice for ~ 2 min. Plain lysogeny broth (LB) media was added to each sample

(500 µL), made according to Molecular Cloning, Laboratory Manual third edition (Sambrook and Russell 2006), and placed on a 37˚C shaker for an hour.

The cultures were spun down and the majority of the LB discarded. The culture pellets were resuspended and spread out on LB plates containing kanamycin. Plates were kept in a 37˚C incubator overnight.

Colonies were grown overnight in 2 mL of LB and kanamycin in a 37˚C incubated shaker. Plasmids were isolated using the PureLink® Quick Plasmid Miniprep Kit (Invitrogen) according to manufacturer's instruction with minor adjustments. The optional wash was not performed and plasmids were eluted in 10 mM Tris-HCl made according to Molecular Cloning, Laboratory Manual third edition (Green and Sambrook 2000).

Samples were prepared for sequencing by adding 11 µL of each plasmid sample in a 1.5 mL Eppendorf tube with 1 µL of either forward or reverse of the specific primer. Samples were then sequenced through Sanger sequencing by AGRF. Sequencing results were downloaded, edited and analysed in Geneious R8.

Phylogenetic Analysis

A phylogenetic tree was constructed using the Tree Drawing function in Geneious R8. The tree was drawn with the same isolates used in the Nyalugwe, Jones et al. (2015) paper. Isolate sequences were downloaded from GenBank into Geneious and aligned using the Multiple Alignment function under the Align/Assemble tab. The tree was drawn using the PHYML Tree function which was downloaded as an add-in and enables the user to draw maximum likelihood trees. The Tamura–Nei substitution model was selected and a bootstrap value of 100 was used. The proportion of invariable sites was fixed, and the number of substitution rates was set to 4. The tree builder used an estimated Gamma distribution parameter and was set to optimise topology/length/rate. The topology search had the fast default seeding of NNI.

Dual Luciferase Assay

The dual luciferase (dual-luc) assay includes a number of methods described above therefore these will not be described in great detail but refered to where needed.

Figure 1: Figure from Moyle *et al*. (2017). The pGrDL_SPb plasmid is used as a quantitative reporter system for validating possible viRNAs target transcripts via transient expression. The host target transcript in cloned into the 3' of the firefly LUC which has a CaMV 35S promoter and terminator. The *Renilla* LUC is expressed using the tomato *ACTIN7* promoter and *NOS* terminator. LB and RB stand for left board and right board, respectively.

The assay is a quick, robust and quantitative technique to validate host transcript sequences which are possibly targeted by viral small RNA (viRNAs). The dual-luc plasmid used to validate the interaction of the viRNAs with the target sequence has a dual reporter system; the host transcript targeted sequence is cloned in the 3' end of the firefly luciferase while the *Renilla* luciferase acts as an internal control (Figure 1).

The process begins with sRNAs data; specifically, that of *A. thaliana* and *N. benthamiana*. viRNAs with 50 copies per million were copied and pasted into the web-based platform; psRNAstarget which was originally designed to predict miRNA-target pairings (Dai and Zhao 2011). This produced a list of gene transcripts with the most likely targets having the lowest except score or highest homology. The following settings were used to run the psRNAstarget program; schema V1 was run with default settings, however no extra weight was placed in the seed region. A different cDNA library was used based on which plant the viRNAs were generated from, for example, the phytozome 12, 167_TAIR10 transcript library was used for *A. thaliana* while the *N. benthamiana*, transcript, Niben101 was used for predicting *N. benthamiana* target transcripts.

The top ten transcripts with the highest *complementarity* were chosen. The target sequence, belonging to the host, was then used to create adapters which were either 21-22 bp in length. The sequence was also mutated at base 6, 8 and 11 reading from 5'-3'. Cloning sites were placed at both endings of the adapter sequence; *Sal*I at the 3' end and *Pst*I at the 5' end. Both the mutated and no-mutated versions were ordered through Integrated DNA Technologies with a 5' phosphorylated modification.

Adapters were ligated into the dual-LUC plasmid pGrDL_SPb (Figure 1) which allows for directional cloning once digested with *Sal*I and *Pst*I. Heat shock transformation, plasmid isolation and Sanger sequencing were performed as above. Sequencing results were screened for plasmids containing the

adapters. These were transformed into Agrobacterium tumefacians (recently referred to as Rhizobium radiobacter) strain GV3101 using electroporation and then grown at 28ºC on LB plates containing Kanamycin (50 mg/ml), Rifampicin (25 mg/ml) and tetracycline (10 mg/ml) for 2 days as described in Molecular Cloning, Laboratory Manual third edition (Green and Sambrook 2000).

N. benthamiana seeds were sown, transplanted and inoculated with TuMV as described above. *Agrobacterium* cultures were prepared two days before infiltration; a 2 mL starter culture was grown first in LB medium containing kanamycin, rifampicin and tetracycline and grown overnight at 28˚C. The following day 30 mL of media was inoculated with approximately 30 µL of starter culture and grown overnight at 28˚C. The cultures were pelleted at 3500 rpm for 15 min. LB media was removed using a vacuum line and cultures were resuspended in pre-made 10 mM magnesium chloride ($MgCl_2$) (2.03 g $MgCl_2$ in 1 L of MilliQ water and autoclaved).

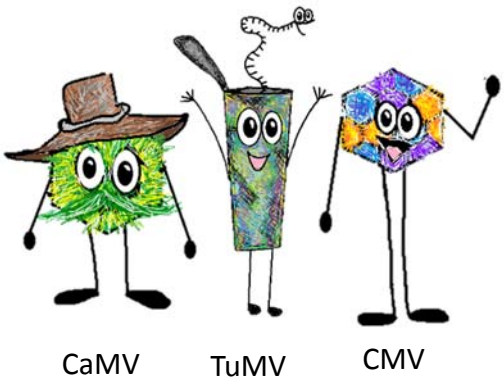
The OD of each culture was adjusted to 0.5 and 200 mM Acetosyringone was added to in a 1:1 ratio. Cultures were left in the dark for four to 24 hours after which time infected and non-infected *N. benthamiana* leaves were infiltrated as previously described in Moyle *et al*. (2017). Plants were left in normal growing condition for three days. Tissue was harvested and ground using ball bearings and a TissueLyser (Qiagen).

The Dual-Luciferase Reporter Assay System (Promega) was used to perform the assay according to manufacturer's instructions. The 5x Passive Lysis Buffer (PLB) was diluted and divided between two sets of 1.5mL Eppendorf tubes; 100 µL was pipetted into the first set and 50 µL into the second set. The ground plant material (ca. 10mg) was placed into the 100 µL of PLB and mix by flicking. Tubes were centrifuged for 1 min at 7,500g and 2.5 µL of this was diluted in the 50 µL of PLB.

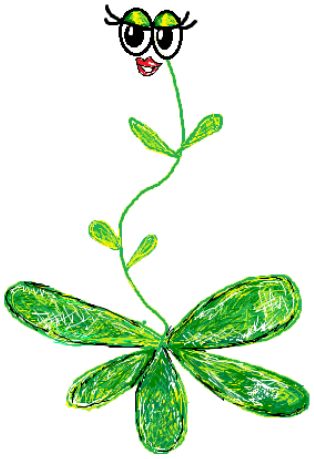
Samples were then analysed using a luminometer; 15 µL of the diluted sample was pipetted into a white 96 well luminescence plate. The plate was placed in a luminometer which was set to dispense the specific reagents and measure the luminescence emitted by each sample. The firefly activity was measured first with the addition of 75µL of Luciferase Assay Buffer II. Similarly, the *Renilla* activity was measured second with the addition of 75µL of the reagent Stop & Glo®.

Data was copied to a USB and analysed in GraphPad Prism 7. A more comprehensive method is found in the following published research article: "An Optimized Transient Dual Luciferase Assay for Quantifying MicroRNA Directed Repression of Targeted Sequences" published in Frontiers in Plant Science. A hyperlink can be found in Appendix 1.

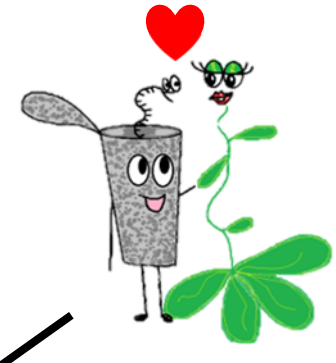
Thesis Overview



Chapter 1 and 2
Getting to know you
Sequencing of viruses

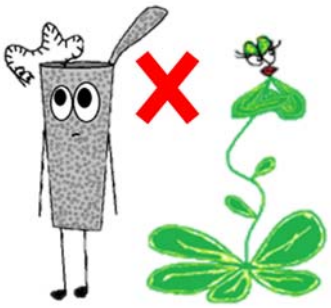


Chapter 3 "Going steady"
"First date"= Infection
Compatibility?
Virus-Induced defences?



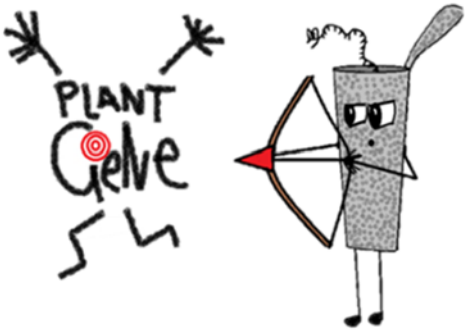
Chapter 4 Trouble in Paradise

"Attack of the Plant"
If the virus is manipulating the plant defensives, what happens when that plant's defences can't be manipulated?
Prevention is the best cure



Chapter 5 Trouble in Paradise

"The Virus Strikes Back"
viRNAs: the consequences of plant RNA silencing
A counter-attack?



Chapter 1: Sequencing and Characterisation of *Turnip mosaic virus* before and after Seven Years of Serial Passage

1.1 Introduction

Turnip mosaic virus belongs to the *Potyvirus* genus, which boasts an impressive 162 species according to The International Committee on Taxonomy on Viruses. It was first characterised in 1921 in the USA by two groups, Gardner and Kendrick (1921) and Schultz (1921).

Turnip mosaic virus is transmitted in a non-persistent manner by aphids. The virions are flexuous filaments and are approximately 700-750nm in length; the genome is roughly 10kb in size. It has a ssRNAs genome, with a single open reading frame (ORF) flanked by two untranslated regions (UTR). The ORF is translated into a polypeptide, then cleaved and processed into eleven viral proteins as depicted in figure 2.

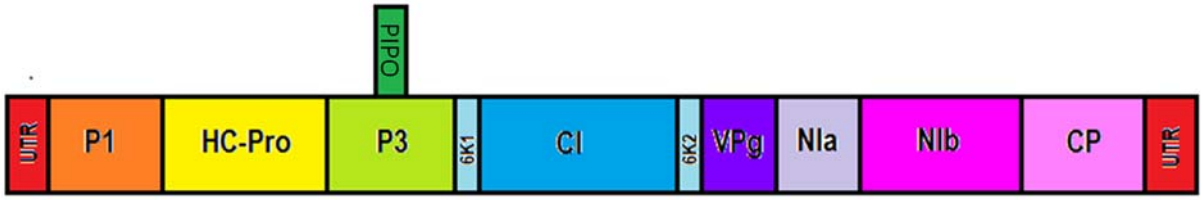


Figure 2: The general genome organisation of *Turnip mosaic virus*, flanked by two UTR in red and coding for 11 proteins processed by viral proteases. P1 (orange) and P3 (green) are protein1and protein3, respectively. The P3N-PIPO is in dark green encoded within the P3. HC-Pro (yellow) helper component protease, 6K1 and 6K2 (light blue) are 6kDa protein1 and 2. CI (blue) cylindrical or cytoplasmic inclusion protein. VPg (purple) virus encoded genome-linked protein. NIa (lilac) and NIb (pink) are nuclear inclusion proteins a and b. CP (rose) coat protein.

Successful infection begins with the virions entering the host cell and releasing genomic RNA that is acted upon by host machinery translating the viral proteins (Ahlquist *et al*. 2003, Simon and Miler 2013). The original genomic RNA is used as a template synthesising a complementary negative strand which will generate proteins and positive RNA; systemically infecting the host (Nagy and Pogany 2012).

Potyvirus replication occurs in vesicles created by the remodelling of the hosts membranous organelles protecting the viral RNA from degradation (Grangeon *et al*. 2010, Grangeon *et al.* 2012). Recent studies suggest that viral replication takes places in the chloroplast as well as in prenuclear structures thought to deliver virus RNA to neighbouring cells using microfilaments (Cotton *et al.* 2009, Wei and Wand 2008, Grangeon, Agbeci *et al.* 2012). .

Once processed the viral polyprotein is translated into 11 mature proteins by three protease domains as seen in figure 1. Studies have recorded 33 protein interactions occurring during infection. The large

number of interactions is not surprising considering the multifunctional nature of many of the viral proteins.

Turnip mosaic virus codes three proteases; P1 protein is a serine protease while the HC-Pro is a cysteine protease, both are auto-proteolytic, cleaving themselves from the polyprotein (Carrington *et al.* 1989, Riechmann *et al.*, 1992, Pan 2016). The NIa protein is the main proteinase and is both trans-proteolytic as well auto-proteolytic and described as a chymotrypsin-like cysteine protease (Rodamilans *et al.* 2018).

The P1 protein has a role in genome amplification. Its interaction with the HC-Pro can increase RNA silencing suppression (Kasschau *et al.* 2003, Valli *et al.* 2007) though P1 has shown to increase infectivity in plants that lack RNA-silencing machinery suggesting an independent role unrelated to RNA-silencing (Parsin *et al.* 2014). The protease domain of the P1 protein is highly conserved while its N-terminal region is the most variable potyviral protein in both size and sequence (Yoshida *et al.* 2102).

HC-Pro name is derived from being the Helper Component for aphis transmission acting as the bridging protein binding the viral particles to the aphid's stylet (Pirone and Blanc 1996, Pirone and Perry 2002). Apart from aphid transmission and RNA silencing the HC-Pro it has been found to play a role in the nucleus of infected cells (Sahana *et al.* 2014) and required to stabilize the CP ensuring infectivity (Valli *et al,* 2014). It has been found to interact with a number of viral and host proteins though relevance of these interaction is still unclear (Revers and García 2015).

P3 is the least well characterised but does has a role in viral amplification, disease development and severity (Rodriguez-Cerezo *et al.* 1993, Riechmann *et al.* 1995, Langenberg and Zhang 1997, Chu *et al.* 1997, Moreno *et al.* 1998). Some more recent studies have shown using transient expression that P3 forms inclusions which collocalize with viral replication complexes (Cui *et al.* 2010). Though it may interact with viral RNA it interacts with several viral proteins specifically CI, NIb and NIa (Zilian and Maiss, 2011).

The P3 also codes for a second viral protein, P3N-PIPO created by a frameshift due to transcriptional slippage (Rodamilans *et al.* 2015). This viral protein plays a vital role in cell-to-cell movement with knockout studies resulting in restricted movement but not viral replication (Wen and hajimorad 2010).

The roles of 6K1 and 6K2 are both relatively unknown and similarly to P3 these proteins do not interact with the viral RNA though studies have shown interactions with other viral proteins (Merits *et al.* 1998, Lin *et al.* 2009).

It is hypothesised that as the spilt between P3-6K1 does not interfere with viral infection; the P3 6K1 may be the main functional viral protein (Riechmann *et al.*1995). Though, the proteolytic splitting of the two affects symptom development suggesting 6K1 may have an independent function (Waltermann and Maiss 2006). . In *Potato virus A* (PVA) and *Tobacco etch virus* (TEV) the 6K2 protein was shown to be membrane bound (Merits *et al.* 2002) and thought to have a role in viral replication (Schaad *et al.* 1997). The 6K2 also forms part of the protein 6K2-VPg-NIaPro found in the viral replication complexes where it has a role in replication (Wei and Wang 2008).

The CI protein is part of the superfamily 2 of proteins identified by their seven conserved blocks of sequence, which are essential to helicase proteins (Kadaré and Haenni 1997). CI helicase activity was first demonstrated by Lain *et al.* (1990) where they performed in vitro assays which clearly displayed CI ability to unwind double strand RNA. More recently the CI has been implicated in interacting with the P3N-PIPO aiding in viral movement (Wei *et al.* 2010). It was also found to interact with the ends of the virions possibly acting as a motor protein transporting virus through plasmodesmata (Gabrenaite-Verkhovskaya *et al.* 2008).

The NIa protein is only partially processed forming the VPg and the NIaPro (Dougherty and Dawn Parks, 1991). The NIaPro as mentioned above is the main protease of the virus and responsible for the proper proteolytic processing. This process is highly regulated and infectivity dependent (Sun *et al.* 2010). It was also found to have DNase activity; degrading host DNA in the nucleus (Anindya and Savithri, 2004). It is location in the nucleus and its DNAse activity have suggested a role in host gene expression (Anindya and Savithri, 2004).

The VPg interacts with almost all the viral proteins and many host proteins (Elena and Rodrigo, 2012). When it is part of the NIa it has been shown to have an important role in viral replication as it interacts with several host eukaryotic translation initiation factor (eIF4E) which is crucial for viral replication (Léonard *et al.* 2000, Wang and Krishnaswamy, 2012).

According to Hong and Hunt (1996) the RdRp of the potyvirus is the NIb protein as it has the classic characteristic GDD sequence as well as replicase activity. It also forms inclusion bodies with NIa in the nucleus and cytoplasm (Knuhtsen et al., 1974). Interactions with several host proteins result in the formation of functional replication complexes (Dufresne *et al.* 2008, Thivierge *et al*. 2008) and it uridylylates the VPg to prime RNA synthesis (Anindya *et al.* 2005).

The final protein to be processed in the polyprotein is the CP which has a variety of important roles including aphid-transmission, cell-to-cell movement, and viral replication though its main function is the encapsidation of the viral RNA genome (Shukla and Ward 1989, Deng *et al.* 2015, Revers and García 2015 Gallo *et al.* 2018). Recent studies also show that it interacts with host Rubisco appearing to be an important factor in infection (Zhao *et al.* 2013).

TuMV, has a fairly high mutation rate as its RdRp has no proof reading ability (Steinhauer *et al.* 1992). Therefore it is not surprising that TuMV isolate sequences have a high degree of variability and quite often a single isolate will exist in a quasi-species which allows for rapid adaption to new hosts (Schneider and Roossinck 2000). Host adaption and virus origins have both been extensively researched with regards to TuMV. One such study by Ohshima *et al.* (2002) deduced that TuMV most likely originated in Europe with the *Brassicaceae* family and then spread throughout the world. The study phylogenetically grouped 76 different isolates based on their host range and from there further grouped them based on molecular analysis of the P1 and CP proteins.

Due to the poor representation of Australian isolates in phylogenetic analysis studies, a paper by Nyalugwe *et al.* (2015) was published focusing specifically on the biological and molecular variation among Australian TuMV isolates. Using seven newly sequenced Australian TuMV isolates as well as one complete genome sequence obtained from GenBank (BRS1) they found that six fell into World-B while the other two form part of Basal-B. The phylogenetic difference between the Australian isolates could suggest two separate introductions of the virus into Australia, one possibly being brought in during the British colonisation (Gibbs *et al.* 2008).

The TuMV isolate used in this study was previously sequenced by PCR circa 2008 and submitted to GenBank under the name BRS1 in 2010 (accession number HM544042), as part of a previous PhD project by a former student of the Schenk lab (Iram 2010). However, limited information was recorded to document the original source of the TuMV isolate used to generate the BRS1 sequence. Four possible candidate isolates were identified after searching the DAF virus database. The four isolates (VIR-0484, VIR-0745, VIR-1280 and VIR-1281) were subjected to molecular analysis to identify which was the original TuMV isolate. A sequence alignment of two short PCR amplified viral fragments suggested VIR-0745 was the original isolate as it had 100% identity to the TuMV isolate used in this study.

Re-sequencing the isolate in 2015 after seven years of serial passaging using deep RNA sequencing reads, revealed a high number of sequence variations when compared to the BRS1 sequence. To determine if these polymorphisms were due to evolution or human error in sequencing BRS1, a decision was taken to re-sequence a 2011 sample of the TuMV isolate and VIR-0745, the original

isolate. The sequencing and analysis of VIR-0745, and samples taken after serial passaging in 2011 and 2015 are presented in this study. A nucleotide alignment between the original and the 2015 isolates revealed 18 SNPs however this did not affect the conserved amino acid motifs within each gene. The original isolate VIR-0745 will be referred to as TuMV-QLD1a while the 2015 RNA sequenced isolate will be referred to as TuMV-QLD1b. Both sequences have been submitted to GenBank under these names with the following accession numbers: KX641465 and KX641466.

1.2 Results

1.2.1 Determining the Origins of the TuMV Isolate Used in This Study

The TuMV isolate used in this study was the subject of an earlier student's PhD project (Iram 2010), which involved sequencing via overlapping PCR. The compiled sequence was submitted to GenBank as isolate BRS1 (HM544042). However, it was unclear what the original source of the TuMV isolate was. Anecdotal evidence suggested the isolate was obtained from DAF and originally isolated from a Chinese cabbage field sample collected in Queensland.

With this information a search on DAF's plant virus database narrowed the possibilities to four different TuMV isolates. Table 1 lists the four isolates with the relevant information.

Table 2: The four isolates obtain from DAF listing the isolate number used to identify samples, the year the sample was collected and the nearest town to the collection point as well as the host the sample was originally isolated from.

Isolate number	**Year collected**	**Nearest town to location**	**Original host**
VIR-0484	1994	Gatton	Chinese Cabbage
VIR-0745	1997	Toowoomba	Chinese Cabbage
VIR-1280	2001	Allora	Turnip weed
VIR-1281	2001	Brisbane	Lettuce

RNA from fresh tissue was extracted and PCR was used to amplify two different fragments with specific primers designed from TuMV-QLD1b. The fragments were blunt-end cloned and sent for sequencing. The fragment sequences were aligned in Geneious with TuMV-QLD1b. Figure 3 shows that both VIR-0484 and VIR-0745 are identical to TuMV-QLD1b while, VIR-1280 and VIR1281 have multiple single nucleotide polymorphisms (SNPs). The result from the alignment of the second fragment (Figure 4) shows that VIR-0484 contained 22 SNPs differing from TuMV-QLD1b while VIR-0745 contained no SNPs. This confirmed VIR-0745 was the original isolate as both sequenced fragments were identical to TuMV-QLD1b.

Table 3: List of primers used to sequence the 2 fragments from the four TuMV isolates

Primer Name	Sequence
F1	GGTGCTTCCTTGCATATCGC
R1	ACGCTGCTTTCTCCATTCCA
F2	GGCGTGACGTGAAAATTCCC
R2	CCCATCGTTCTTCGTGACG

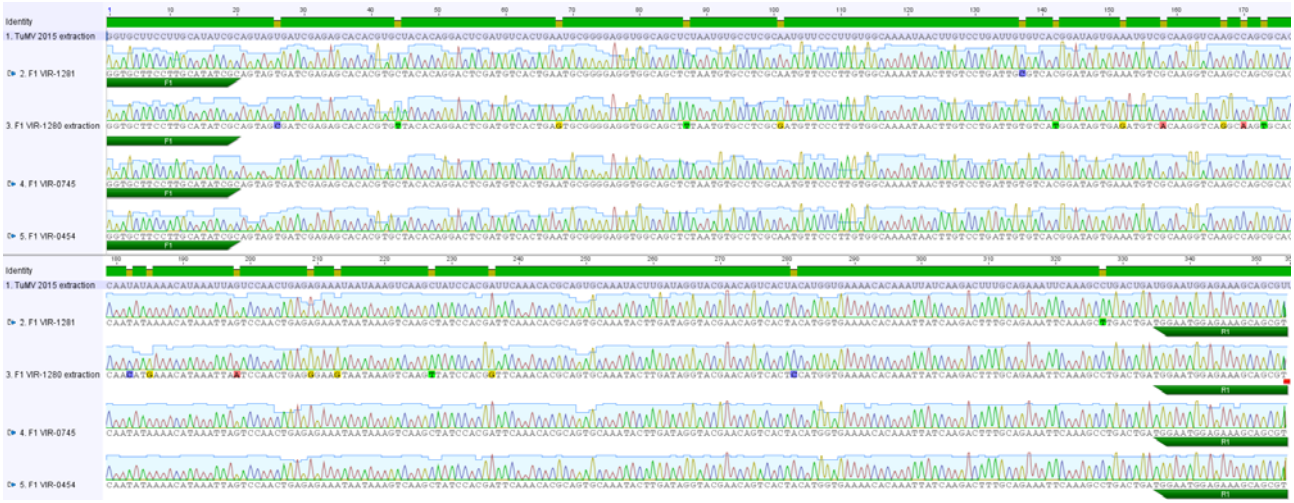


Figure 3: Fragment 1 aligned to the TuMV-QLD1b; the fifth sequence in the identity column named TuMV 2015. F1 and R1 (Table 3) are the primers designed to amplify the fragment. The highlighted base pairs show the SNPs differing from TuMV-QLD1b, particularly throughout VIR-1280 the third sequence in the identity column. The second sequence VIR-0745 showed no mismatches and is identical to TuMV-QLD1b.

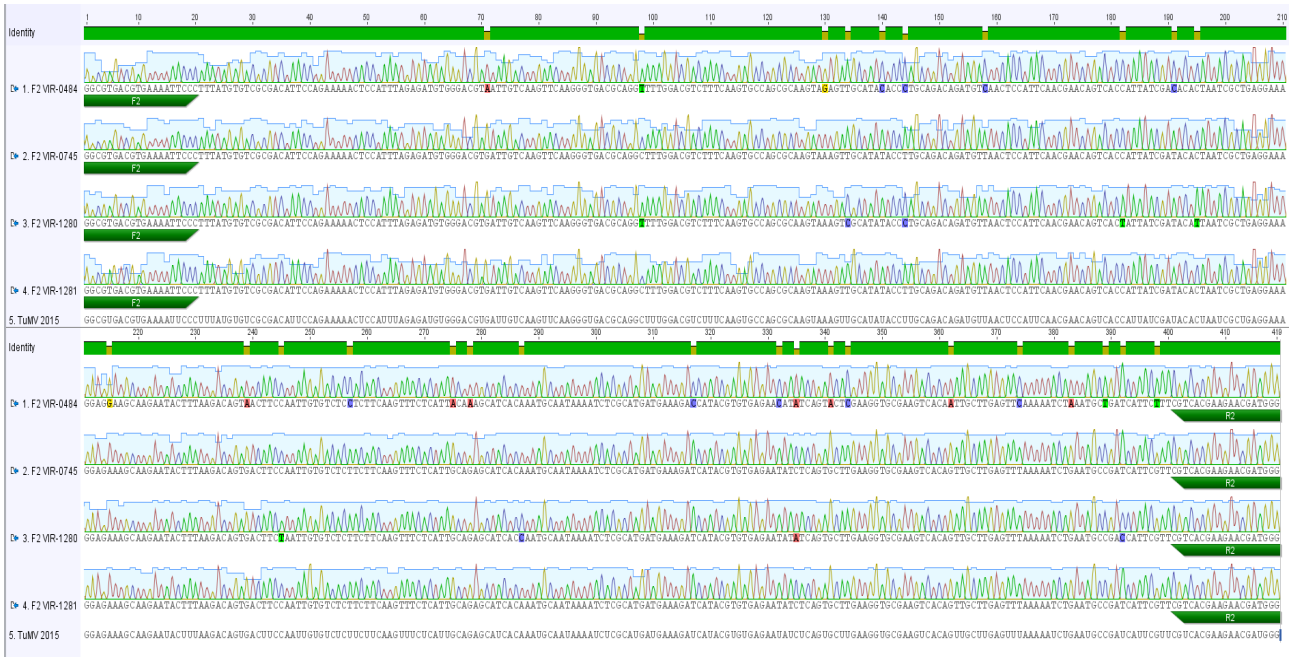


Figure 4: Fragment 2 aligned to the TuMV-QLD1b; the fifth sequence in the identity column named TuMV 2015. F2 and R2 (Table 3) located at the beginning and end of the sequences, respectively, are the primers designed to amplify the fragment. The highlighted bases show the SNPs differing from TuMV_QLD1b, particularly throughout VIR-0484 the first sequence in the identity column. Once again VIR-0745 showed no mismatches.

1.2.2 Sequence Analysis of TuMV-QLD1 before and after Seven Years of Serial Passaging

An alignment of TuMV-QLD1b and BRS1 revealed a high number of sequence variation between the two sequences (Figure 5). A total of 104 variations were identified between the two sequences with the majority being classified as SNPs and indels.

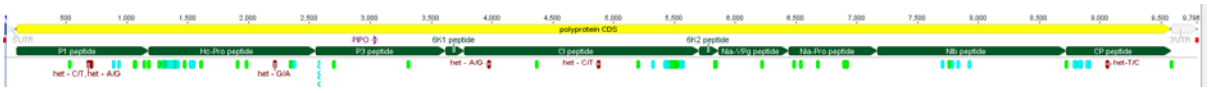


Figure 5: An alignment between TuMV-QLD1b (dark green) and BRS1 (yellow) revealed high variations throughout the genome. Positions of the variations are highlighted below the genome organisation of BRS1; SNPs in bright green, het SNPs in red and indels in light blue. In total there were 104 alterations; 6 het SNPs, 52 SNPs and 46 indels.

As it was unclear whether these variations were due to evolution or merely sequencing errors. A viral sample from 2011 was sequenced via RNA sequencing revealing a higher identity to the TuMV-QLD1b sequence than the BRS1 sequence. The 2011 sample had 99.9% identity to TuMV-QLD1b while it only had a 98.9% identity score to BRS1. The alignment between the 2011 isolate and TuMV-QLD1b reveals only five differences, four of which are possible due to lack of sequence coverage and one being a heterogenous (het) SNP at position 9725 where there was a C for the 2011 sequence but could either be a C or a T for the 2015 sequence. This suggested the high variability between

BRS1 and TuMV-QLD1b was not due to evolution and were more likely sequencing errors. To further confirm that this was true the original isolate VIR-0745 was sequenced and compared to all three of the above-mentioned isolates.

VIR-0745 was sequenced via PCR with ten overlapping fragments by Nasser Hussein, another PhD candidate in the Schenk Lab. Sanger sequencing with two-fold coverage of both forward and reverse orientations were used to ensure the sequence was accurate. A nucleotide alignment of VIR-0745 and TuMV-QLD1b revealed a total of 18 SNPs between the sequences. The result therefore confirms that the variation between TuMV-QLD1b and BRS1 was not due to evolution, but rather sequence errors in the compilation of the BRS1 sequence.

Table 4: Nuclear acid and protein changes found between TuMV-QLD1a and b for the viral various genes. The SNPs and protein changes are from the original isolate sequence to the 2015 version.

Gene	SNPs	Protein Changes
P1	A-G	Silent
	C-Y	Silent or Alanine-Valine
	G-R	Silent
	C-T	Silent
HC-Pro	C-T	Alanine-Valine
	T-C	Silent
	A-R	Silent or Asparagine-Serine
P3	A-G	Silent
	R-A	Silent or Arginine-Glutamine
CI	A-C	Isoleucine-Leucine
	C-Y	Silent
	G-A	Silent
NIA-VPg	R-G	Silent or Histidine-Arginine
	A-G	Silent
NIA-Pro	A-G	Silent
	C-T	Silent
NIB	C-A	Asparagine-Lysine
CP	C-Y	Silent

Of the eighteen SNPs, seven were het SNPs and four of these were in TuMV-QLD1b. More SNPs were found in P1 than other protein coding sequences, where two were SNPs and two het SNPs. HC-Pro and CI had the second highest variation with three SNPs one of which was a het SNP. P3, NIA-VPg and NIA-Pro each had two SNPs and 6K1 and 6K2 had none. An amino acid alignment revealed seven variations suggesting eleven of the SNPs were silent mutations and all conserved amino acid motif were not altered due to these variations. Of the seven variations three resulted in amino acid changes as detailed in Table 4.

1.2.3 Phylogenetic Analysis of TuMV-QLD1b

As mentioned in the introduction, a study was recently conducted on the phylogenetic analysis of Australian TuMV isolates in relation to the other TuMV isolates available on GenBank (Nyalugwe *et al.* 2015). Using the same TuMV isolates as the Nyalugwe *et al.* (2015) study phylogenetic analysis was performed using TuMV-QLD1b in place of BRS1. This tree placed TuMV-QLD1b in the same clade with WA-Ap and TIGA with a boostrap support of 61% and forming part of the basal-B group. One isolate (AB701698 BEL1) which was previously part of the Basal-BR group fell within the World-B group with a bootstrap support of 98%. The Orchis, Asian-BR and Basal-B remained unchanged (Figure 6).

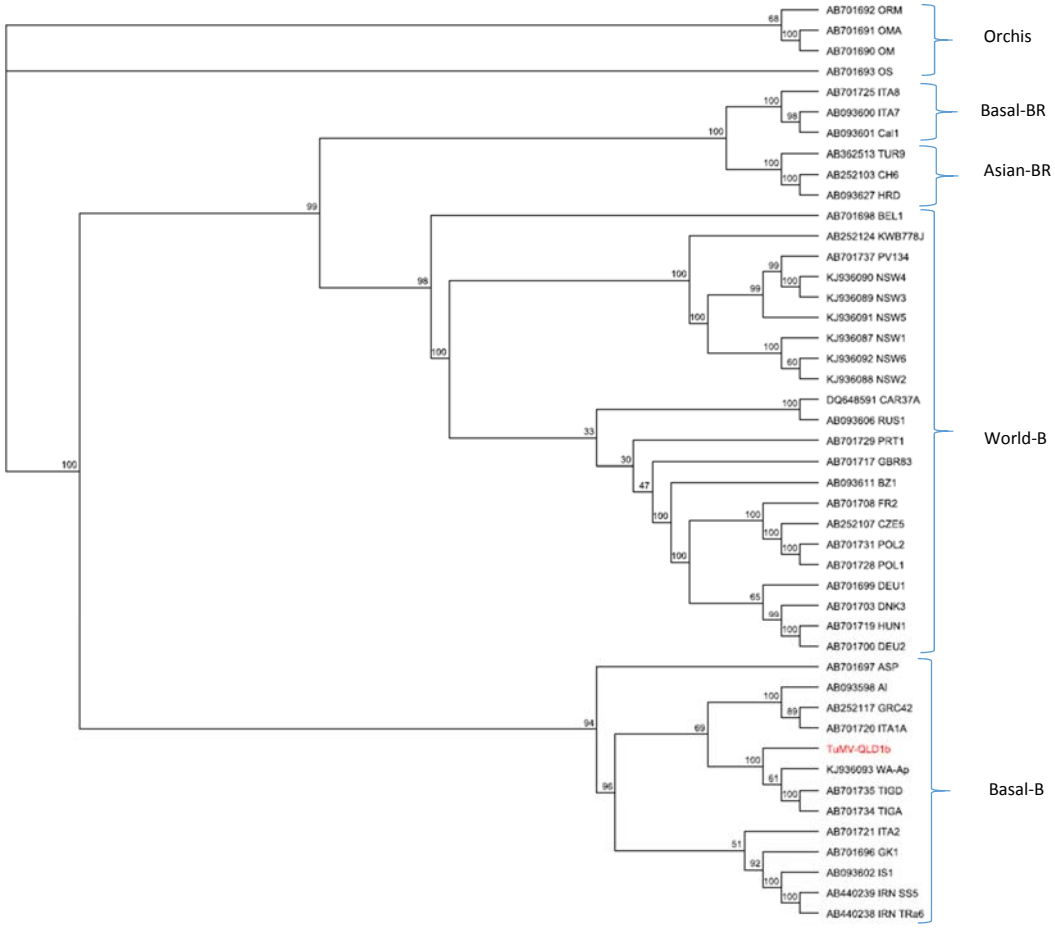


Figure 6: Phylogenetic tree assembled from the alignment of 45 complete genome sequences of TuMV isolates obtained from GenBank. The sequences were aligned in Geneious using Geneious Alignment feature and the PHYML tree builder was used to assemble the tree. TuMV-QLD1b is highlighted in red.

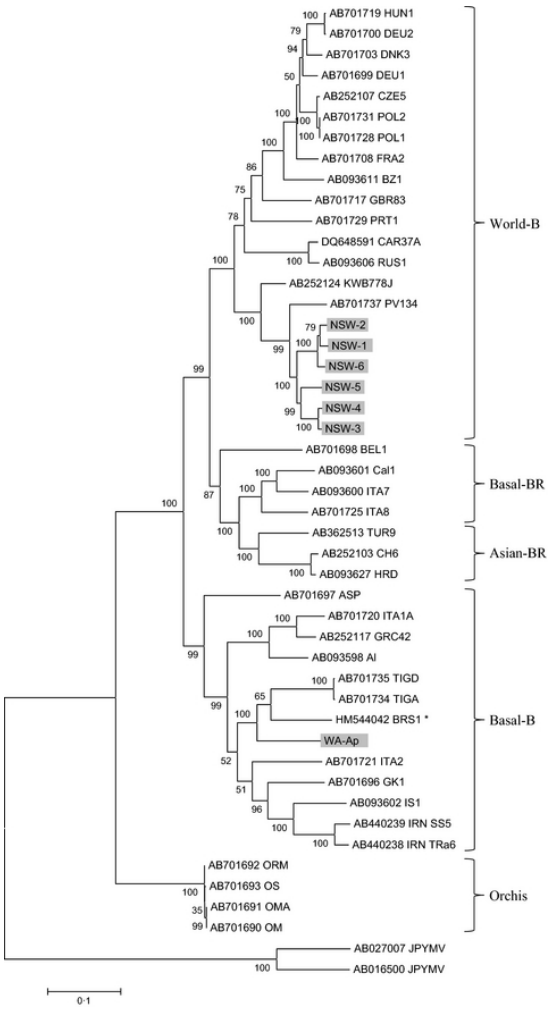


Figure 7: Phylogenetic tree from Nyalugwe *et al.* (2015) paper showing the positioning of BRS1 relative to WA-Ap and the other Australian isolates highlighted in grey.

1.3 Discussion

The origin of a virus is important in understanding how it evolves and if necessary how best to create a management strategy. With regards to the TuMV used in this study it was important to find the original isolate to determine what the original sequence was and whether the virus had evolved during the period it was passaged. As we did not know the isolate number we used other pieces of information to narrow the possibilities.

The original isolate was found in Queensland close to Toowoomba and isolated from Chinese cabbage. It was found by David Carey on 14 May 1997. A few years later an ELISA confirmed the sample was positive for TuMV with further confirmation of flexuous rod particles as seen under an electron microscope.

The virus was given to the Schenk Lab in approximately 2007 as the subject of a PhD project, it was propagated in *N. benthamiana* and studied in *A. thaliana*. The virus was PCR sequenced as part of this project. In 2013 I began my research and propagated the virus in *N. benthamiana* from a frozen sample stored in 2012. The virus isolate was re-sequenced via RNA sequencing. I was surprised to find many sequence variations between the previously PCR sequenced version (~2008) and the deep RNA sequenced version. It was first thought this variation could be due to evolution and mechanically passaging of the virus could have posed certain selective pressure. Though after some research I soon concluded that the differences were most likely sequencing errors as viral evolution is a complex and lengthy process with many influential factors posing varying selective pressure caused by the virus itself, the host or the environment. Though the idea that these errors were due to evolution is explored below.

The idea that RNA viruses can evolve in an unpredictable manner due their RdRp lacking proofreading ability (Elena 2016) than DNA viruses is only partly true. As there are many aspects which drive evolution and the mutation possibly caused by the RdRp is only small factor.

Mutations are a source of genetic diversity within viruses populations enabling the establishment of diverse populations (Roossinck 2003) created by RdRp as well as other cellular enzymes which the virus may encounter in their host causing nucleotide modification, insertion or deletions (Sanjuán and Pilar 2018). The diversity of a virus population is created by certain selective pressures that are brought about by virus- and-host dependent processes (Poirier and Vignuzzi 2017). Once the diversity amongst a population of RNA viruses is stable, it is known as quasi-species and the level of diversity was found to be host dependent (Schneider and Roossinck 2001). It was thought that this could have been a possible driving force behind the variation between the BRS1 PCR sequenced version and 2015 TuMV-QLD1b. As the passaging to different hosts could possibly have influence the virus genome sequence. The paper by Schneider and Roossinck (2001) refers to the diversity within a quasi-species as a cloud size which was stable when passaged between the same host, however changed when the virus was passaged into a new host suggesting the viruses adapts to its new host through sequence alteration.

The fact that diversity and cloud size are host dependent is further confirmed by studies showing that viral mutation rates are also host dependent. A study by Pita *et al.* 2007 found that a CMV strain's replicase fidelity was subject to its host with a higher mutation rate recorded in pepper than in tobacco. This may be due to host specific antiviral factors leading to oxidation or methylation of bases altering the mutation rates (Sanjuán and Pilar 2018).

It was also found that structure of the genome could cause an increase in mutations as replicases paused increasing template slippage certain secondary structures leading to deletions. This ties into the fact certain regions of the genome are more prone to mutations and considered mutation hotspots which is mostly likely due to the nucleotide base arrangement and the resulting secondary structure (Pathak and Temin 1992, Konstantinova *et al.* 2006, Wang *et al.* 2018). The paper by Pita *et al.* 2007 also found that deletions were more frequent that inserts even without the pausing of the replicase. This was achieved by using a "nonstop" region where the ration of insertions to deletion was 1:49. In this study no deletions or insertions were record, only minor base changes.

It is interesting to consider that deletions are more frequent than insertions as this probably relates to the of the size genome being a key factor of the rate of mutation. The main reason RNA viruses are considered to have a higher mutation rate that DNA viruses is due the size of their genome with the size of RNA genome believe to be capped as large genomes experience increased numbers of lethal mutations given the same mutation rate (Belshaw *et al.* 2008, Duffy 2018). This would also be linked to Muller's ratchet were a population crashes due to the accumulation of irreversible deleterious mutations (Novella *et al.* 1999, Holmes 2003). Though RNA viruses are believed to have a mutation rate that is just under the error threshold preventing extinction (Duffy *et al.* 2008).

The environment is another contributing factor to viral mutation as ultraviolet (UV) radiation may have an effect on viral mutation (Duffy *et al.* 2008). Climate change may also have an impact on viral mutation and survival rates with increased temperature and solar ultra violet radiation affecting the hosts and virus ability to survive (Williamson *et al.* 2014, Jones 2016). Though, as this study was performed in growth cabinets the mutations within the viral genome were not due to UV light.

The relationship between mutation and substitution rates must be addressed to fully understand how these two forces are able to shape the genetic structure of a population (Duffy *et al.* 2008). A mutation rate is defined as random genetic errors which can be in the form of single-base pair mutations, deletions or insertions per round of replication (Sanjuán *et al.* 2010). Malpica *et al.* (2002) were the first to determine the mutation spectrum of a plant RNA virus, TMV and found that per genome it was estimated to be 0.10-0.13. Their study showed the 69% of the mutations were either insertions or deletion and that about 35% of those were multiple mutations, meaning the mutation involved 3 or more bases. Though this was an estimate based on the movement protein only and not a genome-wide assessment. Recent research has suggested this may not accurate representation of viral evolution (Jacquemond 2012) but instead a gene specific mutation frequency and would not be a suitable method of assessing the substitution rate. Similarly, the comparison being made in this study could

not accurately assess the substitution rate of the TuMV strains as the genome was only sequenced three times which would not give an accurate result.

Substitutions are those mutations which become fixed in a population either by natural selection or genetic drift usually accessed per year and not per replication round. According to the definition given by Duffy *et al.* (2008) it is the product of four factors: mutation rate, generation time, population size and fitness. Traditional method for studying substitution rates may be flawed leading to over-estimations of substitution. As phylogenetic analysis data are usually a combination of long-term substitutions but also short-term mutations. *Turnip mosaic virus*' average substitution was calculated at 10^{-3} per site per year and considered to be under strong purifying selection (Gibbs *et al.* 2015). Other driving forces of plant virus evolution are recombination and re-assortment (Roossinck 1997). Recombination is an equally important source of genetic variation among virus populations as mutation and has been shown to play a vital role in the speciation of certain viral taxa (Roossinck 1997, Worobey and Holmes 1999, Tromas *et al.* 2014). While re-assortment usually occurs within viruses, which have segmented genomes, like recombination it is also important for the spread of beneficial mutations (Muller 1932) and possibly ridding the genome of deleterious changes (Muller 1964). TuMV is subject to recombination, and a paper by Ohshima *et al.* (2007) showed detailed evidence of hotspots throughout the genome. The study included 92 different TuMV isolates and found the regions of genome with the statistically significant recombination sites were P1, CI, 6K2 and VPg. The comparison of BRS1 to TuMV-QLD1b suggests that approximately 1.07% of the genome had changed over five years with the regions of highest variability being P1 and HC-Pro. Even when considering that both mutation and recombination frequency of TuMV are relatively similar, it seems unlikely the sequence would have altered this drastically during the time period due to evolution. As summarized by García-Arenal *et al.* (2003) no such highly variable plant RNA viruses have been reported and the human viruses which are consider highly variable may be exceptions as the stability of a virus population is considered the rule rather than the exception (Schneider and Roossinck 2000). This was further confirmed once the original viral isolate VIR-0745 was sequenced (TuMV-QLD1a), as well as a sample from 2011, which both confirmed the large number of sequence variations were due to sequencing errors in BRS1, and not due to evolution. Though it is interesting to note that the regions which had the most SNPs were the recombination "hotspots" P1 and the area between C1 and 6K2 (Ohshima *et al.* 2007). The P1 gene had the greatest number of changes as it is under increased purifying pressure with the highest number of sites under positive selective pressure (Gibbs *et al.* 2015).

TuMV-QLD1a and TuMV-QLD1b differed by 0.184% with a total of 18 SNPs. The nucleotide changes between the isolates were found to result in seven amino acid changes in regions of the genome that were not conserved (data not shown). The regions of the genome important for aphid transmission were of particular interest. As the virus was mechanically inoculated during the past five years, it was hypothesised that these regions may experience a loss of selective pressure resulting in a loss of aphid transmissibility. This has previously been reported by Sako (1980). The two conserved areas known to have a role in aphid transmission are the KITC and PKT located in the HC-Pro. Any mutations in these two regions have severe impact on viral-aphid transmission as the KITC motif had a role in transmission efficiency as well as the binding of the HC-Pro to the CP (Huet *et al.* 1994). The PKT formed the other end of the "bridge" binding the HC to the aphid food canal lining (Blanc *et al.* 1998). Upon inspection, both sequence regions were unaltered, including the DAG motif located near the N-terminus of the CP which is equally important for aphid transmission (Lopez-Moya *et al.* 1999), suggesting no loss of transmissibility.

Both isolates are quasi-species displaying a heterogeneous mixture of two different alleles. This as mentioned above allows the virus to quickly adapt to a new host or environment as the specific genome which is best suited to the host may already have been generated by chance (Schneider and Roossinck 2001). More het SNPs are found in the TuMV-QLD1b than TuMV-QLD1a which could be related to its passaging to 3 different hosts over the five years causing its cloud size and mutation rate to change allowing for adaption and creating a more heterogeneous species. Depending on the selective pressure this strain is under one base of the heterogeneous SNPs may become favoured and eventually fixed as a substitution in the population.

This idea leads into the fact that certain base changes within the TuMV genome have been related to host switching as well as symptom severity. A study by Jenner *et al.* (2003) reported that a single amino acid change in the P3 protein of TuMV, previously established as an avirulence determinate, was responsible for the symptom severity of the isolate CDN1 in *Brassica* as well as the resistance breakage of certain genes. The phenotype of both strains is similar suggesting none of the changes resulted in symptom severity and no host switch assay were performed.Though, host switch is also a complex length process and would not have occurred over the time-period of this study. A recent study by Gibbs *et al.* (2015) reviewed how genetic changes resulted in *Turnip mosaic virus'* host switch becoming an import disease agent of brassica crops. The host switching was estimated to have occurred roughly 1000 years ago. The review concluded that though certain genes are known to be host determinates; P1, CI, 6K2 and VPg, it was unlikely that the changing of a single nucleotide led to the switch. Instead it was more likely the switch occurred through the dynamic quasi-species nature

of *Turnip mosaic virus* where several versions of the genome or more specifically genes infected a unique host possibly resulting in a successful infection. As host switching is an important part of viral evolution pinpointing an exact determinate would be difficult. It would be safe to assume that host switching and evolution are driven by the intricate interactions occurring between the viral proteins, the virus and its host factors both of which are further influenced by the environment. Lastly, did the changes discussed in TuMV-QLDa genome alter its position in a relation to the other TuMV strains which it was previously compared to? The TuMV pathotyping system groups strains at a biological level based on their host range. Host type ([B]) and [B] both only infect *Brassica* species and not *Raphanus*; ([B]) latently infects some *Brassica* species while host type [B] infects most *Brassica* species which is evident by the signature mosaic symptoms. Host type [B(R)] is known to infect *Brassica* species symptomatically and *Raphaus* latently. Host type [BR] infects both *Brassica* and *Raphanus* symptomatically. At a molecular level TuMV is grouped phylogenetically which also correlates to its geographical location. The basal-*Brassica* (B) is highly variable found throughout Europe and contains host type [(B)]. World-B is also highly variable and includes strains for other parts of the world excluding Europe. Basal BR is less variable contain host type [BR]. Lastly, the Asian BR also only contains host type [BR] but has the least diversity and isolates in this group all originated in Japan. Similar to the results obtain by Nyalugwe *et al.* (2015) TuMV-QLD1b was placed within the basal-B phylogenetic group (Figure 6 and 7) along with WA-Ap isolate but separate from the other six Australian isolates which formed part of the World-B group. The basal-B and basal-BR group are considered the oldest groups thought to have originated in the Europe–Mediterranean–Asia Minor region (Ohshima *et al.* 2002). This also suggests that these isolates are well established in their infection of *Brassica* crops and mostly likely spread throughout the world with modern agriculture. Considering that both TuMV-QLD1b and the Western Australian isolates are not part of the World-B group they were probably introduced during the British colonisation of Australia and the other six were possibly a more recent quarantine breaches (Gibbs *et al.* 2008). This suggest that changes found in TuMV's genome did not alter its position in relation to the other strain it was previously grouped with.

With these results, I have shown the importance for correct sequencing and knowing the origin of a viral isolate. Without the original isolate it would have appeared that the virus had drastically evolved over the seven-year time frame differing 1.07%. Once VIR0745 was found to be the original virus it was confirmed that the two isolates only differed by 0.184% with most mutations being silent and conserved regions remaining unchanged. I also confirmed that mechanical inoculation did not apply any selective pressure on the virus with no changes to any aphid transmission motifs. Future research

should include determining whether the host range of the virus has increased to included plants from the *Raphanus* species.

Characterising the genome is an important step in the understanding of how the virus is able to overcome host defences and cause disease. The genome sequence enables the characterisation of important genes and can help identify virulence factors owing to the success of the virus-induced defence modulation. The TuMV isolate nucleotide sequencing data was published in the following research article.

Complete Nucleotide Sequence of an Australian Isolate of *Turnip mosaic virus* before and After Seven Years of Serial Passaging

Lara Pretorius, Richard L. Moyle, Jessica Dalton-Morgan, Nasser Hussein, Peer M. Schenk

Nexgen Plants Pty Ltd and School of Agriculture and Food Sciences, The University of Queensland, Brisbane, Queensland 4072, Australia

Abstract

The complete genome sequence of an Australian isolate of *Turnip mosaic virus* was determined by Sanger sequencing. After seven years of serial passaging by mechanical inoculation, the isolate was re-sequenced by RNA-Seq. Eighteen single nucleotide polymorphisms were identified between the original and passaged isolates. Both isolates had 96% identity to isolate AUST10.

Genome Announcement

Turnip mosaic virus is a *Potyvirus* member within the *Potyviridae* family. Due to the wide distribution and extensive host range, *Turnip mosaic virus* (TuMV) is considered to be the second most damaging crop virus across 28 different countries and regions (Walsh and Jenner, 2002). TuMV virions are flexuous filaments of approximately 700-750 nm length. TuMV has a single-strand positive-sense RNA genome, approximately 10 kb in length, encoding a single open reading frame flanked by two untranslated regions. The translated polypeptide is cleaved and processed into ten viral proteins.

Here we report the complete sequence of an Australian isolate of TuMV, designated TuMV-QLD1a, from a field-grown *Brassica pekinensis* plant from Toowoomba, Queensland, in 1997 (Department of Agriculture and Fisheries plant virus collection accession VIR0745). RNA was isolated from mechanically inoculated *Nicotiana benthamiana* and cDNA synthesised as previously described (Moyle *et al.*, 2016). Ten overlapping fragments were PCR-amplified and Sanger-sequenced with at least two-fold coverage in both the forward and reverse orientation. The TuMV isolate was subsequently re-sequenced after being passaged over a seven year period by serial mechanical inoculation of *N. benthamiana* seedlings every two-three months. The re-sequenced isolate, designated TuMV-QLD1b, was resolved by RNA-Seq analysis using the TruSeq RNA library synthesis kit (Illumina) and the MiSeq platform (Illumina). After quality trimming, 33,752,578 RNAseq reads were obtained and Geneious v8.1.7 software was used to map 1,482,837 reads to the TuMV-QLD1a reference genome. Average depth of coverage was 15,905, with maximum coverage of 26,792 and minimum coverage of 578.

Both TuMV-QLD1a and TuMV-QLD1b are 9,796 nt in length. However, 18 single nucleotide polymorphisms (SNPs) were identified between the two isolates. Two heterogeneous SNPs in TuMV-QLD1a were lost in TuMV-QLD1b, whereas five different heterogeneous SNPs emerged in TuMV-

QLD1b. An amino acid alignment revealed ten amino acid changes from TuMV-QLD1a to TuMV-QLD1b. Both isolates share 96% identity with the AUST10 isolate (accession number AB989634), which was also collected in Queensland, Australia, in 1996 (Yasaka *et al.*, 2015). Phylogenetic analysis places TuMV-QLD1a and TuMV-QLD1b in the basal-B group and sub-group basal-B2. This group is unable to infect *Raphanus* plants but does infect *Brassica* plants systemically causing phenotypic symptoms (Tomimura *et al.*, 2003). This is considered to be the most variable group as it emerged paraphyletic to the other lineages (Nguyen *et al.*, 2013). This sub-group is most closely related to four German isolates DEU7, AIIA, TIGD and TIGA (accession numbers: AB701695.1, AB701694.1, AB701735.1and AB701734.1) and is thought to have been introduced into Australia and New Zealand via horticultural material (Yasaka *et al.*, 2015). DEU7 and AIIA have 87% identity to both versions of TuMV-QLD1, while TIGD and TIGA have 86% identity.

Nucleotide accession numbers.

The GenBank accession number for TuMV-QLD1a and TuMV1b are KX641465 and KX641466, respectively.

Acknowledgements

We thank the Department of Agriculture and Fisheries, QLD, Australia for providing the TuMV isolate. The study was funded by Australian Research Council discovery grant (DP140103363).

References

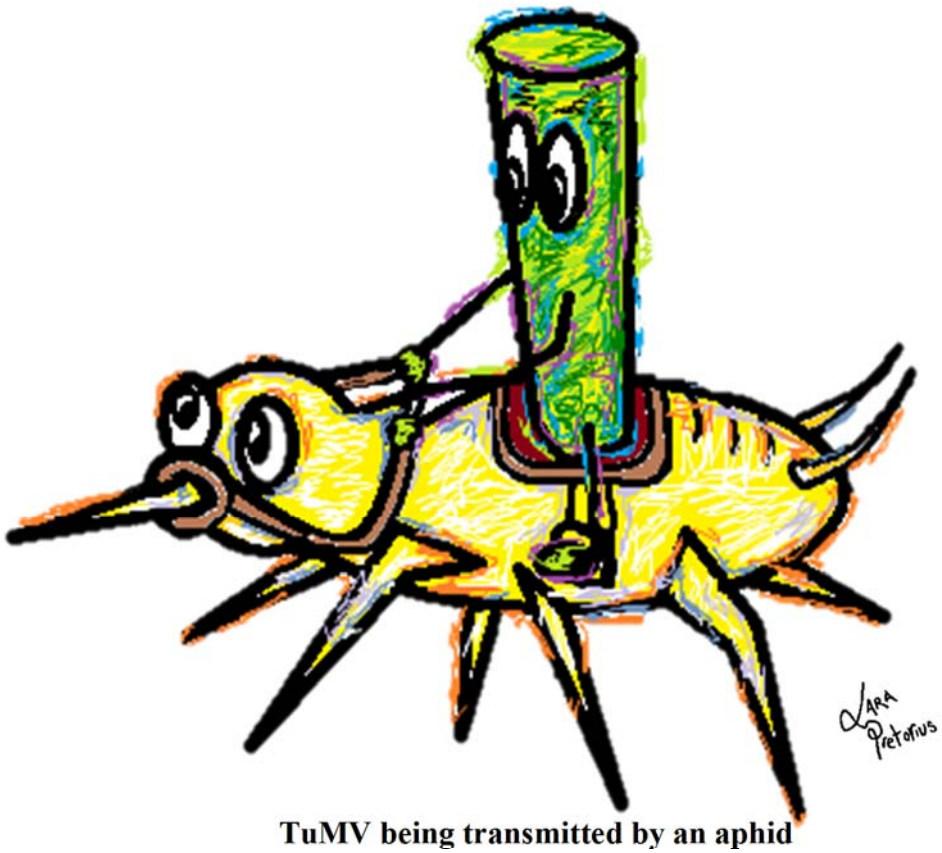
Moyle, R., Pretorius, L.-S., Dalton-Morgan, J., Persley, D., Schenk, P., 2016. Analysis of the first complete genome sequence of an Australian tomato spotted wilt virus isolate. Australasian Plant Pathology, 1-4.

Nguyen, H.D., Tran, H.T.N., Ohshima, K., 2013. Genetic variation of the Turnip mosaic virus population of Vietnam: A case study of founder, regional and local influences. Virus Research 171, 138-149.

Tomimura, K., Gibbs, A., Jenner, C., Walsh, J., Ohshima, K., 2003. The phylogeny of Turnip mosaic virus; comparisons of 38 genomic sequences reveal a Eurasian origin and a recent 'emergence'in east Asia. Molecular ecology 12, 2099-2111.

Walsh, J.A., Jenner, C.E., 2002. Turnip mosaic virus and the quest for durable resistance. Molecular Plant Pathology 3, 289-300.

Yasaka, R., Ohba, K., Schwinghamer, M.W., Fletcher, J., Ochoa-Corona, F.M., Thomas, J.E., Ho, S.Y.W., Gibbs, A.J., Ohshima, K., 2015. Phylodynamic evidence of the migration of turnip mosaic potyvirus from Europe to Australia and New Zealand. Journal of General Virology 96, 701-713.



TuMV being transmitted by an aphid

Chapter 2: Sequencing of Other Viral Isolates

This chapter focuses on the genome sequencing and characterisation of two other distinct plant viruses used in the study, CaMV (Dar78694) and CMV (strain K). CaMV and CMV form part of the top ten list of economically important plant viruses (Rybicki 2015). Isolate sequences were submitted to GenBank and published as research articles. Both virus isolates were used in Chapter 4 to determine whether the virus-induced modulation of defence is affected when *A. thaliana* plants lack a subunit of the mediator complex.

The CaMV isolate was provided by Dr Geering (DAF). The isolate was sequenced by deep RNA sequencing using the AGRF as the service provider. A total of 3,445,961 reads were obtained of which 130,653 reads were mapped to a reference genome with 100% coverage using Geneious mapper. As there are currently no fully sequenced Australian CaMV isolates on GenBank, the sequencing of this virus is significant. The genome sequence has been submitted to GenBank (accession number KX904357). Phylogenetic analysis was also performed in Geneious aligning open reading frames I-V separately from VI the same analysis performed by Yasaka, Nguyen *et al.* (2014). A research note was published in Australasian Plant Pathology titled "First fully sequenced genome of an Australian isolate of Cauliflower mosaic virus" (Pretorius *et al.*). The paper includes the phylogenetic analysis of open reading frames I-V but not VI; this was due to the opening reading frames I-V being a more reliable indicator of evolution.

Figure 8: Maximum-likelihood phylogenetic estimates for ORF VI of Cauliflower mosaic virus PHYML analyses of 91 isolates. Bootstrap analysis was based on 1000 pseudo-replicates. The Australian isolate is highlighted by the red box. Clades were collapsed using FigTree v1.4.3.

The open reading frames were analysed separately as the recombination and substitution rates suggested that the evolutionary history of the open readings frames were different with unique recombination regions found for I-V and VI (Yasaka *et al.* 2014). The phylogenetic analysis for open reading frame VI placed it in the same clade as I-V did, with three Japan isolates though with less confidence resulting in a bootstrap value of 73 instead of 99. However, it was no longer part of the USA sister clade. Instead the nearest sister clade now had a bootstrap value of 30 and this clade included a number of isolates from various geographic regions including the USA. Similar to the analysis by Yasaka *et al.* (2014) two major lineages are produced, though ours was not as clear cut. As seen in Figure 8 the tree appears to be divided into Iran isolates, Turkey isolates and then diverse geographical locations. The Australian isolate is grouped in the same clade in both trees. The three Japanese isolates it is grouped with were described as attenuated as they had very faint symptoms but could infected both *Brassica* and *Raphanus* species (Yasaka *et al.* 2014). Future research should include inoculation of *Raphanus* species to determine if CaMV-Dar78694 is able to infect this

species. However, I would not describe the Australian isolate as having faint symptoms in *Brassica* species, as infected plants produced clear and pronounced symptoms (Figure 9).

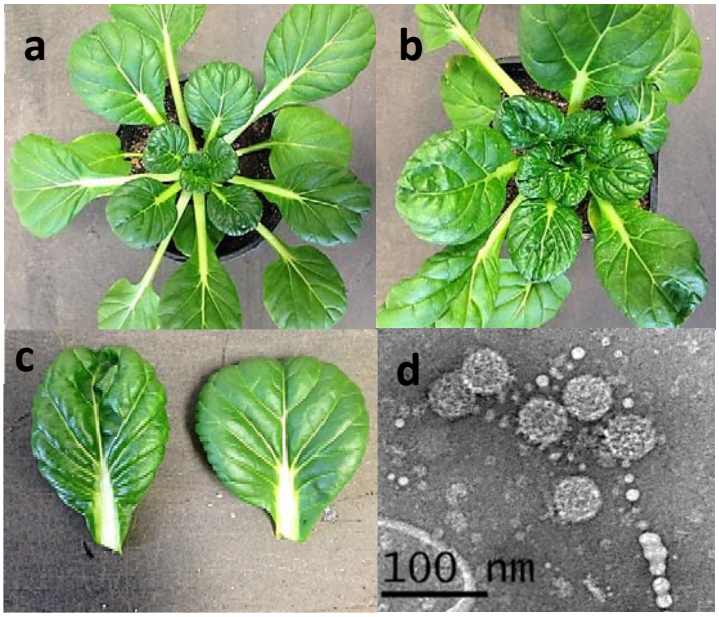


Figure 9: CaMV-Dar78694 symptoms in Bok Choy four weeks after inoculation.

(a) Non-inoculated Bok Choy. (b) CaMV inoculated Bok Choy display rosette distortion, and (c) leaves have a faint mottle and curl, becoming increasing wrinkled and distorted. (d) Electron microscopy of CaMV viral particles. The particles are isometric in shape and approximately 50 nm in diameter.

The second virus isolate sequenced was CMV-K which was kindly provided by John Randles in 2004 for a former PhD student's project. CMV is an ssRNAs virus with a tripartite genome. Segments are denoted RNA 1, 2, and 3, and are roughly 3360bp, 3050bp and 2200bp, respectively. The various segments are all positive sensed and coded for different viral proteins (Roossinck 2002). The viral replicase is coded on RNA 1 and 2 from ORF 1a and 2a. RNA 2b overlaps with 2a while the CP and viral movement protein are both coded for on RNA 3 (Palukaitis *et al.* 1992). This isolate is also used in Chapter 4 due to its host range including *A. thaliana*.

This isolate originated in China and certain parts of its genome, RNA2 and 3, were previously sequenced and submitted to GenBank (S72187 and AF127977). Deep RNA sequencing was performed on tomato infected tissue and the sequences were assembled by Dr Moyle in Geneious using Geneious mapper. Due to considerable *complementarity* of the 3' ends of each segment, assembly required thorough analysis to ensure correct alignment. Sequences were submitted to GenBank with the following accession numbers: MG182148, MG182149 and MG182150. The phylogenetic analysis was also performed in Geneious as outlined in Rabie *et al.* (2017). The same CP sequences used in this paper were downloaded from GenBank, trimmed and aligned in Geneious. A total of 37 isolates from diverse geographic locations were included in this tree. The RNA1, 2 and 3 also used the same sequences as Rabie *et al.* (2017). The previously sequenced RNA 2 and RNA 3 of CMV-K were included in the analysis. It is not surprising that these two sequences had a percentage

identity of near 100% for both RNA 2 and 3. Dr Moyle also performed a pairwise alignment in Geneious of the earlier RNA2 and 3 with the 2016 version to identify any significant changes.

CMV isolates are grouped based on important data related to the strain's serology, peptides of the CP and the recombination or re-assortment of its nucleotides (Roossinck *et al.* 1999). This divided CMV into subgroup I and II, with a further subdivision of group I into IA and IB. More recently a third subgroup has been added containing two Chinese isolates clearly radiate from the same common ancestor which differs from other subgroups (Liu *et al.* 2009). Phylogenetic analysis of CMV strains have shown that each group evolved from the same ancestor (Roossinck 2002) which is also evident in the phylogenetic analysis performed in this study (Figure 10). Previous research placed this isolate in subgroup IB (Roossinck *et al.* 1999).

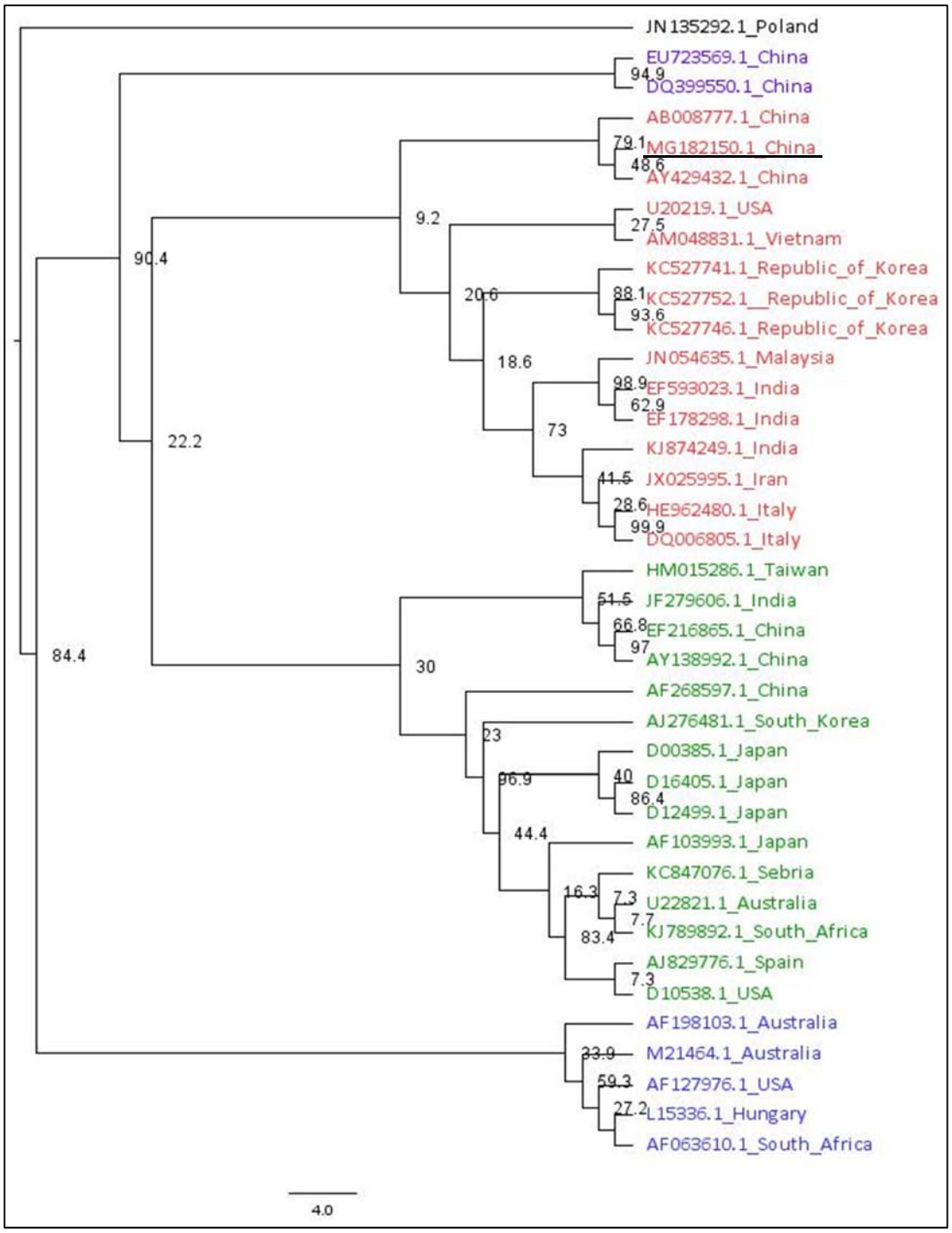


Figure 10: Phylogenetic analysis of CP of CMV isolate from this study (underline in black) in comparison to others from diverse geographic locations. The analysis was performed in Geneious R8 using the maximum likelihood method and bootstrap analysis with 1000 replicates. JN135292_Poland is a Peanut stunt virus (PSV) which was used as the outgroup. The isolates are presented with the accession number and location. The isolates in blue form part of subgroup II, while those in green are subgroup IA and red are IB. The two in purple form a new subgroup III. Bootstrap values are shown in the nodes of the tree and the scale bar indicates four substitutions per site in the alignment.

The CP phylogenetic analysis placed the isolate in a clade with another Chinese isolate with a bootstrap value of 98%. The sister clade had a bootstrap value of 22 and included a geographically diverse range of isolates from; USA, Vietnam, Republic of Korea, Malaysia, India and Italy all belonging to the IB subgroup.

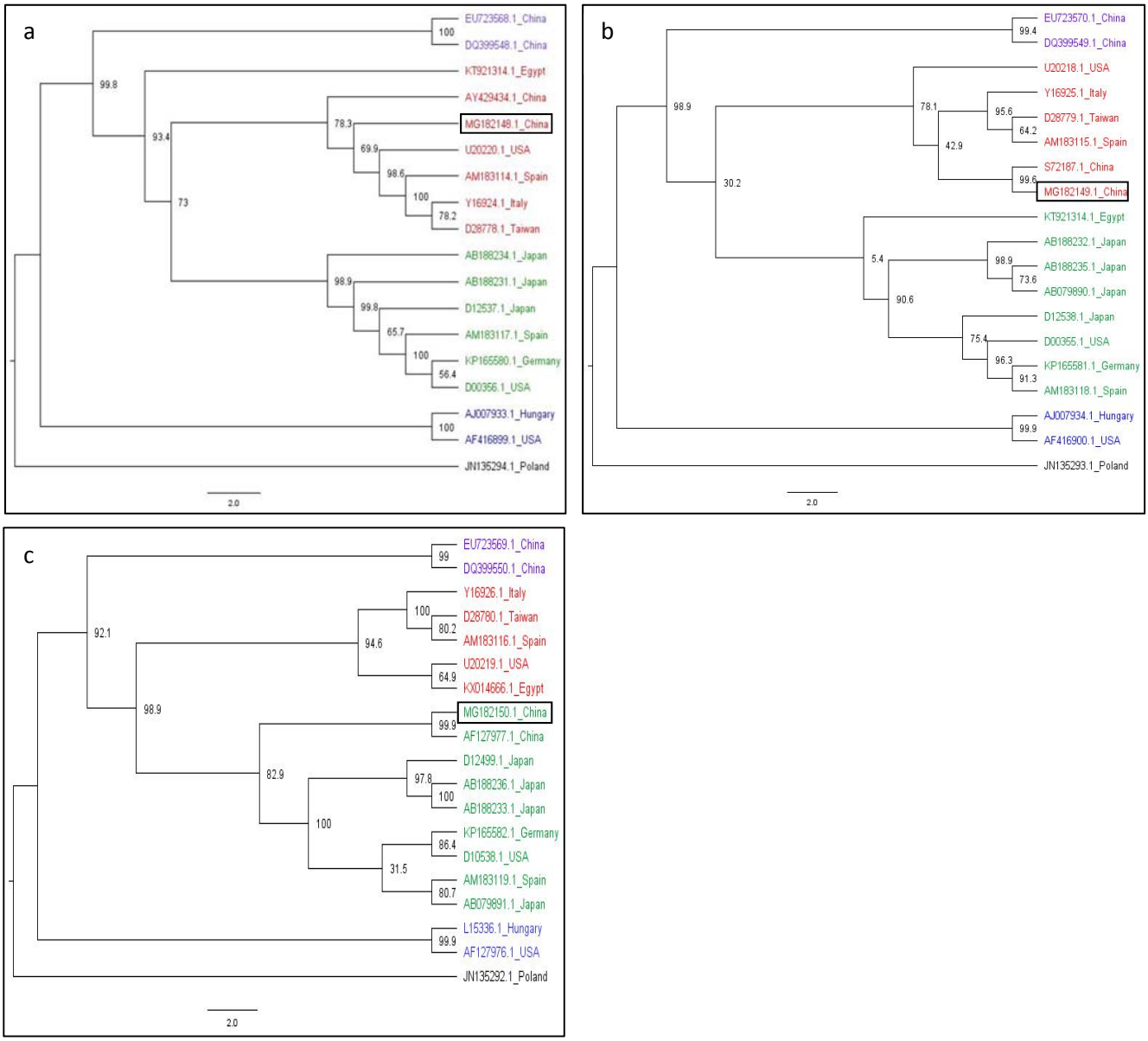


Figure 11: Phylogenetic analysis of RNA 1(a), RNA 2 (b) and RNA 3 (c) of the CMV isolate compared to those previously used in Rabie *et al.* (2017) study. The trees were drawn in Geneious R8 using the maximum-likelihood method with 1000 bootstrap replicates. The outgroup used was a Peanut stunt virus (JN135292) from Poland. The colours denoted the subgroups; subgroup II is blue, subgroup IA is green while IB is red, Subgroup III is purple. The isolate used in this study is indicated by the black box. Bootstrap values are shown in the nodes of the tree and the scale bar indicates two substitutions per site in the alignment.

Interestingly, according to the tree for RNA3 (Figure 11c) the isolate appears to form part of the subgroup IA (green) and not IB (red). This new grouping is possibly due to recombination or re-assortment which are both important in CMV evolution (Roossinck 2002). This also proves that phylogenetic analysis cannot be performed solely on the CP of the virus as this does not provide an accurate presentation of an isolate's evolution (Jacquemond 2012). Natural re-assortment between subgroups has been documented where results suggest that re-assortment between subgroups IA and IB is more common that between subgroup I and II (Bonnet *et al.* 2005). The re-assortment between IA and IB is more common due to the mixed infection being more likely allowing these strains to re-assort. There is a high degree of sequence variation between subgroups I and II making re-assortment less likely. Though reports of re-assortment between subgroups I and II are being reported more frequently (Chen *et al.* 2007, Maoka *et al.* 2010, Nouri *et al.* 2014). In this case it is possible that the CMV-K strain was part of a mixed infection with an isolate from IA resulting in the re-assortment of RNA3. Studies have also shown that recombination events are more frequent in RNA3 (Bonnet *et al.* 2005) which may be due to selective advantage (Jacquemond 2012)

Recombination occurs more frequently as there is a higher cost associated with re-assortment due to the amount of genetic information exchanged (García-Arenal and McDonald 2003, Bonnet *et al.* 2005). Therefore, it is more likely that a recombination event occurred between CMV-K and an isolate from IA as the CP forms part of IB, the movement protein and UTR could originate from an IA isolate. Similar results were found by Nouri *et al.* (2014) where recombination events occur in RNA3 between subgroups IA and IB.

Apart from recombination and re-assortment, mutations are another important driving force of CMV evolution. The alignment of the two previously sequenced RNA 2 and 3 which were first sequenced 22 and 17 years ago, respectively, shows that both sequences are stable over time with most of the mutations occurring in non-coding regions. The polymorphisms in RNA2 lead to 5 amino acid changes in 2a and 1 in 2b. Due to a lower number of polymorphisms found in RNA, 3 there was only 2 amino acid changes in both the movement protein and the CP.

Mutation, recombination and re-assortment events are an important source of genetic variability amongst CMV isolates and can be related to host switching or extending as well as resistance breaking (García-Arenal and McDonald 2003). The understanding of these events are important in improving future management strategies, especially since CMV is such a successful virus with a very wide host range (Palukaitis *et al.* 1992).

The genome sequencing and phylogenetic analysis of the Australian CaMV isolate and the CMV-K isolate were published in the following research articles.

First fully Sequenced Genome of an Australian Isolate of *Cauliflower mosaic virus*

Lara Pretorius[1], Richard L. Moyle[1], Jessica Dalton-Morgan[1], *the late* Mark W. Schwinghamer[3], Kathy Crew[4], Peer M. Schenk[1] and Andrew D.W. Geering[2]

[1]Plant-Microbe Interactions Laboratory, School of Agriculture and Food Sciences and [2]Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, St Lucia, Queensland 4072

[3]School of Environmental and Rural Science, University of New England, Armidale, NSW 2351

[4]Department of Agriculture and Fisheries, Ecosciences Precinct, 41 Boggo Road, Dutton Park, Queensland 4102.

Abstract

We report the first fully sequenced genome of an Australian isolate of *Cauliflower mosaic virus*. The circular genome is 8,022 base pairs in length. A phylogenetic analysis suggests recent common ancestry of this virus isolate with those from Japan and the USA, and origins of the clade in Western Europe.

Cauliflower mosaic virus (CaMV) is type species of both the genus *Caulimovirus* and family *Caulimoviridae* (Geering and Hull, 2012). The *c*. 8 kbp, circular double-stranded DNA genome contains six major open reading frames (ORFs) and is encapsidated within a non-enveloped, isometric virion of 52 nm diameter (Geering, 2014). During replication, an intermediary single-stranded RNA, termed the pregenomic RNA, is transcribed from the 35S promoter and this in turn acts as the template for a virus-encoded reverse transcriptase to convert the RNA back to DNA (Pretorius *et al.*, 2017). A second promoter, the 19S promoter, drives transcription of a monocistronic messenger RNA for the ORF VI protein, which has a dual role as a translational transactivator (TAV) and suppressor of gene silencing (Ryabova *et al.*, 2006). Most, if not all of the remaining open reading frames (ORFs) are transcribed from the pregenomic RNA, and translation of the downstream ORFs is facilitated by the TAV protein. Of the remaining five proteins that have been detected *in planta*, the ORF I product is a cell-to-cell movement protein; the ORF II product is required for aphid transmission; the ORF III product is the virion-associated protein; the ORF IV product is the capsid protein precursor and ORF V product is a polyprotein containing aspartic protease, reverse transcriptase and RNaseH domains (Geering, 2014). CaMV is transmitted in a semi-persistent manner by several aphid species, and mainly infects plant species in the Brassicaceae, with only certain strains able to infect solanaceous species (Qiu and Schoelz, 1992).

It is hypothesized that CaMV originated in the eastern Mediterranean region and from there spread to Western Europe, Japan and the USA (Yasaka *et al.*, 2014). Long distance dispersal of the virus is thought to have occurred as a result of the global trade in broccoli, cauliflower and other brassicaceous species, which were grown as antiscorbutics during the 19th and 20th centuries (Yasaka *et al.*, 2014).

In Australia, information on the status of CaMV is surprisingly scarce. In the 1960s, CaMV was considered the most widespread and important pathogen of cauliflower in South Australia but diagnosis was principally by the assessment of symptoms (Randles and Crowley, 1967). Also in this decade, fundamental research on the purification and aphid transmission of CaMV was conducted in the Australian Capital Territory, presumably using an Australian isolate of the virus, although the provenance of the isolate was not described (Day and Venables, 1960; Day and Venables, 1961). It was not until the turn of the millennium that an unambiguous diagnosis of CaMV in Australia was published; it was detected by double antibody sandwich ELISA in canola crops in south-west Western Australia (Coutts and Jones, 2000). In canola, CaMV causes distinctive yellow rings and mild mottling on the leaves, and often is present as a mixed infection with TuMV (Coutts and Jones, 2000; Hertel *et al.*, 2004).

An examination of the GenBank database (8 August 2017) suggested that neither an entire or partial genome sequence of an Australian isolate of CaMV had been deposited. Here we report the first fully sequenced genome of an isolate of CaMV from New South Wales. Phylogeographic analyses suggest a close relationship with isolates from Japan and the USA.

A wild turnip (*Brassica rapa*) sample with yellow vein clearing, which was collected in Mullaley, New South Wales, in 2007, was lodged in the NSW Plant Pathology Herbarium as accession DAR78694. The virus isolate was propagated in the forage brassica hybrid cv. Hunter, and after three serial transmissions, a sample of leaf was freeze-dried for long term storage. For this study, the virus isolate was recovered onto Bok Choi (*Brassica rapa* subsp. *chinensis*) by mechanical inoculation. To provide a preliminary diagnosis, a sap extract was examined under a JEOL JEM-1400 transmission electron microscope using 1% ammonium molybdate pH 6.8 stain and isometric virions of *c*. 50 nm were observed, suggesting infection with CaMV. No other types of virion were observed.

A total RNA extract, including the viral pregenomic RNA, was prepared using the SV Total RNA Isolation System (Promega). Illumina RNA sequencing was done using the Australian Genome Research Facility as the service provider and a total of 3,445,961 reads obtained. The CaMV genome was assembled with Geneious software v8.1.7 as described previously (Moyle *et al.*, 2016; Pretorius *et al.*, 2016). In total, 130,653 reads aligned to the sequence of the type isolate of CaMV, isolate Cabb B-S (NCBI reference genome accession NC_001497), and 100% sequence coverage was obtained,

with minimum and maximum sequencing depths 33 and 9059 per nucleotide, respectively. CaMV-DAR78694 is 8022 nt in length and has been deposited in GenBank as accession KX904357. CaMV-DAR78694 shared a maximum of 96% nucleotide identity with other CaMV isolates in the GenBank database.

For phylogenetic analyses, all fully sequenced genome sequences of CaMV, as well as that of the outgroup *Horseradish latent virus*, were downloaded from GenBank. The DNA sequences of each ORF were conceptually translated and aligned using the MUSCLE algorithm within MEGA v. 6.06, and then back-translated to generate alignments of the original DNA sequences. Following the approach of Yasaka *et al.* (2014), the ORF I-V alignments were concatenated and a maximum likelihood tree inferred using PHYML v3 in Geneious v8.1.7 with the generalised time-reversible substitution model (Fig. 1) (Guindon *et al.*, 2010). Statistical support for the nodes in the tree was obtained by bootstrapping with 1000 replicates.

CaMV-DAR78694 fell within a clade of virus isolates that are widely dispersed around the world, including Japan and the New World (USA and Argentina) but is rooted by isolates from western Europe, namely France and the UK (Fig. 1). It is hypothesized that the virus hitchhiked within living brassicaceous plants that were carried by mariners during the 18th and 19th centuries to ward off scurvy (Yasaka *et al.*, 2014). During the mid to late 18th century, there were large migrations of people to Victoria and New South Wales from the Western Europe, China and the USA in response to the gold rush (http://www.australia.gov.au/about-australia/australian-story/austn-gold-rush). Large numbers of whaling ships from the USA, Britain and elsewhere also worked the east coast of Australia in the first half of the 18th century (http://www.australia.gov.au/about-australia/australian-story/australias-whaling-industry-and-whales). It is tempting to speculate that CaMV was introduced to Australia from the USA during this period of mass migration, presumably in potted plants that provided fresh vegetables for people on ships, as there are no reports of seed transmission (Geering and Hull, 2012).

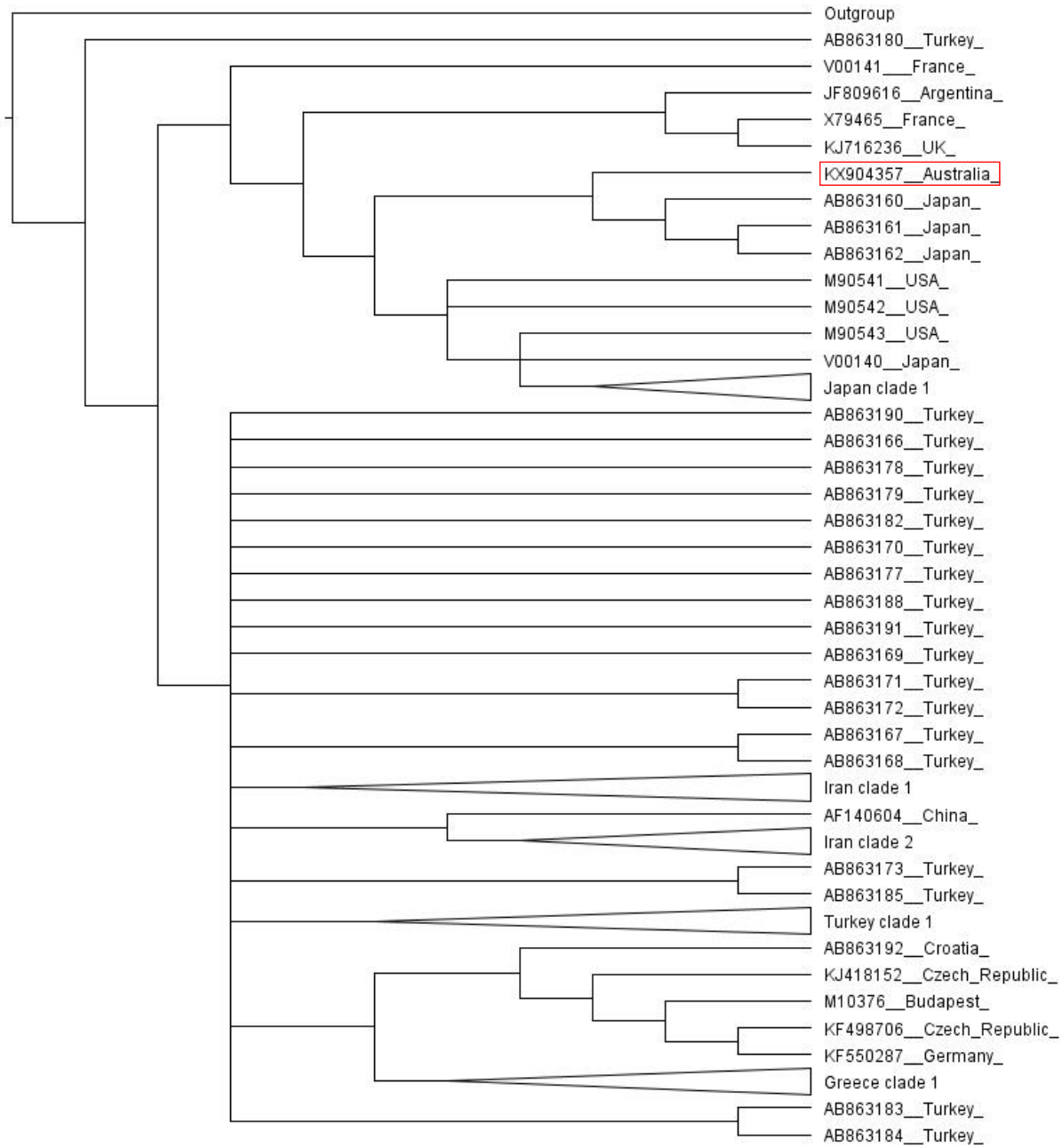


Fig. 1

Maximum-likelihood estimate of the phylogeny of *Cauliflower mosaic virus* based on alignments of open reading frames I to IV. The Australian isolate of CaMV is highlighted by a box and the tree has been rooted using *Horseradish latent virus* (HRLV). To simplify the tree, some clades have been collapsed using FigTree v1.4.3. Bootstrap values are shown in the nodes of the tree and the scale bar indicates two substitutions per site in the alignment.

Funding

This study was funded by the Australian Research Council (Discovery Project grant number DP140103363).

Conflict of interest

The authors declare no conflict of interest.

References

Coutts BA, Jones RAC (2000) Viruses infecting canola (*Bassica napus*) in south-west Australia: incidence, distribution, spread, and infection reservoir in wild radish (*Raphanus raphinistrum*). Australian Journal of Agricultural Research 51: 925-936

Day M, Venables D (1961) The transmission of *Cauliflower mosaic virus* by aphids. Australian Journal of Biological Sciences 14:187-197

Day MF, Venables DG (1960) Purification of *Cauliflower mosaic virus*. Virology 11: 502-505.

Geering ADW (2014) Caulimoviridae (Plant Pararetroviruses). In: eLS. John Wiley & Sons, Ltd. doi:10.1002/9780470015902.a0000746.pub3

Geering ADW, Hull R (2012) Caulimoviridae. In: King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J. (Eds.), Virus Taxonomy. Elsevier, Oxford

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic Biology 59:307-321

Hertel K, Schwinghamer M, Bambach R (2004) Virus diseases in canola and mustard. In: Industries, N.D.o.P. (Ed.), Agnote DPI 495

Moyle R, Pretorius L, Dalton-Morgan J, Persley D, & Schenk P (2016) Analysis of the first complete genome sequence of an Australian *Tomato spotted wilt virus* isolate. Australasian Plant Pathology, 45(5):509-512

Pretorius L, Moyle R, Hussein N, Al-Amery A, Schenk P (2017) Natural and Engineered defenses against plant viruses. Current Biotechnology 6, DOI:10.2174/2213529403666170310143841

Pretorius L, Moyle RL, Dalton-Morgan J, Hussein N, & Schenk PM (2016) Complete nucleotide sequence of an Australian isolate of *Turnip mosaic virus* before and after seven years of serial passaging. Genome Announcements 4:e01269-16

Qiu SG, Schoelz JE (1992) Three regions of *Cauliflower mosaic virus* strain W260 are involved in systemic infection of solanaceous hosts. Virology 190:773-782

Randles JW, Crowley NC (1967) Epidemiology of *Cauliflower mosaic virus* in South Australia. Australian Journal of Agricultural Research 18:289-298

Ryabova LA, Pooggin MM, Hohn T (2006) Translation reinitiation and leaky scanning in plant viruses. Virus Research 119:52-62

Yasaka R *et al.* (2014) The temporal evolution and global spread of *Cauliflower mosaic virus*, a plant pararetrovirus. PLoS One 9:e85641

Analysis of the complete genome sequence of *Cucumber mosaic virus* strain K

Richard Moyle, Lara-Simone Pretorius, Louise S. Shuey, Ekaterina Nowak, Peer M. Schenk[#]

Nexgen Plants Pty. Ltd., School of Agriculture and Food Sciences, The University of Queensland, Brisbane, Queensland, Australia

Running Head: *Cucumber mosaic virus* strain K genome

#Address correspondence to Peer M. Schenk, p.schenk@uq.edu.au.

Abstract

The complete genome sequence of *Cucumber mosaic virus* strain K was determined by deep RNA sequencing. The tripartite genome consists of a 3382 nt RNA1, a 3050 nt RNA2, and a 2218 nt RNA3 segment. Phylogenetic analysis placed RNA1 and RNA2 in subgroup IB. However, RNA3 grouped with subgroup IA isolates, indicating a likely recombination event.

Genome Announcement

Cucumber mosaic virus (CMV) is the type member of the plant virus genus *Cucumovirus* within the *Bromoviridae* family. CMV is distributed world-wide and is primarily vectored by aphids in a non-persistent manner (1). CMV has a wide plant host range, infecting more than 1,200 plant species in over 100 families including vegetables, fruit crops, ornamentals, and weeds (2, 3).

The CMV genome consists of three RNA segments, each individually packaged inside coat protein subunits to form icosahedral particles (4). CMV strains are divided into two major subgroups, I or II, with subgroup I strains further divided into the A or B subgroup (3-5). The CMV-K strain (of subgroup 1B) originates from China (6). RNA2 and RNA3 segment sequences were previously published in 1994 and 1999, respectively (5, 7).

Here we report the complete genome sequence of all three RNA segments of a 2016 version of the CMV-K strain, compiled from Illumina platform deep RNA sequencing (RNA-seq) reads.

Symptomatic tomato (Moneymaker variety) young leaf tissue was harvested 30 days post mechanical inoculation. Typical CMV symptoms of leaf mottling, shoestrings, and filiformity were evident on new leaf growth from ~12 days post-inoculation. Symptomatic leaf tissue was harvested, pulverised, and total RNA extracted using previously described methods (8). Deep RNA-seq was performed using Novogene as the service provider.

After adapter trimming, filtering and subtraction of chloroplast derived sequences, 24,246,450 clean sequencing reads were obtained. The CMV-K 2016 genome was assembled using previously described methods (9-11). To assist assembly, GenBank accessions AB179764, S72187, and AF127977 were used as reference genomes for each RNA segment. In total, 5,646,263 reads were assembled, providing full coverage at an average depth of more than 50,000 sequences per nucleotide across all three RNA segments. Similar to prior plant virus genome sequencing reports that used deep RNA-seq (9, 12), CMV-K 2016 sequencing revealed the virus was present as a quasi-species. The CMV-K 2016 RNA1 segment was 3382 nt in length and one clear sequence variation (A or G) was identified at nucleotide position 3158. The RNA2 segment was 3050 nt in length, while the RNA3 segment was 2218 nt in length and contained clear sequence variations at nucleotide positions 4 and 8 (both A or T).

Alignment of the CMV-K 2016 RNA2 segment with the CMV-K RNA2 sequence released in 1994 revealed that 35 single nucleotide polymorphisms (SNPs) and three indels were introduced during the 22 year interval that the isolate was maintained by serial passage. Similarly, alignment to the CMV-K RNA3 sequence released in 1999 (5) revealed 15 SNPs and one indel were introduced during that 17 year interval.

Phylogenetic analysis of all three RNA segments revealed RNA1 and RNA2 of CMV-K 2016 cluster with subgroup IB. However, RNA 3 appears to form part of subgroup IA. It is likely the CMV-K strain formed after a mixed infection with an isolate from the IA subgroup, resulting in a likely recombination event in RNA3.

Accession Numbers

The GenBank accession numbers for CMV-K 2016 RNA1-3 sequences have been deposited in GenBank with accession numbers MG182148-MG182150.

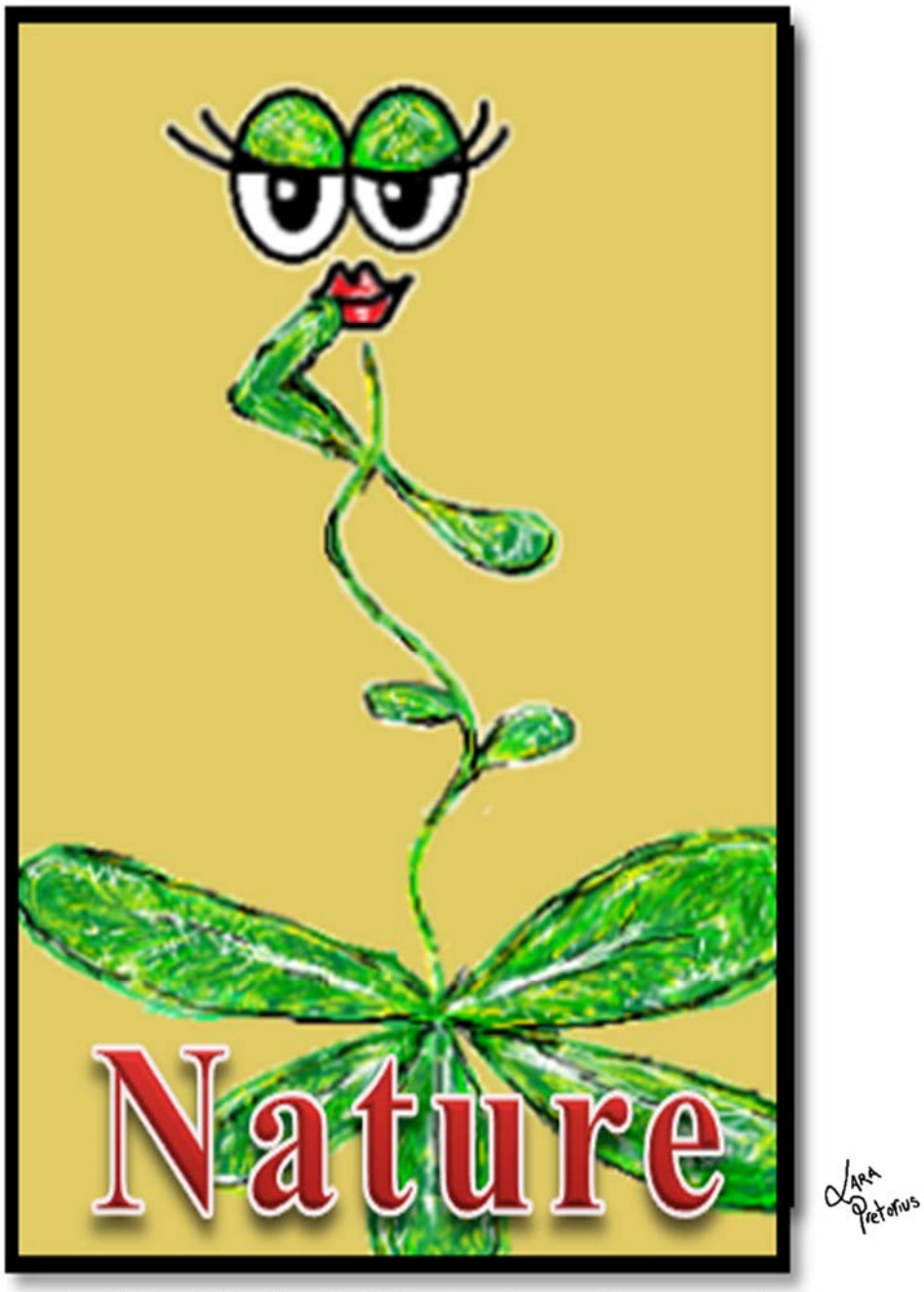
Acknowledgements

We thank John Randles for generously providing the CMV-K isolate. We thank Peter Palukaitis and John Randles for information regarding the provenance and history of the CMV-K isolate.

This study was supported by Nexgen Plants Pty. Ltd. and the Australian Research Council.

References

1. Palukaitis P, Roossinck MJ, Dietzgen RG, Francki RI. 1992. Cucumber mosaic virus. Adv Virus Res 41:281-348.
2. Scholthof KB, Adkins S, Czosnek H, Palukaitis P, Jacquot E, Hohn T, Hohn B, Saunders K, Candresse T, Ahlquist P, Hemenway C, Foster GD. 2011. Top 10 plant viruses in molecular plant pathology. Mol Plant Pathol 12:938-54.
3. Jacquemond M. 2012. Cucumber mosaic virus. Adv Virus Res 84:439-504.
4. Palukaitis P, Garcia-Arenal F. 2003. Cucumoviruses. Adv Virus Res 62:241-323.
5. Roossinck MJ, Zhang L, Hellwald KH. 1999. Rearrangements in the 5' nontranslated region and phylogenetic analyses of cucumber mosaic virus RNA 3 indicate radial evolution of three subgroups. J Virol 73:6752-8.
6. Tienpo, Rao ALN, Hatta T. 1982. Cucumber Mosaic-Virus from Cornflower in China. Plant Disease 66:337-339.
7. Hellwald KH, Palukaitis P. 1994. Nucleotide sequence and infectivity of cucumber mosaic cucumovirus (strain K) RNA2 involved in breakage of replicase-mediated resistance in tobacco. J Gen Virol 75 (Pt 8):2121-5.
8. Sternes PR, Moyle RL. 2015. Deep Sequencing Reveals Divergent Expression Patterns Within the Small RNA Transcriptomes of Cultured and Vegetative Tissues of Sugarcane. Plant Molecular Biology Reporter 33:931-951.
9. Pretorius L, Moyle RL, Dalton-Morgan J, Hussein N, Schenk PM. 2016. Complete Nucleotide Sequence of an Australian Isolate of Turnip mosaic virus before and after Seven Years of Serial Passaging. Genome Announc 4:e01269-16.
10. Moyle R, Pretorius LS, Dalton-Morgan J, Persley D, Schenk P. 2016. Analysis of the first complete genome sequence of an Australian tomato spotted wilt virus isolate. Australasian Plant Pathology 45:509-512.
11. Pretorius L, Moyle RL, Dalton-Morgan J, Schwinghamer MW, Crew K, Schenk PM, Geering ADW. 2017. First fully sequenced genome of an Australian isolate of Cauliflower mosaic virus. Australasian Plant Pathology 46:597-599.
12. Moyle RL, L. P, Carvalhais LC, Schenk PM. 2017. Complete Nucleotide Sequence of Australian Tomato spotted wilt virus Isolate TSWV-QLD2. Genome Announcements 5:e01267-17.



***Arabidopsis* the "model" plant posing on the cover of Nature**

Chapter 3: Early Defence Responses of *Arabidopsis* against TuMV

3.1 Introduction

Plants have a number of preformed and inducible defences, to prevent pathogen infection (Spoel and Dong 2012). Viral pathogens are not able to breach the plant's preformed defences therefore they enter the plant cells via wounding to the cuticle or cell wall (Kombrink and Somssich 1995). The wounding can be due to mechanical damage via gardening implements, herbivores, humans or insects.

The plant's inability to detect a viral pathogen, once it has evaded the plant's preformed defences, is one of the reasons viruses are so successful (Zipfel 2014). Plants can usually identify pathogens by their pathogen-associated molecular patterns (PAMPs) which will be followed a cascade of inducible defences in order to stop the pathogen from further invasion (Nürnberger and Brunner 2002). However, as there have been no identified PAMPs associated with viruses thus far, the plant's main lines of defence are inducible defence-related proteins and RNA silencing (Zvereva and Pooggin 2012). Even though viruses cannot be recognised by PAMPs, viruses still induce pattern-triggered immunity (PTI) (Brault *et al.* 2010) where resistance (R) genes are able to recognise specific non-viral effectors and viral avirulence (avr) factor proteins which will results in effector-triggered immunity (Petersen *et al.* 2006). This is similar to PTI but more intense as this form of immunity will trigger a cascade of signals leading to the hypersensitive response (HR) and programmed cell death (PCD) (Nürnberger and Brunner 2002) in an attempt to locally counteract the viral attack by restricting the viral spread, throughout the plant. This response is also based on viruses being classified as biotrophs and as such should up-regulated the SA pathway in order for the plant to successfully defend itself resulting in HR and PCD, further outlined in Figure 12.

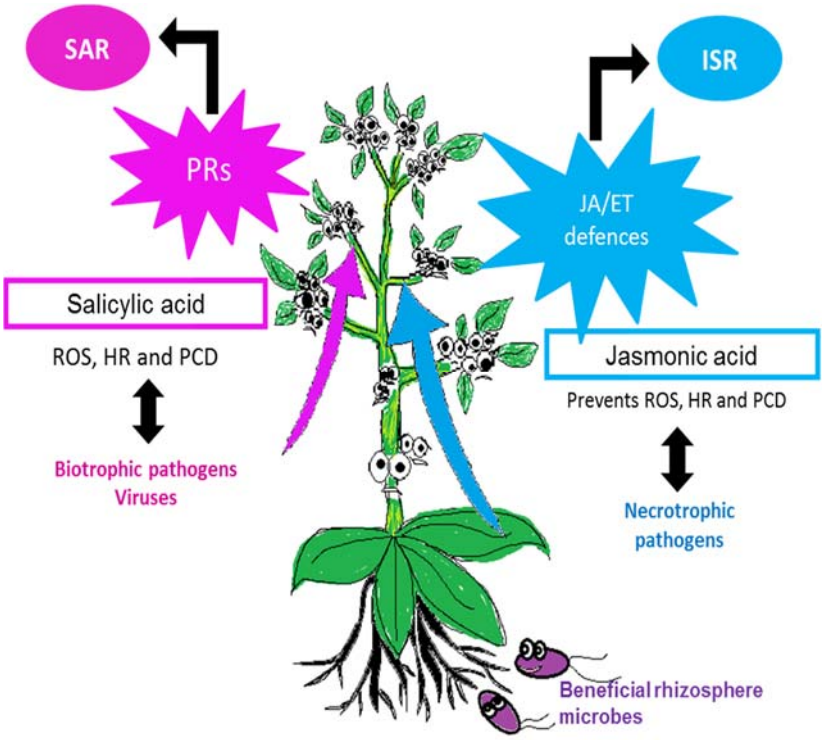


Figure 12: This image was adapted from Pieterse *et al.* (2009) showing the different pathways which induce systemic acquired resistance (SAR) and induced systemic resistance (ISR). For a successful defence against pathogens, the SA pathway is thought to induce SAR through a mobile signal which is activated upon infection by biotrophic pathogens and acts as a warning for distal tissue resulting in reactive oxygen species (ROS), HR and PCD. ISR is thought to be induced by colonisation of soil-borne beneficial microbes which prime the plant's JA and ET defences. The signal is also thought to act over a long distance similar to SAR; however the JA/ET pathways typically prevent the onset of ROS, HR and PCD.

Viruses are also able to interfere with the plant's miRNA biogenesis pathways which can manifest as phenotypic symptoms of the disease (Chapman *et al.* 2004). Chapman *et al.* (2004) proposed that this disruption of the miRNA pathway adds to the pathogenicity of certain viruses. Viruses are also known to target the plants RNA silencing mechanisms which is not surprising as RNA silencing is the plant's main antiviral defence and the high number of viral suppressors of RNA silencing (VSRs) supports this fact (Burgyán and Havelda 2011). In general, a systemic viral infection will lead to the modification of cells and the manipulation of certain processes in order to accommodate the virus and help spread the infection. The molecules that instigate this process can also be referred to as viral effectors.

These will specifically include the up-regulation of pathogenesis-related (PR) proteins, however many non-defence-related genes will also respond due to the changes in cellular processes (Cooper 2001). In some cases it is believed that certain genes are either induced or repressed by the virus to facilitate its infection and pathogenicity (Whitham and Wang 2004).

The virus which this study focuses on is TuMV (*Potyvirus* genus; *Potyviridae* family), which is considered the second most important crop virus of 28 different countries and regions (Walsh and Jenner 2002). It has wide distribution and large host range, with crop plants from the *Brassicaceae* family being its main host.

As typical of potyviruses, TuMV is transmitted in a non-persistent manner by aphids. Its virions are flexuous filaments and are approximately 700-750 nm in length, its genome is roughly 10kb in size and is organised in a similar manner to most monopartite potyviruses. It has a sRNAs positive–sensed genome, with a single open reading frame flanked by two untranslated regions, which is translated into a polypeptide. It's most notable feature is that it is able to infect *A. thaliana* giving us the perfect model system with which to study plant-virus interactions.

A number of studies have used TuMV with the model plant *A. thaliana* to determine the changes in gene expression. A study by Whitham *et al.* (2003) compared the expression changes caused by different RNA viruses while Yang *et al.* (2007) was interested in the spatial distribution of TuMV as well as altered gene expression. From this latter study it was found that the level of gene response was strongly correlated with the amount of virus accumulation in the cell and not the function of the genes specifically. They also report that the genes which were most strongly influenced were consistent with viral symptoms; decreases were found in chloroplast function, sulphate assimilation and cell wall extensibility while protein synthesis and turnover were increased. Figure 13 from Yang *et al.* (2007) shows their interpretation of how TuMV virus accumulation altered gene expression based on mRNA profiles of functional gene groups.

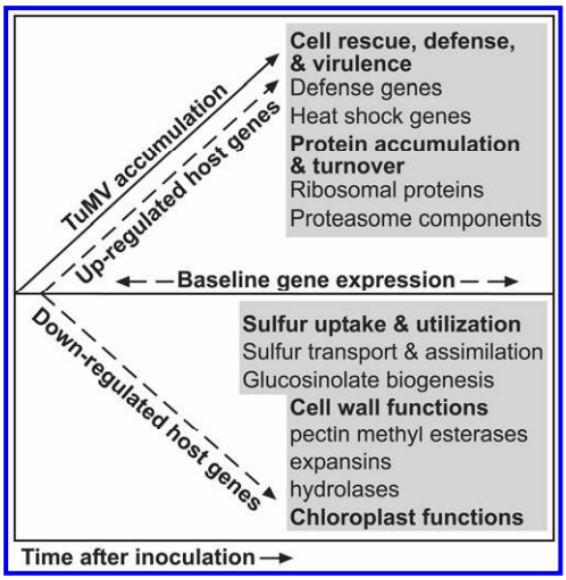


Figure 13: This model from Yang *et al.* (2007) shows the relationship between TuMV accumulation and its effect on the regulation of certain function gene groups. The solid line represents TuMV accumulation while the dashed lines the induction or suppression of certain genes.

The study by Yang *et al.* (2007) had two main focus areas, the first being the spatial analysis which was localised to the leaf to determine the spread of the virus and its effect on genes expression of infected and uninfected neighbouring cells over a 5 day time course. The other focused on the gene expression of three distinct areas; inoculated leaves, rosette leaves and cauline leaves over a 10 day time course. The mRNA profiles were obtained via microarray analysis. They reported that 556 genes had significantly altered expression due to TuMV accumulation. Over the 10 day time course 67 genes stood out from the 338 preselected genes which made up their microarray.

These results were used as a basis for expression changes altered by TuMV infection in my study and 30 of the genes were selected to be used at earlier time points. My study used the TuMV-QLD1b isolate (KX641466) to analyse the expression profile of six marker genes. As the focus of my study was to determine the early defence response of the plant after viral infection, analysis was conducted at 6, 24 and 48 hours after infection. Here, I show that the plant's JA pathway is up-regulated 24 hours after infection.

3.2 Results

A total of 36 genes were used to determine the plant's early responses against TuMV, including six marker genes. The six marker genes used in this assay include *PLANT DEFENSIN 1.2* (*PDF1.2*), *VEGETATIVE STORAGE PROTEIN* (*VSP2*), *PATHOGENESIS-RELATED 5* (*PR5*), *PATHOGENESIS-RELATED 1* (*PR1*), *ETHYLENE RESPONSE FACTOR 1* (*ERF1*) and *ETHYLENE RESPONSE FACTOR 6* (*ERF6*). *PDF1.2* and *VSP2* are both part of the JA pathway (Spoel, Koornneef *et al.* 2003), while *PR5* and *PR1* were found to be responsive to SA (Stein, Molitor *et al.* 2008). As is mentioned in the name, *ERF1* and *6* both respond to ET and were reported to be upregulated as a defence response to *Botrytis cinerea* (Berrocal-Lobo *et al.* 2002). *ERF1* also binds to the *PDF1.2* promoter (Brown *et al.* 2003) while ERF6 is involved in ROS signalling (Sewelam *et al*. 2013).

Interestingly, my results suggest that the virus upregulates the JA pathway 24 hours after inoculation (Figure 14).

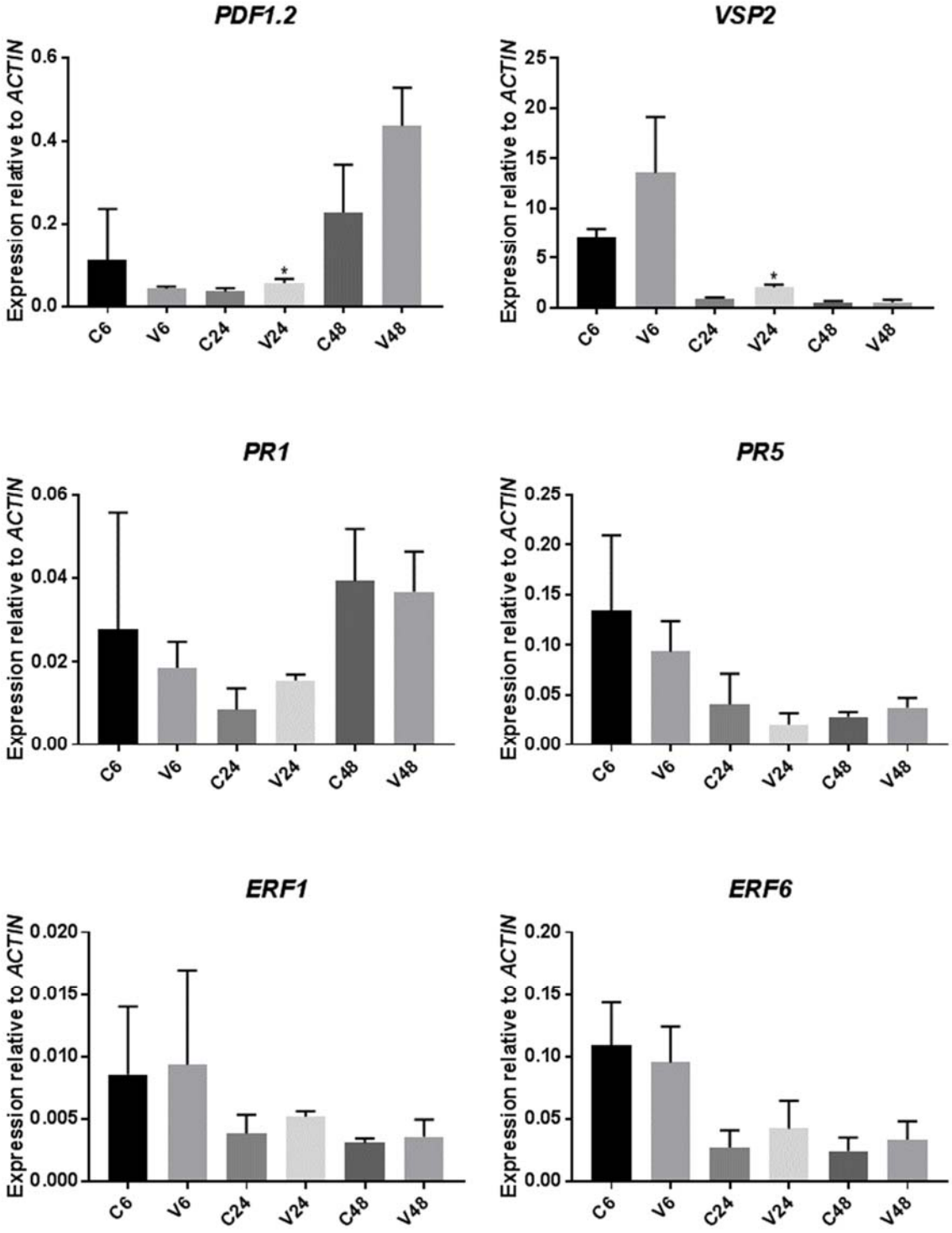


Figure 14: Six marker genes were used to determine which pathway was induced or suppressed by TuMV; two marker genes from the JA, SA and ET pathway were assayed. *A. thaliana* Col-0 plant were mock or virus inoculated and collected 6, 24 and 48 hours afterwards; C6 is the control 6 hours after inoculation and V6 is the virus-inoculated sample, C24, V24, C48 and V48 follow the same naming scheme. PDF1.2 and VSP2, both JA-responsive genes, were induced 24 hours after virus inoculation, no other significant influences were noted for the other four marker genes, though large error bars are possibly the cause. Shown are mean values ± standard deviation of qRT-PCR assays from three biological replicates (containing 20 pooled plants per replicate). Asterisks indicate significant differences to control plants within the same time point ($P<0.05$).

The other marker genes used in this assay showed no significant induction or repression, which in some cases could be due to the large error bars, for example there could be a down-regulation of *PR5* 6 hours after viral inoculation which would correlate with the up-regulation of *PDF1.2* also at the 6 hour time point. However, both error bars were too large to suggest any significant alteration.

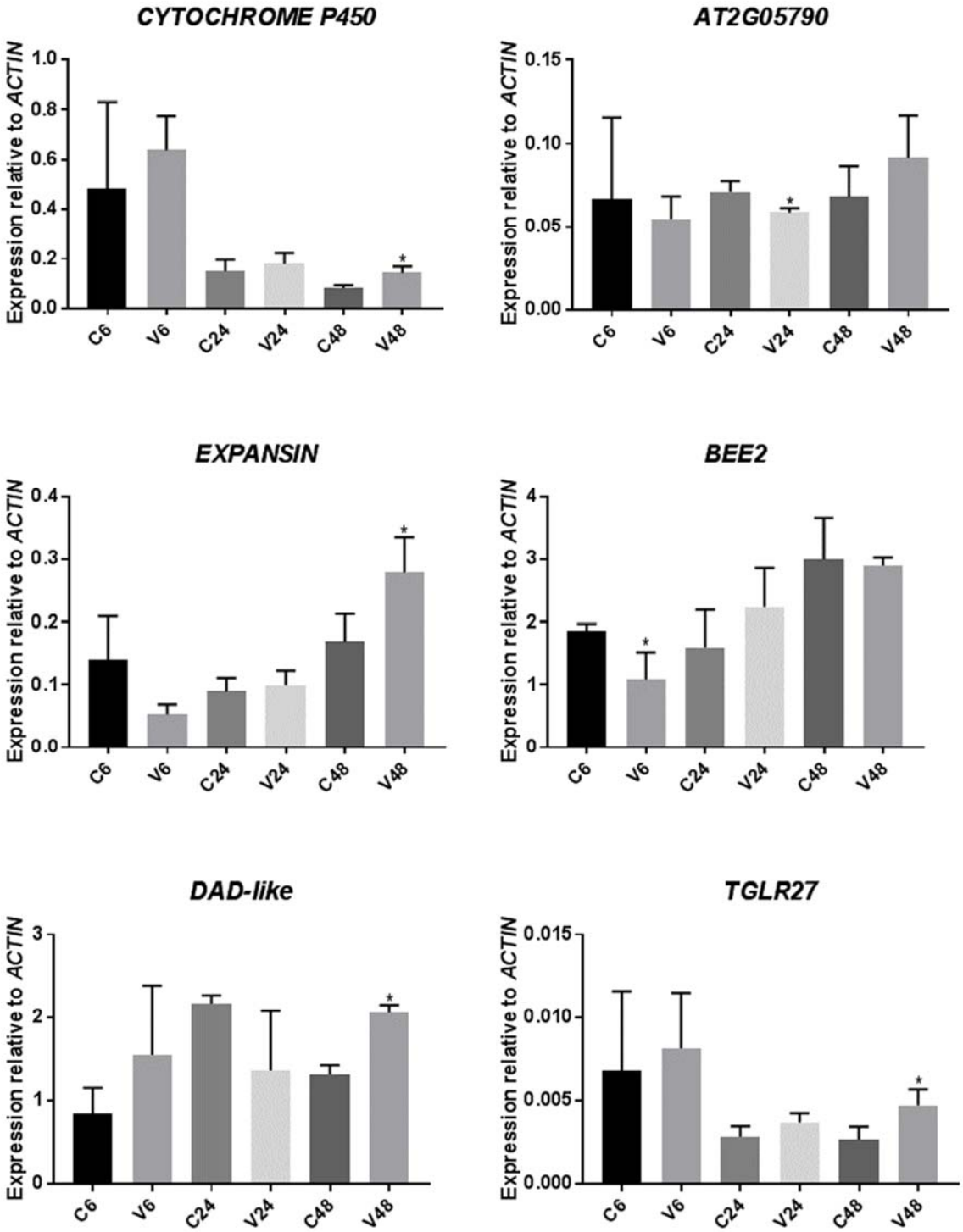
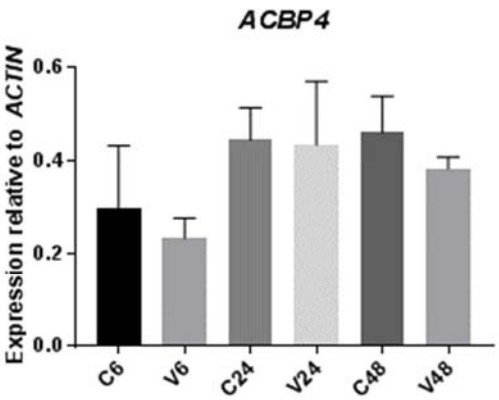
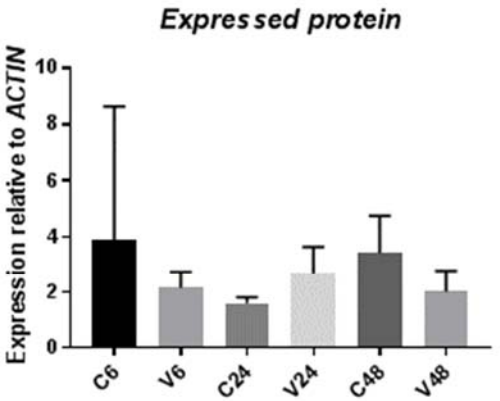
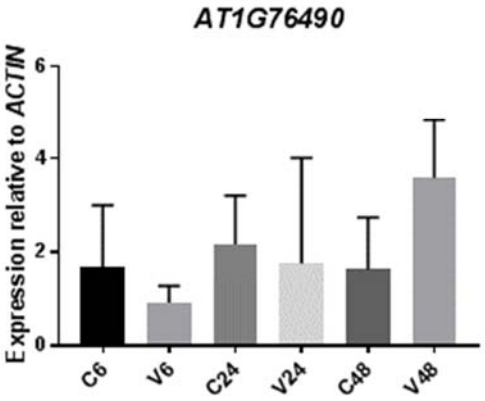


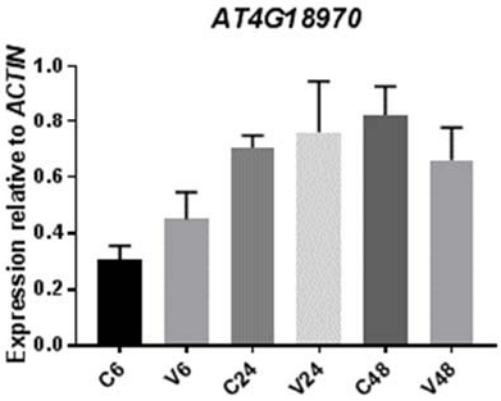
Figure 15: The expression of six genes were significantly altered due to TuMV infection. *A. thaliana* Col-0 plants were either mock or virus inoculated and collected 6, 24 or 48 hours after inoculation. C6 denotes that the plant was mock-inoculated control and collected 6 hours after inoculation, while V6 was virus-inoculated and also collected 6 hours later. The letter represents the treatment while the numbers are the hours after inoculation in which the samples were collected. All gene transcript levels were relative to ACTIN. Similar to the marker gene results, other genes and time points may have been significant if not for the large error bars. Shown are mean values ± standard deviation of qRT-PCR assays from three biological replicates (containing 20 pooled plants per replicate). Asterisks indicate significant differences to control plants within the same time point ($P<0.05$).

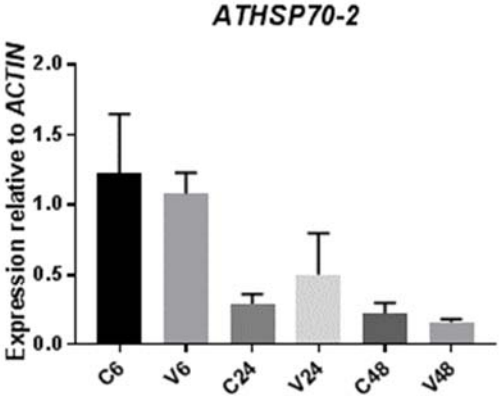
Other genes found to be altered significantly were "*CYTOCHROME P450 FAMILY 79 SUBFAMILY B, POLYPEPTIDE 2*" (*CYP79B2*), *AT2G05790*, *EXPANSIN, BR ENHANCED EXPRESSION 2* (*BEE2*), *DAD1-LIKE LIPASE 3* (*DLL*) and *GLUTAMATE RECEPTOR 2.7* (*TGL27*) (Figure 15). *CYP79B2*, *EXPANSIN*, *DAD_LIKE* and *TGLR27* were up-regulated 48 hours after virus inoculation in comparison to the mock plants. In contrast, *AT2G05790* was down-regulated at 24 hours after inoculation and *BEE2* was also down-regulated but 6 hours after inoculation (Figure 15).

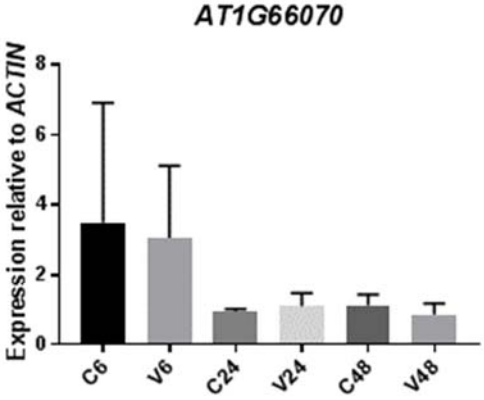

ACBP4
Expression relative to ACTIN
0.6
0.4
0.2
0.0
C6
V6
C24
V24
C48
V48

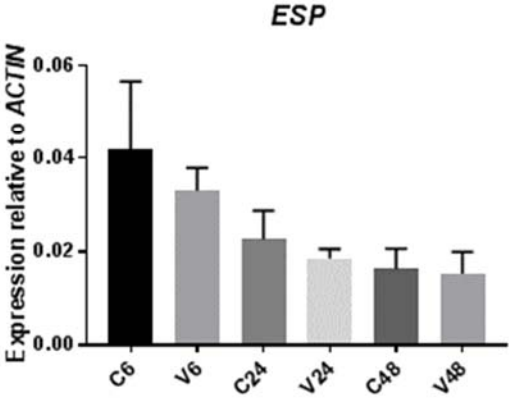

Expressed protein
Expression relative to ACTIN
10
8
6
4
2
0
C6
V6
C24
V24
C48
V48

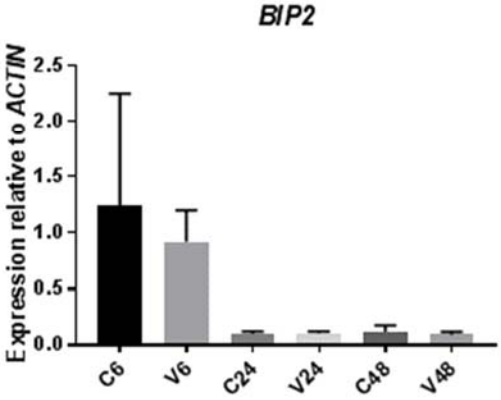

AT1G76490
Expression relative to ACTIN
6
4
2
0
C6
V6
C24
V24
C48
V48

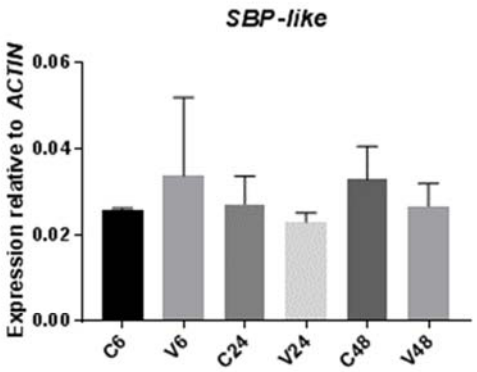

AT4G18970
Expression relative to ACTIN
1.0
0.8
0.6
0.4
0.2
0.0
C6
V6
C24
V24
C48
V48

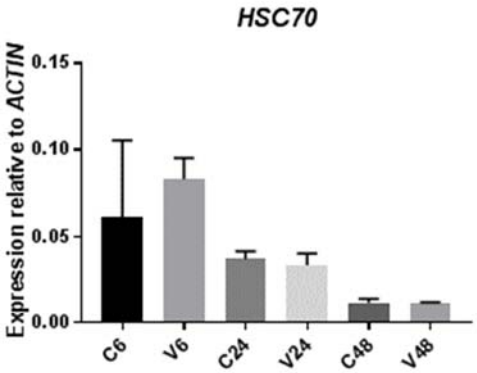

ATHSP70-2
Expression relative to ACTIN
2.0
1.5
1.0
0.5
0.0
C6
V6
C24
V24
C48
V48

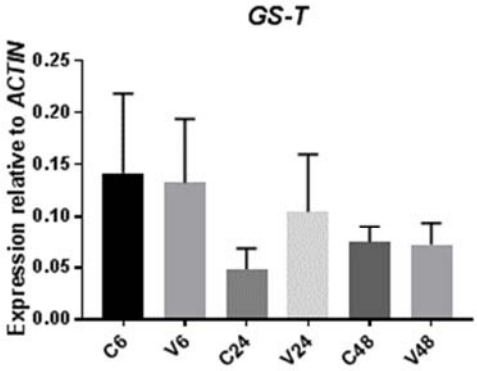

AT1G66070
Expression relative to ACTIN
8
6
4
2
0
C6
V6
C24
V24
C48
V48

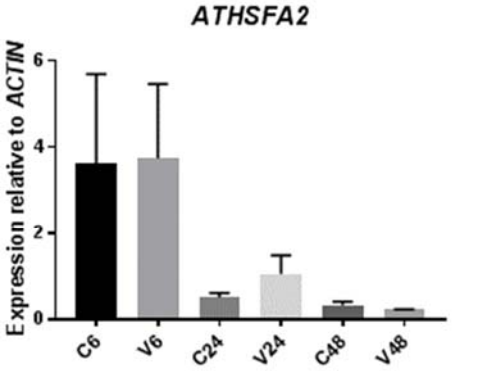

ESP
Expression relative to ACTIN
0.06
0.04
0.02
0.00
C6
V6
C24
V24
C48
V48

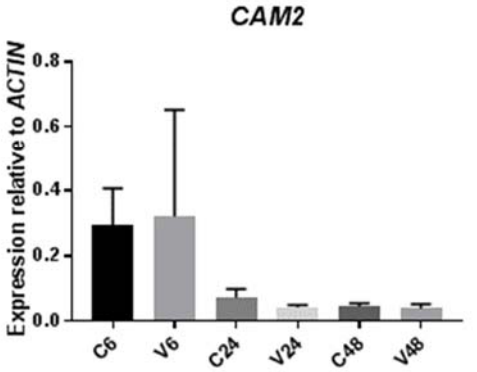

BIP2
Expression relative to ACTIN
2.5
2.0
1.5
1.0
0.5
0.0
C6
V6
C24
V24
C48
V48

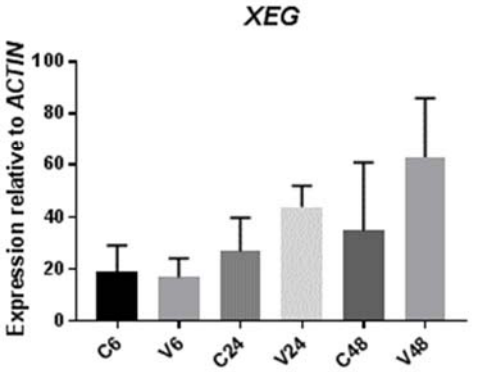

SBP-like
Expression relative to ACTIN
0.06
0.04
0.02
0.00
C6
V6
C24
V24
C48
V48

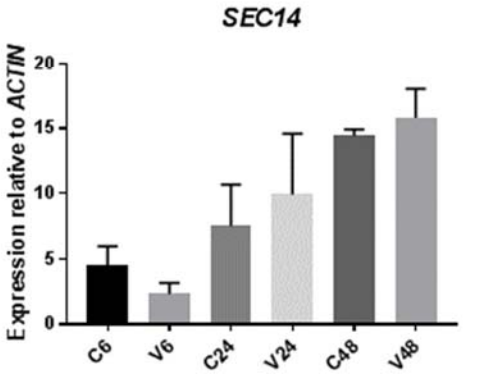

HSC70
Expression relative to ACTIN
0.15
0.10
0.05
0.00
C6
V6
C24
V24
C48
V48

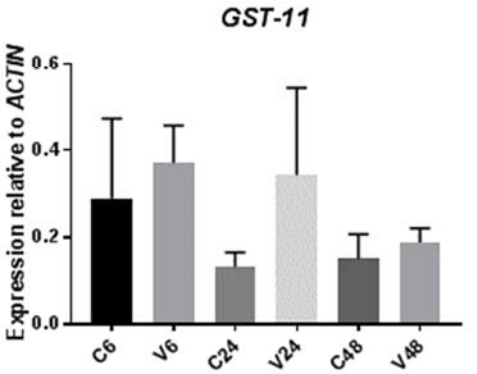

GS-T
Expression relative to ACTIN
0.25
0.20
0.15
0.10
0.05
0.00
C6
V6
C24
V24
C48
V48


ATHSFA2
Expression relative to ACTIN
6
4
2
0
C6
V6
C24
V24
C48
V48


CAM2
Expression relative to ACTIN
0.8
0.6
0.4
0.2
0.0
C6
V6
C24
V24
C48
V48


XEG
Expression relative to ACTIN
100
80
60
40
20
0
C6
V6
C24
V24
C48
V48


SEC14
Expression relative to ACTIN
20
15
10
5
0
C6
V6
C24
V24
C48
V48


GST-11
Expression relative to ACTIN
0.6
0.4
0.2
0.0
C6
V6
C24
V24
C48
V48

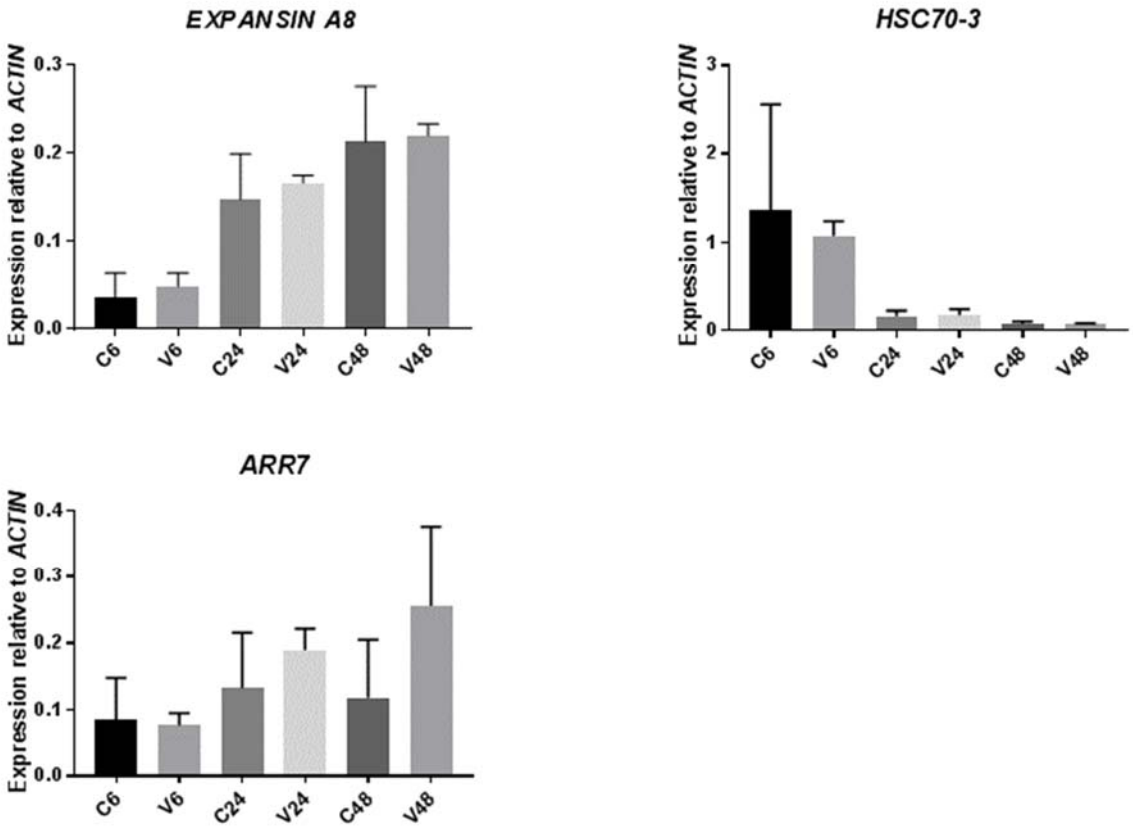


Figure 16: *A. thaliana g*enes which were not significantly affected by virus inoculation at any of the three time points. Shown are mean values ± standard deviation of qRT-PCR assays from three biological replicates (containing 20 pooled plants per replicate).

3.3 Discussion

The JA pathway is typically upregulated during the successful defence against necrotrophic fungi, therefore preventing the induction of ROS, HR and PCD, as these would only facilitate the pathogen's infection that requires dead tissue to survive (Glazebrook 2005). From the results in Figure 14, it can be noted that there is an up-regulation of JA signalling 24 hai, as *VSP2* and *PDF1.2* are both up-regulated. The up-regulation of the incorrect defensive pathway suggests the virus has hi-jacked this pathway, preventing the up-regulation of the SA pathway and possibly facilitating its establishment. The manipulation of the plant's defence pathways is a common strategy (Chung *et al.* 2008) and is looked at in further detail in Chapter 4. *VSP2* is a wound inducible gene (Berger *et al.* 2002) which possibly accounts for the increased expression 6 hai, though the fact that the gene is significantly upregulated in virus-infected plants further suggests the hijacking of the JA pathway.

CYP79B2 is part of the Cytochrome P450 family which is a superfamily of enzymes consisting of heme-containing proteins known to catalyse NADPH- and O_2-dependent hydroxylation reactions (Chapple 1998). *CYP79B2* is involved in the synthesis of camalexin which is a phytoalexin involved in pathogen response (Lemarié, *et al.* 2015). Camalexin accumulation has been upregulated by the

presence of necrotrophic fungi (Kliebenstein *et al.* 2005, Nafisi *et al.* 2007), certain hemi-biotrophs (Bohman *et al.* 2004), biotrophs (Glazebrook *et al.* 1997) and viruses (Callaway *et al.* 1996).

A study by Dempsey *et al.* (1997) found that camalexin accumulation was linked with the HR due to infection by Turnip crinkle virus (TCV). According to Glazebrook (2001), the accumulation of camalexin is able to induce *PDF1.2* even though camalexin accumulation forms part of the SA-dependent resistance pathway. As *PDF1.2* was upregulated before *CYP79B2*, the accumulation of camalexin was a defence response due the presence of the virus. This gene was also part of the list of common genes induced by RNA virus infection (Whitham *et al.* 2003).

Another gene related to wounding is *DLL3*; it is considered to be a JA-biosynthetic lipase specifically induced after wounding (Ruduś *et al.* 2014). However, the inductions of JA and DLL3 are usually within the first few minutes and therefore the upregulation of *DLL3* after 48 hours would not be due to wounding. Ellinger *et al.* (2010) found that the upregulation of this gene, including other members of the same lipase family, could also be due to pathogen-induced jasmonate formation, which is consistent with our results. Interestingly when *A. thaliana* Col-0 plants were infected with CMV, DLL genes were suppressed 2 dai, however this could be due to the upregulation of the SA pathway (Huang *et al.* 2005) though this is not apparent in our results.

TGLR2 was another gene also identified by Whitham *et al.* (2003) as part of the common set of genes induced by a range of RNA viruses. It is involved in the homeostasis of calcium ions and response to light stimulus. Many studies have reported that increases in cytosolic calcium is a crucial step in the cascade which induces pathogen-related responses (Blume *et al.* 2000, Nürnberger and Scheel 2001,. Xu and Heath 1998) found that the elevation of calcium ions is directly link to the HR when cowpea was infected with a rust fungus. Therefore, the up-regulation of *TGLR2* may be due to the presence of the virus and the plant's attempts at a different defence mechanism. Further analysis is needed to confirm this.

The fourth gene up-regulated at 48 hours after virus inoculation was *EXPANSIN 3* which is the only differentially-expressed gene unrelated to defence. This gene has a role in cell growth and is involved in cell wall loosening which is an adaptive and morphological process (Lee *et al.* 2001). As this is not a defence-related gene, its upregulation is unclear especially since this same gene was down-regulated by TuMV-UK1 (Yang *et al.* 2007). Yang *et al.* (2007) suggested that the downregulation of genes such as *EXPANSIN* which had a role in cell wall function were the cause of the severe stunting seen in virus-infected plants. This makes sense, however, this downregulation was only apparent in later stages of infection. It is possible that the virus, while creating its replication complexes from the host lipids and membranes (Schaad *et al.* 1997, Ahlquist *et al.* 2003), requires

increased growth and only once the virus has established an infection will these genes be downregulated.

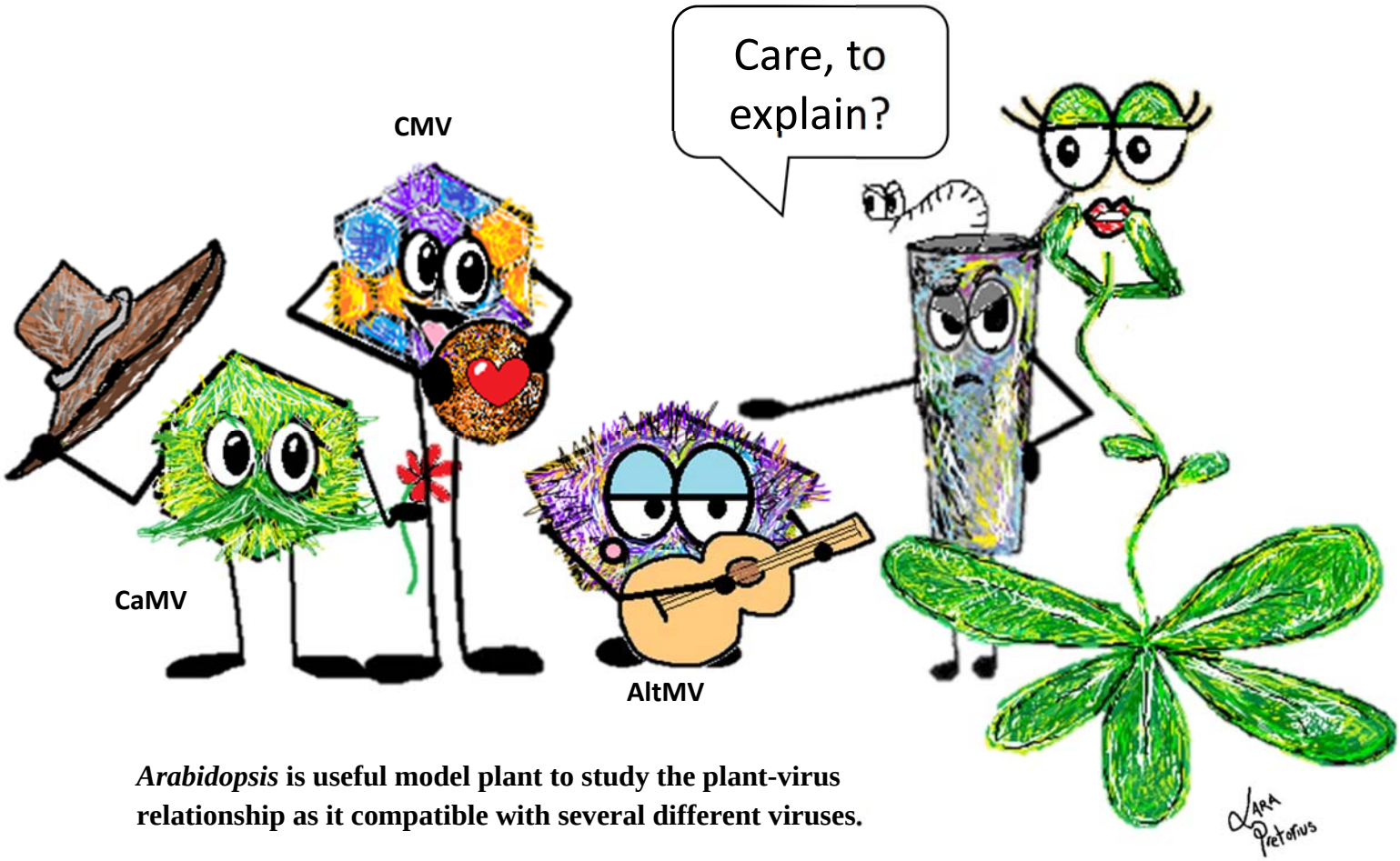
BEE2 and *AT2G05790* were both downregulated after virus inoculation which matched the results of Yang *et al.* (2007). *BEE2* was described as putative protein which was downregulated 3 days after virus inoculation, while our results show a significant downregulation already at 6 hours after inoculation. This gene is thought to be part of the brassinosteroid signalling pathway which has a main role in cell elongation. The downregulation of this gene could be related to the stunting of the plant (Zhu *et al.* 2013, Wang *et al.* 2014). *AT2G05790* was downregulated after 24 hours which was similar to the results Yang *et al.* (2007) reported 3 days after inoculation. The TAIR website described AT2G05790 as having a role in carbohydrate metabolism therefore the downregulation of the gene would correlate with the lowered photosynthetic rate caused by the a viral infection which is a common symptom (Rahoutei *et al.* 2000, Pérez-Bueno *et al.* 2006). There is also a possibility that the suppression of this gene is just adding to the imbalance created by the virus in certain primary metabolites (Balachandran *et al.* 1997). Other studies have suggested links between the accumulation of primary metabolites, such as carbohydrates and sugars, as a way to fuel the energy-intensive defence response of the plant against pathogens (Swarbrick *et al.* 2006). Therefore, the increase in primary metabolism could an early defence response and possibly the beginning of a signal cascade. The study by Herbers *et al.* (2000) found that the SA pathway was mediated by sugars and that carbohydrates acted as signal molecules. The suppression of this gene could be the virus manipulating the plant's metabolism preventing it from "gathering" the energy to fight back, therefore weakening the plant.

The results from this small study and the few genes that were significantly altered after TuMV infection seem to match those reported by Yang *et al.* (2007). It also appears to correlate with their model of which classes of genes are induced or supressed by TuMV (Figure 13). Our results suggest there is an increase of the JA pathway with the up-regulation of both *PDF1.*2 and *VSP2* (Figure 14) which could possibly lead to the upregulation of *CYP79B2* and *DLL* 48 hours after inoculation (Figure 15). The study by Yang *et al.* (2007) did suggest that most of significant changes in expression occurred later in the infection when the virus RNA and protein level had reached a high enough threshold to cause serious effects. Therefore, it is possible that the time points chosen may have been too early to accurately determine the virus-induced modulation of the plant's defences as the viral load would still be low.

Table 5: List of Real-Time Primers used in Chapter 3

Primer Name	Sequence
AT4G36540 BEE2 F	CCA GAT TAC ATT CAT GTT AGG GCT AGA
AT4G36540 BEE2 R	TGC CCT CTC TGC TAA GCT ATG TC
AT4G35750 SEC14 R	TCC ACA CCG GCG TAC AGA
AT4G35750 SEC14 F	TCG CTC GTA GAG CTG AGA TAC CT
AT4G30270 XEG R	TCT CGT GGC GGC GTT T
AT4G30270 XEG F	GCT ACG TTG ACG TCA GTG TTG AA
AT2G41110 CAM2 F	CGC CTC AAG TCC GCA TTC
AT2G41110 CAM2 R	GGG AAA AAC GAA AGA GTT TGT GA
AT2G37640 Expansin F	CCG GGA AAT CCG TCT ATT CTT
AT2G37640 Expansin R	GAG CGA AAT TCG GTG GAC AA
AT1G02920 GS-T11 F	CAA GGA CAT TGC GGG CAT A
AT1G02920 GS-T11 R	GGG TCA AAC TCA TGC GAT TCA
AT3G09440 HSC70-3 F	TTG ACA TTG ATG CAA ACG GTA TC
AT3G09440 HSC70-3 R	CTT CTG CCC CGT TGT CTT GT
AT1G02920 GS-T F	CAA GGA CAT TGC GGG CAT A
AT1G02920 GS-T R	GGG TCA AAC TCA TGC GAT TCA
AT2G30550 DAD1-LIKE LIPASE 3 F	ACG GCG AAA TGG CTC AAG
AT2G30550 DAD1-LIKE LIPASE 3 R	TTT GGA AGC GGG ATC GAA
AT3G12580 HSC70 F	ACT GGG AAC CCG AGA GCT TT
AT3G12580 HSC70 R	CTC TTC GCC CGC TCA CA
AT2G26150 ATHSFA2 F	GGT TCT GTA GCG GCT TCT TCA
AT2G26150 ATHSFA2 R	AGC CCT TCC ATT GGT CTA GGA
AT5G02490 ATHSP70-2 F	GAA ACA CAA CCA TTC CAA CCA A
AT5G02490 ATHSP70-2 R	CAG GCT GGT TGT CCG AAT AAG
AT5G42020 BIP2 F	GAT CGA CGC CAG GAA TGC
AT5G42020 BIP2 R	TGT CGC TCA CTT GGT TCT TCA
AT4G39950 CYTOCHROME P450 F	CGT CGC CGG ATA TCA CAT C
AT4G39950 CYTOCHROME P450 R	CCA GCC CAT ATC GGC TAA GA
AT2G29120 TGLR2.7 F	GAA GAT GTT GTC CAG GCC TCA T
AT2G29120 TGLR2.7 R	CGC TCA CTT GCT CGT TCT TG
AT1G19050 ARR7 F	CTG CTC TCT TTT TTA TTC TGA GTT TGA C
AT1G19050 ARR7 R	CCA CCG GCG GGA ATC T
AT1G09750 Expressed protein F	TCT CAA CAT TGG GAG CGT TTG
AT1G09750 Expressed protein R	TGG CGC CAC ATT CTC GTT
AT2G05790 F	CCG CCG CCA AAA GAA CT
AT2G05790 R	GCT GCT ACG TTG CGT TTG AC
AT1G76490 F	CAC CGG TGG CGT GAC AA
AT1G76490 R	TTC TGT GAT AGT GAC GAC GTG AAG
AT2G40610 EXPANSIN A8 F	CAC CAT GGG CGG AGC TT
AT2G40610 EXPANSIN A8 R	GTC CCG TAA CCT TGG CCA TA
AT4G18970 F	TCG ACT TCC AGT ACG GTC CAA
AT4G18970 R	ACA TCG ACG GTG GTT TTT CC
AT1G54040 ESP F	CAA TCG CTC AAC CCA AAG GA

AT1G54040 ESP R	TGC GCA CGC CTA AGC A
AT5G43720 SBP-like F	CAC GGA GGC TAT GCC AAG AG
AT5G43720 SBP-like R	TGC TTC CCG CCG TCT GT
AT1G66070 F	CGG AGA TTT CTA ATC CAA CGA AA
AT1G66070 R	TTC GCA GAG GAA GAG AAA TCG
AT3G05420 ACBP4 F	GTG GAC TGC ACC CCA AAC A
AT3G05420 ACBP4 R	GCG CCA TGC TCG TAA CG



***Arabidopsis* is useful model plant to study the plant-virus relationship as it compatible with several different viruses.**

Chapter 4: The Role of the Mediator Complex in Virus Infection

4.1 Introduction

Plant viruses are responsible for crop losses throughout the world, with many considered economically important (Pretorius *et al.* 2017). Viral disease phenotypes range from symptomless to severe developmental defects resulting in decreased productivity (Hull 2013). The phenotypic symptoms caused by viruses are often due to the viral manipulation of gene expression either directly through viral proteins interfering with regulatory pathways (Kasschau *et al.* 2003) or indirectly through the production of small viral derived RNAs (Smith, Eamens *et al.* 2011). At the early stages of infection, the plant may recognise the virus or other pathogen based on certain molecular cues related to the pathogen's lifestyle (Koornneef and Pieterse 2008). Viruses are classified as biotrophic pathogens known to induce the SA pathway during successful defence responses, while necrotrophic pathogens induce the JA/ET pathways. The abscisic acid (ABA) pathway is not part of the classic defence pathways, however there is increasing evidence suggesting that ABA plays an important role in disease development; with its effects increasing or decreasing resistance depending on the pathogen (Ton, Flors *et al.* 2009). The role of ABA in disease resistance is possibly indirectly related to its regulation of abiotic signalling such as stomatal closing and fortified cell walls (Melotto, Underwood *et al.* 2006).

Interestingly, abiotic stress signalling pathways, such as the ABA pathway, have the potential to override biotic stress pathways such as those mediated by SA and JA (Audenaert *et al.* 2002, Anderson *et al.* 2004). This trade-off can be disadvantageous, especially when pathogens can exploit this weakness and use it to manipulate host defence responses. A well-documented example of this is where *Pseudomonas syringae* up-regulates the plant's JA pathway with molecules (coronatine) that mimic the plant's natural JA signalling molecules. These actions could supress the SA pathway making the plant more susceptible to this particular pathogen (Nomura *et al.* 2005). Similarly, the hemi-biotroph *F. oxysporum* also up-regulates the JA pathway by producing JA-related compounds (Miersch *et al.* 1999). Interestingly, the study by Thatcher *et al.* (2009) showed that JA insensitive mutant *coi1* plants had less phenotypic chlorosis when infected with *F. oxysporum*. Foliar tissue from *coi1* also showed less chlorosis when treated with *F. oxysporum*culture filtrate, while wild-type leaves had extensive chlorosis and localised cell death. This suggested that the fungus was producing a metabolite which the *coi1* plants were unresponsive to, possibly due to their insensitivity to JA.

Similar experiments have been performed with mediator mutants that also show high levels of resistance to *F. oxyporum* (Kidd *et al.* 2009). In the study by Kidd *et al.* (2009), mediator subunit MEDIATOR25 was researched. It was found that *med25* plants also showed increased resistance to

*F. oxysporum*and that MED25 had a role in the JA defence pathway. Other mediator mutants assayed with *F. oxysporum*showed varying levels of resistance to this pathogen as well as increased levels of susceptibility to necrotrophic pathogens. *med25* and *med8* both showed resistance to *F. oxysporum*but increased susceptibility to *Alternaria brassicicola* and *Botrytis cinerea* (Kidd *et al.* 2009, Lai *et al.* 2014). Other mediator mutants displaying similar properties were *med20* and *med18* (Fallath *et al*. 2017).

The mediator complex is evolutionary conserved among eukaryotes and has been extensively studied in mammals and was only recently discovered in plants (Bäckström *et al.* 2007). It is described as "a central co-regulator of transcription". It acts as an interface between RNA dependent RNA polymerase 2 (RDRP2) and transcription factors (TF) in the cell, creating a bridge and co-ordinating interactions between RDRP2 and distal transcription factors (Malik and Roeder 2005). It was also shown by Holstege *et al.* (1998) that 93% of all genes expressed in yeast were dependent on the mediator subunit.

The mediator complex in plants is made up of approximately 21 conserved subunits and is divided into three sections; head, middle and tail (Figure 17). These denote what the subunits interact with; subunits in the tail sections are thought to interact with the TF, while the middle and head sections bind to the C-terminal of the RDRP (Malik and Roeder 2005).

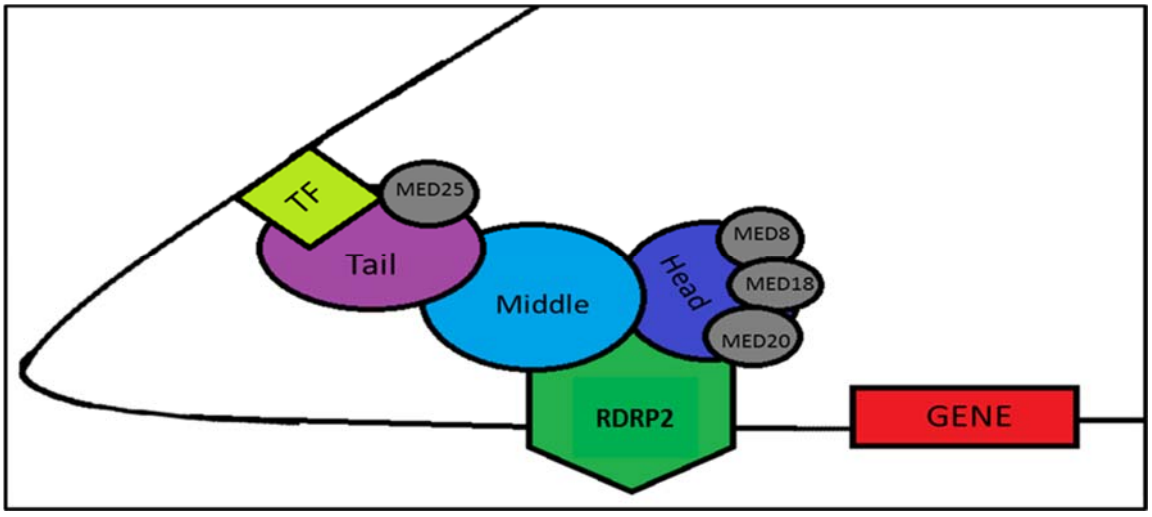


Figure 17: The arrangement of the mediator complex is divided into 3 parts: the head, middle and tail. The head and middle sections are thought to act with the C-terminal of the RDRP2 while the tail is thought to primarily interact with the transcription factors (TF). The mediator subunits mentioned above are located as follows: MED8, MED18 and MED20 form part of the head while MED25 forms part of the tail. Interestingly, all mutant lines show resistance to *F. oxysporum*.

MED18 forms part of the head molecule of the mediator complex and is considered to be a multifunctional protein with roles in plant immunity, development and response to hormones, specifically JA and ABA (Lai *et al.* 2014). It was also confirmed that MED18 has a role in microRNA biogenesis (Kim *et al.* 2011).

The various subunits of the mediator complex with relation to fungal and bacterial resistance have been well documented, although their effects on virus disease development are lacking. Due to this fact and my previous results, where the JA pathway was upregulated 24 hours after viral inoculation (Chapter 2), I hypothesised that *med18* which showed resistance to *F. oxysporum*, could also be less susceptible to TuMV. Results showed that *med18* plants had significantly less viral RNA when compared to WT plants, suggesting that these plants were less susceptible (Figure 19). However, phenotypically there were no differences between *med18* and WT plants as symptoms are mild in *Arabidopsis.* To determine whether this result was specific to TuMV or possibly broad-range resistance affecting viruses from various families, the same assay was performed using CMV, AltMV and CaMV. TuMV, CMV and AltMV are all RNA viruses, while CaMV is a DNA virus. All three viruses were chosen as they can infect *Arabidopsis*. The *med18* plants showed varying levels of resistance to all viruses used in the assay with CMV showing a significant decrease in viral load. This work will contribute to our understanding in plant-virus interactions and help determine which pathways are important for virus disease development.

4.2 Results

Previous research performed by T. Fallath in the Schenk Lab (UQ) reported that *med8*, *med18*, *med20* and *med25* plants all displayed varying levels of resistance to *F. oxyporum*. This coincides with the theory that the fungus is hijacking the JA pathway, using the host's defence responses to its own benefit as many of these *med* mutants have reduced JA signalling in comparison to wild-type plants.

Here we tested whether this theory would lead to similar results when using TuMV which was also found to up-regulate the JA pathway. qRT-PCR was performed to determine the gene expression of several marker genes as well as viral load compared to a reference gene. *PDF1.2, VSP2, PR1, PR2* and *PR5* were used as marker genes for specific pathways. *PDF1.2* responds to both JA and ET (Manners *et al.* 1998) while *VSP2* only responds to JA (Utsugi *et al.* 1998). *PR1* is downregulated by JA and is part of the SA pathway and has a role in systemic acquired resistance (Zhang *et al.* 2010). *PR2* and *PR5* both have roles in systemic acquired resistance (Van Loon and Van Strien 1999); *PR5* was found to respond to JA, SA, ET and ABA (Zhang *et al.* 2010) while *PR2* is induced by SA (Kariola *et al.* 2005). These marker genes were analysed 24 hours after virus inoculation to determine which pathways were affected. While certain marker genes (*PDF1.2*, *VSP2* and *PR1*) expression was analysed 14 days after viral inoculation when the virus is systemic.

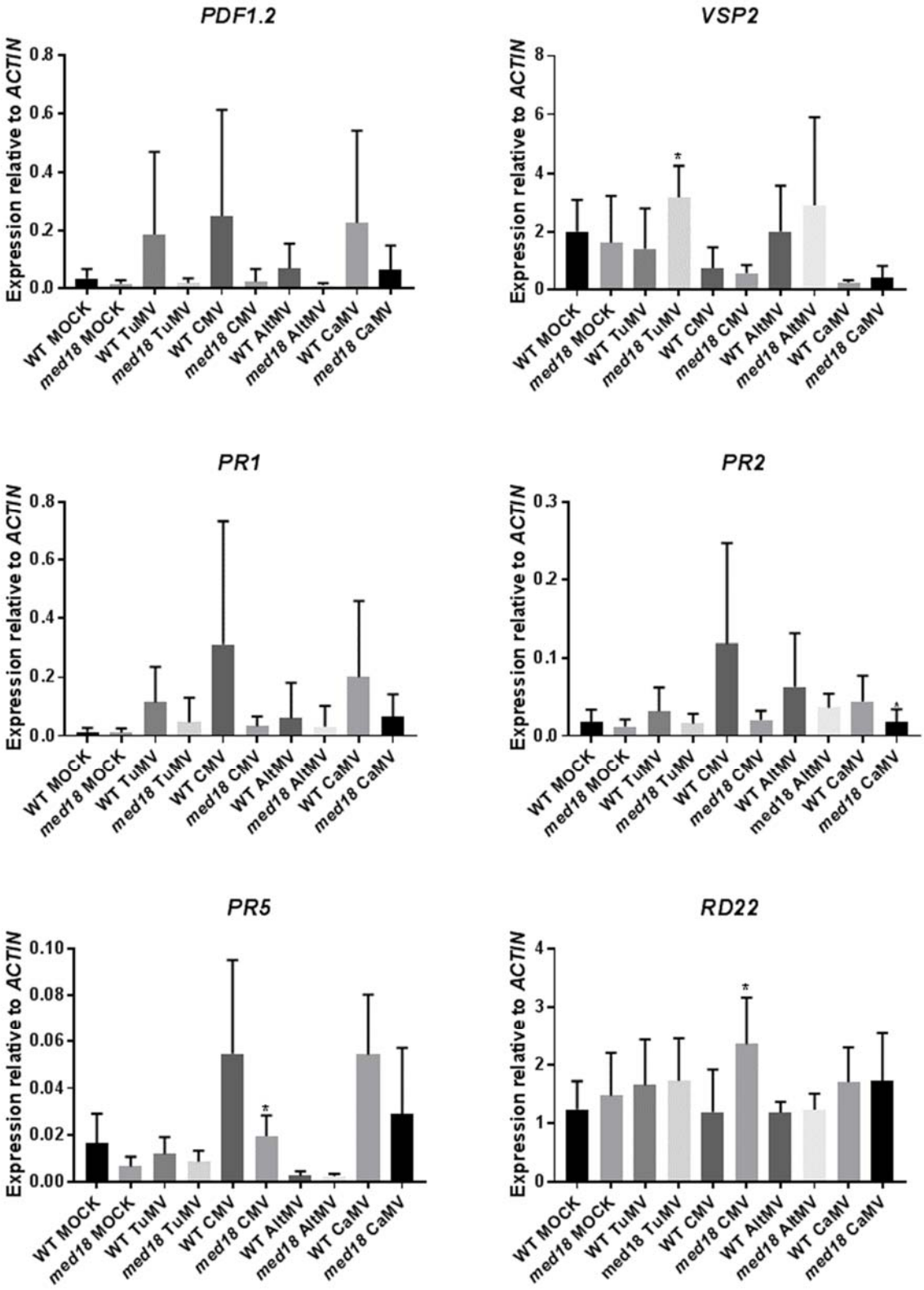


Figure 18: Marker gene analysis for the JA, SA and ABA pathways relative to *ACTIN* transcripts 48 h after viral inoculation of various *Arabidopsis* genotypes. JA, SA and ABA pathway maker gene expression were calculated relative to *ACTIN* using the delta CT method. T-tests were performed comparing WT infected to *med18* infected plants. Significant differences are indicated by the asterisks. Errors bars were calculated using standard deviations between replicates.

The first assay was performed with TuMV infected WT and *med18* plants, collecting tissue at 7 and 14 days after inoculation. The viral expression level or viral load was significantly higher in WT plants than *med18* plants (Figure 19).

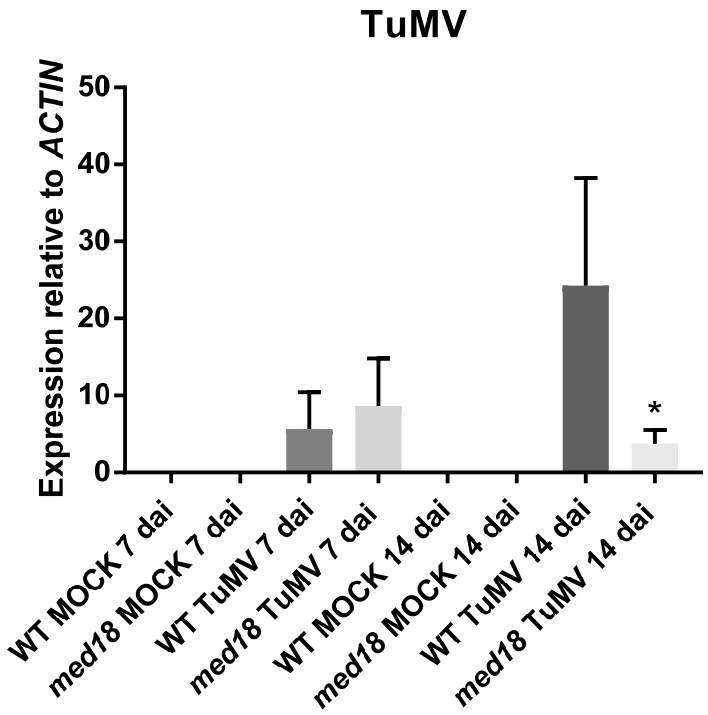


Figure 19: The expression or abundance of TuMV RNA was quantified in relation to *ACTIN* using specifically designed RT primers. The averages of three biological replicates were used to create the graphs while the error bars are the standard deviation between the three replicates. A t-test was performed comparing WT infected to *med18* infected plants.

To determine whether this result was reproducible with other viruses, *med18* plants were inoculated with TuMV, CMV, AltMV and CaMV and collected after 14 days (Figure 20). This assay also differed from the first, as, besides only have one time point, individual plants were collected, and replicates were increased from three (with pooled plants) to ten (using individual plants). The assay was altered to prevent grouping of infected and non-infected plants together; this would also ensure that gene expression results were more specific.

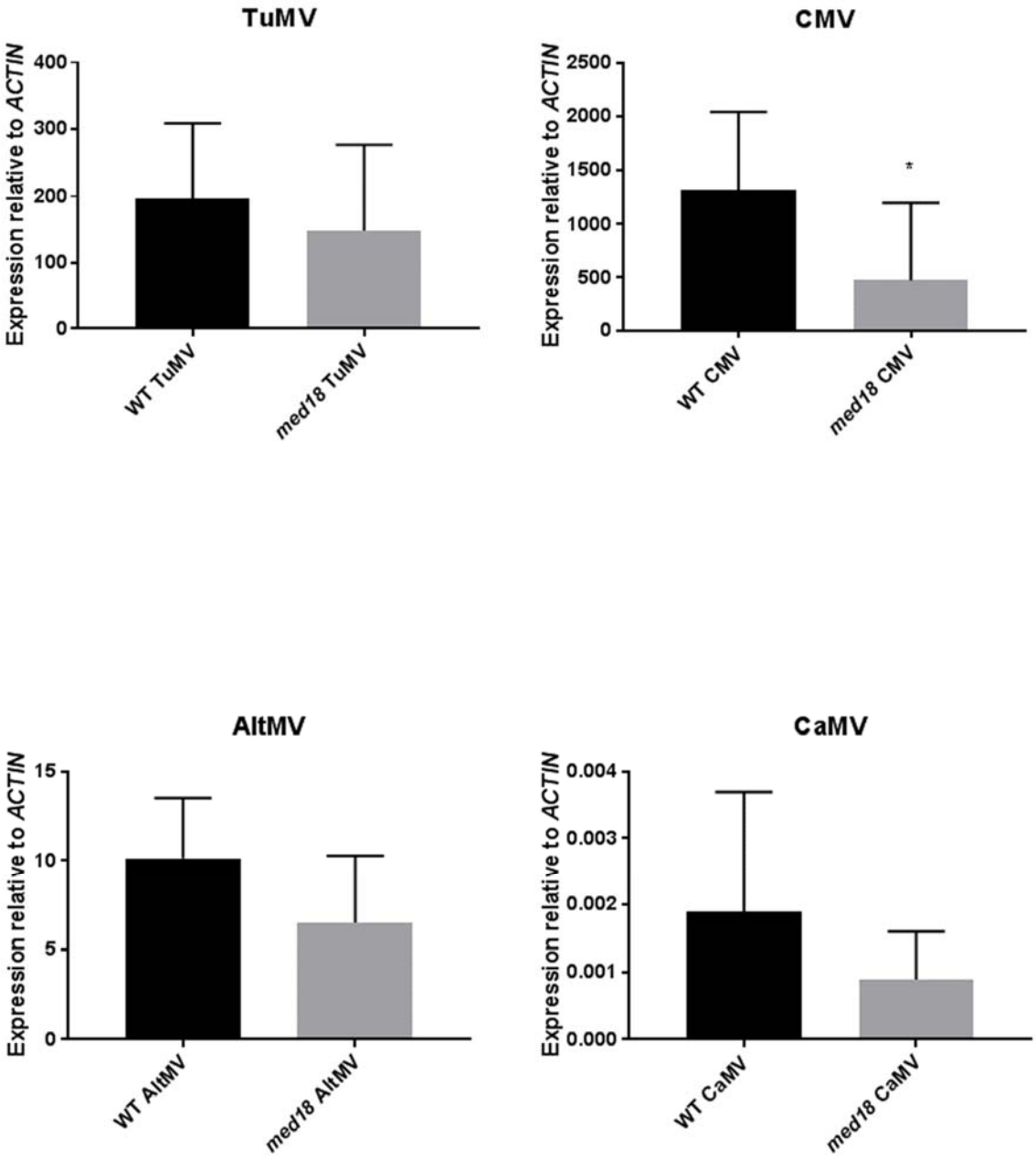


Figure 20: The viral load of TuMV, CMV, AltMV and CaMV relative to *ACTIN* transcripts 14 days after inoculation. The averages of ten biological replicates, excluding outliers identified using GraphPad Prism 7.03, were used to create the graphs. The error bars represent the standard deviation between the replicates. Viral load differs quite drastically between plants which accounts for the large error bars.

It was hypothesised that the reason for the difference in viral load between WT and *med18* plants was possibly due to the *med18* plants having decreased JA signalling capability, as the JA pathway does not function correctly in these plants (Fallath 2016). As my previous early defence response experiment showed that TuMV upregulates the JA pathway after 24 hours (Chapter 2, Figure 14), an early defence response experiment was performed with WT and *med18* plants infected with TuMV, CMV, AltMV or CaMV. Non-inoculated leaves were collected 24 hours after infection and qRT-PCR was performed with the following genes: *PDF1.2*, *VSP2*, *PR1*, *PR2* and *PR5* (Figure 18).

The expression of the JA and SA marker genes was varied although some appear to follow similar trends. *PDF1.2* and *VSP2* are both JA marker genes, *PDF1.2* is upregulated by all viruses when compared to the mock-inoculated plants but it is also strongly induced in comparison to the *med18* plants. In the *med18* plants, the JA pathway appears to be unaffected by the virus and was only slightly higher when infected with CaMV. However, due to the large error bars there was no significant difference between or within groups. *VSP2* in WT plants show no trends when infected although *med18* infected with CMV was significantly downregulated, while this gene in plants infected with AltMV was significantly upregulated in comparison to the mock-treated plants. When infected with CaMV there was a significant downregulation for both WT and *med18* plants, in comparison to the mock-treated plants.

The SA marker genes *PR1, PR2* and *PR5* were all strongly upregulated by CMV in WT plants when compared to the mock control plants. *PR1* was also significantly upregulated in WT plants infected with TuMV. *PR2* and *PR5* were significantly upregulated in *med18* plants infected with CMV. In AltMV-infected plants, *PR2* was significantly upregulated in *med18* plants, while *PR5* was significantly downregulated for both WT and *med18* plants, in comparison to the mock-treated plants.

The expression of *PDF1.2*, *VSP2*, *PR1* and *RD22* were also analysed 14 days after inoculation to determine if the significant down-regulation of CMV viral load in *med18* plants was possibly related to one of these defence pathways (Figure 21). Unsurprisingly, *PDF1.2* was significantly down-regulated in *med18* plants inoculated with CMV. However, the ABA pathway or *RD22* marker gene was significantly up-regulated in these plants. This suggests the decrease in *PDF1.2* is possibly causing an increase in the ABA as these two pathways act antagonistically. Interestingly, *VSP2* which is also JA inducible gene has increased expression in these plants though not significantly. *VSP2* is wound inducible and its increase may be related to the increase in oxidative stress resulting in cell damage and death.

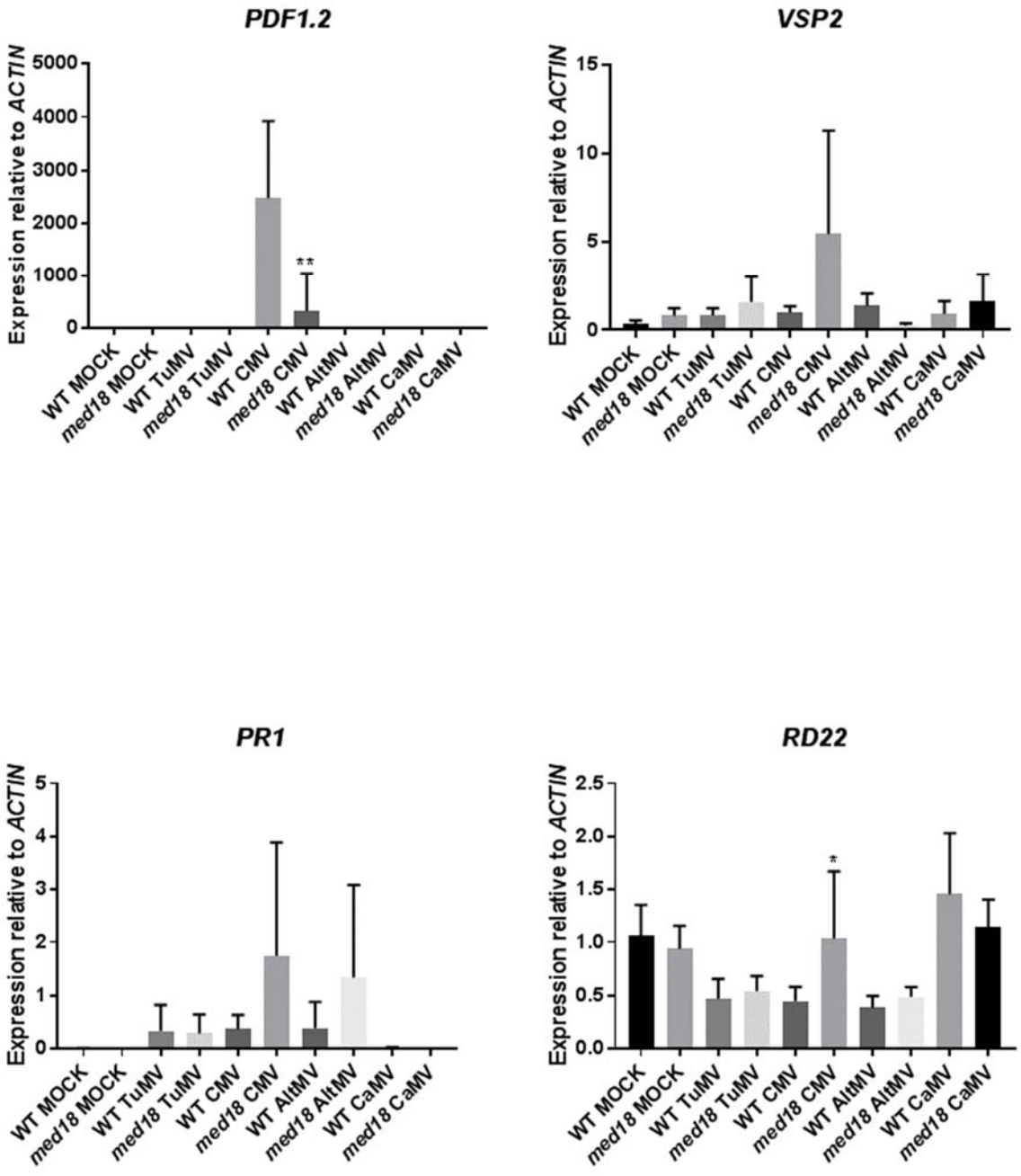


Figure 21: Marker gene analysis for the JA, SA and ABA pathways relative to *ACTIN* transcripts 14 days after viral inoculation of various *Arabidopsis* genotypes. The averages of ten biological replicates, excluding outliers identified using GraphPad Prism 7.03, were used to create the graphs. Viral load differs between plants effecting gene expression which accounts for the large error bars. Error bars represent the standard deviations between replicates. T-tests were performed comparing infected WT to infected *med18* plants.

As the MED18 subunit is thought to have a role in regulating ABA (Lai *et al.* 2014) and ABA effects RNA silencing, then the decrease in viral load seen in *med18* plants may also be due to this link. I tested three RNA silencing genes (Figure 22). *AGO1* was tested as miR168 which regulates its

expression is controlled by ABA. I analysed *Dicer-like (DCL) 2* and *4* as these are the main DCLs known to have a role in viral defence (Parent *et al.* 2015). The expression of *AGO1* and *DCL2* was higher in *med18* plants, although not significantly due to the variability of the expression. *AGO1* expression was decreased in all plants after inoculation. The expression of *DCL2* and *DCL4* followed a similar pattern; where there was an increase in *med18* plants when inoculated with TuMV, CMV or AltMV but the opposite when inoculated with CaMV. Though no significant up or down regulation were evident.

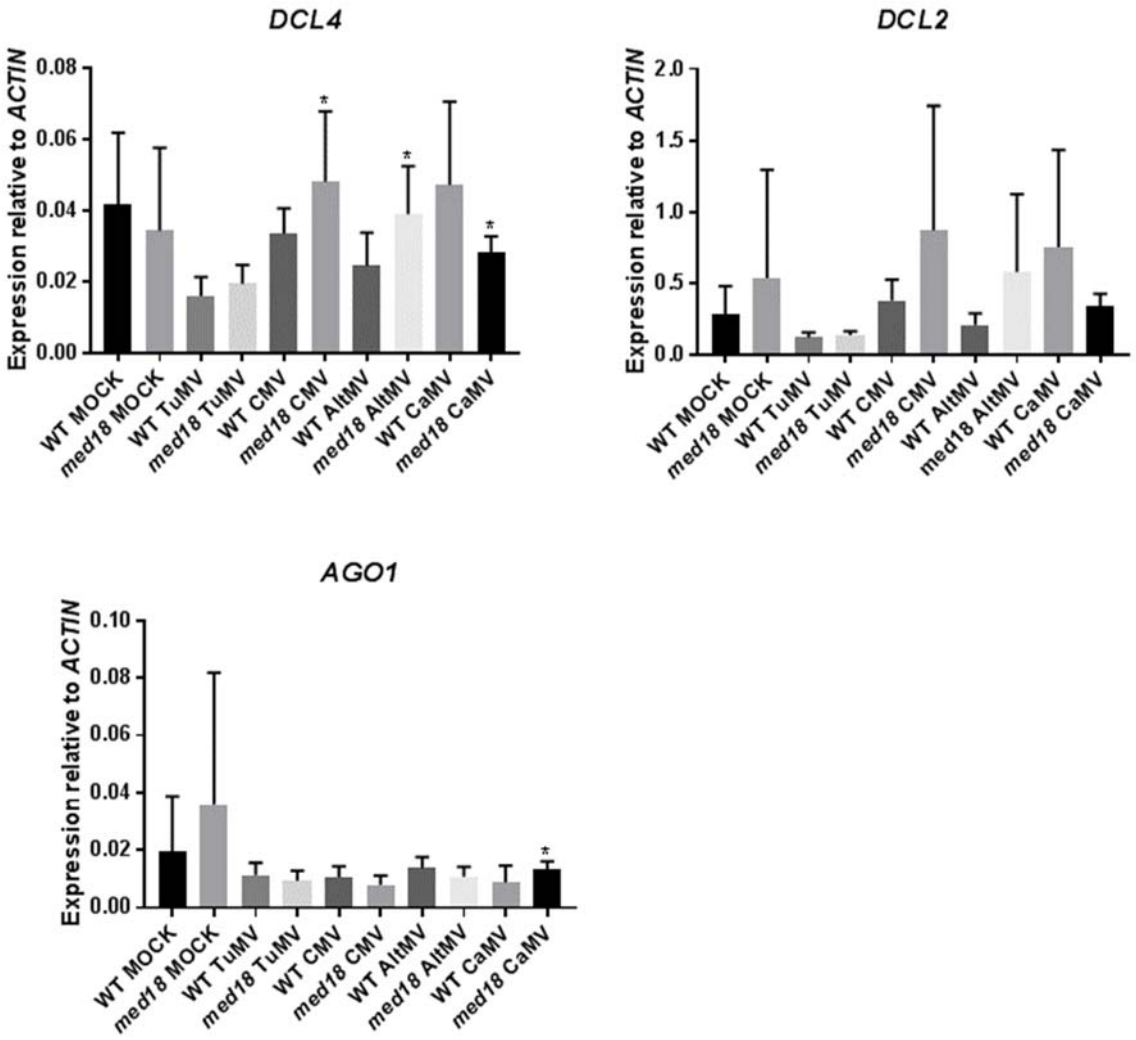


Figure 22: The expression changes 14 dai of important RNA silencing genes of various A t-test was performed to determine significant changes within treatments between WT and *med18* plants. The averages of ten biological replicates were used. Error bars represent the standard deviation between replicates.

4.3 Discussion

Results of the virus challenge show a trend suggesting that *med18* plants had a reduced viral load compared to WT plants, although due the highly variable levels of viral load between individual plants the differences were not significant (Figure 20). The only significant difference in viral load between *med18* and WT plants was when plants were infected with CMV. The possible reasons for this difference will be the focus of the discussion, first looking at the mediator complex itself and how the virus may benefit for targeting certain subunits.

In mammalian cells the mediator subunit MED8 was found to be targeted by the second non-structural protein of the *Bunyamwera virus* (BUNV) resulting in a significant decrease of host cell protein synthesis and the inhibition of the innate immune response (Léonard *et al.* 2006). BUNV is a single-strand negative-sense RNA virus with a tripartite genome; it replicates in the cytoplasm, though some of its proteins have been found in the nucleus of infected cells. The viral protein directly interacts with the MED8 subunit preventing the specific phosphorylation of RNA Pol II required for correct mRNA transcription. They also found that the interaction was necessary for efficient viral replication as a virus lacking the interaction domain was susceptible to the cells innate inhibition of viral replication. Similarly, plant RNA viruses also have non-structural proteins which are found in the nucleus and could possibly be preventing certain plant defence responses inhibiting viral replication. For example, TuMV has a number of proteins which are translocated to the nucleus (Ivanov *et al.* 2014); the viral-genome-linked protein of the nuclear inclusion protein interacts with the potyviral VP-g interacting protein which has a role in transcriptional control and this interaction may interfere with host gene expression (Dunoyer *et al.* 2004). TuMV also has another non-structural protein which accumulates in the nucleus and is known to interfere with RNA silencing, helper component proteinase (Mallory *et al.* 2002, Kasschau *et al.* 2003), which is similar to the CMV 2b which is described as a strong inhibitor of RNA silencing and RNA-directed DNA methylation (Lucy *et al.* 2000, Wang *et al.* 2004, Diaz-Pendon *et al.* 2007). AltMV protein triple gene block 1 is located in the nucleus and is also described as a suppressor of RNA silencing (Nam *et al.* 2013). Lastly, CaMV, though it is a DNA virus also replicates in the cytoplasm via an RNA intermediate. Its protein 6 is also known to accumulate in the nucleus with one of its role being supressing RNA silencing though interacting with the host nuclear protein DRB4 (Love *et al.* 2012). This suggests that all viruses used in this study can interfere with the mediator complex subunits and its functionality.

The MED8 subunit in plants appears to have a similar function to MED18 and has been linked to the JA-dependent defence pathway with mutants showing increased resistance to *F. oxyporum* but increased susceptibility to necrotrophic fungi (Fallath *et al.* 2017). MED8 interacts with MED18 and MED20 attaching these other two subunits to the mediator complex (Larivière *et al.* 2008) and the

domains of each subunit were found to be dependent on each other for proper formation (Shaikhibrahim, Rahaman *et al.* 2009). As all three subunits are in the head section of the molecule they are considered to have a key role in regulating RDR2 leading to the induction or repression of certain genes through mRNA transcription (Chadick and Asturias 2005, Takagi *et al.* 2006, Imasaki *et al.* 2011). Subunits MED18, 17 and 20a were found to have a role miRNA biogenesis (Kim *et al.* 2011). Kim *et al.* (2011) found that these subunits were involved in promoting the transcription of miRNA genes and were also responsible for recruiting RDR2 to specific promotors. The mediator complex also has a role in the transcription of long non-coding RNAs and transcriptional gene silencing (TGS) regulated by small interfering RNAs (Kim *et al.* 2011, Allen and Taatjes 2015). The complexity of the mediator complex and the initiation of RDR2 transcription has been linked to a type of ancient cellular defence in preventing DNA and RNA elements such as transposons and viruses from hijacking the host transcriptional machinery (Madhani 2013). This could possibly explain why the mediator subunits could be targeted by viruses; the mediator has a known role in plant defence, it is a key regulator in host gene expression with specific roles in mRNA transcription, miRNA gene transcription and transcriptional genes silencing - in other words preventing the correct function of the mediator complex would significantly impact the plant's production on many levels and the lacking subunit may also impact negatively on viral replicate as seen in the BUNV study.

While the mediator complex is an important part of transcription, certain subunits have been linked to defence pathways. qRT-PCR results from this thesis suggest that the marker which has the most striking difference between WT and *med18* plants is *PDF1.2*, a marker gene for the JA defence pathway. The expression of *PDF1.2* was upregulated in all virus-infected WT plants, while all *med18* plants appeared to be unresponsive, with the expression of *PDF1.2* remaining almost unchanged from the mock treatment to virus inoculations (Figure 18). *PDF1.2* is a plant defensin gene which is induced in response to certain pathogens via the JA/ET pathway (Brown *et al.* 2003). It was found that both pathways needed to be upregulated simultaneously to induce *PDF1.2* (Penninckx *et al.* 1998). The upregulation of *PDF1.2* possibly suggests that the virus may be hijacking the JA pathway, making the plant more susceptible which may be the reason for the lower viral load in *med18* plants. Comparable findings were demonstrated when *coi1* plants were infected with *Tobacco mosaic virus*, resulting in reduced viral replication and increased plant resistance (Oka *et al.* 2013). JA defence pathway upregulation is linked to a downregulation of SA and ROS signalling (Niki *et al.* 1998, Kunkel and Brooks 2002), both of which are effective in anti-viral defence via programmed cell death, thus limiting the virus spread throughout the plant.

Interestingly, MED18, as well as a number of other mediator subunits, were found to be required for the JA induced expression of *PDF1.2* (Wang *et al.* 2016). This correlates with my results, as all

med18 plants displayed reduced expression of *PDF1.2*. A recent study performed in the Schenk Lab also found that *med18* plants showed significantly reduced JA-associated gene expression, confirming the notion that MED18 has a key role in the JA pathway (Fallath 2016).

Though the role of JA is clear with regards to necrotrophic pathogens, there is still much debate as to what its role is against biotrophic pathogens, such as viruses. A study by Kovač *et al.* (2009) found that there was a significant increase in JA in resistant cultivars approximately 1 h after virus infection. However, as this increase was only in the infected tissue it was suspected that the increase in JA was an early local defence against *Potato virus* Y^{NTN}. Similar to my results, the early upregulation of JA may be an attempt of the virus to prevent early HR and the accumulation of ROS. There is an example where JA was found to accumulate during HR in response to TMV infection, though only at temperatures below 25ºC and if cultivars carried the resistance *N* gene (Seo *et al.* 2001).

The upregulation of JA typically results in a decrease in SA as these two pathways are known to act antagonistically (Niki *et al.* 1998, Kunkel and Brooks 2002). Though there are some examples where they act in a co-ordinated manner (Schenk *et al.* 2000, Mur *et al.* 2006, Kidd *et al.* 2009). This seems to be the case with our results as there is a strong increase in JA expression in infected WT plants but no significant decreased expression for any of the SA marker genes. However, we do see an antagonistic interaction between the JA and ABA pathway in *med18* plants specifically (Anderson *et al.* 2004). Interestingly, the SA pathway seems to be non-reactive in *med18* plants though this could be due to the increase in ABA as seen in Figure 20 for *RD22*. The ABA pathway was significantly upregulated by CMV in *med18* plants compared to the mock-treated and the WT plants infected with CMV.

ABA has a role in regulating development (Rock and Quatrano 1995), the plant's response to abiotic stress, specifically salinity and drought (Bartels and Sunkar 2005) as well as plant defence (Mauch-Mani and Mauch 2005). The recent review by Alazem and Lin (2015) highlights the importance of ABA in viral defence and interactions. Firstly, ABA upregulation is linked to the increase of callose deposits which was found to resist the movement of viruses. Endogenous application of ABA also reduced disease symptoms and decreased the viral load of bean plants (*Phaseolus vulgaris*) infected with *Tobacco necrosis virus* (TNV) (Iriti and Faoro 2008). This resistance varied between hosts and viruses and appeared to be dependent on specific R-genes (Alazem and Lin 2015). A study by Alazem *et al.* (2013) found that both the SA and ABA pathways were up-regulated after viral infection.

Similarly, there was also more *RD22* transcript abundance in *med18* mock-treated than in WT plant at 48 hours after infection and 14 days after infection (Figure 18 and 21). I propose that apart from hijacking of the JA pathway this is also a possible reason for decreased viral load in *med18.*

Furthermore, ABA was also found to have a role in RNA silencing (Chinnusamy *et al.* 2008) which is the main defence against viruses. ABA levels were directly correlated to miR168 which regulates *ARGONAUTE1* (*AGO1*), which is a vital part of the siRNAs pathway resulting in the silencing of certain genes (Li *et al.* 2012) and viral defence (Wang *et al.* 2011). Though other AGO proteins have important roles in RNA silencing against viruses, AGO1 was chosen due to its regulation through ABA. The RNA silencing pathway is triggered by double-stranded RNA which is then acted upon by dicer-like nucleases (DCL) creating small RNA fragments usually 21-24 nt in length. These are then loaded into an AGO protein forming the RNA silencing complex, silencing transcripts in a sequence-specific manner. DCL2 and DCL4 are vital for effective viral silencing against single-stranded RNA viruses (Mlotshwa *et al.* 2008). Looking at the expression of *AGO1*, *DCL2* and *DCL4* at 14 days after infection, *med18* mock-infected plants had slightly increased expression of all three over WT plants, though not significantly (Figure 22).

Expression of *DCL2* was significantly decreased in both, WT and *med18* plants, when infected with TuMV. CMV and AltMV. CaMV caused a slight increase in expression for *DCL2* in WT plants and a decrease in *med18* plants. In the comparison between infected WT and *med18* plants, both, CMV and AltMV, had increased *DCL2* expression in *med18* plants, while TuMV and CaMV caused higher *DCL2* expression in WT plants. The expression of *DCL4* is significantly higher in *med18* plants infected with TuMV, while it was downregulated in WT plants. The infection of CMV and AltMV also caused a decrease in expression in WT and *med18* plants. Even though the expression of *DCL4* was decreased by viral infection for both, WT and *med18* plants, the expression of *DCL4* in infected *med18* plants was higher than in WT plants, with the exception of CaMV. These results show a trend for TuMV, CMV and AltMV with an increase in expression in *med18* compared to WT plants.

Due to the variability of the viral load results it is difficult to determine whether the *med18* subunit is manipulated by viruses for replication as seen in the BUNV study. Alternatively, its absence may lead to downstream effects preventing the virus from manipulating the plants' defences in another manner. From this study it can be concluded that the virus does manipulate the plant's defences at the initial stages of infection with the upregulation of the JA pathway. A possible cause for the decrease in viral load is that the non-functional JA pathway may allow for a constant increase in ABA. Which could result in the *med18* plants i) having an increase in callose deposition and limiting the spread of the virus and/or ii) increasing the expression of miR168, resulting in more RNA silencing machinery and more efficient silencing of the virus. This could be particularly true for CMV as its relative abundance was significantly decreased in *med18* plants (Figure 20) which could correlate with the significant induction of the ABA pathway (Figure 18 and 21).

Table 6: List of real-time PCR primers used in Chapter 4

Primer Name	Sequence
rtTuMV_F	TCGAGCGTTACGGAATTTCAG
rtTuMV_R	GATGATCATACAGCGCTTGCA
Rt-CMV-KCP-B	AGTACCGGTGAGGCTCCGTC
Rt-CMV-KCP-A	TGATTCTACCGTGTGGGTGAC
rtAltMV_F	GGACGCTTTCACCCCACAT
rt_CaMVF	TGAAATCCTCAGTGACCAAAAATC
rt_CaMVR	TACAAGGACAATCATTGATGAGC

I always imagine two Luke Skywalkers whenever I hear "Dual-Luc Assay".

Chapter 5: Analysis of TuMV Derived Small Interfering RNAs and Their Role in Plant-Virus Interactions

5.1 Introduction

sRNAs can be thought of as David in the David and Goliath story, as their tiny size packs a massive punch. sRNAs are small pieces of RNA, usually between 20-24 nucleotides in length, that act as gene regulators (Ramachandran and Chen 2008). This mechanism of regulation is called RNA silencing; implicated in regulating a number of different biological processes including development, response to biotic and abiotic stresses, defence against various pathogens, as well as chromatin condensation and suppression of transposable elements within the plant's genome (Baulcombe 2004, Zvereva and Pooggin 2012). RNA silencing is mediated by double-stranded RNA (dsRNAs), which is then processed by RNase III-like Dicer (DCL) protein, forming short 20-24 nucleotide sRNAs. These sRNAs are classified as either short interfering RNAs (siRNAs) or microRNAs (miRNAs) depending on the fragments they were derived from (Balmer and Mauch-Mani 2013). There is much debate around the difference between siRNAs and miRNAs, in function as well as where they are derived from. According to Voinnet (2009) the distinctions made in past between siRNAs and miRNAs are hazy and the only distinctions which still hold true is that miRNAs are derived from an imperfect stem-loop precursor while siRNAs are derived from perfect RNA duplexes.

With regards to function it is believed that siRNAs have a crucial role in plant antiviral defence (Wang *et al.* 2012) while miRNAs are involved in development, signal transduction, the degradation of protein, as well as response to both biotic and abiotic stressors (Lu *et al.* 2008). The silencing process for both sRNAs are similar; after the dsRNAs or hairpin RNA is cleaved by DCL, it is loaded into an enzyme which is part of the AGO family, forming the RNA-induced silencing complex (RISC). The sRNAs sequence is then used as a guide to direct the silencing of the target sequence based on complementarity (Baulcombe 2004, Voinnet 2009, Wang *et al.* 2012, Balmer and Mauch-Mani 2013).

5.1.1 miRNA Biogenesis

The biogenesis of miRNAs is different to that of siRNAs, as it begins with the transcription of a nuclear encoded miRNA gene by the RDRP 2. According to Lee *et al.* (2004), RDRP 2 is the only polymerase responsible for transcribing miRNA genes, to produce the pri-miRNA. Next the enzyme DAWDLE is thought to stabilise the pri-miRNA allowing the orchestrated action of C2H2-zinc finger protein SERRATE (SE), double-stranded RNA-binding protein HYPONASTIC LEAVES1 (HYL1), as well as DCL1 and the nuclear cap-binding complex (CBC) to form D-bodies, creating the stem-loop structure (Han *et al.* 2004, Yang *et al.* 2006). The pre-miRNA or mature miRNA are produced by DCL1 and are then exported to the cytoplasm by the enzyme HASTY, a plant exportin 5 ortholog,

(Bollman *et al.* 2003) where they are methylated by HEN1 (Yang *et al.* 2006) protecting them from degradation by endonucleases. One strand from the mature miRNA is then selected and incorporated in to an AGO protein forming the RISC complex, which is then directed by the miRNA sequence to the specific target. A simplified diagram can be seen in figure 23.

5.1.3 Different Enzymes in sRNAs Processing

Four DCL proteins have been identified in *Arabidopsis thaliana*, each producing different sized sRNAs; DCL1 produces 18-21 nt sRNAs while the remaining three, DCL2, DCL3 and DCL4, produce 22, 24 and 21 nucleotide long sRNAs, respectively (Chapman and Carrington 2007). Depending on the sRNAs function, it will either remain in the nucleus where it was first processed or exported to the cytoplasm where it will be involved in posttranscriptional gene silencing. The respective DCL proteins act on and process different forms of dsRNAs. Firstly DCL1 is known to process fold-back precursors releasing miRNAs (Bartel 2004). DCL3 processes siRNAs from DNA repeats which then guide heterochromatin formation (Xie *et al.* 2004). DCL4 processes siRNAs which have a role in posttranscriptional gene silencing of endogenous genes (Xie *et al.* 2005) and transgenes (Dunoyer *et al.* 2005). DCL2 was found to act on stress related natural antisense siRNAs (natsiRNAs) (Borsani *et al.* 2005). Deleris *et al.* (2006) reported that DCL2 could act as a substitute for DCL4, as dcl4 plants revealed DCL4 function was performed as well as the DCL2 function by DCL2.

As mentioned above, the next step after cleavage is for the sRNAs to be incorporated into an AGO protein forming the RISC complex. Ten AGO proteins paralogues have been identified in *Arabidopsis thaliana*, only four of which have been properly characterised; AGO1, AGO4, AGO6 and AGO7 have roles in RNA silencing (Voinnet 2009). According to Vaucheret (2008), AGO protein can be placed into three different clades base on their phylogeny; the first clade includes AGO1, AGO5 and AGO10, while AGO2, AGO3 and AGO7 make up the second clade, and lastly AGO4, AGO6, AGO8, and AGO9 are in the third.

AGO proteins have a central role in RNA silencing, as all three RNA silencing pathways require these proteins to ensure correct regulation of targeted mRNAs. For example, it has been found that *Arabidopsis ago1* mutants have malfunctioning cytoplasmic RNA silencing pathways, meaning that there was a decrease in miRNA and an increase in the corresponding mRNA. AGO1 is consider the most important AGO protein as it required to ensure the miRNA pathways integrity and its absence results in severe developmental defects in comparison to other *ago* mutants. AGO1 is also suspected to play a role in siRNAs viral resistance as *ago1* mutants are also highly susceptible to CMV (Morel *et al.* 2002). Interestingly, there seems to be a redundancy of function between AGO1 and PINHEAD/ZWILLE or AGO10, especially in postembryonic developmental aspects. It was observed

that a single mutation of *ago1* or *ago10* acted dominantly in a homozygous mutation of *ago10* or *ago1*, respectively (Vaucheret 2008). Lynn *et al.* (1999) found that both AGO1 and AGO10 are from the same gene family known for mediating protein-protein interactions and mRNA translation. They also found that AGO1 and AGO10 function together ensuring correct growth and gene regulation during embryogenesis.

The second closest paralogue to AGO1 is AGO5, which like AGO1 acts in both the cytoplasm and the nucleus. However, unlike AGO1, *ago5* mutants do not show any severe growth defects. Immunoprecipitation experiments showed that AGO5 is associated with sRNAs that have a 5'-terminal cytosine (Takeda *et al.* 2008), the relevance of this specific association has not yet been determined.

AGO7 is part of the second clade of AGO proteins which was found to play a role in correct timing between the juvenile and adult leaf phase change (Hunter *et al.* 2003). It is associated with trans-acting siRNAs (tasiRNAs) which were discovered roughly at the same time (Vaucheret 2006). These siRNAs do not necessarily differ in function, as they still have a role in gene regulation, they differ in their processing. Firstly, tasiRNAs are processed via the miRNA pathway before commencing their role in gene regulation. Specific miRNAs are processed to cleave tasiRNAs precursor RNAs (TAS) which are then stabilized by SGS3 (Peragine *et al.* 2004). Thereafter one of the two cleaved TAS fragments are transcribed by RDR6 to produce dsRNAs, which is then cleaved again by DCL4 into 21-nt tasiRNAs (Vaucheret 2005). The percentage of complementation to the target sequence will determine which AGO protein they are coupled with to produce a RISC structure; if near-perfect complementation, like most miRNAs, the tasiRNAs will associate with AGO1 to guide silencing through cleavage, while less perfect complementation will mean the tasiRNAs will interact with AGO7 (Vaucheret 2008). However, TAS3-derived tasiRNAs usually interact with AGO7, which is known to control the phase change form juvenile to adult leaves (Adenot *et al.* 2006).

Also in the second clade are the proteins AGO2 and AGO3, both displaying similar genetic structure and presumably function suggesting they recently originated from duplications (Vaucheret 2008). Similar to *ago5*, there were no obvious phenotypic defects in *ago2* or *ago3*. In accordance with the redundancy hypothesis, *ago2* and *ago3* were not rescued through forward genetic screens. AGO2 is one of the few AGO which has an identified affinity for siRNAs with a 5'-terminal adenosine, discovered through immunoprecipitation experiments by Takeda *et al.* (2008). The reason for this association is still unclear, as when the uridine is modified in artificial miRNAs to an adenosine, the miRNA associates with AGO2 instead of AGO1 which eliminates its regulatory functions altogether (Mi *et al.* 2008).

AGO4 was identified as being associated with transcriptional gene silencing (TGS), which is a process of histone modification usually through direct DNA methylation coordinated by homologous siRNAs which originated from transposable or repeated elements. As a result, *ago4* mutants had a decrease in chromatin silencing (Baulcombe 2004). Besides AGO4, the other proteins which are known to have a key role in histone modification are DCL3, RDR3, as well as two different forms of nuclear Pol IV (Pontes *et al.* 2006). Similar to *ago4*, mutation in any of these genes would also result in decreased chromatin methylation. AGO4, like AGO1, has slicer activity; Qi *et al.* (2006) showed that AGO4 cleaved its target mRNA however it's catalytic and methylation function were found to act separately, depending on the target mRNA. AGO6 also has a role in heterochromatin methylation, with similar functions to AGO4; the results from Zheng *et al.* (2007) experiments suggested that there was high redundancy between AGO6 and AGO4 function with regards to both siRNAs gene regulation and DNA methylation.

The functions of the remaining two AGO proteins, AGO8 and AGO9, are relatively unknown; mutants of either do not exhibit any obvious developmental defects and forward genetics screening has not recovered either mutation. However, AGO9 was found to have a role in female gamete formation and was found to mainly interact with 24-nt sRNAs. Both AGO8 and AGO9 are physically separated from one another by 50 kb, and it is assumed, like AGO2 and AGO3, they originated from a recent duplication; which also suggests they most likely have functional redundancy (Vaucheret 2008). AGO8 is also thought to be a pseudogene as its transcripts levels were very low in all tissues tested (Mallory and Vaucheret 2010).

5.1.4 RNA Silencing and Antiviral Defences

As well as having a regulatory role in gene expression, RNA silencing is the plant's main defence against viruses (Voinnet 2001, Waterhouse *et al.* 2001). dsRNAs can be thought of as a signalling molecule; evoking the plant's gene silencing machinery in an attempt to destroy the foreign RNA and defend itself against the perceived viral attack. These sRNAs produced by the plant silencing machinery are called viRNAs (Molnár *et al.* 2005).

The main DCL proteins which are known to have a role in producing viRNAs are DCL2 and DCL4, as virus replication was at its maximum in double *Arabidopsis dcl2dcl4* mutants. This was true for RNA viruses such as CMV (Bouché *et al.* 2006) as well as *Cabbage leaf curl virus* (CaLCuV) (Blevins *et al.* 2006). DCL1 has a negligible role in viral defence against certain RNA viruses as the silencing of the other 3 DCL mutants was comparable to silencing all DCL and the viRNAs which DCL1 should have been producing were almost undetectable (Ding and Voinnet 2007). It is however thought that DCL1 has a role in defence against DNA viruses. Similar to its role in the tasiRNAs pathway, DCL1 processes hair-pins making sRNAs which have complementarity to cytoplasmic viral

mRNA. RDRP will create a secondary dsRNAs which will then be acted upon by another DCL proteins and used to silence viral mRNAs (Wang and Metzlaff 2005). DCL3 also has an important role against DNA viruses, specifically, through DNA methylation of the viral genome it is able to suppress viral replication (Bian *et al.* 2006). Figure 23 outlines the basic antiviral RNA silencing processing of a plant which clearly depends on the viral genome molecular structure, which also dictates where it replicates.

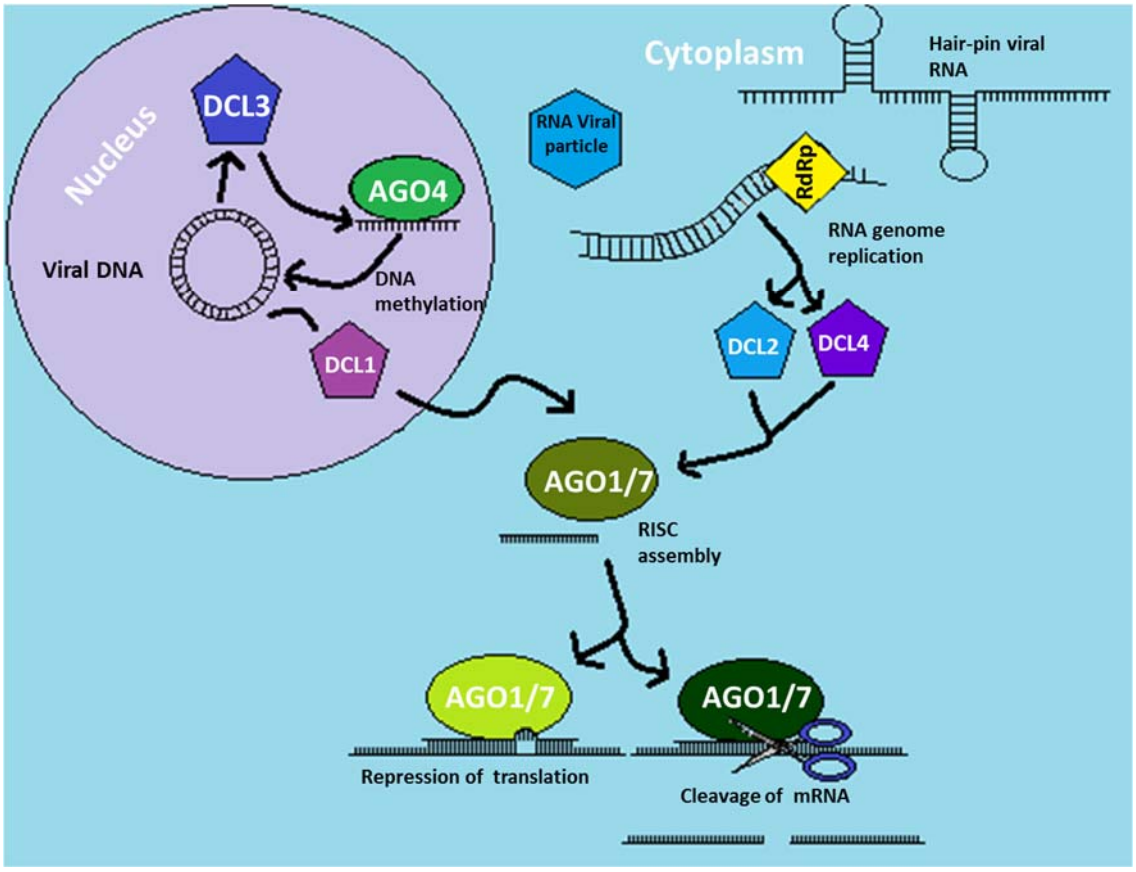


Figure 23: RNA silencing of viral products occurs in both the nucleus and cytoplasm depending on where the virus replicates; DNA viruses usually replicate in the nucleus while RNA viruses replicate in the cytoplasm. DCL3 and 1 were found to have a role in antiviral activity against DNA viruses, with DCL3 creating 24-nt siRNAs which methylate regions of the DNA virus' genome. While DCL1's role in antiviral defence is minor, it is thought to produces viral dsRNAs which are presumed to be processed similarly to the endogenous miRNAs; exported to the cytoplasm, incorporate with an AGO to form RISC and then used to targeted viral mRNA. dsRNAs from RNA viruses can be produced either through hair-pins in the viral genome or dsRNAs formed during viral replication. These are then acted upon by either DCL2 or DCL4 which produce 22-nt and 21-nt siRNAs, respectively. DCL4 is considered to be more dominant as 21-nt siRNAs are more abundant. These viral derive small RNAs are then loaded into an AGO protein forming the RISC molecule which will target viral mRNA, resulting in silencing either through cleavage or inhibition of translation.

As RNA silencing is the plant's main defence against viruses, it is not surprising that viruses have evolved their own counter defence strategies, specifically viral suppressors of RNA silencing (VSRs).

Counter defence mechanisms are intended to disrupt and inhibit the plant's RNA silencing pathways which ultimately prevent the plant from defending itself. VSRs have been shown to prevent RNA silencing in a number of ways, ranging from inhibiting crucial RNA silencing enzymes, such as DCL (Mérai *et al.*, 2006) and AGO (Azevedo *et al.*, 2010), to modification of its own as well as the plant's genome (Kanazawa *et al.*, 2011). Through these actions and others, viruses are able to weaken the plant's RNA silencing pathways, and use it to its own advantage as this interference will affect the host's mRNA production and ultimately its productivity.

This inference can be seen phenotypically in viral symptoms. For example Smith *et al.* (2011) reported that the yellowing symptoms from CMV were the result of a siRNAs derived from a viral satellite which silenced the chlorophyll biosynthesis gene, *CHLI*. As viral satellites and viroids do not encode any protein coding sequences, it would seem they act through RNA silencing, expressing sRNAs with high complementarity to target genes within the host genome, thereby inhibiting their expression (Wang *et al.*, 2004). Conversely, some plant viruses may purposely induce certain host miRNAs. For example, the p19 protein of *Cymbidium ring spot virus* was reported as an inducer of the host microRNA miR168 which prevents the expression of AGO1, a vital protein for the antiviral defence and the correct functioning of RISC (Várallyay *et al.*, 2010).

Some viruses are also able to express their own miRNAs; viral miRNAs are common amongst animal viruses and have been shown to actively regulate the host sRNAs metabolism (Sullivan and Ganem, 2005). Viral miRNAs have been identified for plant DNA viruses and it is still a matter of debate whether sRNAs derived from RNA viruses can be classified as miRNA (Pfeffer *et al.*, 2005). Some speculate that plant RNA viruses do not produce miRNA because their replication takes place in the cytoplasm and not in the nucleus, therefore they do not have access to the correct cellular machinery.

Interestingly, VSRs are reconditioned by the plant which elicits a counter-counter defence that is thought to work in a gene-for-gene manner, according to Li *et al.* (1999). In their paper they demonstrate that the 2b VSRs encoded by CMV prevent the plant from defending itself by initiating RNA silencing. This counterattack of the virus provokes an independent defence mechanism in the form of a counter-counter defence. Similarly, Sansregret *et al.* (2013) describe a type of extreme resistance which is triggered when the p19 protein of *Tombusviruses* attempts to suppress the plant's RNA silencing defence by binding to sRNAs, preventing loading into the AGO. This action can be recognized by the plant as a threat which can then trigger a defence mechanism similar to HR, however a more intense version which almost heightens the plant's senses inducing a board-spectrum antiviral state, similar to 'priming' mediated by beneficial microbes. Sansregret *et al.* (2013) propose that plants have an ongoing monitoring system for their RNA silencing machinery due to its

importance in growth, development and defence, and if threatened the plant is able to suppress the suppressor preventing any further damage. They also suggest that this higher level of defence will impose greater pressure on viruses to evolve in order to overcome these defences allowing the evolutionary arms race between host and pathogen to continue. This area of host counter-counter defence is still relatively new and much research is needed to determine whether this next layer of defence acts against viral pathogens only, or whether the plant is able to assess other pathogenic threats and choose to activate a more potent defence mechanism.

The results presented in this chapter focus on the viRNAs generated from TuMV in two model plant systems, *A. thaliana* and *N. benthamiana*. The viRNAs profile generated, the changes to the endogenous sRNAs profile, and the identity of viRNAs predicted to target host gene transcripts were investigated. The sRNAs were trimmed and filtered by size to create a library of 20-24nt sRNAs. These were then mapped to the host and TuMV genome to identify where the sRNAs originated. TuMV viRNAs profiling revealed that 21-nt viRNAs were most abundant followed by 22-nt viRNAs, which was expected for viRNAs produced by DCL4 and DCL2, respectively, which have major roles in antiviral defence (Dunoyer *et al.* 2005, Bouché *et al.* 2006, Deleris *et al.* 2006, Ding and Voinnet 2007). The endogenous siRNAs profile changed as a result of TuMV infection, with 24nt siRNAs abundance decreasing noticeably. Coverage of 100% was obtained after mapping viRNAs to the TuMV genome, with the coverage pattern also identifying hotspots along the genome suggesting certain areas are more prevalent to be processed by host RNA silencing factors. The sense and antisense profile of the virus was roughly 50:50 with no bias shown to either strand. This suggests that the majority of the processing is performed while the virus is replicating and not predominantly due to folding of the single stranded genome. The most abundant viRNAs (>50 copies per million) were searched for complementarity to transcripts from the respective host, using the web-based program psRNAstarget (Dai and Zhao 2011). The possible targeted transcripts with the lowset expect score (and therefore highest complementarity) were tested using a novel transient assay to determine which transcript sequences may be targeted by viRNAs. My results revealed one possibly true transcript target sequence from *N. benthamiana*, transcript 3160g02007, which encodes for a protein of unknown function.

The following draft manuscript "Analysis of *Turnip mosaic virus* Derived Small RNAs Generated in *Arabidopsis thaliana* and *Nicotiana benthamiana* and Identification of a Viral Small RNA Interaction with a Host Transcript Sequence" presents the main findings.

5.2 Analysis of *Turnip mosaic virus* Derived Small RNAs Generated in *Arabidopsis thaliana* and *Nicotiana benthamiana* and Identification of a Viral Small RNA Interaction with a Host Transcript Sequence

5.2.1 Introduction

Proteins encoded by plant viruses are known to contribute to virus proliferation and disease symptom development (Atreya *et al.* 1992, Suzuki *et al.* 1995, Lucy *et al.* 2000). But viroids, for example, do not encode for proteins yet still cause disease symptoms in plants (Wang *et al.* 2004). As a consequence, it is theorised that interactions at the RNA level must also contribute to viral disease symptoms (Zhang *et al.* 2015, Shi *et al.* 2016). RNA silencing is the main plant defence mechanism against viruses (Waterhouse *et al.* 2001). Yet RNA silencing also occurs in healthy plants and is used to control host gene expression (Baulcombe 2004). Consequently, it has been hypothesised that the small RNA products of viral RNA silencing may act on host gene transcript expression and contribute to the variety of disease symptoms displayed by virus infected plants (Smith *et al.* 2011).

In healthy plants, RNA silencing is triggered by dsRNAs which is then processed by RNA silencing machinery into small interfering RNA molecules (siRNAs), 21-24 nt in length, resulting in sequence specific targeting of RNA transcripts (Hamilton and Baulcombe 1999, Castel and Martienssen 2013, Matzke and Mosher 2014). Targeted genes are silenced either through methylation (He *et al.* 2011, Xie and Yu 2015), inhibition of translation, or cleavage leading to degradation of the target (Li *et al.* 2013).

In virus infected plants, RNA silencing is triggered by viral dsRNAs formed either through hairpins in single-stranded viral genomes or during the replication phase. The viral dsRNAs will be acted upon by Dicer-like nucleases (DCL) and loaded into Argonaute (AGO) proteins forming the RNA silencing complex (RISC) which target and silence the virus RNA (Baulcombe 2015). DCL4 and DCL2 are the main DCL's involved in the generation of primary viral derived small RNAs (viRNAs), creating 21 and 22 nt viRNAs respectively. *A. thaliana* mutants revealed that DCL4 was the main DCL known to act against viruses with DCL2 being redundant in comparison (Parent *et al.* 2015). There are 10 known AGOs in *Arabidopsis thaliana* of which AGO1, 2, 3, 5, 7, and 10 have all found to be implicated in viral defence (Li *et al.* 2016). AGO1, 2 and 10 were found to enhance the plants viral defence against *Turnip mosaic virus* (TuMV) in certain tissues (Garcia-Ruiz *et al.* 2015). Secondary viRNAs are generated or synthesised by the host's RNA-dependent RNA polymerases (RdRp) which uses the primary viRNAs as primers and the viral genome as a template to create a dsRNAs template (Donaire *et al.* 2009, Ghildiyal and Zamore 2009). The dsRNAs is usually acted upon by DCL4 creating secondary siRNAs (Axtell *et al.* 2006). Currently, RdRp1 and 6 are thought to have roles in defence specifically against viruses (Dalmay *et al.* 2001, Xie *et al.* 2001) and due to the sense and

antisense nature of the viRNAs there is strong evidence that RdRps are involved in their biosynthesis (Mourrain *et al.* 2000, Yu *et al.* 2003, Qi *et al.* 2009). Secondary viRNAs are more abundant than primary and are possibly the long-distance signalling molecule priming uninfected host cells (Schwach *et al.* 2005). Interestingly, there is also evidence of viRNAs with host sequence similarity are used as primers by RdRp to create dsRNAs of host genes, known as transitive silencing (Baulcombe 2015).

As RNA silencing relies on sequence similarity, viRNAs with extensive *complementarity* to host genes have the potential to regulate host gene expression (Llave 2010, Zhu and Guo 2012). A number of studies have predicted numerous host targets of viRNAs through deep sequencing (Qi *et al.* 2009, Catalano *et al.* 2012, Miozzi *et al.* 2013, Xia *et al.* 2014). There is also growing experimental evidence that viRNAs may regulate the expression of host genes. For example, a seminal study found yellowing symptoms associated with *Cucumber mosaic virus* (CMV) were caused by viRNAs from a viral satellite targeting a key host gene involved in chlorophyll synthesis (Shimura *et al.* 2011). Similarly, a viRNAs from a potato spindle tuber viroid was found to target and repress a callose synthesising gene (Adkar-Purushothama *et al.* 2015) which was also related to specific phenotypic symptoms. A recent study used 5'RACE to demonstrate cleavage events occurring in transcripts predicted to be targeted by *Tomato spotted wilt virus* viRNAs (Moyo *et al.* 2017). Identifying viRNAs host targets will lead to a better understanding of the plant-virus relationship, symptom development, and highlight potential resistant strategies. However, the approaches in those prior studies are laborious and not easily scalable to the validation of multiple predicted viRNAs-target transcript sequence interactions.

This study investigates the viRNAs generated from TuMV-QLDb (Pretorius *et al.* 2016) in *A. thaliana* and *N. benthamiana*, and predicts viRNAs interactions with the host transcriptome. TuMV is an economically important member of the genus *Potyvirus*, infecting many crops throughout the world. Through deep small-RNA sequencing I characterise and profile the viRNAs generated from TuMV. A recent report described an optimised transient dual luciferase assay used to assess predicted miRNA-target sequence interactions (Moyle *et al.* 2017). I investigated the utility of a similar transient dual luciferase assay system to experimentally validate predicted target sequences of viRNAs. A predicted viRNAs interaction with a target sequence of an *N. benthamiana* gene transcript of unknown function was validated using the transient dual luciferase assay.

5.2.2 Materials and Methods

5.2.2.1 Plant Materials for RNA Isolation

Arabidopsis thaliana seeds were germinated and grown for ~4 weeks in UC mix soil in a growth chamber under fluorescent tube lighting with an 8 h photoperiod and a constant 23°C. Seedlings were either mechanically inoculated with sodium phosphate buffer (mock inoculated) or mechanically inoculated with isolate TuMV-QLD1b in sodium phosphate buffer (Pretorius *et al.* 2016). Sampled tissue was snap frozen under liquid nitrogen, pulverised using a mortar and pestle, and stored at -80°C prior to RNA extraction. Aerial tissue was harvested from 20 individual plants per treatment. Equal amounts of tissue were pooled from 20 individual plants, then RNA was extracted using previously described methods (Sternes and Moyle 2015). The sampling times consisted of 7 days post mock inoculation (M7), 14 days post mock inoculation (M14), 7 days post TuMV inoculation, and 14 days post TuMV inoculation.

N. benthamiana seeds were germinated and grown for ~4 weeks in UC mix soil in a growth chamber under fluorescent tube lighting with an 8h photoperiod and at a constant 22°C. Seedlings at the four leaf stage of development were mechanically inoculated with isolate TuMV-QLD1b in sodium phosphate buffer (Pretorius *et al.* 2016). Leaf tissue was harvested 14 days post-inoculation, snap frozen under liquid nitrogen, pulverised using a mortar and pestle, and stored at -80°C prior to RNA extraction using previously described methods (Sternes and Moyle 2015).

5.2.2.2 sRNAs Library Construction and Deep sRNAs Sequencing

sRNAs libraries were prepared from the *A. thaliana* RNA samples using the TruSeq Small RNA kit (Illumina). The libraries were 35 base sequenced on the Illumina platform using the Australian Genome Research Facility as the service provider. The *N. benthamiana* sRNAs library construction and deep small RNA sequencing were performed using Novogene as the service provider.

5.2.2.3 Bioinformatics Pipeline

The *A. thaliana* raw sequences from each library were subjected to an initial filtering step to remove empty sequences from adapter-trimmed FASTA formatted read sequence files, using the Geneious version 8.1.7 software package. The remaining sequences were filtered on size to exclude reads less than 16 bases or more than 27 bases in length. *A. thaliana* chloroplast genome matches in the sense and antisense direction were then excluded. Similarly, the remaining sequences were filtered against matches to known tRNA and rRNA sequences, which included *A. thaliana* rRNA and snoRNA entries in the SILVA database. The *N. benthamiana* raw sequences were trimmed and filtered by the service provider (Novogene).

The sequences were subsequently filtered by size to build a list of sRNAs that are between 20-24 nt in length. The 20-24 nt subset from each library were mapped to the TuMV-QLD1b genome using

previously described methods (Moyle *et al.* 2016). Mapped sequences were extracted and listed by relative abundance in Excel. Those with a relative abundance of at least 50 counts per million sequences in a given library were entered into the psRNAstarget resource (Dai and Zhao 2011) and a list of computationally predicted target gene transcripts generated using schema V1 and default settings, with the exception of clearing the extra weight in seed region option. For predicting *A. thaliana* target transcripts, the phytozome 12, 167_TAIR10 transcript library was searched. For predicting *N. benthamiana* target transcripts, transcriptome assembly v5.1 was selected.

5.2.2.4 Target Sequence Cloning

Selected predicted transcript target sequences were cloned into the dual luciferase plasmid pGrDL_SPb (Moyle *et al.* 2017) (GenBank accession KX758648.1, available from Addgene as plasmid #83205). Plasmid pGrDL_SPb allows directional cloning of target sequences in the 3'UTR of the firefly LUC expression cassette, via ligation between unique *Sal*I and *Pst*I restriction sites. Cloning was facilitated by first designing complementary 5' phosphorylated primer pairs that when annealed formed adaptors with *Sal*I and *Pst*I overhangs (Supplementary Table 1a and b). Mutated versions of the target sequence primer pairs, with 3-5introduced nucleotide changes, were also designed (Supplementary Table 1a and b). Approximately 7 nM of each primer pair were combined into a 1.5 mL microfuge tube in a volume of ~ 50-60 µL and annealed by heating to 95°C in a waterbath, followed by slowly cooling down to room temperature (at a rate of ~ 1 degree per minute). The annealed adaptors were diluted 100-fold before cloning. The pGrDL_SPb plasmid was prepared by *Sal*I & *Pst*I double restriction digestion, followed by Antarctic Phosphatase treatment (New England Biolabs). Each annealed adaptor was ligated into the prepared pGrDL_SPb plasmid using T4 DNA ligase (New England Biolabs). Each ligation was heat-shock transformed into α-Select Silver Competent Cells, following the manufacturer's instructions (Bioline), and subsequently plated on LB agar plates with 50 µg/ml kanamycin selection. After overnight incubation at 37°C, resulting colonies were grown and plasmid extracted using the PureLink Quick Plasmid Miniprep Kit (Invitrogen), following the manufacturer's instructions. Correct cloning of each adaptor was verified by Sanger sequencing using previously described primer lucqRT5' (Chou and Moyle 2014).

5.2.2.5 Agroinfiltration

Dual luciferase assays were performed using a modified transient Agroinfiltration system based on the protocol described by Moyle *et al.* (2017). Plasmids were transformed into Agrobacterium strain GV3101 and preparation of cultures to OD 0.5 for Agroinfiltration was performed as previously described (Moyle *et al.* 2017).

N. benthamiana seedlings were grown at room temperature under 400W metal halide lighting with a 16 h photoperiod. Young seedlings were mechanically inoculated with isolate TuMV-QLD1b in

sodium phosphate buffer. Seedlings were grown for approximately two weeks post-inoculation to allow systemic spread of the virus to newly emerging leaves and to allow time for symptom development. Three symptomatic expanded/expanding leaves per plant were Agroinfiltrated by applying pressure on the abaxial surface of the leaf with a disposable 5 mL syringe containing the *Agrobacterium* suspension. Typically, three plants were infiltrated with each *Agrobacterium* suspension to provide nine replicate leaves per treatment. Agroinfiltrated plants were incubated for three days in a growth chamber set to 22°C with 16 h photoperiod. Agroinfiltrated leaf tissues were harvested individually, snap-frozen in liquid nitrogen, powdered using a ball mill tissuelyser (Retsch), and stored at -80°C prior to measurement in dual LUC assays.

5.2.2.6 Dual Luciferase Assays

Dual luciferase (LUC) assay extracts were prepared using the Dual Luciferase Reporter Assay System (Promega), with the following modifications. Typically ~ 5 mg of powdered tissue was added to 100 µL of passive lysis buffer (PLB) and the cellular debris pelleted by centrifugation at 7,500 x g for 1 min. The supernatant was diluted 20-fold in PLB and 15 µL loaded into a well of a white flat bottom Costar 96 well plate (Corning). The assay was performed using a GloMax 96 microplate luminometer (Promega). The dual injectors were used to introduce 75 µL of LAR and Stop & Glo reagent, respectively, per well.

Statistical analysis of the resulting data was performed using GraphPad Prism 6 software.

5.2.3 Results

5.2.3.1 Analysis of sRNAs Complexity in Libraries Constructed from Mock or TuMV Inoculated *A. thaliana* Aerial Tissues

Small RNA libraries were sequenced to investigate the diversity of the small RNA component of the transcriptome of TuMV inoculated and mock inoculated aerial tissues of *A. thaliana*. Chloroplast and ribosomal RNA derived sequences were among the most abundant contaminant sequences filtered out from each library (Table 7). A high proportion of contaminant sequences within leafy small RNA libraries has previously been reported, including a recently published study where 90% were filtered out from a sugarcane leaf small RNA library, also in part due to a high proportion of chloroplast derived sequences (Sternes and Moyle 2015).

Table 7: Description of small RNA sequence libraries from mock and TuMV infected *A. thaliana* aerial tissue

Library	raw sequence reads	filtered reads (16-27 nt)
A. thaliana 7 day post mock inoculation	44,485,063	5,476,845
A. thaliana 14 day post mock inoculation	36,831,787	6,455,366
A. thaliana 7 day post TuMV inoculation	35,385,473	8,079,267
A. thaliana 14 day post TuMV inoculation	37,241,755	9,772,417

The size distributions of filtered reads from each library were further examined (Figure 23a). The majority of small RNAs were found to be between 21-24 nt in size, in agreement with other small RNA sequencing studies of angiosperms. The 7 and 14 day post mock inoculation libraries exhibited similar small RNA size distributions to each other, with 21 nt small RNAs the most abundant size class, followed closely by the 23 and 24 nt small RNAs (Figure 24a). The 7 and 14 days post inoculation (dpi) with TuMV libraries also displayed small RNA distributions similar to each other. However, there was a substantial shift in the small RNA distribution between the mock inoculated and TuMV inoculated libraries. The TuMV inoculated libraries contained approximately three fold more 21 nt sRNAs than 24 nt sRNAs (Figure 24a). The size distributions of the non-redundant sequences revealed there are substantially more unique 24 nt sRNAs sequences than unique 21 nt sequences (Figure 24b), indicating the 24 nt class are rich in sequence diversity whereas the 21 nt class are largely comprised of a smaller set of highly abundant sequences.

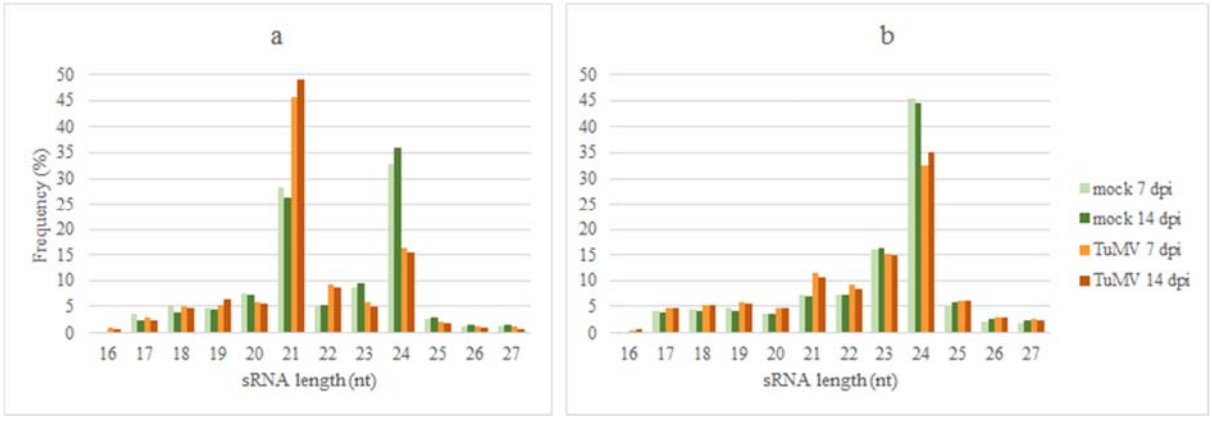


Figure 24: Analysis of small RNA profiles from TuMV and mock-inoculated *A. thaliana* aerial tissue. a Size distribution of the filtered 16-27 nt sRNAs subset. b Size distribution of the filtered non-redundant 16-27 nt sRNAs subset. The number of sequences is presented as a percentage of the total number of sequences in the 16-27 nt sRNAs subset in each library.

Reads in the range of 20-24 nt constituted 87-93% of the mapped 16-27 nt sRNAs reads and the composition and complexity of this subset was analysed further. The 20-24 nt sequences were mapped to both the *A. thaliana* and TuMV genomes. The percentage of redundant mapped 20-24 nt sequences in the 7 day and 14 day post mock inoculation libraries was 72% and 75%, respectively (Table 8). The redundancy in the TuMV inoculated libraries was higher, with 82% and 87% redundancy in the 7 day and 14 post inoculation libraries respectively (Table 8).

Table 8: Description of the 20–24 nt sequence subset from each *A. thaliana* small RNA library

Library	filtered reads (20-24 nt)	Mapped reads (20-24)	non-redundant mapped reads (20-24 nt)	% redundancy (20-24 nt)
7 day post mock inoculation (M7)	2,872,481	2,238,681	625,560	72.1%
14 day post mock inoculation (M14)	4,344,669	4,797,761	931,790	75.4%
7 day post TuMV inoculation (T7)	5,457,189	3,791,683	862,105	82 %
14 day post TuMV inoculation (T14)	7,254,814	6,830,250	878,556	87.1 %

Inspection of the 20-24 nt subset sRNAs that mapped to the *A thaliana* genome revealed a shift in the ratio of 21:24 nt size classes. There were slightly more *A. thaliana* 21 nt sRNAs than 23 or 24 nt sRNAs in the mock inoculated libraries. However, this ratio increased substantially in the TuMV inoculated libraries, predominantly due to the frequency of *A. thaliana* 24 nt sRNAs more than halving in the TuMV inoculated libraries relative to the mock inoculated libraries (Figure 25).

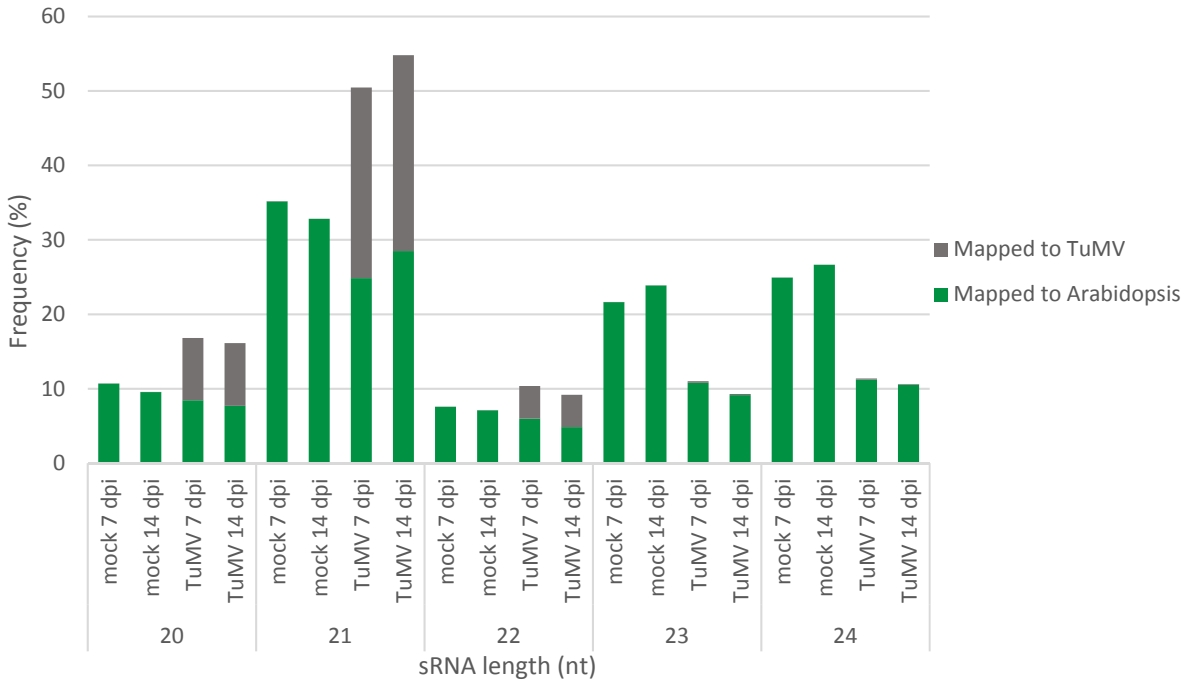


Figure 25: Analysis of the 20-24 nt subset of small RNA that map to either the *A. thaliana* or TuMV genomes. The number of sequences are presented as a percentage of the total number of sequences in the 20-24 nt sRNAs subset in each library.

5.2.3.2 Analysis of 20-24 nt TuMV Derived sRNAs in *A. thaliana*

Of the 20-24 nt subset of filtered small RNAs that mapped to the *A. thaliana* or TuMV-QLD1b genome, 39% of 7 or 14 dpi library reads mapped to the TuMV genome. A very small proportion (~0.06%) of the sRNAs from the mock inoculated libraries also mapped by chance to the TuMV genome. Of the 20-24 nt small RNA mapping to the TuMV genome, the 21 nt length class were clearly dominant, followed by the 20 and 22 nt length class (Figure 24). Approximately 50% of the 21 and 22 nt small RNA from the inoculated mapped to the virus. Depth of coverage was 100%, with minimum coverage of 49 and 50, and maximum coverage of 50,336 and 75,857 from the 7 dpi and 14 dpi libraries respectively.

Relatively few of the TuMV derived sRNAs were present as a single copy in either library (0.7% & 0.8% in the 7 dpi and 14 dpi libraries respectively), indicating the depth of sequencing of TuMV derived sRNAs was near saturation. The most abundant TuMV derived sRNAs in the 7 dpi library had a copy number of 11458; the same sRNAs was also the most abundant in the 14 dpi library with a copy number of 17278.

There was a 51:49 percent split between those sRNAs that mapped to the TuMV sense and antisense strands respectively in both the 7 dpi and 14 dpi libraries. Small RNA sequencing coverage was not distributed evenly across the TuMV genome. There were clear “hotspots” at regions where coverage

spiked due to the mapping of abundant sRNAs (Figure 26). The coverage maps generated from the 7 and 14 dpi *A .thaliana* libraries were near identical (Figure 26 A and B).

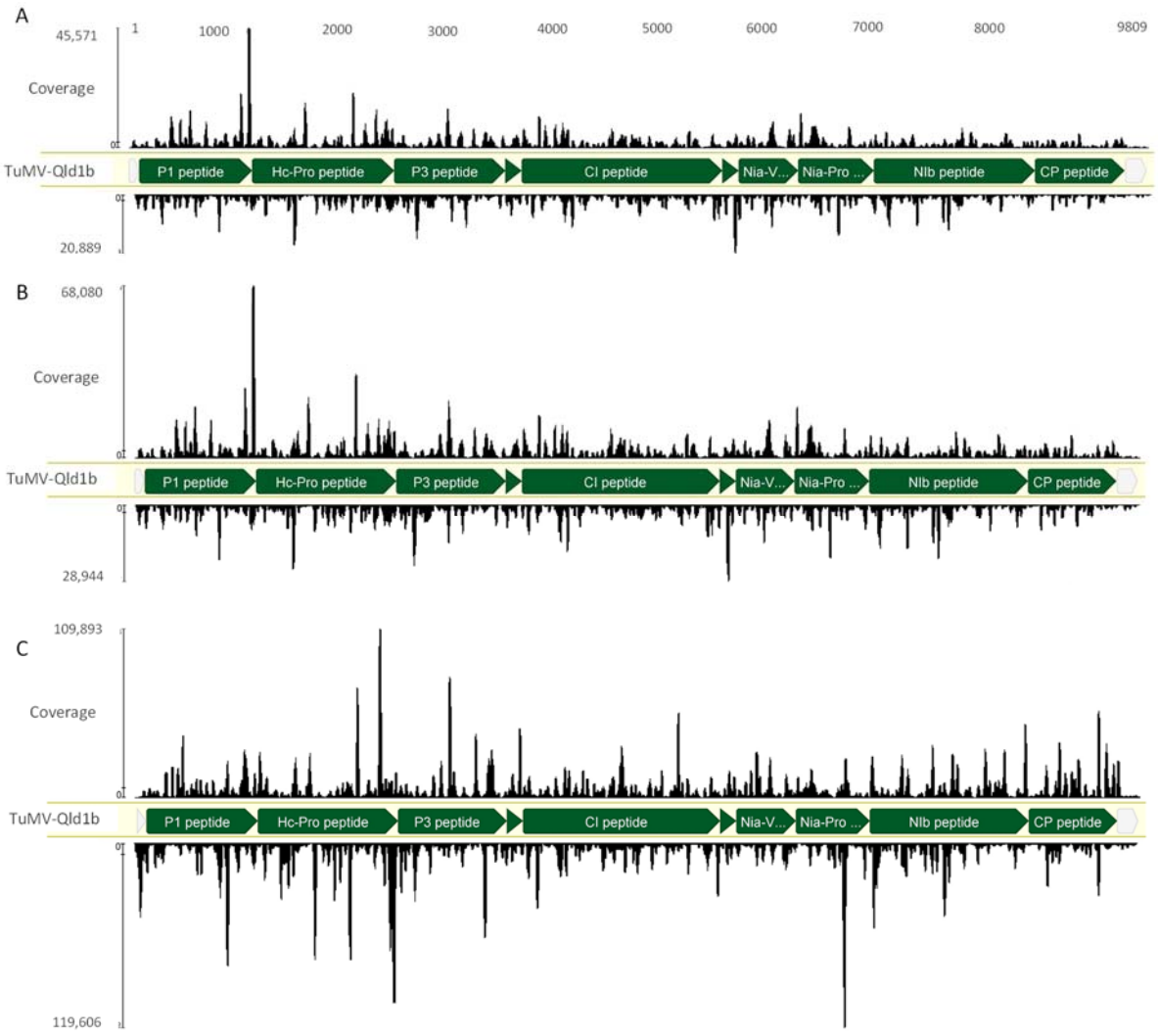


Figure 26: Coverage of sRNAs mapping in the sense and antisense direction of the TuMV-Qld1b genome.

A. coverage of sRNAs from the 7 dpi *A.thaliana* library.

B. coverage of sRNAs from the 14 dpi *A.thaliana* library.

C. coverage of sRNAs from the 14 dpi *N. benthamiana* library.

5.2.3.3 Analysis of 20-24 nt TuMV Derived sRNAs in *N. benthamiana*

For comparison, a 14 day post TuMV infection *N. benthamiana* library was also subject to deep small RNAseq analysis. However, the *N. benthamiana* library preparation and filtering of contaminant sequences was undertaken using the service provider's in house pipeline (Novogene). Of the 20-24 nt subset of filtered small RNAs obtained, 43% mapped to the TuMV-QLD1b genome (Table 9). Relatively few of the TuMV derived sRNAs were present as a single copy (1.3%) indicating the depth of sequencing of TuMV derived sRNAs was near saturation. The 21 nt length class were clearly dominant, followed by the 22 nt and 20 nt length class (Figure 27).

Table 9: Description of the *N. benthamiana* TuMV infected sRNAs library

Library	number of raw sequence reads	number of filtered reads (16-27 nt)	number of filtered reads (20-24 nt)	reads mapped to TuMV (20-24 nt)
N. benthamiana 14 day post TuMV inoculation	18,973,821	17,176,941	14,457,979	6,235,908

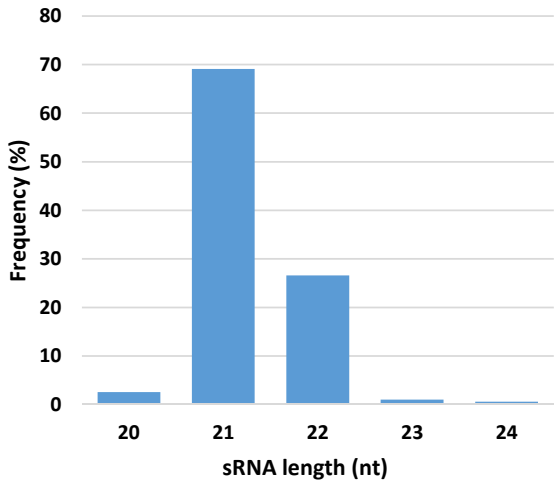


Figure 27: Analysis of the *N. benthamiana* library 20-24 nt subset of small RNA that map to the TuMV genome. The number of sequences are presented as a percentage of the total number of 20-24 nt sRNAs sequences.

Depth of coverage of the TuMV genome was 100%, with minimum coverage of 204, and maximum coverage of 135,128. The most abundant TuMV derived sRNAs from the *N. benthamiana* library had a copy number of 29622. There was a 49:51 percent split between those sRNAs that mapped to the TuMV sense and antisense strands respectively. Small RNA sequencing coverage was not distributed evenly across the TuMV genome. There were clear "hotspots" at regions where coverage spiked due to the mapping of abundant sRNAs (Figure 26). Although there were some hotspots in common

between the mapping of reads derived in the *A. thaliana* and *N. benthamiana* hosts, the overall TuMV derived small RNA mapping pattern was different between the host plants.

5.2.3.4 Prediction of Viral Derived sRNAs with High Complementarity to Host Transcript Target Sequences

Recent studies have predicted and experimentally proved that viral derived small RNA can target host genes. Here we investigated which plant transcripts contain sequences with high complementarity to TuMV viRNAs sequences and are therefore possible targets of the virus derived viRNAs. The most abundant viral derived 20-24nt sRNAs (>50 copies per million) were run through the free-to-use web based platform psRNAstarget (Dai and Zhao 2011), to generate a list of predicted transcript targets from *A. thaliana* and *N. benthamiana* highly complementary targets were selected for validation testing using a dual luciferase reporter assay.

5.2.3.5 Validation of Predicted TuMV viRNAs Interaction with Host Transcript Target Sequences

Recent studies coupled the transient nature of *Agrobacterium* mediated transformation of *N. benthamiana* leaves with the quantitative dual LUC report system to validate microRNA-target sequence interactions (Liu and Axtell 2015). The quantitative and transient nature of the assay provides an important advantage over other qualitative reporter assays. Briefly, the assay involves cloning predicted target transcripts sequences directly 3' of a firefly luciferase coding sequence within a dual LUC T-DNA cassette (Figure 28). The dual LUC plasmid is co-Agroinfiltrated with a miRNA precursor expression plasmid. The firefly luciferase is used to report miRNA directed repression, while the Renilla luciferase expression cassette is used as an internal standard to normalise for Agroinfiltration and leaf to leaf variability.

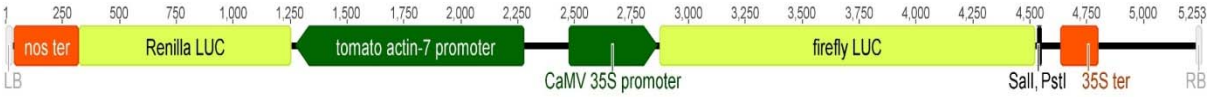


Figure 28: Figure from Moyle, Carvalhais *et al.* (2017). The pGrDL_SPb plasmid is used as a quantitative reporter system for validating possible viRNAs target transcripts via transient expression. The host target transcript in cloned into the 3' of the firefly LUC which has a CaMV 35S promoter and terminator. The Renilla LUC is expressed using the tomato *ACTIN7* promoter and *NOS* terminator. LB and RB stand for left board and right board, respectively.

We modified this dual luciferase assay to assess viRNAs-target sequence interaction. We cloned the predicted transcript target sequence directly 3' of the firefly luciferase coding sequence, but instead of co-Agroinfiltrating with a miRNA precursor, we simply infiltrated TuMV infected *N. benthamiana* leaves for the source of the TuMV viRNAs. As a control, we also infiltrated a dual LUC plasmid containing a mutated version of the predicted transcript target sequence. The dual LUC ratio generated from the plasmid containing the predicted target sequence was compared to the dual LUC ratio generated from the plasmid containing the mutated target sequence. A reduced dual LUC ratio from the target sequence, relative the mutated target sequence, indicated an interaction between the complementary viRNAs. An additional control where non-infected *N. benthamiana* plant leaves were Agroinfiltrated with the target or mutated target plasmids was included to ensure the sequence variations were not responsible for any observed difference in dual LUC ratios.

Systemically TuMV infected *N. benthamiana* leaves were agro-infiltrated approximately 2 weeks after inoculation. The dual LUC assay was performed on 10 predicted *N. benthamiana* and 10 *A. thaliana* transcript target sequences and their mutated versions. Only one *N. benthamiana* transcript, Niben3160g02007, was experimentally validated as a target of TuMV viRNAs (Figure 29A). The target sequence contains 19 perfectly complementary bases to the TuMV viRNAs, 2 wobble bases, and 1 mismatch (Figure 28B). The function of the protein is unknown, nor does it contain any known conserved protein domains.

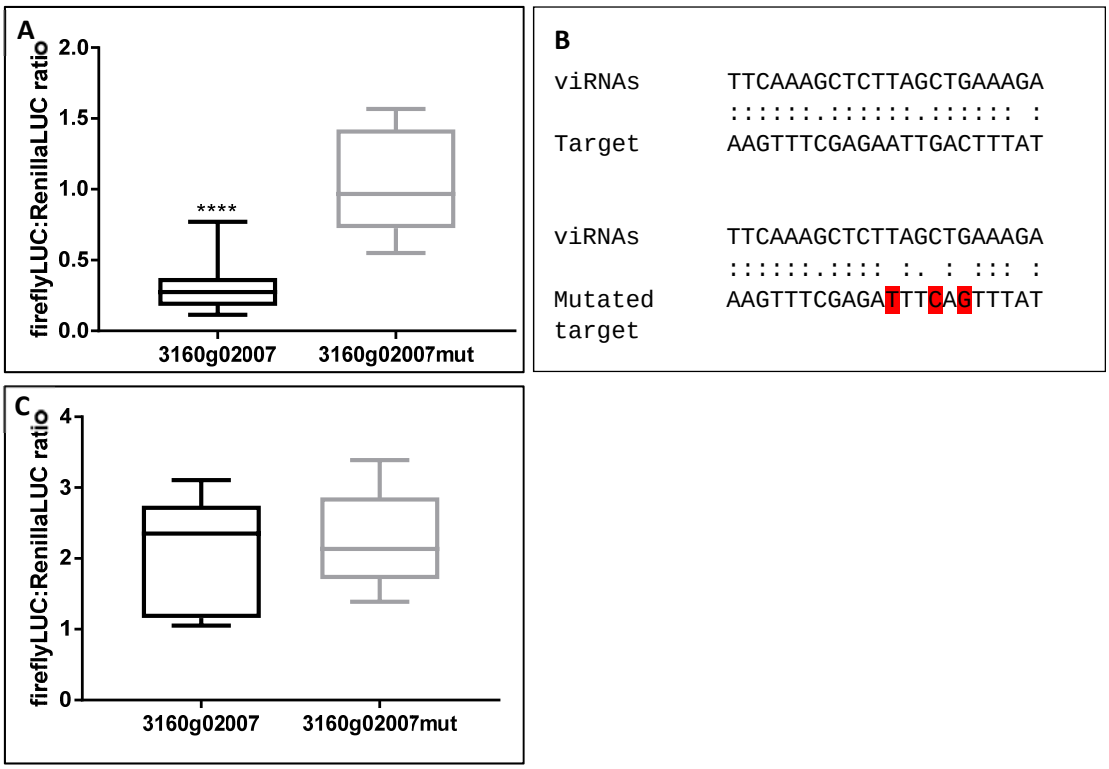


Figure 29: Small viral derived RNAs are able to target certain host gene transcripts. A. There is a clear knockout of the gene when agro-infiltrated in TuMV infected *N. benthamiana*. When the host gene target is mutated at bases 6, 8 and 11 (highlighted in red) the expression is restored. B. The alignment of the viRNAs and target reveal 2 wobble bases and a mismatch. The mutated target introduces 3 more mismatches to prevent binding. C. When non-infected plants are agro-infiltrated with the same target and mutated target plasmids no knock-down effect is evident.

5.2.4 Discussion

Few examples have been reported of virus or viroid derived sRNAs that negatively regulate a host transcript, in turn causing symptoms associated with disease (Shimura *et al.* 2011, Smith *et al.* 2011, Navarro *et al.* 2012). While there are an increasing number of studies predicting viral small RNA interaction with host transcript target sequences, the lack of a simple, robust, quantitative method has prevented experimental validation of interactions. To the best of our knowledge, there are no reports in the literature of the small RNA profile generated by TuMV infected *A. thaliana* or *N. benthamiana*, and no prediction or validation interactions between TuMV-derived small RNA and host transcripts. In this study we used deep small RNA sequencing to investigate the small RNA profile generated in TuMV infected *A. thaliana* and *N. benthamiana*. We predict which host transcripts may be targeted by the most abundant TuMV derived small RNA and apply a modified transient quantitative dual luciferase assay to validate predicted viRNAs-target sequence interactions.

In mock inoculated *A. thaliana* libraries, there were slightly more 24 nt small RNAs than 21 nt small RNAs. This is consistent with other studies were various other RNA viruses have produced a similar shift (Xia *et al.* 2014, Li *et al.* 2016, Margaria *et al.* 2016, Qiu *et al.* 2017). However, in TuMV infected libraries, there were three times as many 21 nt sRNAs than 24 nt sRNAs. This is consistent with reported studies where other RNA virus infected libraries cause a shift in the small RNA distribution to increase the ratio of 21 nt small RNA to 24 nt small RNA (Xia *et al.* 2014, Li *et al.* 2016, Margaria *et al.* 2016, Qiu *et al.* 2017).

The 20-24 nt subset of each library was mapped to the *A. thaliana* genome. The TuMV inoculated libraries mapped a significantly lower proportion of 23/24 nt small RNAs to the *A. thaliana* genome than the mock inoculated libraries. Comparable results have been reported for various other RNA viruses (Kreuze *et al.* 2009, Ruiz-Ruiz *et al.* 2011, Herranz *et al.* 2015). Similar results were found when plants were inoculated with an active potyviral HC-Pro, a known viral suppressor of RNA silencing (VSR), where decreased levels of endogenous siRNAs were evident (Garcia-Ruiz *et al.* 2015). It is hypothesised that the decreased frequency of the endogenous siRNAs could be related to several factors. Firstly, a study using the VSR P25 from the Potato virus X found that it interacted with a number of AGO proteins including AGO4 which binds 24 nt siRNAs (Chiu *et al.* 2010). Another study found a similar result with the CMV 2b VSR which bound to 24 nt siRNAs and inhibited AGO4 function (Hamera *et al.* 2012). It is possible the decreased proportion of 23/24 nt sRNAs in the TuMV infected *A. thaliana* libraries could due to a VSR (either HC-Pro or Vpg) disrupting this siRNAs biogenesis pathway. Potyviral Vpg has been extensively studied and found to interact with many host proteins such as eukaryotic initiation factor 4E (eIF4E) and its isoform eIF(iso)4E (Lellis *et al.* 2002), eIF4G and its isomer eIF(iso)4G (Michon *et al.* 2006), and more recently RNA silencing factors RdRp6 and Suppressor of Gene Silencing 3 (SGS3) (Cheng and Wang 2017). Interestingly, its interaction with RdRp6 could be the link to the decrease in 24 nt siRNAs as this host factor is a crucial part of this biogenesis pathway and considered to have a role in regulating the noncoding RNAs as well as DNA methylation (Brosnan *et al.* 2007, Zhang *et al.* 2014).

The small RNA libraries were also mapped to the TuMV genome. The 20-22nt viRNAs were overwhelmingly the most abundant size classes mapping to the TuMV genome (Figure 25). A similar size distribution was evident after mapping small RNA generated from TuMV infected *N. benthamiana* (Figure 26). This result is consistent with distributions generated from other RNA viruses, where viRNAs are overwhelmingly 20-22 nt in size and only a very low proportion of 23/24 nt map to the virus genome (Donaire *et al.* 2009, Qi *et al.* 2009, Garcia-Ruiz *et al.* 2010, Li *et al.* 2012). The relative absence of 24 nt viRNAs is thought to be due to compartmentalization. If RNA viruses do not enter the nucleus then they cannot be acted upon by DCL3, which produces 24 nt siRNAs in the nucleus (Pontes *et al.* 2006, Pretorius *et al.* 2017). As RNA viruses replicate in the

cytoplasm, the 21/22 nt viRNAs are predominately generated by DCL4 and DCL2 respectively. Processing foreign RNAs by DLC4 and DCL2 are central to plant viral defences (Parent *et al.* 2015). Interestingly, DCL2 was found to have a role in intracellular post-transcriptional gene silencing with 22 nt viRNAs possibly being signalling molecules priming neighbouring cells against subsequent viral invasion (Qin *et al.* 2017). This study also suggested that DCL4 inhibits DCL2 action and DCL2 intercellular defence is considered a second line of defence if DCL4 is inhibited by a viral suppressor of RNA silencing. The viRNAs generated from TuMV support this theory, as the 21 nt viRNAs was the most abundant size fraction.

The coverage of viRNAs mapping to the TuMV genome was near identical at 7 and 14 dpi, in *A. thaliana* (Figure 26 A and B). This suggests the processing of the viRNAs did not change over time. The coverage pattern revealed several "hotspot" regions where abundant viRNAs were generated. It has been proposed that hotspots of the most abundant viRNAs are possibly link to GC-richness (Ho *et al.* 2007), or secondary structures created by viral single stranded RNA (Molnár *et al.* 2005). The GC richness theory is based on the preference of DCL enzymes to target regions of increased GC content (Ho *et al.* 2007). This study also suggested that GC bias of DCL enzymes is highly conserved in dicots and considered an ancient mechanism. Interestingly, the proportion of viRNAs mapping to the sense and antisense strands was nearly 50:50. Near 50:50 ratios of viRNAs mapping to the sense or antisense strand have been reported for some RNA viruses (Silva *et al.* 2011, Xia *et al.* 2014, Xia *et al.* 2016). This suggests processing of dsRNAs replication intermediates may be the main source of viRNAs, rather than secondary hairpin structures formed by folding within the genomic RNA strand. Alternative dsRNAs templates could possibly be produced through the RdRp dependent pathway which processes secondary siRNAs (Axtell *et al.* 2006, Diaz-Pendon *et al.* 2007, Qi *et al.* 2009, Chen *et al.* 2010, Baulcombe 2015, Borges and Martienssen 2015).

The coverage map of viRNAs generated in *N. benthamiana* did not match the coverage map generated in the *A. thaliana* host (Figure 26 C). This suggests the processing of the virus can be host dependent. This result is contrary to previous studies with other viruses where the profiling of viRNAs from different hosts had the same mapping pattern (Xu *et al.* 2012, Mitter *et al.* 2013).

RNA silencing is the plant adaptive antiviral defence mechanism. Foreign viral RNA is targeted and processed by various RNA silencing factors resulting in abundant viRNAs. However, the abundance and frequency of viRNAs being generated by the plant could have unwanted secondary effects. If the plant is unable to distinguish plant siRNAs from viRNAs it is possible that these viRNAs could be used in the RNA silencing pathways targeting complementary host transcripts leading to the

unwanted regulation of host genes. These off target-effects have been linked to phenotypic symptoms typically associated with viral infections. For example viRNAs derived from certain viroids and viral satellites have been linked to the yellowing mottled symptoms associated with certain viral infection due to the targeting of host genes (Shimura *et al.* 2011, Smith *et al.* 2011, Navarro *et al.* 2012). A number of studies have implicated viRNAs in gene silencing through predictions, though no experimental validation has been completed to confirm whether the viRNAs do target the predicted host transcripts (Angell and Baulcombe 1997, Xia *et al.* 2016). Recently, there have been an increasing number of reports of experimentally validated viRNAs targeting host transcripts (Qi *et al.* 2009, Miozzi *et al.* 2013, Avina-Padilla *et al.* 2015, Wang *et al.* 2016, Moyo *et al.* 2017).
Ramesh *et al.* (2017) predicted and experimentally validated several host transcript targets of TSWV viRNAs. The study used quantitative real-time PCR (qRT-PCR) and rapid amplification of cDNA ends (RACE) to confirm down regulation by cleavage of the targeted transcripts. Using different tomato varieties, they proposed that resistance was possibly due to presence of the NBS-LRR resistance genes which may prevent the off-target silencing of viRNAs. Similarly, a paper using *Cotton leaf curl virus* also predicted and validated viRNAs host targets through qRT-PCR and RACE (Wang *et al.* 2016). The off-target effects of viRNAs have also been shown phenotypically. Using the pathosystem *Rice stripe virus* infected *N. benthamiana*, researchers found that chlorosis and twisting of leaf symptoms were linked to the targeting and down regulation of specific host genes. The idea that viRNAs are linked to phenotypic effects and are possible pathogenicity determinants suggests identification as well as validation is important to fully understand virus-host interactions.

We used the psRNAstarget web-based platform to predict a number of possible viRNAs-host transcript targets (Supplementary Table 1a and b). To validate the predicted interactions, we modified a transient quantitative dual luciferase reporter assay that was originally developed to report miRNA-target sequence interaction. Here, we cloned predicted transcript target sequences or mutated versions, into the dual luciferase cloning vector. Comparison of the dual luciferase assay ratios generated from the target sequence vector versus the mutated target sequence vector in TuMV infected *N. benthamiana* was used to report if the complementary TuMV viRNAs was capable of interacting with the predicted target sequence. Out of ten predicted TuMV viRNAs-transcript target sequence pairings from each of the *A. thaliana* and *N. benthamiana* plant hosts, only one predicted transcript target sequence generated a repeatable knockdown of expression relative to the mutated target sequence control. A similarly low validation rate of predicted targets was found in SCMV infected maize using qRT PCR as a validation method (Xia *et al.* 2014). Northern analysis of predicted target transcripts of SCMV viRNAs showed that most predicted targets were not down-regulated in virus infected samples, relative to mock-inoculated controls, inferring many factors

might affect the functionality of viRNAs and restrict their potential to regulate host transcripts *in vivo* (Xia *et al.* 2014).

Past research has suggested that siRNAs incorporation into AGO proteins to form the RISC is a specific interaction with the 5'-terminal nucleotide being a strong indicator as to which AGO protein will interact with which siRNAs (Mi *et al.* 2008, Takeda *et al.* 2008). It is unclear whether the viRNAs generated in the host can be incorporated into all AGO proteins to form an active RISC (Zhu and Guo 2012). Recent findings by Schuck *et al.* (2013) has shown the viRNAs are able to be incorporated and guide cleavage of most AGO proteins. The cleavage patterns observed suggested that certain parts of the viral genome were more accessible for targeting and only certain viRNAs were efficient at guiding cleavage (Schuck *et al.* 2013). Similarly, a degradome analysis reported that only a small percentage of viRNAs were active in antiviral defence and that hotspots did not correlate with identified cleavage sites (Miozzi *et al.* 2013). We speculate that if not all viRNAs are incorporated into AGO proteins and form efficient RISC molecules to target the virus genome, then presumably not all viRNAs can target complementary host transcripts either. This could explain the low percentage of predicted viRNAs-target sequence interactions that were experimentally validated by the dual LUC reporter assay.

The transcript sequence that was targeted in TuMV infected *N. benthamiana* originates from a gene of unknown function. A BLASTP search using the protein coding sequenced failed to detect any putative conserved domains. Further work is required to determine if TuMV viRNAs targeting of this gene transcript contributes to disease symptoms in *N. benthamiana*.

Despite using 9 reps per treatment, the assay was more variable than the related dual LUC assay used to report miRNA-target sequence interaction. A number of other predicted viRNAs-target sequence interactions initially produced a measurable knockdown. However, upon replication of the experiment, the knockdown effect was lost. Therefore it is possible that variability within the assay parameters may mean the dual LUC assay is not the best methodology for detecting viRNAs-target sequence interactions. It is possible further optimisation of the assay parameters may help to reduce the variation. Future work should involve a direct comparison between 5'RACE and the dual LUC assay for detection of viRNAs-target sequence interactions. The differential processing of the TuMV in the different host plant species may hinder the use of the *N. benthamiana* based assay to validate viRNAs-target pairings predicted in other host species. The impact that the reduction in host derived 24 nt sRNAs identified in this study has on methylation patterns could also be the subject of future investigation. The deep small RNA sequencing dataset could also be applied to analysing the impact TuMV infection has on endogenous small RNA populations, including the various microRNAs and their targets.

Supplementary Table 1a: Detailing Alignment of the Host Transcripts and the viRNAs

List of *A. thaliana* genes which were chosen as targets due to high sequence complementarity to an abundant viRNAs. The table shows the gene name, the alignment for the target to the viRNAs and the adapters ordered to validate the interaction. Letters in red are mutated bases and the parts of the adapters that are underline are the cloning sites.

Gene	Alignment	Adapters
AT2G40000	Target ACGATGCGTTTTGTTTCAGA viRNAs TGCTACGCTAAACGGAGTCT	**AT2G40000_F** TCGACACGATGCGTTTTGTTTCAGACTGCA **AT2G40000_R** GTCTGAAACAAAACGCATCGTG
AT5G47370	Target AAAGAGTCCTAACTTCGAGTT viRNAs TTTCTCACGATTGTAGCTCAA	**AT5G47370_F** TCGACACGATGCGTTTTGTTTCAGACTGCA **AT5G47370_R** GTCTGAAACAAAACGCATCGTG
AT1G21920	Target GAATCGAAAGTTTGTTTCAGA viRNAs CTTAGTTTTCACATAAAGTCT	**AT1G21920_F** TCGACGAATCGAAAGTTTGTTTCAGACTGCA **AT1G21920_R** GTCTGAAACAAACTTTCGATTCG
AT1G21920 Mutated 1	Target GAATCGAAAGATTGTTTGTCT viRNAs CTTAGTTTTCACATAAAGTCT	**AT1G21920_F_mut1** TCGACGAATCGAAAGATTGTTTGTCTCTGCA **AT1G21920_R_mut1** GTCTGAAACAAACTTTCGATTCG
AT1G21920 mutated 2	Target CATTGGATACTTTGTATGACA viRNAs CTTAGTTTTCACATAAAGTCT	**AT1G21920_F_mut2** TCGACCATTGGATACTTTGTATGACACTGCA **AT1G21920_R_mut2** GTGTCATACAAAGTATCCAATGG
AT1G21920 mutated 3	Target TTTGACTTAGTTTGTTTCAGA viRNAs CTTAGTTTTCACATAAAGTCT	**AT1G21920_F_mut3** TCGACTTTGACTTAGTTTGTTTCAGACTGCA **AT1G21920_R_mut3** GTCTGAAACAAACTAAGTCAAAG
AT1G21920 mutated 4	Target GAATCGAAAGTTTTAAGACAC viRNAs CTTAGTTTTCACATAAAGTCT	**AT1G21920_F_mut4** TCGACGAATCGAAAGTTTTAAGACACTGCA **AT1G21920_R_mut4** GTGTCTTAAAACTTTCGATTCG
AT1G21920 Perfect Match	Target GAATCAAAAGTGTATTTCAGA viRNAs CTTAGTTTTCACATAAAGTCT	**AT1G21920_F_PM** TCGACGAATCAAAAGTGTATTTCAGACTGCA **AT1G21920_R_PM** GTCTGAAATACACTTTTGATTCG
AT5G39660	Target AGTCGTTTCTCTCTCTCTTT viRNAs TCAGCGCGGAGAGAGAGAAA	**AT5G39660_F** TCGACAAGTCGTTTCTCTCTCTCTTTCTGCA **AT5G39660_R** GAAAGAGAGAGAGAAACGACTTG
AT5G39660 Mutated 1	Target Mut1 AGTCGTTTCCATCTCAGATT viRNAs TCAGCGCGGAGAGAGAGAAA	**AT5G39660_F_mut1** TCGACAAGTCGTTTCCATCTCAGATTCTGCA **AT5G39660_R_mut1** GAATCTGAGATGGAAACGACTTG
AT5G39660 Mutated 2	Target Mut2 ACTGGTTTGTCTGTGAGTAT viRNAs TCAGCGCGGAGAGAGAGAAA	**AT5G39660_F_mut2** TCGACACTGGTTTGTCTGTGAGTATCTGCA **AT5G39660_R_mut2** GATACTCACAGACAAACCAGTG
AT5G39660 Mutated 3	Target Mut3 CTCAATTTCTCTCTCTCTTT viRNAs UCAGCGCGGAGAGAGAGAAA	**AT5G39660_F_mut3** TCGACCTCAATTTCTCTCTCTCTTTCTGCA **AT5G39660_R_mut3** GAAAGAGAGAGAGAAATTGAGG
AT5G39660 Mutated 4	Target Mut4 AGTCGTTTCTCTACAGGAGG viRNAs TCAGCGCGGAGAGAGAGAAA	**AT5G39660_F_mut4** TCGACAGTCGTTTCTCTACAGGAGGCTGCA **AT5G39660_R_mut4** GCCTCCTGTAGAGAAACGACTG
AT5G39660 Perfect Match	Target PM AGTCGCGCCTCTCTCTCTTT viRNAs TCAGCGCGGAGAGAGAGAAA	**AT5G39660_F_PM** TCGACAGTCGCGCCTCTCTCTCTTTCTGCA **AT5G39660_R_PM** GAAAGAGAGAGAGGCGCGACTG
AT4G00660	Target GATCTCTCTCTGTCTCGCTTTCT viRNAs CGAGAGAGAGACA-AGCGAAGGA	**AT4G00660_F** TCGACGATCTCTCTCTGTCTCGCTTTCTCTGCA **AT4G00660_R** GAGAAAGCGAGACAGAGAGAGATCG

AT4G00660 Mutated	Target GATCTCTCTCTCTCTGGGTTTCT viRNAs CGAGAGAGAGACA-AGCGAAGGA	**AT4G00660mut_F** TCGACGATCTCTCTCTCTCTGGGTTTCTCTGCA **AT4G00660mut_R** GAGAAACCCAGAGAGAGAGAGATCG
AT1G58250	Target GTTCTTCTTCGGGTTTCTCAT viRNAs AGAGAAGAAGTTCAAAGAGTA	**AT1G58250_F** TCGACGTTCTTCTTCGGGTTTCTCATCTGCA **AT1G58250_R** GATGAGAAACCCGAAGAAGAACG
AT1G58250 Mutated	Target GTTCTTCTTCTGGATGCTCAT viRNAs AGAGAAGAAGTTCAAAGAGTA	**AT1G58250mut_F** TCGACGTTCTTCTTCTGGATGCTCATCTGCA **AT1G58250mut_R** GATGAGCATCCAGAAGAAGAACG
AT5G25780	Target ATGAGTTTTCAATTGGAGGAC viRNAs CACTCAAGAGTTAACCTCCTC	**AT5G25780_F** TCGACATGAGTTTTCAATTGGAGGACCTGCA **AT5G25780_R** GGTCCTCCAATTGAAAACTCATG
AT5G25780 Mutated	Target ATGAGTTTTCTATAGCAGGAC viRNAs CACTCAAGAGTTAACCTCCTC	**AT5G25780mut_F** TCGACATGAGTTTTCTATAGCAGGACCTGCA **AT5G25780mut_R** GGTCCTGCTATAGAAAACTCATG
AT3G45140	Target CACCTCACTCATTACTTGGGA viRNAs CTGGAGTGAGTCATGAACCTT	**AT3G45140_F** TCGACCACCTCACTCATTACTTGGGACTGCA **AT3G45140_R** GTCCCAAGTAATGAGTGAGGTGG
AT3G45140 Mutated	Target CACCTCACTCTTTTCATGGGA viRNAs CTGGAGTGAGTCATGAACCTT	**AT3G45140mut_F** TCGACCACCTCACTCTTTTCATGGGACTGCA **AT3G45140mut_R** GTCCCATGAAAAGAGTGAGGTGG
AT5G42810	Target ACTTAAAATTTGTTGGAACTA viRNAs GGAGTTTCAAACAACCTTGAT	**AT5G42810_F** TCGACACTTAAAATTTGTTGGAACTACTGCA **AT5G42810_R** GTAGTTCCAACAAATTTTAAGTG
AT5G42810 Mutated	Target ACTTAAAATTGGTAGCAACTA viRNAs GGAGTTTCAAACAACCTTGAT	**AT5G42810mut_F** TCGACACTTAAAATTGGTAGCAACTACTGCA **AT5G42810mut_R** GTAGTTGCTACCAATTTTAAGTG
AT3G55610	Target ACCATCATAAGAAGATCTCAT viRNAs CGGTAGTGTTCTTCTGGAGTG	**AT3G55610_F** TCGACACCATCATAAGAAGATCTCATCTGCA **AT3G55610_R** GATGAGATCTTCTTATGATGGTG
AT3G55610 Mutated	Target ACCATCATAACAACAGCTCAT viRNAs CGGTAGTGTTCTTCTGGAGTG	**AT3G55610mut_F** TCGACACCATCATAACAACAGCTCATCTGCA **AT3G55610mut_R** GATGAGCTGTTGTTATGATGGTG
AT3G55610 Mutated	Target ACCATCATAACAACAGCTCAT viRNAs CGGTAGTGTTCTTCTGGAGTG	**AT3G55610mut_F** TCGACACCATCATAACAACAGCTCATCTGCA **AT3G55610mut_R** GATGAGCTGTTGTTATGATGGTG
AT4G37460	Target CTGCTAATATGGAGTTCACTG viRNAs CACGATTGTAGCTCAAGTGAC	**AT4G37460_F** TCGACCTGCTAATATGGAGTTCACTGCTGCA **AT4G37460_R** GCAGTGAACTCCATATTAGCAGG
AT4G37460 Mutated	Target CTGCTAATATGCACTACACTG viRNAs CACGATTGTAGCTCAAGTGAC	**AT4G37460mut_F** TCGACCTGCTAATATGCACTACACTGCTGCA **AT4G37460mut_R** GCAGTGTAGTGCATATTAGCAGG
AT4G34710	Target TTTTTCAGTTTTTGCTCTGT viRNAs AAGAAGTTAAGAACGAGACA	**AT4G34710_F** TCGACTTTTTCAGTTTTTGCTCTGTCTGCA **AT4G34710_R** GACAGAGCAAAAACTGAAAAAG
AT4G34710 Mutated	Target TTTTTCAGTGTTGGGTCTGT viRNAs AAGAAGTTAAGAACGAGACA	**AT4G34710mut_F** TCGACTTTTTCAGTGTTGGGTCTGTCTGCA **AT4G34710mut_R** GACAGACCCAACACTGAAAAAG

Supplementary Table 1b: Detailing Alignment of the Host Transcripts and the viRNAs

List of *N. benthamiana* genes which were chosen as target due to high sequence complementarityF to an abundant viRNAs. The table shows the gene name, the alignment for the target and the viRNAs and the adapter primers ordered to validate the interaction. Letters in red are mutated bases and the parts of the adapters that are underline are the cloning sites.

Gene	Alignment	Adapters
Niben101Scf03979g02010	Target 5' TTTTAGTTTCGGACTGACTAT 3' viRNAs 3' AAATTCAAAGCCTGACTGATG 5'	**3979g02010_F** TCGACTTTTAGTTTCGGACTGACTATCTGCA **3979g02010_R** GATAGTCAGTCCGAAACTAAAAG
Niben101Scf03979g02010 Mutated	Target 5' TTTTAGTTTCCGAGTCACTAT 3' viRNAs 3' AAATTCAAAGCCTGACTGATG 5'	**3979g02010_mutF** TCGACTTTTAGTTTCCGAGTCACTATCTGCA **3979g02010_mutR** GATAGTGACTCGGAAACTAAAAG
Niben101Scf02615g00003	Target 5' GTGTTCCTAGAGGAGTTCATGT 3' viRNAs 3' CA-AAGGATCTCCTCAAGTACG 5'	**2615g00003_F** TCGACGTGTTCCTAGAGGAGTTCATGTCTGCA **2615g00003_R** GACATGAACTCCTCTAGGAACACG
Niben101Scf02615g00003 Mutated	Target 5' GTGTTCCTAGACGACTACATGT 3' viRNAs 3' CA-AAGGATCTCCTCAAGTACG 5'	**2615g00003_mutF** TCGACGTGTTCCTAGACGACTACATGTCTGCA **2615g00003_mutR** GACATGTAGTCGTCTAGGAACACG
Niben101Scf02360g04001	Target 5' ATACATTTAGAAGGTCTTTAG 3' viRNAs 3' TATGTAGATCTTCCAGAAATT 5'	**2360g04001_F** TCGACATACATTTAGAAGGTCTTTAGCTGCA **2360g04001_R** GCTAAAGACCTTCTAAATGTATG
Niben101Scf02360g04001 Mutated	Target 5' ATACATTTAGTAGCTGTTTAG 3' viRNAs 3' TATGTAGATCTTCCAGAAATT 5'	**2360g04001_mutF** TCGACATACATTTAGTAGCTGTTTAGCTGCA **2360g04001_mutR** GCTAAACAGCTACTAAATGTATG
Niben101Scf03160g02007	Target 5' AAGTTTCGAGAATTGACTTTAT 3' viRNAs 3' TTCAAAGCTCTTAGCTGAAAGA 5'	**3160g02007_F** TCGACAAGTTTCGAGAATTGACTTTATCTGCA **3160g02007_R** GATAAAGTCAATTCTCGAAACTTG
Niben101Scf03160g02007 Mutated	Target 5' AAGTTTCGAGATTTCAGTTTAT 3' viRNAs 3' TTCAAAGCTCTTAGCTGAAAGA 5'	**3160g02007**_mutF TCGACAAGTTTCGAGATTTCAGTTTATCTGCA **3160g02007_mutR** GATAAACTGAAATCTCGAAACTTG
Niben101Scf00538g05002	Target 5' AATTCTCTACACCCTGCACT 3' viRNAs 3' TTTAGAGATGTGGGACGTGA 5'	**0538g05002_F** TCGACAATTCTCTACACCCTGCACTCCTGCA **0538g05002_R** GGAGTGCAGGGTGTAGAGAATTG
Niben101Scf00538g05002 Muated	Target 5' AATTCTCTACTCCGTCCACTC 3' viRNAs 3' TTTAGAGATGTGGGACGTGAT 5'	**0538g05002_mutF** TCGACAATTCTCTACTCCGTCCACTCCTGCA **0538g05002_mutR** GGAGTGGACGGAGTAGAGAATTG
Niben101Scf03450g02025	Target 5' CATCCAGTCTGAACAAATTTT 3' viRNAs 3' GTAGGTCAGGCTTGTTTAACA 5'	**3450g02025_F** TCGACCATCCAGTCTGAACAAATTTTCTGCA **3450g02025_R** GAAAATTTGTTCAGACTGGATGG
Niben101Scf03450g02025 Mutated	Target 5' CATCTAGGCTCAACATATTTT 3' viRNAs 3' GTAGGTCAGGCTTGTTTAACA 5'	**3450g02025_mutF** TCGACCATCTAGGCTCAACATATTTTCTGCA **3450g02025_mutR** GAAAATATGTTGAGCCTAGATGGG

Niben101Scf08470g06007	Target TTTCTTGTGTAACTTATATGATTT | ||||||| |||||| |||| || viRNAs ACAGAACAC-TTGAAT- TACTCAA	**8470g06007_F** TCGACTTTCTTGTGTAACTTATATGATTTCTGCA **8470g06007_R** GAAATCATATAAGTTACACAAGAAAG
Niben101Scf08470g06007 Mutated	Target TTTCATGTCTAAGTTCTAAGATTT | || ||| || || | || || viRNAs ACAGAACAC-TTGAAT- TACTCAA	**8470g06007_mutF** TCGACTTTCATGTCTAAGTTCTAAGATTTCTGCA **8470g06007_mutR** GAAATCTTAGAACTTAGACATGAAAG
Niben101Scf04506g02002	Target CCACTTGACTTGTCGTAAGAATG ||||||||||||| ||| ||||| viRNAs GGTGAACTGAACA-CAT-CTTAC	**4506g02002_F** TCGACCCACTTGACTTGTCGTAAGAATGCTGCA **4506g02002_R** GCATTCTTACGACAAGTCAAGTGGG
Niben101Scf04506g02002 Mutated	Target CCAGTTCACATGACGTTAGAATG ||| || || || || ||||| viRNAs GGTGAACTGAACA-CAT-CTTAC	**4506g02002_mutF** TCGACCCAGTTCACATGACGTTAGAATGCTGCA **4506g02002_mutR** GCATTCTAACGTCATGTGAACTGGG
Niben101Scf01414g02022	Target TTGGGGTAGAGTTTTGCTGAGT | ||| |||||||| |||||| viRNAs AGCTCTATCTCAAAA-GACTCA	**1414g02022_**F TCGACTTGGGGTAGAGTTTTGCTGAGTCTGCA **1414g02022_R** GACTCAGCAAAACTCTACCCCAAG
Niben101Scf01414g02022 Mutated	Target TTGGCGTTGAGATTAGGTGAGT | | | ||| || ||||| viRNAs AGCTCTATCTCAAAA-GACTCA	**1414g02022_mutF** TCGACTTGGCGTTGAGATTAGGTGAGTCTGCA **1414g02022_mutR** GACTCACCTAATCTCAACGCCAAG
Niben101Scf04015g02030	Target ATTTGATCTCAAATGGCACGG |||||||||| ||| ||| | viRNAs AAAACTAGAGTCTACTGTGAC	**4015g02030_F** TCGACATTTGATCTCAAATGGCACGGCTGCA **4015g02030_R** GCCGTGCCATTTGAGATCAAATG
Niben101Scf04015g02030 Mutated	Target ATTTCAACTCTAAAGGGACGG ||| | ||| | | || | viRNAs AAAACTAGAGTCTACTGTGAC	**4015g02030_mutF** TCGACATTTCAACTCTAAAGGGACGGCTGCA **4015g02030_mutR** GCCGTCCCTTTAGAGTTGAAATG

Conclusion

The advances in plant virology have rapidly improved over the past few years due to the implementations of next-generation sequencing (NGS) (Barba et al. 2014, Hadidi et al. 2016). These sequencing platforms have provided rapid, efficient and low-cost sequencing helping in many areas of plant virology. We can now quickly sequence whole virus genomes aiding in discovery and detection. Knowing a virus sequence gives us increase understanding on many levels; we can study its evolution in relation to other viruses and organisms, we gain insight of its structure and protein arrangement and it allows us to study its interactions with its host and other viral factors on a nucleotide level.

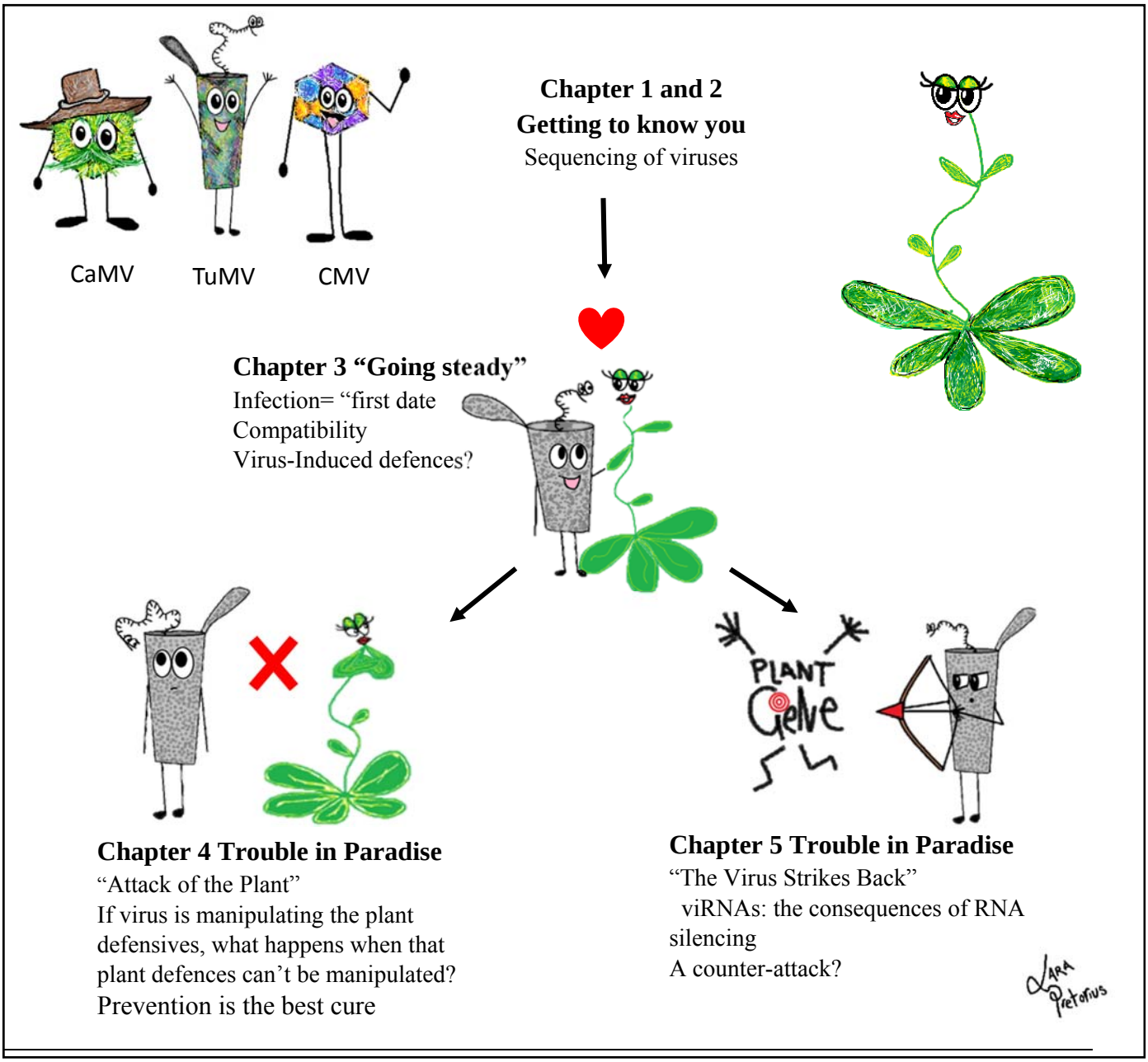


Figure 30: Thesis overview figure depicts how each chapter is related.

These were all important aspects and aims of my thesis which would have been impossible without current advances such as NGS. Using NGS tools it was possible to obtain the sequences of the 2011 and 2015 versions of the TuMV strain as well as CMV strain K and the Australian CaMV strain (Figure 30 “First date” section).

The sequencing of the TuMV strains helped identify the original isolate. Once the original isolate sequence was known further research into this strain’s evolution and relationship with other TuMV strains was accessed. Turnip mosaic virus evolution has been extensively researched tracing its host switch to brassica crops roughly 1000 years ago (Nguyen et al. 2013). The results from this study revealed that though the viral sequence did not alter drastically from the original isolate the sequence changes did follow all the “rules” with most variation found within regions which were known to be recombination hotspots (Gibbs et al 2015) and its heterogenous state was possibly due to its passaging between different hosts (Schneider and Roossinck 2001, Pita et al. 2007). Though, *Turnip mosaic virus* is a virus that has a quasi-species state with a few of its genes being defined as quasi-genes giving it greater genetic flexibility to quickly adapt to new hosts (Gibbs et al. 2015). Phylogenetic analysis placed it in the basal-B group suggesting it may have originated in Europe and possibly introduced into Australia during the British colonisation (Gibbs et al. 2008).

The sequencing of the virus also enabled further research into its interactions with its host as the viral load could be quantified as well as the subsequent impact this may have on the plant. From a phenotypic aspect the virus interferes with the plants normal functioning causing abnormal growth and deformities (Kasschau et al. 2003). These deformities usual appear once the virus has successfully infected the plant and spread systemically. The changes that occur on a nucleotide level in the early stages of infection such as manipulation of gene expression, could allow us to understand how the virus is able to successful overcome the plant’s natural defences and cause disease (Whitham and Wang 2004). Using the model plant *A. thaliana* it was concluded that within the first 24 hours of infection the virus is able to manipulate the plants hormones preventing a proper defence response. TuMV triggered the upregulation of the JA pathway which is usually the response to necrotrophic pathogens. This manipulation could favour viral replication and modification resulting in successful infection and overcoming plant defences at the early stages of infection (Figure 30 “going steady” section).

This was further confirmed by using a mediator mutant known to have a non-function JA pathway and showing increased resistance to other biotrophic pathogens (Fallath et al. 2017). The viral load was decreased in med18 plants suggesting that the JA pathway has an important role in viral

replication. The viral load of CMV had a significant decrease in med18 plants compared to WT though these plants also had an increased ABA expression.

This was particularly interesting as ABA was link to viral defence in a few ways firstly an upregulation of ABA was linked to an increase in callose deposits possibly reducing viral spread (Lin 2015). Secondly, ABA has a role in RNA silencing which is the plant's main defence against viruses (Chinnusamy et al. 2008). The levels of ABA correlate to miR168 which regulates AGO1, a vital part of the RNA silencing pathway. Research has shown that ABA treatment increases AGO1 levels (Li et al. 2012). This suggests that apart from the JA pathway not being manipulated by the virus at the early stages of infection the virus may also experience increased RNA silencing both of which could result in a decrease in viral load as seen in CMV infected med18 plants (Figure 30 "Attack of the Plant").

The idea of RNA silencing and viral processing is the focus of the last chapter where it is hypothesised that abundant viRNAs may have coincidental sequence homology to host-transcripts causing unwanted silencing resulting in phenotypic symptoms. This idea has been confirmed by other studies where viRNAs target host genes resulting in viral symptoms (Chapman et al. 2004, Shimura et al. 2011, Smith et al. 2011, Navarro et al. 2012).

One of the reasons viruses are so successful is their rapid and extreme replication which will also result in highly abundant viRNAs which if sequence homology is high enough an unwanted off-target effect could occur (Figure 30 The Virus Strikes Back"). On paper this sounds plausible and possibly even likely though after many attempts at trying to find and confirm these events using a dual reporter system only one was confirmed. The low level of confirmed targets could possibly be due to the assay not being robust enough to detect these events.

Through small RNA sequencing it was confirmed that there was a significant increase of the 21 and 22 nt sRNAs in plants inoculated with TuMV in comparison to mock inoculated plants which is consistent with other studies (Xia et al. 2014, Li et al. 2016, Margaria et al. 2016, Qiu et al. 2017). The decrease in the 24 nt subset of sRNAs in TuMV inoculated plants was possibly due to viral proteins interferring with the 24 nt biogensis pathway which has an important role in DNA methylation (Chiu et al. 2010, Garcia-Ruiz et al. 2015, Cheng and Wang 2017). This is one of the areas where further investigation is needed; to determine whether this decrease affects host DNA methylation in way that could benefit the virus?

Interestingly, the coverage pattern observed in the two hosts used, *A. thaliana* and *N. benthamiana*, did not match which was contrary to previous results (Xu et al. 2012, Mitter et al. 2013). This

suggest that the processing of the virus is host dependent which in my opinion makes sense. As discussed in Chapter 1 the level of sequence diversity was host dependent as well as the viral mutation rates (Schneider and Roossinck 2001, Pita et al. 2007) therefore the different hosts may pose a different selective pressure on the virus causing changes within the sequence resulting in different processing and hotspots to occur. It should also be noted that the variety of *N. benthamiana* used in this study is known to have a disruption in its *Rdr1* gene making it more susceptible to viral infection (Bally et al. 2015). The *Rdr1* was found to have an important role RNA silencing against virus and therefore this may be another contributing factor to why the coverage pattern observed between this variety of *N. bethamiana* used and *A. thaliana* are different.

The data and research present in this thesis successfully confirmed all aims and research questions stated at the beginning of this "research journey". Aim 1 was completed with the successful sequencing of TuMV genomes; the original isolate as well as the 2011 and 2015 versions, where minor sequence alterations were observed. An alignment to other Australian isolates was performed allowing further phylogenetic analysis.

Studying the plant-viral interactions through gene expression, completed the second aim. This revealed that the virus manipulates the plant's hormones which are linked to specific defence pathways ultimately preventing a proper defence modulation at early stages in the infection process.

The manipulation of the defence pathways was further confirmed in the third chapter; results showed that without certain hormones or pathways to manipulate the virus was not as successful. Though the fact that the lack of the MED18 subunit decreased viral load may not suggest that it specifically is linked to viral replication or infection and the decrease may just be due to the downstream effects caused by the missing subunit. Therefore, it cannot be confirmed whether the MED18 specifically influences viral infection as stated in aim 3.

Aim 4 was confirmed as I successfully confirmed one host transcript targeted by an abundant viRNAs that was highly homologous to the N. benthamiana gene; Niben3160g02007 of unknown function.

Like most research the manipulation of the plant defence pathways and possibly host transcript sequence targeting may just be the tip of the plant-virus relationship iceberg. It is safe to assume that there are still many interactions and processes occurring that we are not even aware of. However, with the ever-improving advances in technology we will no doubt look back and think once again how far we have come. I can only hope that this research may help in getting us a

fraction closer to changing the plant-virus relationship status from “It’s complicated” to “In a relationship”.

The current advance in science has allowed us to study the plant-virus relationship in greater detailer revealing that on some rare occasion plants and viruses are able to live happily ever after. These are rare events though all the research and advances has improved our greater understanding; closing knowledge gaps and making the plant-virus relationship a little less complicated.

References

Adenot, X., *et al.* (2006). "DRB4-Dependent TAS3 trans-Acting siRNAs Control Leaf Morphology through AGO7." Current Biology **16**(9): 927-932.

Adkar-Purushothama, C. R., *et al.* (2015). "Small RNA Derived from the Virulence Modulating Region of the Potato spindle tuber viroid Silences callose synthase Genes of Tomato Plants." The Plant Cell **27**(8): 2178-2194.

Ahlquist, P., *et al.* (2003). "Host factors in positive-strand RNA virus genome replication." Journal of virology **77**(15): 8181-8186.

Alazem, M. and N.-S. Lin (2015). "Roles of plant hormones in the regulation of host–virus interactions." Molecular Plant Pathology **16**(5): 529-540.

Alazem, M., *et al.* (2013). "The Abscisic Acid Pathway Has Multifaceted Effects on the Accumulation of Bamboo mosaic virus." Molecular plant-microbe interactions **27**(2): 177-189.

Allen, B. L. and D. J. Taatjes (2015). "The Mediator complex: a central integrator of transcription." Nat Rev Mol Cell Biol **16**(3): 155-166.

Anderson, J. P., *et al.* (2004). "Antagonistic Interaction between Abscisic Acid and Jasmonate-Ethylene Signaling Pathways Modulates Defence Gene Expression and Disease Resistance in Arabidopsis." The Plant Cell **16**(12): 3460-3479.

Angell, S. M. and D. C. Baulcombe (1997). "Consistent gene silencing in transgenic plants expressing a replicating potato virus X RNA." EMBO J **16**(12): 3675-3684.

Anindya, R., and Savithri, H. S. (2004). "Potyviral NIa proteinase: a proteinase with novel deoxyribonuclease activity." Journal of Biological Chemistry.

Anindya, R., Chittori, S., and Savithri, H. S. (2005). "Tyrosine 66 of Pepper vein banding virus genome-linked protein is uridylylated by RNA-dependent RNA polymerase." Virology 336(2): 154-162.

Atreya, C. D., *et al.* (1992). "Site-directed mutations in the potyvirus HC-PRO gene affect helper component activity, virus accumulation, and symptom expression in infected tobacco plants." Virology **191**(1): 106-111.

Audenaert, K., *et al.* (2002). "Abscisic Acid Determines Basal Susceptibility of Tomato toBotrytis cinerea and Suppresses Salicylic Acid-Dependent Signaling Mechanisms." Plant Physiology **128**(2): 491-501.

Avina-Padilla, K., *et al.* (2015). "In silico prediction and validation of potential gene targets for pospiviroid-derived small RNAs during tomato infection." Gene **564**(2): 197-205.

Axtell, M. J., *et al.* (2006). "A Two-Hit Trigger for siRNAs Biogenesis in Plants." Cell **127**(3): 565-577.

Azevedo J ,*et al*. (2010). "Argonaute quenching and global changes in Dicer homeostasis caused by a pathogen-encoded GW repeat protein." Genes Develop **24**: 904-15

Bäckström, S., *et al.* (2007). "Purification of a Plant Mediator from Arabidopsis thaliana Identifies PFT1 as the Med25 Subunit." Molecular cell **26**(5): 717-729.

Balachandran, S., *et al.* (1997). "Concepts of plant biotic stress. Some insights into the stress physiology of virus-infected plants, from the perspective of photosynthesis." Physiologia Plantarum **100**(2): 203-213.

Bally, J., *et al*. (2015). "The extremophile Nicotiana benthamiana has traded viral defence for early vigour." Nature plants 1(11): 15165.

Balmer, D. and B. Mauch-Mani (2013). "Small Yet Mighty MicroRNAs in Plant-Microbe Interactions." MicroRNA **2**(1): 73-80.

Barba, M., *et al*. (2014). "Historical perspective, development and applications of next-generation sequencing in plant virology." Viruses 6(1): 106-136.

Bartel, D. P. (2004). "MicroRNAs: Genomics, Biogenesis, Mechanism, and Function." Cell **116**(2): 281-297.

Bartels, D. and R. Sunkar (2005). "Drought and salt tolerance in plants." Critical reviews in plant sciences **24**(1): 23-58.
Baulcombe, D. (2004). "RNA silencing in plants." Nature **431**(7006): 356-363.
Baulcombe, D. C. (2015). "VIGS, HIGS and FIGS: small RNA silencing in the interactions of viruses or filamentous organisms with their plant hosts." Current Opinion in Plant Biology **26**: 141-146.
Belshaw, R., Gardner, A., Rambaut, A., & Pybus, O. G. (2008). "Pacing a small cage: mutation and RNA viruses." Trends in Ecology & Evolution 23(4): 188-193.
Berger, S., *et al.* (2002). "Local and differential control of vegetative storage protein expression in response to herbivore damage in Arabidopsis thaliana." Physiologia Plantarum **114**(1): 85-91.
Berrocal-Lobo, M., *et al.* (2002). "Constitutive expression of ETHYLENE-RESPONSE-FACTOR1 in Arabidopsis confers resistance to several necrotrophic fungi." The Plant Journal **29**(1): 23-32.
Bian, X.-Y., *et al.* (2006). "Analysis of silencing escape of tomato leaf curl virus: an evaluation of the role of DNA methylation." Molecular plant-microbe interactions **19**(6): 614-624.
Blanc, S., *et al.* (1998). "Mutations in the potyvirus helper component protein: effects on interactions with virions and aphid stylets." Journal of General Virology **79**(12): 3119-3122.
Blevins, T., *et al.* (2006). "Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing." Nucleic acids research **34**(21): 6233-6246.
Blume, B., *et al.* (2000). "Receptor-mediated increase in cytoplasmic free calcium required for activation of pathogen defence in parsley." The Plant Cell **12**(8): 1425-1440.
Bohman, S., *et al.* (2004). "Characterisation of an Arabidopsis–Leptosphaeria maculans pathosystem: resistance partially requires camalexin biosynthesis and is independent of salicylic acid, ethylene and jasmonic acid signalling." The Plant Journal **37**(1): 9-20.
Bollman, K. M., *et al.* (2003). "HASTY, the Arabidopsis ortholog of exportin 5/MSN5, regulates phase change and morphogenesis." Development **130**(8): 1493-1504.
Bonnet, J., *et al.* (2005). "Role of recombination in the evolution of natural populations of Cucumber mosaic virus, a tripartite RNA plant virus." Virology **332**(1): 359-368.
Borges, F. and R. A. Martienssen (2015). "The expanding world of small RNAs in plants." Nat Rev Mol Cell Biol **16**(12): 727-741.
Borsani, O., *et al.* (2005). "Endogenous siRNAs Derived from a Pair of Natural cis-Antisense Transcripts Regulate Salt Tolerance in Arabidopsis." Cell **123**(7): 1279-1291.
Bouché, N., *et al.* (2006). "An antagonistic function for Arabidopsis DCL2 in development and a new function for DCL4 in generating viral siRNAs." The EMBO journal **25**(14): 3347-3356.
Brault, V., *et al.* (2010). "Aphids as transport devices for plant viruses." Comptes Rendus Biologies **333**(6–7): 524-538.
Brosnan, C. A., *et al.* (2007). "Nuclear gene silencing directs reception of long-distance mRNA silencing in Arabidopsis." Proceedings of the National Academy of Sciences **104**(37): 14741-14746.
Brown, R. L., *et al.* (2003). "A Role for the GCC-Box in Jasmonate-Mediated Activation of the PDF1.2 Gene of Arabidopsis." Plant Physiology **132**(2): 1020-1032.
Burgyán, J. and Z. Havelda (2011). "Viral suppressors of RNA silencing." Trends in plant science **16**(5): 265-272.
Callaway, A., *et al.* (1996). "Characterization of cauliflower mosaic virus (CaMV) resistance in virus-resistant ecotypes of Arabidopsis." MPMI-Molecular Plant Microbe Interactions **9**(9): 810-818.
Carrington, J. C., Freed, D. D., & Sanders, T. C. (1989). "Autocatalytic processing of the potyvirus helper component proteinase in Escherichia coli and in vitro." Journal of Virology 63(10): 4459-4463.
Castel, S. E. and R. A. Martienssen (2013). "RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond." Nat Rev Genet **14**(2): 100-112.

Catalano, D., *et al.* (2012). "In silico prediction of virus-derived small interfering RNAs and their putative host messenger targets in Solanum lycopersicum infected by different potato virus Y isolates." 2012 **18**.
Chadick, J. Z. and F. J. Asturias (2005). "Structure of eukaryotic Mediator complexes." Trends in biochemical sciences **30**(5): 264-271.
Chapman, E. J. and J. C. Carrington (2007). "Specialization and evolution of endogenous small RNA pathways." Nature Reviews Genetics **8**(11): 884-896.
Chapman, E. J., *et al.* (2004). "Viral RNA silencing suppressors inhibit the microRNA pathway at an intermediate step." Genes & development **18**(10): 1179-1186.
Chapple, C. (1998). "Molecular-genetic analysis of plant cytochrome P450-dependent monooxygenases." Annual review of plant biology **49**(1): 311-343.
Chen, H.-M., *et al.* (2010). "22-nucleotide RNAs trigger secondary siRNAs biogenesis in plants." Proceedings of the National Academy of Sciences **107**(34): 15269-15274.
Chen, Y., *et al.* (2007). "Molecular evidence and sequence analysis of a natural reassortant between Cucumber mosaic virus subgroup IA and II strains." Virus genes **35**(2): 405-413.
Cheng, X. and A. Wang (2017). "The Potyvirus Silencing Suppressor Protein VPg Mediates Degradation of SGS3 via Ubiquitination and Autophagy Pathways." Journal of virology **91**(1).
Chinnusamy, V., *et al.* (2008). "Abscisic Acid-mediated Epigenetic Processes in Plant Development and Stress Responses." Journal of Integrative Plant Biology **50**(10): 1187-1195.
Chiu, M.-H., *et al.* (2010). "The silencing suppressor P25 of Potato virus X interacts with Argonaute1 and mediates its degradation through the proteasome pathway." Molecular Plant Pathology **11**(5): 641-649.
Chou, T.-C. and R. L. Moyle (2014). "Synthetic versions of firefly luciferase and Renilla luciferase reporter genes that resist transgene silencing in sugarcane." BMC Plant Biology **14**(1): 92.
Chu, M., *et al.* (1997). "Two separate regions in the genome of the tobacco etch virus contain determinants of the wilting response of Tabasco pepper." Molecular plant-microbe interactions **10**(4): 472-480.
Chung, K.-M., *et al.* (2008). "New perspectives on plant defence responses through modulation of developmental pathways." Mol Cells **26**(2): 107-112.
Cooper, B. (2001). "Collateral gene expression changes induced by distinct plant viruses during the hypersensitive resistance reaction in Chenopodium amaranticolor." The Plant Journal **26**(3): 339-349.
Cotton, S., *et al.* (2009). "Turnip Mosaic Virus RNA Replication Complex Vesicles Are Mobile, Align with Microfilaments, and Are Each Derived from a Single Viral Genome." Journal of virology **83**(20): 10460-10471.
Cui, X., *et al.* (2010). "The Tobacco etch virus P3 protein forms mobile inclusions via the early secretory pathway and traffics along actin microfilaments." Virology 397(1): 56-63.
Dai, X. and P. X. Zhao (2011). "psRNAsTarget: a plant small RNA target analysis server." Nucleic acids research **39**(Web Server issue): W155-W159.
Dalmay, T., *et al.* (2001). "SDE3 encodes an RNA helicase required for post-transcriptional gene silencing in Arabidopsis." The EMBO journal **20**(8): 2069-2077.
Deleris, A., *et al.* (2006). "Hierarchical Action and Inhibition of Plant Dicer-Like Proteins in Antiviral Defence." Science **313**(5783): 68-71.
Dempsey, D., *et al.* (1997). "Identification of an Arabidopsis locus required for resistance to turnip crinkle virus." The Plant Journal **11**(2): 301-311.
Deng, P., Wu, Z., and Wang, A. (2015). "The multifunctional protein CI of potyviruses plays interlinked and distinct roles in viral genome replication and intercellular movement." Virology journal 12(1): 141.
Diaz-Pendon, J. A., *et al.* (2007). "Suppression of Antiviral Silencing by Cucumber Mosaic Virus 2b Protein in Arabidopsis Is Associated with Drastically Reduced Accumulation of Three Classes of Viral Small Interfering RNAs." The Plant Cell **19**(6): 2053-2063.

Ding, S.-W. and O. Voinnet (2007). "Antiviral Immunity Directed by Small RNAs." Cell **130**(3): 413-426.
Donaire, L., *et al.* (2009). "Deep-sequencing of plant viral small RNAs reveals effective and widespread targeting of viral genomes." Virology **392**(2): 203-214.
Dougherty, W. G., and Parks, T. D. (1991). "Post-translational processing of the tobacco etch virus 49-kDa small nuclear inclusion polyprotein: identification of an internal cleavage site and delimitation of VPg and proteinase domains." Virology 183(2): 449-456.
Dufresne, P. J., Thivierge, K., Cotton, S., Beauchemin, C., Ide, C., Ubalijoro, E., ... & Fortin, M. G. (2008). "Heat shock 70 protein interaction with Turnip mosaic virus RNA-dependent RNA polymerase within virus-induced membrane vesicles." Virology 374(1): 217-227.
Dunoyer, P., *et al.* (2004). "A Cysteine-Rich Plant Protein Potentiates Potyvirus Movement through an Interaction with the Virus Genome-Linked Protein VPg." Journal of virology **78**(5): 2301-2309.
Dunoyer, P., *et al.* (2005). "DICER-LIKE 4 is required for RNA interference and produces the 21-nucleotide small interfering RNA component of the plant cell-to-cell silencing signal." Nature genetics **37**(12): 1356-1360.
Elena, S. F. (2016). "Evolutionary transitions during RNA virus experimental evolution." Phil. Trans. R. Soc. B **371**(1701): 20150441.
Elena, S. F., and Rodrigo, G. (2012). "Towards an integrated molecular model of plant–virus interactions." Current opinion in virology 2(6): 719-724.
Ellinger, D., *et al.* (2010). "DONGLE and DEFECTIVE IN ANTHER DEHISCENCE1 Lipases Are Not Essential for Wound- and Pathogen-Induced Jasmonate Biosynthesis: Redundant Lipases Contribute to Jasmonate Formation." Plant Physiology **153**(1): 114-127.
Fallath, T. A. (2016). Investigating the role of the Mediator complex in plant defence, The University of Queensland, School of Agriculture and Food Sciences.
Fallath, T., *et al.* (2017). "MEDIATOR18 and MEDIATOR20 confer susceptibility to Fusarium oxysporum in Arabidopsis thaliana." PLoS ONE **12**(4): e0176022.
Gabrenaite-Verkhovskaya, R., Andreev, I. A., Kalinina, N. O., Torrance, L., Taliansky, M. E., & Mäkinen, K. (2008). "Cylindrical inclusion protein of potato virus A is associated with a subpopulation of particles isolated from infected plants." Journal of General Virology, 89(3): 829-838.
Gallo, A., *et al*. (2018). "A functional link between RNA replication and virion assembly in the potyvirus Plum pox virus." Journal of virology JVI-02179.
García-Arenal, F. and B. A. McDonald (2003). "An analysis of the durability of resistance to plant viruses." Phytopathology **93**(8): 941-952.
García-Arenal, F., *et al.* (2003). "Variation and evolution of plant virus populations." International Microbiology **6**(4): 225-232.
Garcia-Ruiz, H., *et al.* (2010). "Arabidopsis RNA-Dependent RNA Polymerases and Dicer-Like Proteins in Antiviral Defence and Small Interfering RNA Biogenesis during Turnip Mosaic Virus Infection." The Plant Cell **22**(2): 481-496.
Garcia-Ruiz, H., *et al.* (2015). "Roles and programming of Arabidopsis ARGONAUTE proteins during Turnip mosaic virus infection." PLoS Pathog **11**(3): e1004755.
Gardner, M. W. and J. B. Kendrick (1921). "TURNIP MOSAIC1." Journal of Agricultural Research **22**(3): 121.
Ghildiyal, M. and P. D. Zamore (2009). "Small silencing RNAs: an expanding universe." Nat Rev Genet **10**(2): 94-108.
Gibbs, A., *et al.* (2008). "The potyviruses of Australia." Archives of virology **153**(8): 1411-1420.
Glazebrook, J. (2001). "Genes controlling expression of defence responses in Arabidopsis — 2001 status." Current Opinion in Plant Biology **4**(4): 301-308.
Glazebrook, J. (2005). "Contrasting mechanisms of defence against biotrophic and necrotrophic pathogens." Annu. Rev. Phytopathol. **43**: 205-227.

Glazebrook, J., *et al.* (1997). "Phytoalexin-deficient mutants of Arabidopsis reveal that PAD4 encodes a regulatory factor and that four PAD genes contribute to downy mildew resistance." Genetics **146**(1): 381-392.
Grangeon, R., *et al*. (2012). "Impact on the endoplasmic reticulum and Golgi apparatus during Turnip mosaic virus infection." Journal of virology JVI-01146.
Grangeon, R., *et al.* (2010). "A model for the biogenesis of turnip mosaic virus replication factories." Communicative & integrative biology **3**(4): 363-365.
Grangeon, R., Jiang, J., & Laliberté, J. F. (2012). "Host endomembrane recruitment for plant RNA virus replication." Current opinion in virology 2(6): 683-690.
Green, M. R. and J. Sambrook (2000). Molecular Cloning: A Laboratory Manual Cold Spring Harbor Laboratory Press.
Hadidi, A., Flores, R., Candresse, T., & Barba, M. (2016). "Next-generation sequencing and genome editing in plant virology." Frontiers in microbiology 7: 1325.
Hamera, S., *et al.* (2012). "Cucumber mosaic virus suppressor 2b binds to AGO4-related small RNAs and impairs AGO4 activities." The Plant Journal **69**(1): 104-115.
Hamilton, A. J. and D. C. Baulcombe (1999). "A Species of Small Antisense RNA in Posttranscriptional Gene Silencing in Plants." Science **286**(5441): 950-952.
Han, M.-H., *et al.* (2004). "The Arabidopsis double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation." Proceedings of the National Academy of Sciences of the United States of America **101**(4): 1093-1098.
He, X.-J., *et al.* (2011). "Regulation and function of DNA methylation in plants and animals." Cell research **21**(3): 442-465.
Herbers, K., *et al.* (2000). "Regulation of carbohydrate partitioning during the interaction of potato virus Y with tobacco." Molecular Plant Pathology **1**(1): 51-59.
Herranz, M. C., *et al.* (2015). "Comparative analysis among the small RNA populations of source, sink and conductive tissues in two different plant-virus pathosystems." BMC Genomics **16**(1): 117.
Ho, T., *et al.* (2007). "Evidence for targeting common siRNAs hotspots and GC preference by plant Dicer-like proteins." FEBS letters **581**(17): 3267-3272.
Holmes, E. C. (2003). "Error thresholds and the constraints to RNA virus evolution." Trends in microbiology **11**(12): 543-546.
Holstege, F. C., *et al.* (1998). "Dissecting the regulatory circuitry of a eukaryotic genome." Cell **95**(5): 717-728.
Hong, Y. and A. G. Hunt (1996). "RNA polymerase activity catalyzed by a potyvirus-encoded RNA-dependent RNA polymerase." Virology **226**(1): 146-151.
Huang, Z., *et al.* (2005). "Salicylic acid-dependent expression of host genes in compatible Arabidopsis-virus interactions." Plant Physiology **137**(3): 1147-1159.
Huet, H., *et al.* (1994). "Mutations in the helper component protease gene of zucchini yellow mosaic virus affect its ability to mediate aphid transmissibility." Journal of General Virology **75**(6): 1407-1414.
Hull, R. (2013). Plant virology, Academic press.
Hunter, C., *et al.* (2003). "The Arabidopsis Heterochronic Gene ZIPPY Is an ARGONAUTE Family Member." Current Biology **13**(19): 1734-1739.
Imasaki, T., *et al.* (2011). "Architecture of the Mediator head module." Nature **475**(7355): 240-243.
Iram, S. (2010). Role of viral and host microRNAs in plant virus interaction. School of Biological Sciences, The University of Queensland.
Iriti, M. and F. Faoro (2008). "Abscisic acid is involved in chitosan-induced resistance to tobacco necrosis virus (TNV)." Plant Physiology and Biochemistry **46**(12): 1106-1111.
Ivanov, K. I., *et al.* (2014). "Molecular and cellular mechanisms underlying potyvirus infection." Journal of General Virology **95**(7): 1415-1429.
Jacquemond, M. (2012). Chapter 13 - Cucumber Mosaic Virus. Advances in virus research. G. Loebenstein and H. Lecoq, Academic Press. **84:** 439-504.

Jenner, C. E., *et al.* (2003). "The Dual Role of the Potyvirus P3 Protein of Turnip mosaic virus as a Symptom and Avirulence Determinant in Brassicas." Molecular plant-microbe interactions **16**(9): 777-784.
Kadaré, G. and A.-L. Haenni (1997). "Virus-encoded RNA helicases." Journal of virology **71**(4): 2583.
Kanazawa A,*et al*.(2011). "Virus-mediated efficient induction of epigenetic modifications of endogenous genes with phenotypic changes in plants." The Plant Journal. **65**: 156-68.
Kariola, T., *et al.* (2005). "Chlorophyllase 1, a damage control enzyme, affects the balance between defence pathways in plants." The Plant Cell **17**(1): 282-294.
Kasschau, K. D., *et al.* (2003). "P1/HC-Pro, a Viral Suppressor of RNA Silencing, Interferes with Arabidopsis Development and miRNA Function." Developmental Cell **4**(2): 205-217.
Kidd, B. N., *et al.* (2009). "The mediator complex subunit PFT1 is a key regulator of jasmonate-dependent defence in Arabidopsis." The Plant Cell **21**(8): 2237-2252.
Kim, Y. J., *et al.* (2011). "The role of Mediator in small and long noncoding RNA production in Arabidopsis thaliana." The EMBO journal **30**(5): 814-822.
Kliebenstein, D. J., *et al.* (2005). "Secondary metabolites influence Arabidopsis/Botrytis interactions: variation in host production and pathogen sensitivity." The Plant Journal **44**(1): 25-36.
Knuhtsen, H., *et al*. (1974). "Partial purification and some properties of tobacco etch virus induced intranuclear inclusions." Virology 61(1): 200-209.
Kombrink, E. and I. E. Somssich (1995). "Defence responses of plants to pathogens." Advances in botanical research **21**: 2-34.
Konstantinova, P., *et al*. (2006). "Hairpin-induced tRNA-mediated (HITME) recombination in HIV-1." Nucleic acids research 34(8): 2206-2218.
Koornneef, A. and C. M. Pieterse (2008). "Cross talk in defence signaling." Plant Physiology **146**(3): 839-844.
Kovač, M., *et al.* (2009). "Multiple hormone analysis indicates involvement of jasmonate signalling in the early defence of potato to potato virus YNTN." Biologia Plantarum **53**(1): 195-199.
Kreuze, J. F., *et al.* (2009). "Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: A generic method for diagnosis, discovery and sequencing of viruses." Virology **388**(1): 1-7.
Kunkel, B. N. and D. M. Brooks (2002). "Cross talk between signaling pathways in pathogen defence." Current Opinion in Plant Biology **5**(4): 325-331.
Lai, Z., *et al.* (2014). "MED18 interaction with distinct transcription factors regulates multiple plant functions." Nat Commun **5**.
Lain, S., *et al.* (1990). "RNA helicase: a novel activity associated with a protein encoded by a positive strand RNA virus." Nucleic acids research **18**(23): 7003-7006.
Langenberg, W. G. and L. Zhang (1997). "Immunocytology shows the presence of tobacco etch virus P3 protein in nuclear inclusions." Journal of structural biology **118**(3): 243-247.
Lannou, C. (2012). "Variation and selection of quantitative traits in plant pathogens." Annual review of phytopathology **50**: 319-338.
Larivière, L., *et al.* (2008). "Structure–system correlation identifies a gene regulatory Mediator submodule." Genes & development **22**(7): 872-877.
Lee, Y., *et al.* (2001). "Expansins: ever-expanding numbers and functions." Current Opinion in Plant Biology **4**(6): 527-532.
Lee, Y., *et al.* (2004). "MicroRNA genes are transcribed by RNA polymerase II." The EMBO journal **23**(20): 4051-4060.
Lellis, A. D., *et al.* (2002). "Loss-of-Susceptibility Mutants of< i> Arabidopsis thaliana</i> Reveal an Essential Role for eIF (iso) 4E during Potyvirus Infection." Current Biology **12**(12): 1046-1051.
Lemarié, S., *et al.* (2015). "Camalexin contributes to the partial resistance of Arabidopsis thaliana to the biotrophic soilborne protist Plasmodiophora brassicae." Frontiers in Plant Science **6**: 539.
Léonard, S., *et al.* (2000). "Complex formation between potyvirus VPg and translation eukaryotic initiation factor 4E correlates with virus infectivity." Journal of virology **74**(17): 7730-7737.

Léonard, V. H., *et al.* (2006). "Interaction of Bunyamwera Orthobunyavirus NSs protein with mediator protein MED8: a mechanism for inhibiting the interferon response." Journal of virology **80**(19): 9667-9675.
Li, H. W., *et al*. (1999). "Strong host resistance targeted against a viral suppressor of the plant gene silencing defence mechanism." The EMBO journal, **18**(10), 2683-2691.
Li, M.-L., *et al.* (2016). "The evolving world of small RNAs from RNA viruses." Wiley Interdisciplinary Reviews: RNA **7**(5): 575-588.
Li, R., *et al.* (2012). "Deep Sequencing of Small RNAs in Tomato for Virus and Viroid Identification and Strain Differentiation." PLoS ONE **7**(5): e37127.
Li, S., *et al.* (2013). "microRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in Arabidopsis." Cell **153**(3): 562-574.
Li, W., *et al.* (2012). "Transcriptional Regulation of Arabidopsis MIR168a and ARGONAUTE1 Homeostasis in Abscisic Acid and Abiotic Stress Responses." Plant Physiology **158**(3): 1279-1292.
Li, Y., *et al.* (2016). "Characterization of siRNAs derived from cucumber green mottle mosaic virus in infected cucumber plants." Archives of virology **161**(2): 455-458.
Lin, L., *et al*. (2009). "Protein–protein interactions in two potyviruses using the yeast two-hybrid system." Virus research 142(1-2): 36-40.
Liu, Q. and M. J. Axtell (2015). Quantitating Plant MicroRNA-Mediated Target Repression Using a Dual-Luciferase Transient Expression System. Plant Functional Genomics: Methods and Protocols. Alonso, J. M. and Stepanova, A. N. New York, NY, Springer New York**:** 287-303.
Liu, Y. Y., *et al.* (2009). "Molecular variability of five Cucumber mosaic virus isolates from China." Acta Virol **53**(2): 89-97.
Llave, C. (2010). "Virus-derived small interfering RNAs at the core of plant–virus interactions." Trends in plant science **15**(12): 701-707.
Lopez-Moya, J., *et al.* (1999). "Context of the coat protein DAG motif affects potyvirus transmissibility by aphids." Journal of General Virology **80**(12): 3281-3288.
Love, A. J., *et al.* (2012). "Cauliflower mosaic virus Protein P6 Inhibits Signaling Responses to Salicylic Acid and Regulates Innate Immunity." PLoS ONE **7**(10): e47535.
Lu, Y.-d., *et al.* (2008). "Roles of microRNA in plant defence and virus offense interaction." Plant Cell Reports **27**(10): 1571-1579.
Lucy, A. P., *et al.* (2000). "Suppression of post-transcriptional gene silencing by a plant viral protein localized in the nucleus." The EMBO journal **19**(7): 1672-1680.
Lynn, K., *et al.* (1999). "The PINHEAD/ZWILLE gene acts pleiotropically in Arabidopsis development and has overlapping functions with the ARGONAUTE1 gene." Development **126**(3): 469-481.
Madhani, H. D. (2013). "The Frustrated Gene: Origins of Eukaryotic Gene Expression." Cell **155**(4): 744-749.
Malik, S. and R. G. Roeder (2005). "Dynamic regulation of pol II transcription by the mammalian Mediator complex." Trends in biochemical sciences **30**(5): 256-263.
Mallory, A. and H. Vaucheret (2010). "Form, function, and regulation of ARGONAUTE proteins." The Plant Cell **22**(12): 3879-3889.
Mallory, A. C., *et al.* (2002). "A viral suppressor of RNA silencing differentially regulates the accumulation of short interfering RNAs and micro-RNAs in tobacco." Proceedings of the National Academy of Sciences **99**(23): 15228-15233.
Malpica, J. M., *et al.* (2002). "The rate and character of spontaneous mutation in an RNA virus." Genetics **162**(4): 1505-1511.
Manners, J. M., *et al.* (1998). "The promoter of the plant defensin gene PDF1. 2 from Arabidopsis is systemically activated by fungal pathogens and responds to methyl jasmonate but not to salicylic acid." Plant Molecular Biology **38**(6): 1071-1080.
Maoka, T., *et al.* (2010). "Mixed infection in tomato to ensure frequent generation of a natural reassortant between two subgroups of Cucumber mosaic virus." Virus genes **40**(1): 148-150.

Margaria, P., *et al.* (2016). "Comparison of small RNA profiles in Nicotiana benthamiana and Solanum lycopersicum infected by polygonum ringspot tospovirus reveals host-specific responses to viral infection." Virus Research **211**: 38-45.
Matzke, M. A. and R. A. Mosher (2014). "RNA-directed DNA methylation: an epigenetic pathway of increasing complexity." Nat Rev Genet **15**(6): 394-408.
Mauch-Mani, B. and F. Mauch (2005). "The role of abscisic acid in plant–pathogen interactions." Current Opinion in Plant Biology **8**(4): 409-414.
Melotto, M., *et al.* (2006). "Plant Stomata Function in Innate Immunity against Bacterial Invasion." Cell **126**(5): 969-980.
Merits, A., *et al.* (2002). "Proteolytic processing of potyviral proteins and polyprotein processing intermediates in insect and plant cells." Journal of General Virology **83**(5): 1211-1221.
Merits, A., Guo, D., & Saarma, M. (1998). "VPg, coat protein and five non-structural proteins of potato A potyvirus bind RNA in a sequence-unspecific manner." Journal of General Virology 79(12): 3123-3127.
Mi, S., *et al.* (2008). "Sorting of Small RNAs into Arabidopsis Argonaute Complexes Is Directed by the 5′ Terminal Nucleotide." Cell **133**(1): 116-127.
Michon, T., *et al.* (2006). "The potyviral virus genome-linked protein VPg forms a ternary complex with the eukaryotic initiation factors eIF4E and eIF4G and reduces eIF4E affinity for a mRNA cap analogue." FEBS Journal **273**(6): 1312-1322.
Miersch, O., *et al.* (1999). "Jasmonates and related compounds from Fusarium oxysporum." Phytochemistry **50**(4): 517-523.
Miozzi, L., *et al.* (2013). "Genome-wide identification of viral and host transcripts targeted by viral siRNAs in Vitis vinifera." Molecular Plant Pathology **14**(1): 30-43.
Mitter, N., *et al.* (2013). "Differential Expression of Tomato Spotted Wilt Virus-Derived Viral Small RNAs in Infected Commercial and Experimental Host Plants." PLoS ONE **8**(10): e76276.
Mlotshwa, S., *et al.* (2008). "Small RNAs in viral infection and host defence." Trends in plant science **13**(7): 375-382.
Molnár, A., *et al.* (2005). "Plant Virus-Derived Small Interfering RNAs Originate Predominantly from Highly Structured Single-Stranded Viral RNAs." Journal of virology **79**(12): 7812-7818.
Morel, J.-B., *et al.* (2002). "Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance." The Plant Cell Online **14**(3): 629-639.
Moreno, M., *et al.* (1998). "Resistance in plants transformed with the P1 or P3 gene of tobacco vein mottling potyvirus." Journal of General Virology **79**(11): 2819-2827.
Mourrain, P., *et al.* (2000). "Arabidopsis SGS2 and SGS3 Genes Are Required for Posttranscriptional Gene Silencing and Natural Virus Resistance." Cell **101**(5): 533-542.
Moyle, R. L., *et al.* (2017). "An Optimized Transient Dual Luciferase Assay for Quantifying MicroRNA Directed Repression of Targeted Sequences." Frontiers in Plant Science **8**(1631).
Moyle, R., *et al.* (2016). "Analysis of the first complete genome sequence of an Australian tomato spotted wilt virus isolate." Australasian Plant Pathology **45**(5): 509-512.
Moyo, L., *et al.* (2017). "The effects of potato virus Y-derived virus small interfering RNAs of three biologically distinct strains on potato (Solanum tuberosum) transcriptome." Virology Journal **14**: 129.
Muller, H. J. (1932). "Some genetic aspects of sex." The American Naturalist **66**(703): 118-138.
Muller, H. J. (1964). "The relation of recombination to mutational advance." Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis **1**(1): 2-9.
Mur, L. A. J., *et al.* (2006). "The Outcomes of Concentration-Specific Interactions between Salicylate and Jasmonate Signaling Include Synergy, Antagonism, and Oxidative Stress Leading to Cell Death." Plant Physiology **140**(1): 249-262.
Nafisi, M., *et al.* (2007). "Arabidopsis cytochrome P450 monooxygenase 71A13 catalyzes the conversion of indole-3-acetaldoxime in camalexin synthesis." The Plant Cell **19**(6): 2039-2052.
Nagy, P. D., and Pogany, J. (2012). "The dependence of viral RNA replication on co-opted host factors." Nature Reviews Microbiology 10(2): 137.

Nam, J., *et al.* (2013). "AltMV TGB1 Nucleolar Localization Requires Homologous Interaction and Correlates with Cell Wall Localization Associated with Cell-to-Cell Movement." The Plant Pathology Journal **29**(4): 454-459.
Navarro, B., *et al.* (2012). "Small RNAs containing the pathogenic determinant of a chloroplast-replicating viroid guide the degradation of a host mRNA as predicted by RNA silencing." Plant J **70**(6): 991-1003.
Niki, T., *et al.* (1998). "Antagonistic effect of salicylic acid and jasmonic acid on the expression of pathogenesis-related (PR) protein genes in wounded mature tobacco leaves." Plant and cell physiology **39**(5): 500-507.
Nomura, K., *et al.* (2005). "Suppression of host defence in compatible plant–Pseudomonas syringae interactions." Current Opinion in Plant Biology **8**(4): 361-368.
Nouri, S., *et al.* (2014). "Genetic Structure and Molecular Variability of Cucumber mosaic virus Isolates in the United States." PLoS ONE **9**(5): e96582.
Novella, I. S., *et al.* (1999). "Exponential fitness gains of RNA virus populations are limited by bottleneck effects." Journal of virology **73**(2): 1668-1671.
Nürnberger, T. and D. Scheel (2001). "Signal transmission in the plant immune response." Trends in plant science **6**(8): 372-379.
Nürnberger, T. and F. Brunner (2002). "Innate immunity in plants and animals: emerging parallels between the recognition of general elicitors and pathogen-associated molecular patterns." Current Opinion in Plant Biology **5**(4): 318-324.
Nyalugwe, E. P., *et al.* (2015). "Biological and molecular variation amongst Australian Turnip mosaic virus isolates." Plant Pathology **64**(5): 1215-1223.
Ohshima, K., *et al.* (2002). "Molecular evolution of Turnip mosaic virus: evidence of host adaptation, genetic recombination and geographical spread." Journal of General Virology **83**(6): 1511-1521.
Ohshima, K., *et al.* (2007). "Patterns of recombination in turnip mosaic virus genomic sequences indicate hotspots of recombination." Journal of General Virology **88**(1): 298-315.
Oka, K., *et al.* (2013). "Jasmonic acid negatively regulates resistance to Tobacco mosaic virus in tobacco." Plant and cell physiology **54**(12): 1999-2010.
Palukaitis, P., *et al.* (1992). "Cucumber mosaic virus." Advances in virus research **41**: 281-348.
Pan, Q. (2016). “Molecular Characterization of the Potyviral First Protein (P1 Protein)”.
Parent, J. S., *et al.* (2015). "Respective contributions of Arabidopsis DCL2 and DCL4 to RNA silencing." The Plant Journal **81**(2): 223-232.
Pasin, F., *et al.* (2014). “The hypervariable amino-terminus of P1 protease modulates potyviral replication and host defense responses.” PLoS pathogens, **10**(3): e1003985.
Pathak, V. K., and Temin, H. M. (1992). “5-Azacytidine and RNA secondary structure increase the retrovirus mutation rate.” Journal of virology 66(5): 3093-3100.
Penninckx, I. A. M. A., *et al.* (1998). "Concomitant Activation of Jasmonate and Ethylene Response Pathways Is Required for Induction of a Plant Defensin Gene in Arabidopsis." The Plant Cell **10**(12): 2103-2113.
Peragine, A., *et al.* (2004). "SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in Arabidopsis." Genes & development **18**(19): 2368-2379.
Pérez-Bueno, M. L., *et al.* (2006). "Imaging viral infection: studies on Nicotiana benthamiana plants infected with the pepper mild mottle tobamovirus." Photosynthesis research **90**(2): 111-123.
Petersen, C. P., *et al.* (2006). "Short RNAs repress translation after initiation in mammalian cells." Molecular cell **21**(4): 533-542.
Pfeffer, S.,*et al.* (2005). “Identification of microRNAs of the herpesvirus family”. Nature methods, **2**(4), 269.
Pieterse, C. M. J., *et al.* (2009). "Networking by small-molecule hormones in plant immunity." Nat Chem Biol **5**(5): 308-316.

Pirone, T. P. and K. L. Perry (2002). "Aphids: non-persistent transmission." Advances in botanical research **36**: 1-19.
Pirone, T. P. and S. Blanc (1996). "Helper-dependent vector transmission of plant viruses." Annual review of phytopathology **34**(1): 227-247.
Pita, J. S.,*et al*. (2007). "Environment determines fidelity for an RNA virus replicase." Journal of virology 81(17): 9072-9077.
Poirier, E. Z., and Vignuzzi, M. (2017). "Virus population dynamics during infection." Current opinion in virology 23: 82-87.
Pontes, O., *et al.* (2006). "The Arabidopsis Chromatin-Modifying Nuclear siRNAs Pathway Involves a Nucleolar RNA Processing Center." Cell **126**(1): 79-92.
Pretorius, L., *et al.* "First fully sequenced genome of an Australian isolate of Cauliflower mosaic virus." Australasian Plant Pathology: 1-3.
Pretorius, L., *et al.* (2016). "Complete nucleotide sequence of an Australian isolate of Turnip mosaic virus before and after seven years of serial passaging." Genome Announcements **4**(6): e01269-01216.
Pretorius, L.-S., *et al.* (2017). "Natural and engineered defences against plant viruses." Current Biotechnology **6**(4): 339-348.
Qi, X., *et al.* (2009). "Small RNA Deep Sequencing Reveals Role for *Arabidopsis thaliana* RNA-Dependent RNA Polymerases in Viral siRNAs Biogenesis." PLoS ONE **4**(3): e4971.
Qi, X., *et al.* (2009). "Small RNA Deep Sequencing Reveals Role for Arabidopsis thaliana RNA-Dependent RNA Polymerases in Viral siRNAs Biogenesis." PLoS ONE **4**(3): e4971.
Qi, Y., *et al.* (2006). "Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation." Nature **443**(7114): 1008-1012.
Qin, C., *et al.* (2017). "Roles of Dicer-Like Proteins 2 and 4 in Intra- and Intercellular Antiviral Silencing." Plant Physiology **174**(2): 1067-1081.
Qiu, Y., *et al.* (2017). "Characterization of siRNAs derived from cucumber mosaic virus in infected tobacco plants." Archives of virology **162**(7): 2077-2082.
Rabie, M., *et al.* (2017). "Phylogeny of Egyptian isolates of Cucumber mosaic virus (CMV) and Tomato mosaic virus (ToMV) infecting Solanum lycopersicum." European Journal of Plant Pathology **149**(1): 219-225.
Rahoutei, J., *et al.* (2000). "Inhibition of photosynthesis by viral infection: effect on PSII structure and function." Physiologia Plantarum **110**(2): 286-292.
Ramachandran, V. and X. Chen (2008). "Small RNA metabolism in Arabidopsis." Trends in plant science **13**(7): 368-374.
Ramesh, S. V., *et al.* (2017). "Transcriptome-wide identification of host genes targeted by tomato spotted wilt virus-derived small interfering RNAs." Virus Research **238**(Supplement C): 13-23.
Revers, F., and García, J. A. (2015). "Molecular biology of potyviruses." In Advances in virus research Academic Press. 92:101-199.
Riechmann, J. L.,*et al*. (1995). "Processing of the plum pox virus polyprotein at the P3-6K1 junction is not required for virus viability." Journal of General Virology 76(4): 951-956.
Riechmann, J. L., *et al.* (1995). "Processing of the plum pox virus polyprotein at the P3-6K1 junction is not required for virus viability." Journal of General Virology **76**(4): 951-956.
Rock, C. D. and R. S. Quatrano (1995). The role of hormones during seed development. Plant hormones, Springer**:** 671-697.
Rodamilans, B.,*et al*. (2018). "Plant viral proteases: Beyond the role of peptide cutters." Frontiers in plant science 9.
Rodamilans, B., *et al*. (2015). "RNA polymerase slippage as a mechanism for the production of frameshift gene products in plant viruses of the Potyviridae family." Journal of virology 89(13): 6965-6967.
Rodriguez-Cerezo, E., *et al.* (1993). "Association of the non-structural P3 viral protein with cylindrical inclusions ion potyvirus-infected cells." Journal of General Virology **74**: 1945-1949.

Roossinck, M. J. (1997). "Mechanisms of plant virus evolution." Annual review of phytopathology **35**(1): 191-209.
Roossinck, M. J. (2002). "Evolutionary history of Cucumber mosaic virus deduced by phylogenetic analyses." Journal of virology **76**(7): 3382-3387.
Roossinck, M. J. (2003). "Plant RNA virus evolution." Current Opinion in Microbiology **6**(4): 406-409.
Roossinck, M. J., *et al.* (1999). "Rearrangements in the 5′ nontranslated region and phylogenetic analyses of cucumber mosaic virus RNA 3 indicate radial evolution of three subgroups." Journal of virology **73**(8): 6752-6758.
Ruduś, I., *et al.* (2014). "Wound-induced expression of DEFECTIVE IN ANTHER DEHISCENCE1 and DAD1-like lipase genes is mediated by both CORONATINE INSENSITIVE1-dependent and independent pathways in Arabidopsis thaliana." Plant Cell Reports **33**(6): 849-860.
Ruiz-Ruiz, S., *et al.* (2011). "Citrus tristeza virus infection induces the accumulation of viral small RNAs (21–24-nt) mapping preferentially at the 3′-terminal region of the genomic RNA and affects the host small RNA profile." Plant Molecular Biology **75**(6): 607-619.
Rybicki, E. P. (2015). "A Top Ten list for economically important plant viruses." Archives of virology **160**(1): 17-20.
Sako, N. (1980). "Loss of aphid transmissibility of turnip mosaic virus." Phytopathology **70**(7): 647-649.
Sambrook, J. and D. W. Russell (2006). "Preparation of genomic DNA from mouse tails and other small samples." Cold Spring Harbor Protocols **2006**(1): pdb. prot4038.
Sanjuán, R.,*et al*. (2010). "Viral mutation rates". Journal of virology 84(19): 9733-9748.
Sanjuán, Rafael, and Pilar Domingo-Calap. (2016) "Mechanisms of Viral Mutation." Cellular and Molecular Life Sciences 73(23): 4433–4448.
Sansregret, R., *et al.* (2013). "Extreme resistance as a host counter-counter defense against viral suppression of RNA silencing." PLoS pathogens, **9**(6), e1003435.
Schaad, M. C., *et al.* (1997). "Formation of plant RNA virus replication complexes on membranes: role of an endoplasmic reticulum-targeted viral protein." The EMBO journal **16**(13): 4049-4059.
Schaad, M. C., *et al.* (1997). "Formation of plant RNA virus replication complexes on membranes: role of an endoplasmic reticulum-targeted viral protein." The EMBO journal **16**(13): 4049-4059.
Schenk, P. M., *et al.* (2000). "Coordinated plant defence responses in Arabidopsis revealed by microarray analysis." Proceedings of the National Academy of Sciences **97**(21): 11655-11660.
Schneider, W. L. and M. J. Roossinck (2000). "Evolutionarily related Sindbis-like plant viruses maintain different levels of population diversity in a common host." Journal of virology **74**(7): 3130-3134.
Schneider, W. L. and M. J. Roossinck (2001). "Genetic diversity in RNA virus quasispecies is controlled by host-virus interactions." Journal of virology **75**(14): 6566-6571.
Schuck, J., *et al.* (2013). "AGO/RISC-mediated antiviral RNA silencing in a plant in vitro system." Nucleic acids research **41**(9): 5090-5103.
Schultz, E. (1921). A transmissible mosaic disease of Chinese cabbage, mustard, and turnip, US Government Printing Office.
Schwach, F., *et al.* (2005). "An RNA-Dependent RNA Polymerase Prevents Meristem Invasion by Potato Virus X and Is Required for the Activity But Not the Production of a Systemic Silencing Signal." Plant Physiology **138**(4): 1842-1852.
Seo, S., *et al.* (2001). "Transient accumulation of jasmonic acid during the synchronized hypersensitive cell death in tobacco mosaic virus-infected tobacco leaves." Molecular plant-microbe interactions **14**(2): 261-264.
Sewelam, N., *et al.* (2013). "Ethylene response factor 6 is a regulator of reactive oxygen species signaling in Arabidopsis." PLoS ONE **8**(8): e70289.

Shaikhibrahim, Z., *et al.* (2009). "Med8, Med18, and Med20 subunits of the Mediator head domain are interdependent upon each other for folding and complex formation." Proceedings of the National Academy of Sciences **106**(49): 20728-20733.
Shi, B., *et al.* (2016). "Identification and regulation of host genes related to Rice stripe virus symptom production." New Phytol **209**(3): 1106-1119.
Shimura, H., *et al.* (2011). "A Viral Satellite RNA Induces Yellow Symptoms on Tobacco by Targeting a Gene Involved in Chlorophyll Biosynthesis using the RNA Silencing Machinery." PLoS pathogens **7**(5): e1002021.
Shukla, D. T. and C. Ward (1989). "Structure of potyvirus coat proteins and its application in the taxonomy of the potyvirus group." Adv Virus Res **36**: 273-314.
Silva, T. F., *et al.* (2011). "Profile of small interfering RNAs from cotton plants infected with the polerovirus Cotton leafroll dwarf virus." BMC Molecular Biology **12**(1): 40.
Simon, A. E., and Miller, W. A. (2013). "3′ cap-independent translation enhancers of plant viruses." Annual review of microbiology 67:21-42.
Smith, N. A., *et al.* (2011). "Viral Small Interfering RNAs Target Host Genes to Mediate Disease Symptoms in Plants." PLoS Pathog **7**(5): e1002022.
Spoel, S. H. and X. Dong (2012). "How do plants achieve immunity? Defence without specialized immune cells." Nature Reviews Immunology **12**(2): 89-100.
Spoel, S. H., *et al.* (2003). "NPR1 modulates cross-talk between salicylate-and jasmonate-dependent defence pathways through a novel function in the cytosol." The Plant Cell **15**(3): 760-770.
Stein, E., *et al.* (2008). "Systemic resistance in Arabidopsis conferred by the mycorrhizal fungus Piriformospora indica requires jasmonic acid signaling and the cytoplasmic function of NPR1." Plant and cell physiology **49**(11): 1747-1751.
Steinhauer, D. A., *et al.* (1992). "Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase." Gene **122**(2): 281-288.
Sternes, P. R. and R. L. Moyle (2015). "Deep Sequencing Reveals Divergent Expression Patterns Within the Small RNA Transcriptomes of Cultured and Vegetative Tissues of Sugarcane." Plant Molecular Biology Reporter **33**(4): 931-951.
Sullivan, C. S., and Ganem, D. (2005). "MicroRNAs and viral infection." Molecular cell. **20**(1), 3-7.
Sun, P.,*et al*. (2010). "Structural determinants of tobacco vein mottling virus protease substrate specificity." Protein Science, 19(11): 2240-2251.
Suzuki, M., *et al.* (1995). "Point mutations in the coat protein of cucumber mosaic virus affect symptom expression and virion accumulation in tobacco." Journal of General Virology **76**(7): 1791-1799.
Swarbrick, P. J., *et al.* (2006). "Metabolic consequences of susceptibility and resistance (race-specific and broad-spectrum) in barley leaves challenged with powdery mildew." Plant, Cell & Environment **29**(6): 1061-1076.
Takagi, Y., *et al.* (2006). "Head Module Control of Mediator Interactions." Molecular cell **23**(3): 355-364.
Takeda, A., *et al.* (2008). "The mechanism selecting the guide strand from small RNA duplexes is different among argonaute proteins." Plant and cell physiology **49**(4): 493-500.
Takeda, A., *et al.* (2008). "The mechanism selecting the guide strand from small RNA duplexes is different among argonaute proteins." Plant Cell Physiol **49**(4): 493-500.
Thatcher, L. F., *et al.* (2009). "Fusarium oxysporum hijacks COI1-mediated jasmonate signaling to promote disease development in Arabidopsis." The Plant Journal **58**(6): 927-939.
Thivierge, K., *et al*. (2008). "Eukaryotic elongation factor 1A interacts with Turnip mosaic virus RNA-dependent RNA polymerase and VPg-Pro in virus-induced vesicles." Virology 377(1): 216-225.
Ton, J., *et al.* (2009). "The multifaceted role of ABA in disease resistance." Trends in plant science **14**(6): 310-317.

Tromas, N., *et al.* (2014). "Estimation of the in vivo recombination rate for a plant RNA virus." Journal of General Virology **95**(3): 724-732.
Utsugi, S., *et al.* (1998). "Arabidopsis thaliana vegetative storage protein (VSP) genes: gene organization and tissue-specific expression." Plant Molecular Biology **38**(4): 565-576.
Valli, A.,*et al*. (2014).” A novel role of the potyviral helper component proteinase contributes to enhance the yield of viral particles.” Journal of virology JVI-01010.
Valli, A.,*et al*. (2014). “A novel role of the potyviral helper component proteinase contributes to enhance the yield of viral particles.” Journal of virology, JVI-01010.
Valli, A.,*et al*. (2007). “Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family Potyviridae.” Journal of General Virology 88(3): 1016-1028.
Van Loon, L. C. and E. A. Van Strien (1999). "The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins." Physiological and Molecular Plant Pathology **55**(2): 85-97.
Várallyay É, *et al*. (2010). “Plant virus-mediated induction of miR168 is associated with repression of ARGONAUTE1 accumulation.” EMBO J. **29**: 3507-19.
Vaucheret, H. (2005). "MicroRNA-dependent trans-acting siRNAs production." Science Signaling **2005**(300): pe43.
Vaucheret, H. (2006). "Post-transcriptional small RNA pathways in plants: mechanisms and regulations." Genes & development **20**(7): 759-771.
Vaucheret, H. (2008). "Plant argonautes." Trends in plant science **13**(7): 350-358.
Voinnet, O. (2001). "RNA silencing as a plant immune system against viruses." Trends in Genetics **17**(8): 449-459.
Voinnet, O. (2009). "Origin, Biogenesis, and Activity of Plant MicroRNAs." Cell **136**(4): 669-687.
Walsh, J. A. and C. E. Jenner (2002). "Turnip mosaic virus and the quest for durable resistance." Molecular Plant Pathology **3**(5): 289-300.
Waltermann, A., and Maiss, E. (2006). “Detection of 6K1 as a mature protein of 6 kDa in plum pox virus-infected Nicotiana benthamiana.” Journal of general virology 87(8): 2381-2386.
Wang, A., and Krishnaswamy, S. (2012). “Eukaryotic translation initiation factor 4E-mediated recessive resistance to plant viruses and its utility in crop improvement.” Molecular plant pathology 13(7): 795-803.
Wang, C., *et al.* (2016). "The Mediator Complex Subunits MED14, MED15, and MED16 Are Involved in Defence Signaling Crosstalk in Arabidopsis." Frontiers in Plant Science **7**: 1947.
Wang, J., *et al.* (2016). "Cotton Leaf Curl Multan Virus-Derived Viral Small RNAs Can Target Cotton Genes to Promote Viral Infection." Front Plant Sci **7**: 1162.
Wang, M. B., *et al.* (2012). "RNA silencing and plant viral diseases." Mol Plant Microbe Interact **25**(10): 1275-1285.
Wang, M.-B. and M. Metzlaff (2005). "RNA silencing and antiviral defence in plants." Current Opinion in Plant Biology **8**(2): 216-222.
Wang, M.-B., *et al.* (2004). "On the role of RNA silencing in the pathogenicity and evolution of viroids and viral satellites." Proceedings of the National Academy of Sciences of the United States of America **101**(9): 3275-3280.
Wang, W., *et al.* (2014). "The brassinosteroid signaling network — a paradigm of signal integration." Current Opinion in Plant Biology **21**: 147-153.
Wang, X.-B., *et al.* (2011). "The 21-nucleotide, but not 22-nucleotide, viral secondary small interfering RNAs direct potent antiviral defence by two cooperative argonautes in Arabidopsis thaliana." The Plant Cell **23**(4): 1625-1638.
Wang, Y., *et al.* (2004). "Functional analysis of the Cucumber mosaic virus 2b protein: pathogenicity and nuclear localization." Journal of General Virology **85**(10): 3135-3147.
Wang, Y.,*et al*. (2018). “RNA 3-dimensional structural motifs as a critical constraint of viroid RNA evolution.” PLoS pathogens 14(2): e1006801.
Waterhouse, P. M., *et al.* (2001). "Gene silencing as an adaptive defence against viruses." Nature **411**(6839): 834-842.

Wei, T., and Wang, A. (2008). “Biogenesis of cytoplasmic membranous vesicles for plant potyvirus replication occurs at endoplasmic reticulum exit sites in a COPI-and COPII-dependent manner.” Journal of virology 82(24):12252-12264.
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): e1000962.
Wen, R. H., and Hajimorad, M. R. (2010). “Mutational analysis of the putative pipo of soybean mosaic virus suggests disruption of PIPO protein impedes movement.” Virology 400(1): 1-7.
Whitham, S. A. and Y. Wang (2004). "Roles for host factors in plant viral pathogenicity." Current Opinion in Plant Biology **7**(4): 365-371.
Whitham, S. A., *et al.* (2003). "Diverse RNA viruses elicit the expression of common sets of genes in susceptible Arabidopsis thaliana plants." The Plant Journal **33**(2): 271-283.
Williamson, C. E., *et al*. (2014). “Solar ultraviolet radiation in a changing climate.” Nature Climate Change 4(6): 434.
Worobey, M. and E. C. Holmes (1999). "Evolutionary aspects of recombination in RNA viruses." Journal of General Virology **80**(10): 2535-2543.
Xia, Z., *et al.* (2014). "Characterization of Small Interfering RNAs Derived from Sugarcane Mosaic Virus in Infected Maize Plants by Deep Sequencing." PLoS ONE **9**(5): e97013.
Xia, Z., *et al.* (2016). "Synergistic infection of two viruses MCMV and SCMV increases the accumulations of both MCMV and MCMV-derived siRNAs in maize." Scientific Reports **6**: 20520.
Xie, M. and B. Yu (2015). "siRNAs-directed DNA Methylation in Plants." Current Genomics **16**(1): 23-31.
Xie, Z., *et al.* (2001). "An important role of an inducible RNA-dependent RNA polymerase in plant antiviral defence." Proceedings of the National Academy of Sciences **98**(11): 6516-6521.
Xie, Z., *et al.* (2004). "Genetic and Functional Diversification of Small RNA Pathways in Plants." PLoS Biol **2**(5): e104.
Xie, Z., *et al.* (2005). "DICER-LIKE 4 functions in trans-acting small interfering RNA biogenesis and vegetative phase change in Arabidopsis thaliana." Proceedings of the National Academy of Sciences of the United States of America **102**(36): 12984-12989.
Xu, H. and M. C. Heath (1998). "Role of calcium in signal transduction during the hypersensitive response caused by basidiospore-derived infection of the cowpea rust fungus." The Plant Cell **10**(4): 585-597.
Xu, Y., *et al.* (2012). "Population Diversity of Rice Stripe Virus-Derived siRNAs in Three Different Hosts and RNAi-Based Antiviral Immunity in Laodelphgax striatellus." PLoS ONE **7**(9): e46238.
Yang, C., *et al.* (2007). "Spatial analysis of Arabidopsis thaliana gene expression in response to Turnip mosaic virus infection." Molecular plant-microbe interactions **20**(4): 358-370.
Yang, L., *et al.* (2006). "SERRATE is a novel nuclear regulator in primary microRNA processing in Arabidopsis." The Plant Journal **47**(6): 841-850.
Yang, Z., *et al.* (2006). "HEN1 recognizes 21–24 nt small RNA duplexes and deposits a methyl group onto the 2′ OH of the 3′ terminal nucleotide." Nucleic acids research **34**(2): 667-675.
Yasaka, R., *et al.* (2014). "The Temporal Evolution and Global Spread of <italic>Cauliflower mosaic virus</italic>, a Plant Pararetrovirus." PLoS ONE **9**(1): e85641.
Yoshida, N.,*et al*. (2012). “Variability in the p1 gene helps to refine phylogenetic relationships among leek yellow stripe virus isolates from garlic.” Archives of virology 157(1): 147-153.
Yu, D., *et al.* (2003). "Analysis of the Involvement of an Inducible Arabidopsis RNA-Dependent RNA Polymerase in Antiviral Defence." Molecular plant-microbe interactions **16**(3): 206-216.
Zhang, C., *et al.* (2015). "Biogenesis, Function, and Applications of Virus-Derived Small RNAs in Plants." Frontiers in Microbiology **6**: 1237.
Zhang, H., *et al.* (2014). "An Rrp6-like Protein Positively Regulates Noncoding RNA Levels and DNA Methylation in Arabidopsis." Molecular cell **54**(3): 418-430.
Zhang, J., *et al.* (2010). "Expression of pathogenesis related genes in response to salicylic acid, methyl jasmonate and 1-aminocyclopropane-1-carboxylic acid in Malus hupehensis (Pamp.) Rehd." BMC Research Notes **3**: 208-208.

Zheng, X., *et al.* (2007). "Role of Arabidopsis AGO6 in siRNAs accumulation, DNA methylation and transcriptional gene silencing." The EMBO journal **26**(6): 1691-1701.
Zhu, H. and H. Guo (2012). "The role of virus-derived small interfering RNAs in RNA silencing in plants." Science China Life Sciences **55**(2): 119-125.
Zhu, J.-Y., *et al.* (2013). "Brassinosteroid signalling." Development **140**(8): 1615-1620.
Zilian, E., & Maiss, E. (2011). "Detection of plum pox potyviral protein–protein interactions in planta using an optimized mRFP-based bimolecular fluorescence complementation system." Journal of general virology 92(12): 2711-2723.
Zipfel, C. (2014). "Plant pattern-recognition receptors." Trends in Immunology **35**(7): 345-351.
Zvereva, A. S. and M. M. Pooggin (2012). "Silencing and innate immunity in plant defence against viral and non-viral pathogens." Viruses **4**(11): 2578-2597.