INTERACTIONS BETWEEN PARARETROVIRUSES AND THEIR PLANT HOSTS

Melanie Kalischuk
MSc. 2004

A Thesis
Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfilment of the Requirements for the Degree

DOCTORATE OF PHILOSOPHY, BIOMOLECULAR SCIENCE

Department of Biological Sciences
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

INTERACTIONS BETWEEN PARARETROVIRUSES AND THEIR PLANT
HOSTS

MELANIE KALISCHUK

Date of Defence: 15 April 2015

Dr. D. Johnson Supervisor	Professor	Ph.D.
Dr. S. Rood Thesis Examination Committee Member	Professor	Ph.D.
Dr. J. Thomas Thesis Examination Committee Member	Professor	Ph.D.
Dr. D. Gaudet Internal Examiner	Research Scientist AAFC	Ph.D.
Dr. K. Eastwell External Examiner	Professor WSU	Ph.D.
Dr. A. Hontela Examination Committee Chair	Professor	Ph.D.

Dedicated

*To my ever supportive husband Larry, son Nicholas
and parents Vic and Ruth.*

Abstract

To defend themselves against all types of pathogens, plants have evolved an array of defense strategies to prevent or attenuate invasion by potential attackers. *Brassica rapa* exposed to 50 ng purified *Cauliflower mosaic virus* (CaMV; Family Caulimoviridae, genus *Caulimovirus*) virions prior to the bolting stage produced significantly larger seeds and greater CaMV resistance than mock-inoculated treatment. Differences in defense pathways involving fatty acids, primary and secondary metabolites were detected in pathogen resistant and susceptible progeny. To extend the interplay of host and pathogen interactions involving members of the dsDNA plant viruses, the *Rubus yellow net virus* (RYNV) genome was characterised and contained numerous nucleic acid binding motifs, multiple zinc finger-like sequences and domains associated with cellular signaling. Silencing as a mechanism to combat virus accumulation was indicated by an uneven genome-wide distribution of 22-nt length virus-derived small RNAs with strong clustering to small regions distributed over both strands of the RYNV genome.

Acknowlegements

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AGO	argonaute
AzA	azelaic acid
BAK1	brassinosteroid insensitive 1-associated kinase
BLAST	basic alignment search tool
BCVBV	*Bougainvillea spectabilis chlorotic vein-banding virus*
BRNV	*Black raspberry necrosis virus*
BSV	*Banana streak virus*
CaMV	*Cauliflower mosaic virus*
CDD	Conserved protein domain database
CLCV	*Cabbage leaf curl geminivirus*
CP	coat protein
CSSV	*Cacao swollen shoot virus*
CTAB	hexadecyltrimethylammonium bromide
CYMV	*Citrus yellow mosaic virus*
DA	dehydroabiential
DCL	dicer-like protein
DEGS	differentially expressed genes
DBV	*Dioscorea bacilliform virus*
DIR1	defective in induced resistance 1
DMV	*Dracaena mottle virus*
DRB	double-stranded RNA-binding protein
DRM2	defective in RNA-directed DNA methylation 2
ETI	effector-triggered immunity
FDR	false discovery rate
FLS2	flagellin sensing 2
flg22	bacterial flagellin 22
FPKM	fragments per kilobase of transcript per million mapped reads
GO	Gene Ontology database
GVBV	*Gooseberry vein banding virus*
HEN1	methyltransferase Hua enhancer 1
HR	hypersensitive response
HST	exportin-5 homolog; synonymous with HASTY
JA	jasmonic acid
KTSV	*Kalanchoe top-spotting virus*
KEGG	Kyoto Encyclopedia of Genes and Genomes pathway database
LPS	lipopolysaccharides
LRR	leucine-rich repeat
MAMP	microbe-associated molecular patterns
MeSA	methyl salicylate
MET1	methyltransferase 1
miRNA	micro-RNA
MP	movement protein
natsi RNA	natural-antisense transcript-derived small RNA

NB-LRR	nucleotide-binding site/leucine-rich repeat
NES	nuclear export signal
NPR1	transcription co-factor non-expression of PR gene 1
ORF	open reading frame
PAMP	pathogen-associated molecular patterns
PCNA	proliferating cell nuclear antigen
Pfam	Protein families database
PBV	*Pineapple bacilliform virus*
PLANT	plant cis-acting regulatory DNA elements database
PlantCARE	plant cis-acting regulatory element database
PEST	proline, glutamic acid, serine and threonine sequence
piRNA	Piwi interacting short RNAs
PR	aspartic protease
PRR	plant recognition receptor
PRs	pathogenesis-related proteins
Pst	*Pseudomonas syringae* pathovar *tomato* strain DC3000
PTI	pattern-triggered immunity
PTGS	post-transcriptional gene silencing
PVX	*Potato virus X*
RLMV	*Raspberry leaf mottle virus*
RdRP	RNA dependent RNA polymerase, synonymous with RDP
R-gene	resistance gene
RISC	RNA-induced silencing complex
RMD	raspberry mosaic disease
RNAi	RNA interference
RNase	ribonuclease H
ROS	reactive oxygen species
RT	reverse transcriptase
RTBV	*Rice tungro bacilliform virus*
RYNV	*Rubus yellow net virus*
SA	salicylic acid
SAR	systemic acquired resistance
siRNA	small interferring RNA
SMART	simple modular architecture research tool
SCBV	*Sugarcane bacilliform virus*
SYLSV	*Spiraea yellow leaf spot virus*
TBV	*Taro bacilliform virus*
TE	transposable element
TGMV	*Tomato golden mosaic virus*
TGS	transcriptional gene silencing
TMV	*Tobacco mosaic virus*
TRV	*Tobacco rattle virus*
T3SS	type III protein secretion system
Ve	Verticilium transmembrane receptor-like protein

VIGS   virus induced gene silencing
vsRNA   viral small RNAs

1. Interactions between pararetroviruses and their hosts

1.1 Thesis overview

To defend themselves against pathogens, plants have evolved an array of defense strategies to prevent or attenuate invasion by potential attackers (Mithöfer & Boland 2012; Pieterse et al. 2014). There were several significant historical breakthroughs in the understanding of host and plant pathogen interactions. In 1942, Harold Henry Flor proposed the gene-for-gene relationship for explaining a form of disease resistance in plants. This discovery became paramount in that it directed research towards identifying and studying the structure of host plant resistance (R) genes and pathogen avirulent (avr) genes in almost every type of economically important field crop (Boller & Felix 2009; Rempel et al. 2014). The identification of R genes eventually lead to the use of vertical and horizontal disease resistance strategies for many crop plants severely impacted by disease (Senthil-Kumar & Mysore 2013; Rempel et al. 2014). The gene-for-gene relationship further provided a model and the infrastructure for studying cell signalling including those involved in the hypersensitive response (HR), reactive oxygen species (ROS) production, phytohormone activity and systemic acquired resistance (SAR) (Wi et al. 2013; Rosli et al. 2013). A greater understanding of innate immunity in plants has been gained, and recognition of the differences between compatible (i.e., host susceptible) and incompatible (i.e., host resistance) host pathogen interactions is now possible (Kathiria et al. 2010).

The small interfering RNA pathway (siRNA) is another form of innate immunity in plants (Ruiz-Ferrer & Voinnet 2009). The siRNA pathway is one of the multiple pathways in RNA silencing, a larger more complex pathway that is conserved in plants, humans, animals and other life-forms. It is responsive during virus protection, regulation of gene expression and genome stability through methylation and chromatin formation and modification (Fusaro et al. 2006). Mello and Fire (2005) demonstrated that dsRNA was more than ten times more effective at down regulating mRNA than single-stranded RNA. Subsequently, Fire and Mello received the 2006 Nobel Prize in Physiology and Medicine for their work on the identification of dsRNA as a trigger of gene silencing. Another milestone in understanding plant innate immunity was the identification of virus-encoded silencing suppressors. For a successful infection to occur, a virus must overcome the plant's gene silencing defense mechanism (Baulcombe 2004). Such a complex system for RNA silencing provides ample opportunity for viruses to develop ways to avoid the host's defense machinery (Voinnet 2001; Baulcombe 2004). The combative relationship between the host and pathogen is sculptured by natural selection whereby viruses have evolved silencing suppressors to overcome the silencing machinery in the host plant (Baulcombe 2004). It is estimated that every virus contains at least one silencing suppressor, and that for many of them the exact mode of action remains unknown (Voinnet 2001). However, the broad level at which the suppressor proteins act within the gene silencing pathway is often known (Voinnet 2001).

Gene silencing and effector triggered resistance (i.e., gene-for-gene relationship) are associated with SAR (Lu et al. 2003; Durrant & Dong 2004; Carrillo-Tripp et al. 2006). A signal sends a message to distal tissue rendering the entire organism resistant to further infection of pathogens (Lindbo et al. 1993; Baulcombe 1996). SAR is a form of broad-spectrum resistance that protects the plant by producing an abundance of salicylic acid (SA) and pathogenesis-related proteins (PRs) in areas distal to the initial infection site (Conrath 2011). With the onset of SAR, biochemical signaling turns on defense faster and stronger (Conrath 2011). Along with alterations in cell signaling the onset of SAR is also associated with massive transcriptional re-programming (Wang et al. 2005; Wang & Dong 2011).

The small interfering RNA pathway and R-gene mediated resistance are examples of within-generation pathogen defense strategies. More recently, attention has been directed towards across-generation pathogen defense (Kathiria et al. 2010; Luna et al. 2012; Slaughter et al. 2012). Pathogen resistance was demonstrated to be a transgenerational trait (Boyko et al. 2007; Kathiria et al. 2010). Transgenerational pathogen resistance is similar to SAR in that it produces a broad-spectrum resistance and it produces a faster response of protection from further pathogen attacks (Kathiria et al. 2010). Both abiotic and biotic stress types are known to induce transgenerational responses; however, responses induced by biotic stress remains to be explored in more detail (Luna et al. 2012). In fact, only a limited number of pathogen types such as ss(+)RNA

viruses, Gram negative bacteria and synthetic chemical elicitors have been examined for the onset of a transgenerational effect (Kathiria et al. 2010; Slaughter et al. 2012; Luna et al. 2012). Another limitation of these studies is that most transgenerational studies have been performed in model plants (e.g., *Arabidopsis thaliana* and *Nicotiana tabacum*) (Kovalchuk et al. 2003; Boyko et al. 2007; Kathiria et al. 2010). It is unclear if economically important crop plants would produce similar transgenerational responses as examined in common laboratory plants. To strengthen the scope of transgenerational response research, it would be advantageous to examine whether transgenerational effects occur in economically important crop plants. Broadening the scope of transgenerational research would be enhanced by exposure to other groups of plant pathogens, such as dsDNA viruses.

The first chapter of this thesis provides an overview of the literature encompassing innate immunity in plants. It provides the background on various disease resistance strategies such as pattern triggered immunity (PTI), effector triggered immunity (ETI), systemic acquired resistance (SAR) and transgenerational inheritance. All of these strategies are used by plants to avoid or ameliorate the effects of pathogens. The second chapter of this thesis focuses on the transgenerational effects of the economically important plant, *Brassica rapa* following exposure to biotic stress of a dsDNA virus, *Cauliflower mosaic virus* (CaMV). Results from this study suggests that *B. rapa* exposed to 50 ng purified cauliflower mosaic virions at four weeks following germination produces

stress resistance in progeny. The next chapter of this thesis attemps to broaden the application of transgenerational inheritance to include all dsDNA plant viruses including the second major genus of the family *Caulimoviridae,* genus *Badnavirus* (Kalischuk et al. 2013). In this chapter an unknown virus displaying distinct leaf mottling and mosaic symptoms was characterized and identified as *Rubus yellow net virus*. Results from this study suggests that the host plant (*Rubus idaeus* L.) exhibits gene silencing activity against RYNV attack but that there appears to be an evolutionary arms race between the host and pathogen.

2. Literature Review: Disease resistance in plants

2.1. Introduction

Plants have evolved an array of defense strategies to prevent or attenuate attack or invasion (reviews include Mithöfer & Boland 2012). Physical barriers and constitutive antimicrobial metabolites are the first lines of defense that plants use to prevent invasion by foreign attackers, mainly pathogens (Senthil-Kumar & Mysore 2013). A waxy epidermis cuticle and components of primary and secondary cell walls like the fatty acid substances cutin, suberin and waxes act as barriers to prevent the access of pests to plant tissue. Physical structures such as high densities of trichomes, thorns, spines or prickles effectively deter pests (Fordyce & Agrawal 2001). Many secondary metabolites produced by plants display antimicrobial and/or insecticidal properties including, essential oil terpenoids, phenolic compounds and alkaloids (Lu et al. 2013; Radulovic et al. 2013). Plants of the family Brassicaceae contain glucosinolates which are sulfur containing compounds that are stored separately in an uninjured plant cell from the enzyme myrosinase (Agerbirk & Olsen 2013). Upon mechanical damage to a plant cell, such as during insect feeding or pathogen infection, glucosinolates are broken down by myrosinase to produce nitriles, isothiocyanates, thiocyanates, epithionitriles and vinyl ozazolidinethione (Buxdorf et al. 2013). Isothiocyanates, known as mustard oil are commonly associated with crushing *Brassica* species tissue. The compound becomes volatile taking on additional roles to initial pathogen defense such as cell signaling (Conti et al. 2008). In some cases, plant

breeders use constitutive defenses to enhance the performance of plants that are important in agriculture. For example, "Hairy Canola" was developed as a crop plant that was resistant to flee beetle feeding by increasing trichome density on seedling canola with transformation of the *Arabidopsis* GLABRA3 (GL3) gene (Gruber et al. 2006; Soroka et al. 2011).

Although the production of physical barriers and constitutive metabolites prevent most pathogens from infecting plants, this is costly (Rojas et al. 2014). If pests and pathogens are absent in the environment, a plant that invests largely into structure or continuous production of metabolites for pathogen defense will be shorter on resouces for other biological processes such as those involved in growth, photosynthesis and reproduction (Meldau et al. 2012; Huot et al. 2014). It is the equilibrium in these trade-offs that become important in the evolution of pathogen defense strategies and fitness and phenotype. Inducible defenses are an alternative to constitutive defense strategies that reduce fitness costs by minimizing the expenditures by only being produced upon pathogen attack or defense demand (Pieterse et al. 2014). This avoids costly allocation of resources to unneccesary pathogen defense and ensures energy expenditures for growth and development. This review will provide an overview of the recent advances in induced pathogen defense strategies that contribute to innate immunity including basal defense, effector-triggered immunity and short interfering RNA gene silencing. SAR is another form of induced disease resistance and the recent advances in the knowledge of it observed within and across generations (i.e.,

transgenerational effects) will be explored. These analyses highlight a need for additional studies to examine more types of pathogens (i.e., dsDNA viruses) in establishing specific pathogen and non-specific beneficial transgenerational effects. Moreover, the exploration of transgenerational effects in host plants should expand to include economically important crop species. The next stage in transgenerational research involves the evaluation of generating disease resistance through a transgenerational means, thus offering a novel method for producing disease resistance in economically important plants.

2.2. Innate immunity in plants

Some pathogens are adapted to the host and able to overcome physical and structural barriers or constitutive metabolites. Disease resistance in plants occurs by innate immunity with resistance (R) genes or RNA silencing playing an important role (Jones & Dangl 2006; Lee et al. 2009; Padmanabhan et al. 2009). Innate immunity involves pathogen recognition followed by a rapid activation of defense responses to produce an incompatible response (Torres et al. 2006).

2.2.1. Pattern triggered immunity

On the external surface of pathogens are elicitors also known as pathogen-associated molecular patterns (PAMPs) or in bacteria, microbe-associated molecular patterns (MAMPs) (Zipfel & Felix 2005). PAMPs are essential components to a group of pathogens and some examples include bacterial flagellin, lipopolysaccharides (LPS) from Gram negative bacteria, peptidoglycans from Gram positive bacteria or fungal chitin (Gómez-Gómez & Boller 2000;

Petutschnig et al. 2010). Plants use pattern recognition receptors (PRRs) to recognize PAMPs which induce pattern-triggered immunity (PTI) (Lee et al. 2009). The two classes of PRR are transmembrane receptor kinases and transmembrane receptor-like proteins (Boller & Felix 2009). All PRR contain highly conserved domains (Kawchuk et al. 2009). The transmembrane receptor kinases have an extracellular leucine-rich repeat (LRR), a transmembrane domain and a cellular kinase. The transmembrane receptor-like proteins are similar to the transmembrane receptor kinases but lack the cellular signaling domain. Most of the information known about PTI comes from studies on FLAGELLIN SENSING 2 (FLS2) initially isolated from the model plant *Arabidopsis thaliana* (Gómez-Gómez & Boller 2000). FLS2 is a LRR receptor kinase that binds to bacterial flagellin (i.e., flg22) and then interacts with the related protein BRASSINOSTEROID INSENSITIVE 1-ASSOCIATED KINASE 1 (BAK1) to initiate further signaling cascades involved in PTI (reviewed in Boller & Felix 2009).

Similarly, Ve1 and Ve2 are transmembrane receptor-like proteins that were isolated from tomato (Diwan et al. 1999; Kawchuk 2001). Ve1 was associated with resistance to *Verticillium albo-atrum*, the causal agent of verticillium wilt disease that affects tomato, potato, strawberry, sunflower, cucurbits and eggplant (Kawchuk et al. 2001). The involvement of Ve2 in the interaction for verticillium wilt and early dying resistance is likely but, the exact mechanism remains to be characterized (Kawchuk, unpublished). Ve receptors initially appeared to be

9

different from other known transmembrane receptor-like proteins because, in addition to the conserved domains found in other transmemebrane receptor-like proteins, Ve contains an extracellular coiled-coil domain and a cellular endocytosis domain (Kawchuk et al. 2001). Although the involvement of endocytosis in defense signaling for plants remains to be fully documented, endocytosis was demonstrated to be important in recycling of specific mammalian cell-surface receptors (Zanoni et al. 2011). Ve receptor-mediated endocytosis provides a mechanism through which cells selectively capture ligands and remove signaling receptors from their surfaces, thereby actively responding to changing disease pressures (Kawchuk, personal communication). Although many PAMPs and MAMPs have been recognized, the number of PRR identified and isolated in plants remains relatively limited (Zipfel 2009).

2.2.2. Gene-for-gene resistance: Effector triggered immunity

Effector-triggered immunity (ETI), also known as gene-for-gene resistance, is another important component of resistance in plants (Jones & Dangl 2006; Dodds & Rathjen 2010). ETI occurs with pathogens that are able to suppress PTI by injecting race-specific effectors into the host cell (Dangl & Jones 2001; Jones & Dangl 2006; Feng & Zhou 2012). To activate ETI, plants have intracellular R-genes that respond directly or indirectly to the effectors (van der Biezen & Jones 1998). Most R-genes involved in ETI belong to the nucleotide-binding site/leucine-rich repeat (NB-LRR) family (reviewed in Kawchuk et al. 2009) . *Pseudomonas syringae* pathovar (pv.) *tomato* strain DC3000 (Pst) is a

Gram negative bacterium and produces approximately 30 effectors that are injected into the host cell using a type III protein secretion system (T3SS) (Lindeberg et al. 2012). AvrPtoB is one of the effectors secreted by Pst and it has an E3 ligase function that targets the flagellin receptor FLS2 for degradation through the proteasome (i.e., PTI) (Mathieu et al. 2014). To counter this virulence strategy put forth by the pathogen, the plant contains the Prf R-protein that detects the AvrPtoB activity, triggering ETI and rendering the pathogen avirulent (Xing et al. 2007; Gutierrez et al. 2010).

In most cases, R-gene mediated resistance is accompanied by an oxidative burst, consisting of rapid production of reactive oxygen species (ROS) (Wi et al. 2012; Rosli et al. 2013). ROS production is associated with a hypersensitive cell death response (i.e., hypersensitive response (HR)), which is a form of programmed cell death that limits the access of pathogens to water and nutrients (Greenberg & Yao 2004). Downstream of the HR, phytohormones like salicyclic acid (SA), jasmonic acid (JA) or ethylene signal activation of pathogenesis-related (PR) proteins or other molecules involved in an active defense response. For example, the SA dependent signaling pathway leading to activation of pathogenesis-related protein-1 (PR1) is thought to be involved in resistance to biotrophic pathogens and are parts of the cascade of events during an incompatible interaction (reviewed by Glazebrook 2005; Vasyukova & Ozeretskovskaya 2007; Foyer & Noctor 2013). PTI and ETI are examples of the continuous arms race between pathogens and plants whereby in this example,

pathogens interfere with plant PTI using effectors and plants mount strong ETI responses upon recognition of the pathogen, shaping plant and pathogen evolution.

2.2.3. Systemic acquired resistance

Effector triggered immunity induces SAR that is a form of broad-spectrum resistance that protects the plant by producing an abundance of SA and defense proteins in areas distal to the initial infection site (Metraux et al. 1990; Malamy et al. 1990; Durrant & Dong 2004). With the onset of SAR, biochemical signaling turns on defenses faster and stronger making further infections of pathogens difficult or impossible (Conrath 2011). The onset of SAR is also associated with massive transcriptional re-programming, with dependence on the transcription co-factor NON-EXPRESSION OF PR GENE 1 (NPR1) and its associated transcription factors (Wang et al. 2005; Wang & Dong 2011).

Communication from the initial site of infection, through the vasculature to distal plant tissues is necessary for SAR to occur. Initially it was thought that SA might be the signal; however, the reverse was demonstrated with the accumulation of SA at the distal tissue and not at the site of infection required for a SAR response (Vernooij et al. 1994). Although the long-distance signal responsible for SAR remains unknown, there are recent interests in exploring the involvement of the lipid-transfer protein DEFECTIVE IN INDUCED RESISTANCE 1 (DIR1), and long-distance metabolites such as methyl salicylate (MeSA),

dehydroabietinal (DA) and azelaic acid (AzA) in SAR signaling (reviewed in Shah
& Zeier 2013; Fu & Dong 2013).

2.2.4. Gene silencing

Post-transcriptional gene silencing (PTGS) and transcriptional gene silencing
(TGS) are other forms of innate immunity in plants, although the latter remains to
be characterized (Ruiz-Ferrer & Voinnet 2009). The PTGS pathway is one of the
multiple pathways in RNA silencing, a larger, more complex pathway that is
conserved in plants, humans, animals and other life-forms, and is responsive
during virus protection, regulation of gene expression and genome stability
through methylation and chromatin formation and modification (Fusaro et al.
2006).

2.2.5. Post-transcriptional gene silencing as an antiviral mechanism

The first biological function of gene silencing was the discovery of its antiviral
mechanism in plants (Lindbo et al. 1993; Baulcombe 1996). There is significant
support for PTGS as an antiviral mechanism in plants. First it was discovered
that PTGS was induced by a transgene-containing viral sequence that then
targeted homologous viral RNAs for silencing to confer virus resistance (Lindbo
et al. 1993). This key study was followed by a number of other important studies
supporting this model (for reviews see Lomonossoff 1995; Baulcombe 1996;
Beachy 1997). Another line of evidence for PTGS as an antiviral mechanism
includes a vast number of studies that used reverse genetics to knock out
important components to the gene silencing pathways (Mourrain et al. 2000;

Dalmay et al. 2001; Xie et al. 2001; Morel et al. 2002). In these studies the inactive silencing components rendered plants more susceptible to virus infections. Probably the most convincing piece of evidence supporting PTGS as an antiviral mechanism comes from demonstration that almost all viruses contain virulence factors that influence the immunity of plants to viruses (Anandalakshmi et al. 1998; Brigneti et al. 1998; Kasschau & Carrington 1998). After PTGS was discovered in plants, it was also found to be an antiviral mechanism in other organisms and it is called quelling in fungi (Cogoni et al. 1996) and RNA interference (RNAi) in *Drosophila* (Li et al. 2002; Wang et al. 2006), mammals (Pfeffer et al. 2004; Sullivan et al. 2005; Li & Ding 2005) and nematodes (Lu et al. 2005; Wilkins et al. 2005).

As an antiviral mechanism, double-stranded RNA is targetted from either replicating DNA or RNA viruses. Over 70% of plant viruses have single-stranded ss(+) RNA genomes that replicate by a virus encoded RNA-dependent RNA polymerase (RdRP) producing a dsRNA replicative intermediate. Some viruses have double-stranded (ds) RNA in which the genome itself is a source of dsRNA. Double-stranded DNA and ssDNA viruses are also abundant and they produce DNA and RNA replicative intermediates. Efficiency of PTGS is increased with nucleic acids containing hairpins, inverted repeats and other high secondary structure (Smith et al. 2000; Fusaro et al. 2006). Antisense and hairpin RNA technologies often rely on these secondary structures for improvements in their products. In addition, viruses are often used in virus-induced gene silencing

(VIGS) because they are efficient at triggering gene silencing (reviewed in Lu et al. 2003; Carrillo-Tripp et al. 2006). VIGS relies on the use of viral vectors carrying a transgene that can trigger PTGS causing the degradation of its homologue within the plant. *Tobacco rattle virus* (TRV) has been used in VIGS to elucidate mechanisms of floral scent production in petunia (Spitzer et al. 2007) and to facilitate the dissection of the flavonoid biosythesis pathway (Nagamatsu et al. 2007). DNA viruses are also used in VIGS, for example *Tomato golden mosaic virus* (TGMV) was used to silence a meristematic gene called proliferating cell nuclear antigen (PCNA) in *Nicotiana benthamiana* (Peele et al. 2001). Plant transcripts or aberrant RNA may also act as templates with endogenous RNA-dependent RNA polymerase (RdRP) activity producing dsRNA (Xie et al. 2001).

Production of small RNA involves the cleavage of a dsRNA precursor by RNase III-like proteins known as DICER-like protein (Bernstein et al. 2001). *A. thaliana* contains four Dicer-like (DCL1-DCL4) proteins. DCL1 in *A. thaliana* processes 18-21-nt-long miRNAs; DCL2 processes 22-nt natural-antisense transcript-derived small RNAs (natsiRNA) and some viral siRNAs; DCL3 processes 24-nt siRNAs that mediate transcriptional gene silencing and maintenance of DNA methylation, and DCL4 processes 21-nt trans-acting small-interfering RNAs (ta-siRNA) along with the majority of viral siRNAs (Xie et al. 2004). Interestingly, dicer-like proteins display functional redundancy (Fusaro et al. 2006; Bouche et al. 2006; Deleris et al. 2006; Diaz-Pendon et al. 2008).

These studies demonstrated that the loss of function of both DCL4 and DCL2 are

necessary and sufficient to make plants highly susceptible to ssRNA viruses.

DCL4 appears to be the main dicer-like protein involved in PTGS; however,

DCL2 will compensate in the production of virus derived siRNA when DCL4 is

inactive (Fusaro et al. 2006). DNA viruses like *Cauliflower mosaic virus* (CaMV)

or *Cabbage leaf curl virus* (CLCV) are exceptions where all four DCLs are

involved in siRNA biogenesis in infected hosts (Blevins et al. 2006; Moissiard &

Voinnet 2006). Double-stranded DNA plant viruses have a long intergenic region

(e.g., CaMV 35 S promoter region) and therefore DCL1 plays a role in antiviral

defense by cleaving these areas that resemble the secondary structure of

miRNAs. DCL3 is involved in antiviral mechanisms during the nuclear phase of

plant DNA virus multiplication.

Humans and nematodes have one dicer-like protein responsible for the

production of both siRNAs and miRNAs (Lau et al. 2012; Gao et al. 2014). There

are two dicer-like proteins in *Drosophila* (DCL1 and DCL2) that dice pre-miRNA

and dsRNA, respectively. Dicing produces sRNAs with a characteristic 2-nt

overhang at the 3' ends (Bernstein et al. 2001). Double-stranded RNA-binding

proteins (DRBs) facilitate DCLs in the dicing process (Hiraguri et al. 2005;

Nakazawa et al. 2007). DRBs do not contain hierarchical redundancy as do

DCLs (Curtin et al. 2008). DRB4 interacts with DCL4 in the *A. thaliana* antiviral

defence silencing pathway (Qu et al. 2008).

Upon dicing in *A. thaliana*, the sRNA 3' overhanging ends are 2'-O-methylated by methyltransferase HUA ENHANCER 1 (HEN1), and this protects them from degradation (Li et al. 2005). It is known that HEN1 participates in the antiviral RNA silencing pathways because *Arabidopsis hen1* mutant exhibited increased susceptibility to *Cucumber mosaic virus* (CMV) and accumulated a five-fold increase in virus titer in comparison to wild type (Boutet et al. 2003). Once dicing is completed, the stabilized sRNA duplexes are then retained in the nucleus for transcriptional gene silencing (TGS) at the chromatin-level or exported to the cytoplasm, possibly via the exportin-5 homolog HASTY (HST), for PTGS (Bollman et al. 2003). The sRNAs are incorporated into a large ribonucleaprotein complex, the RNA-induced silencing complex (RISC) (Hannon 2002).

The RISC contains an ARGONAUTE (AGO) protein that has a sRNA-binding PAZ domain and a PIWI domain (Parker et al. 2005). The PIWI domain has structural similarity to RNaseH and has endonucleolytic activity to digest the target RNA with use of a guide strand, a process called slicing. *A. thaliana* contains 10 predicted family members of AGO (AGO1-AGO10) with established roles for AGO1, AGO4, AGO6 and AGO7 in sRNA directed silencing. Slicer activity has been demonstrated for AGO1, AGO4 and AGO7 (Song et al. 2004; Qu et al. 2008). AGO4 and AGO6 are required in the TGS pathway (Ziberman et al. 2003). AGO1 and AGO7 are required in PTGS. AGO1 is the main slicer for viral RNAs because it has a higher affinity than AGO7 for more compact structures (Qu et al. 2008). AGO1 has additional roles in miRNA and other

17

siRNA pathways (Vaucheret et al. 2004). During the slicing process, the two RNA strands in the sRNA unwind and become separated. One strand is preferentially incorporated into the RISC to guide the complex to degrade transcripts or viral genomes, whereas the other is rapidly degraded. Preference for the guide strand by AGO is based on the weakest base-pairing interaction at the 5' terminus (Khvorova et al. 2003; Eamens et al. 2009).

2.2.6. Transcriptional gene silencing

Transposable elements (TE) and foreign nucleic acids can be regulated by DNA methylation, through the process of TGS. Transposable elements are DNA sequences that have the capacity to insert and excise within a genome (Haas et al. 2009). They have been identified in all organisms, from prokaryotes to eukaryotes and can contribute a substantial amount to the size of a genome. For example, TE comprise approximately 12% of *C. elegans* genome (Stein et al. 2003), 37% of mouse genome (Waterston et al. 2002), 45% of human genome (Lander et al. Nature 2001) and up to >80% of plant genomes (SanMiguel 1996). From bacteria to humans, TEs have accumulated over time and continue to be a main player in genomic evolution. The activity of TE can positively or negatively impact a genome (reviewed in Bennetzen & Wang 2014). For example, TEs can play a significant role in genomic evolution by promoting gene inactivation, modulating gene expression or inducing homologous and/or non-homologous recombination. However, TEs are also able to produce various genetic alterations upon insertion, excision, duplication or translocation, rendering

deleterious effects or disease in the host. It is not surprising that organisms have

evolved mechanisms for controlling the translocation of TEs. One of the

mechanisms is mediated through TGS, involving DNA methylation. Generally,

silencing of transposons occurs through methylation whereas, activation of

transposons occurs through loss of methylation (Ito & Kakutani 2014).

The heterochromatin formation pathway produces 24-nt sRNA (hcRNA) that

mediate TGS through maintenance of DNA methylation and chromatin structure

(Hamilton et al. 2002). In *A. thaliana*, the pathway uses DCL3 to cleave dsRNA

derived from endogenous transcripts. TEs provide dsRNA templates used in

TGS through mechanisms that are not fully understood (for models see Matzke

et al. 2000 and Waterhouse et al. 2001). Double stranded RNA derived from

endogenous transcripts can also be generated through the pathway involving

RNA dependent RNA polymerase 2 (RDR2) and RDR6 (also known as

SGS2/SDE1) (Dalmay et al. 2000; Tang et al. 2003; Voinnet 2008). AGO4 is the

main AGO involved in TGS (Qi et al. 2006). It binds 24-nt siRNA that either

guide cleavage or *de novo* DNA methylation. At a site homologous to the 24-nt

sRNAs *de novo* methylation is carried out through RNA-directed DNA

methylation (RdRM) with various DNA methyltransferases such as

METHYLTRANSFERASE1 (MET1) or DEFECTIVE IN RNA-DIRECTED DNA

METHYLATION2 (DRM2) and others (see Matzke et al. 2009). Other key

enzymes involved in TGS are DNA glycosylases, lysases, chromatin remodelling

proteins and RNA polymerases (reviewed in Matzke et al. 2009). *De novo* DNA

methylation is the methylation of cytosines in all sequence contexts (CG, CNG and CNN where N is A,T or C). Promoters and sometimes coding regions are the targets for DNA methylation. DNA methylation at a promoter prevents binding of factors necessary for transcription.

In *Drosophila* and vertebrate germ lines, TE silencing relies on Piwi Argonaute proteins and a class of sRNAs known as Piwi interacting short RNAs (piRNAs) (Aravin et al. 2007; Hartig et al. 2007; Klattenoff & Theurkauf 2008). Interestingly dicer-like proteins are not involved in producing piRNAs. In *Drosophila* the majority of piRNAs target TEs, while in vertebrates most piRNAs target repetitive sequences with only a minority complementary to TEs. The ping-pong model has been proposed in vertebrates as a cyclic feedback process that alternatively cleaves sense and antisense TEs (for a review see Klattenhoff & Theurkauf 2008).

2.2.7. Transitive silencing

Fungi, plants and nematodes encode eukaryotic RNA-dependent RNA polymerase (RdRP or RDR) that can generate new sources of dsRNA leading to amplification of silencing in the organism. Transitive gene silencing occurs in both plants and nematodes (Sijen et al. 2001). In this process, the virus or TE derived RNA pool is amplified using RdRP. This leads to the propagation and spread of the silencing signal beyond the region initially targeted for gene silencing. *A. thaliana* encodes six RdRPs that work with DCLs to control the biogenesis of sRNAs. The function of RDR2 is required for the production of 24

nt siRNAs by DCL3, which are involved mainly in hcRNA pathway and sometimes antiviral gene silencing. RDR6 is involved in the production of siRNAs by DCL1, DCL2 or DCL4. RdRP1 influences virus replication and is involved in antiviral defense in plants (Xie et al. 2001). During the loss of function of RdRP1, *A. thaliana* expressed enhanced accumulation of viral RNAs and increased susceptibility to viral infections (Xie et al. 2001; Yu et al. 2003). These studies are consistent with RdRP being involved in virus defense.

2.2.8. Silencing suppressors - A molecular arms race between viruses and hosts

RNA silencing is a highly complex system with numerous proteins and processes. The multiple pathways makes defense more difficult to evolve. The mechanics of a plant virus to infect a host relies on the ability of a virus to overcome the plant's gene silencing defense mechanisms. Such a complex system for RNA silencing provides ample opportunity for viruses to develop ways of avoiding the host's defense machinery. Viruses encode suppressor proteins capable of interfering with various steps of the PTGS and TGS pathways. It is estimated that every virus contains at least one silencing suppressor, and for many of them, the exact mode of action remains unknown. However, the level at which they act within the gene silencing pathway is often known.

A major class of silencing suppressors are dsRNA-binding proteins, which usually have a high affinity for binding duplex siRNAs and long dsRNAs. Examples include *Closterovirus* P21 (Chapman et al. 2004), *Cucumber mosaic*

21

virus 2b (Li et al. 1999), *Nodavirus* B2 (Li et al. 2002) and *Influenza virus A* NS1

(Li et al. 2004). The *Tombusvirus* P19 is also a dsRNA-binding protein (Silhavy

et al. 2002) but unique to this suppressor, it selects its substrate on the basis of

length of the RNA duplex region (Vargason et al. 2003; Ye at al. 2003). P19

binds 21-nt duplex siRNAs with a much higher affinity than 22-nt dsRNA

duplexes. Another class of suppressors limits cell-to-cell and systemic spread of

the silencing signal and this includes *Potexvirus* P25 and *Cucumovirus* 2b. One

of the first experiments suggesting that P25 inhibits the movement of the signal

showed that systemic silencing did not occur with a deactivated form of P25 with

the reverse also being demonstrated (Voinnet et al. 2000). Later it was shown

that P25 suppression of RNA silencing was required for cell-to-cell movement of

Potato virus X (PVX) (Angell et al. 1996; Bayne et al. 2005). Cucumovirus 2b

inhibits systemic movement in a slightly different manner. It silences the

systemic signal by physically interfering with AGO proteins of the RISC complex

(Zhang et al. 2006). The polerovirus P0 also interferes with AGO1 but the

silencing suppressor mechanism is different. P0 encodes a F-box protein that

promotes ubiquitine-dependent proteolysis of AGO1 and this mechanism of

suppression avoids inhibiting the systemic signal, as it may be important for this

virus that is mostly located in the phloem tissue (Baumberger et al. 2007). DNA

viruses also encode silencing suppressors. For example, CaMV P6 encodes a

silencing suppressor that binds to DRB4 protein in the nucleus of cells (Haas et

al. 2008). By binding, DRB4 is inactivated and unavailable as a cofactor involved

22

in DCL4 dicing activities. The geminiviruses contain AC4, a silencing suppressor that competes with AGOs by binding to single-stranded siRNA and thereby preventing RISC assembly (Chellappan et al. 2005).

From the above examples, it is clear that a viral encoded silencing suppressor is universal; however, they appear to have evolved independently of one another. In fact, it appears that there is a constant arms race between host plants and the pressures of foreign nucleic acids from either viruses or transposition of transposable elements, and this battle is seen across kingdoms. Over the last 15 years since the discovery of gene silencing, we have gained enormous knowledge about this mechanism's involvement in the innate immunity response. However, we are only beginning to understand the mechanics behind the gene silencing components and the parasite's mechanisms that are used to overcome the host's silencing mechanisms. For example, viroids are 350-nt ssRNA plant pathogens that do not encode protein however, they contain an unusually high level of secondary structure. Interestingly, viroids are able to bypass the silencing defense mechanism of their host's but the mechanism is unknown. A viroid's ability to overcome the host's defense is just one example of an unanswered question, amongst many other exciting discoveries that remain to be made. The combative relationship between the host and pathogen is sculptured by the pressures of natural selection whereby viruses have evolved silencing suppressors to overcome the silencing machinery in the host plant (Baulcombe 2004).

2.2.9. Gene silencing and systemic signaling

Gene silencing is associated with cell-to-cell and non-cell autonomous (systemic) signals in plants and nematodes. The systemic signal sends a message to distal tissue rendering the entire organism resistant to further infection by the same virus. In nematodes, the systemic signal requires SID-1, a transmembrane protein that efficiently transports dsRNA that is longer than 100-nt (Winston et al. 2002; Feinberg & Hunter 2003).

In plants, the cell-to-cell signal is believed to move 15-20 cells through the plasmadesmata whereas the systemic signal is believed to pass through the phloem. Presence of a systemic silencing signal is supported by a grafting experiment that showed that a silencing signal for GUS can move from the roots to new shoots through a GUS expressing scion, and that the transmission of the signal was up to 30 cm through wild-type tissue (Palauqui et al. 1997). Biological evidence for this signal is further supported by the systemic action of silencing suppressors p25 of potato virus X (Voinnet et al. 2000) and 2b of cucumber mosaic virus (Diaz-Pendon et al. 2007) and that sRNA has been found in phloem tissue (Sasaki et al. 1998; Ruiz-Medrano et al. 1999).

Many suggestions have been made regarding what the signals might be, although, convincing evidence has been difficult to obtain. Some of the candidates have been long dsRNA (i.e., precursor siRNA), dsRNA molecules or products of dsRNA, which might be produced through a dicer-independent pathway (Brosnan et al. 2007), modified product from the methylated target gene

(Mallory et al. 2001) or 21-24-nt dsRNA. Recently, the most convincing evidence

suggests that a 21-nt siRNA duplex as opposed to their long precursor molecules

is the mobile signal between plant cells (Dunoyer et al. 2010). Also recently, a

systemic signal was identified through grafting experiments supplemented with

sRNA high-throughput sequencing (Molnar et al. 2010). The results from these

experiments suggest that 24-nt sRNA produced using the DCL3/AGO4 pathway

is a systemic signal and may be acting as a signal in the hcRNA pathway (Molnar

et al. 2010).

2.3. Across-generation resistance

Most recently, with the study of disease resistance in plants, attention has

been directed towards across-generation pathogen defense (Kathiria et al. 2010;

Luna et al. 2012; Slaughter et al. 2012). Transgenerational inheritance is a

commonly reported phenomenon whereby it involves pre-treating (i.e., priming) a

plant with a stress to obtain reduced losses associated with subsequent stress

events (Conrath 2011). The priming stress can be abiotic or biotic with

examples of the later arising by pathogen infection or herbaceous insect feeding

(Luna et al. 2012). Examples of abiotic stress producing transgenerational

effects include salinity, shortwave ultraviolet radiation, flood and drought (Boyko

et al. 2010). Although controversy exists among observations of

transgenerational responses, it is noteworthy that the onset of the effect is

dependent upon host species tested, developmental timing and the amount of

stress exposed to the host plant. The remainder of this review focuses on

priming with a biotic stress to produce inherited pathogen resistance.

2.3.1. Evidence for pathogen resistance as a transgenerational response

Plant pathogens can be broadly classified into three main groups: viruses,

bacteria and fungi (including the oomycetes). So far, transgenerational effects

have been observed with stresses in the parent generation being a ss(+)RNA

virus, the bacterium *Pseudomonas syringae DC3000* pv *tomato* (Pst) and a

chemical that mimics a salicylic acid pathogenic defense response (Kathiria et al.

2010; Luna et al. 2012; Slaughter et al. 2012). Typically, disease resistance is

measured directly as pathogen titer or indirectly as a marker by gene expression

of pathogen response genes such as PR1 or phytoalexin deposit (Kathiria et al.

2010). *Nicotiana tabacum* was primed by *Tobacco mosaic virus* (TMV), a

ss(+)RNA virus and the progeny of the treated *N. tabacum* had lower TMV titer,

up-regulation of salicyclic acid pathway marker PRs1 and more abundant callose

deposition than the mock-treated control group (Kathiria et al. 2010). That study

further showed the resistance to be broad-spectrum against a ss(+)RNA virus

(i.e., TMV), the bacterium Pst and the oomycete *Phytophthora brassicae*.

Transgenerational pathogen resistance to virulent bacterial Pst and up-regulation

of pathogen defense genes were observed in the progeny of *Arabidopsis thaliana*

primed with β-aminobutyric acid or an avirulent isolate of Pst (Slaughter et al.

2012). In this study, transgenerational pathogen resistance was measured as

fewer colonies of bioluminiscent Pst in the primed plants than the non-treated

control plants. These studies clearly demonstrate that pathogens can trigger a transgenerational response, however, it is unclear if all pathogen types and all host plants, including economically important ones, can bare a similar transgenerational response.

2.3.2. Mechanisms of transgenerational inheritance

The mechanisms behind pathogen resistance as a transgenerational response remain elusive but the meiotic inheritance of epigenetic signatures, such as DNA methylation, acetylation of histone tails, chromatin remodelling or small RNAs have been suggested to give rise to transgenerational responses (Jablonka & Raz 2009; Jablonka 2013). Genetic epigenic marks change the openness and repressiveness of chromatin to alter cellular transcription during plant development or in response to specific environmental conditions.

Environmental conditions can change the stability and epigenetic state of the genome of an organism and in some cases the changes can be inherited by progeny. Stressful events give rise to dsDNA strand breaks that creates genomic instability in the organism. Homologous recombination is used by the organism to repair the damaged DNA. Interestingly, it was demonstrated that a part of the DNA repair mechanism was remembered by the organism and it was passed on to offspring rendering them more prepared for similar stressful events (Kovalchuk et al. 2003). DNA methylation also plays an important role in pathogen defense. For example, the abundance of DNA methylation was correlated to disease resistance in plants, both within and across generations (Kathiria et al. 2010;

Slaughter et al. 2012). Moreover, genome wide DNA methylation was observed in *A. thaliana* exposed to Pst, avirulent bacteria or SA and it revealed a direct relationship with plant-based immunity (Dowen et al. 2012). Upon pathogen exposure, hypomethylation was detected in areas of the genome associated with pathogen defense response whereas, other regions of the genome, such as areas associated with transposable elements or DNA repeats, were hypermethylated. Repression of transcription is associated with the addition of methyl groups to DNA and provides protection to the genome upon exposure to environmental stress. DNA methylation is often coupled with the establishment of other epigenetic signatures and therefore it is imaginable that histone modifications, silencing of coding or non-coding regions of the genome or chromatin remodeling may be involved with further complexities of defense responses to pathogens. In addition to methylation, small RNAs (21-24-nt) have a role in defense response of *A. thaliana* to bacteria pathogens (Dowen et al. 2012).

For a trait to be transgenerational, a signal must be passed from the parent to the progeny during meiosis (Kovalchuk et al. 2003). At the present time, DNA methylation is the most likely form of epigenetic modification to be passed from the parent to the offspring. Unlike animals in which re-setting of methylation occurs indefinitely, in plants the re-setting of methylation is less definitive (Feng 2010). Methylation is an important process during the early embryonic stage in plants. For example, transposable elements are silenced, genomic imprinting

28

occurs and in flowering plants there is hypomethylation of the endosperm in seeds (Xiao et al. 2006). During the early embryonic stage in sexually reproducing flowering plants, a parent-of-origin effect on seed size was demonstrated (Xiao et al. 2006). The parent with the hypomethylated genome, the gametophyte and both the maternal and paternal genomes of the F1 seed became hypomethylated. If the hypomethylated genome were female, the seed produced was large whereas, if the hypomethylated genome was male, the seed produced was small. With the array of methylation activities occurring at early embryogenesis in plants it is difficult to imagine that the gametophyte stage would be the only stage for resetting methylation (Xiao et al. 2006). Methylation is often coupled with establishment of other epigenetic signatures such as histone modifications, silencing of coding or non-coding regions of the genome or chromatin remodeling. It is possible that along with methylation, some of the other epigenetic signatures are passed through the gametophyte and early embryonic stages thus passing from parent to offspring and influencing phenotype.

2.4. Framework for exploring transgenerational resistance: The host plant *Brassica rapa*

   With the world dependent on plants as a food source, plant biologists continue to search for novel methods for increasing yields while reducing loss of yield from pests and pathogens. Disease resistance developed in a transgenerational manner offers a strategy to obtain disease resistance without genetic modifications and therefore worthwhile testing in economically important crops. The Brassicaceae is an economically important family of flowering plants and consist of 338 genera and 3710 species, which include crops, ornamentals and many weeds. The genus Brassica belongs to the subtribe *Brassicinae* of the Brassicaceae family and comprises approximately 159 species, including both cultivated and wild species. The cultivated species have a worldwide distribution and include oilseed rape, cabbage, cauliflower, broccoli, Brussel sprouts, turnip, kale, mustards and other leafy vegetables. *Brassica napus* and *B. rapa* are economically important crops in Canada and are grown for seed oil, high grade animal feed and biofuels (Rempel et al. 2014). They are major crops in Canada, contributing $2.5 billion per year to the economy with a five-year average of 11 million acres harvested each year and it is ranked third after wheat and barley in terms of acres seeded (Genome Canada and Genome Prarie 2010 http://www.brassicagenomics.ca/ECTG/study.htm).

   The genetic relationship among different Brassica species was established nearly a century ago (Nagaharu 1935) and the classical work continues to be

30

ascribed as the Triangles of U **(Figure 2.1)**. The corners of the U's triangles include three diploid species: *B. rapa* L. (2n = 20; AA), *B. nigra* L. Koch (2n = 16; BB) and *B. oleracea* L. (2n = 18; CC). The other three species in the middle of the triangles are the amphidiploid species and include: *B. juncea* (L.) Czern. (2n = 36; AABB); *B. napus* L. (2n = 38; AACC) and *B. carinata* Braun (2n = 34; BBCC). *Arabidopsis thaliana* was the first plant to have a fully sequenced genome. With a small sized (i.e., 135 megabases) diploid genome *A. thaliana* is extensively used as a model system for studying genetics. Since transgenerational inheritance was previously detected in *A. thaliana*, a member of Family *Brassicaceae*, the next step in the exploration of transgenerational inheritance in the complex allotetraploid *B. napus* would be to examine one of its progenitor diploid species, *B. rapa* or *B. oleracea*.

2.4.1. Priming *Brassica rapa* with a dsDNA virus

Viruses reduce the production and quality of food, fiber and biofuel with estimated yield losses between 10-40% (Kawchuk et al. 2001; Kalischuk et al. 2013). All dsDNA viruses are classified as members of family Caulimoviridae and they are also referred to as pararetroviruses because they use reverse transcription during their replication. Members of *Caulimoviridae* contain a 7.6-8.2-kbp circular dsDNA genome and each strand of DNA has one to three discontinuities. Pararetroviruses have three to seven open reading frames (ORFs) but the organization of the ORFs is the differentiating factor among the
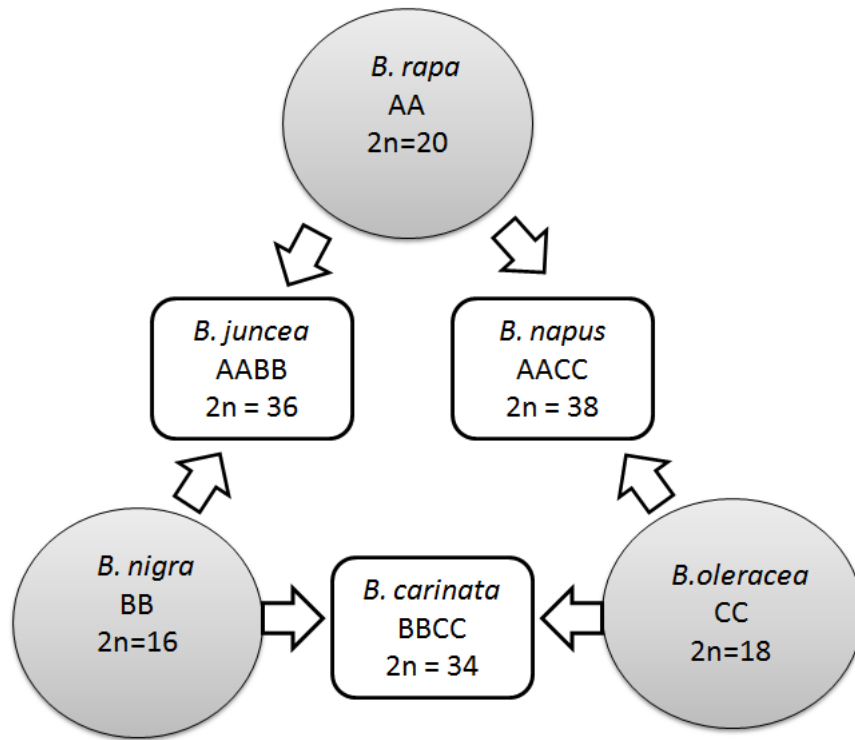
Figure 2.1. The Triangle of U diagram showing genetic relationships between the *Brassica* species.

eight genera. Most *Caulimoviridae* viruses belong to the genera *Badnavirus* or

Caulimovirus.

   Badnaviruses have a 120-150 x 30-nm non-enveloped bacilliform capsid

containing a singular circular dsDNA genome that is 7.3-8.0-kb in size (Lockhart,

1990). Virions enter a host in a non-circulative semi-persistent manner via

feeding by a mealybug or aphid vector. The typical genomic structure for

badnaviruses consists of three open reading frames (ORFs), and all of the genes

are encoded on the same DNA strand (Xu et al. 2011). The first two ORFs

encode proteins that are 17 and 15 kDa, respectively, and little is known about their function. Most information regarding the badnaviruses is derived from the characterization, organization and highly homologous nature of ORF 3 which encodes a 216 kDa polyprotein that is cleaved post-translationally by the viral-encoded aspartic protease to produce a movement protein, a coat protein and a replicase comprised of a reverse transcriptase and ribonuclease H (Laney et al. 2012; Sether et al. 2012). Some examples of badnaviruses include *Rubus yellow net virus*, *Banana streak virus* and *Cacao swollen shoot virus*. Despite the devastating economic impact that badnaviruses can have on plants, the host-pathogen relationships of this genus remains to be explored in detail.

Caulimoviruses differ from badnaviruses in genome organization, particle morphology and transmission is by aphids using an aphid transmission factor (i.e., ORF 2). Virus particle morphology is icosahedral and bacilliform for *Caulimovirus* and *Badnavirus*, respectively. Transmission occurs by aphids and mealybugs and host-specific aphids for *Caulimovirus* and *Badnavirus*, respectively. Replication of CaMV requires the viral translational transactivator protein P6 that is translocated to the nucleus of the host cell and is essential for CaMV infectivity (Haas et al. 2008). P6 is a multifunctional protein with one function demonstrated to be involved in signaling involving SA, JA and ethylene (Love et al. 2012). Another nuclear function of P6 is to suppress RNA silencing, a gene regulation mechanism that provides antiviral capabilities by inactivating the nuclear dsRNA-binding protein (DRB4) that is required by the major plant

antiviral silencing dicer 4 complex (DCL4). Besides a regulatory role, RNA silencing confers a sequence-specific antiviral immunity to plants through virus-derived short interfering RNA (reviewed in Ding & Voinnet 2007). Localization of CaMV components such as P6 in the nucleus of the plant cell provides an opportunity for genome and transcriptome modifications. These modifications may provide transgenerational protection to other pathogens (Kathiria et al. 2010) and influence other phenotypic characteristics.

2.5. Concluding remarks

Plants have many strategies for defending themselves against attack from pests and pathogens. With so many forms of disease resistance (i.e., physical barriers, constitutive metabolites, pattern triggered immunity, effector triggered immunity, gene silencing), it is important to intimately understand and describe the underlying life cycle strategies set forth by the pathogen that could influence the form of disease resistance in host plants. Interestingly, there appears to be long-distant systemic signals that provide another layer of complexity to understanding disease resistance in plants. Under precise stressful situations, a signal is inherited by the progeny contributing to a form of resistance of the same stress type experienced by the parent. Although most of the transgenerational responses are documented in model systems (i.e., *Arabidopsis thaliana*, *Nicotiana tabacum*), there is a need to begin exploring this type of resistance in economically important plants. Expanding transgenerational research to include

crop plants provides a promising approach for developing disease resistance

while avoiding genetic engineering.

3. PRIMING WITH A DOUBLE-STRANDED DNA VIRUS ALTERS *BRASSICA RAPA* SEED ARCHITECTURE AND FACILITATES A DEFENSE RESPONSE

3.1. Background

Plants have developed protection and defense strategies for dealing with adverse environmental conditions and biological stresses. Induced resistance is one of the strategies that plants use to combat pathogens and it involves pre-exposure of a plant to a stress to obtain reduced losses associated with subsequent stressful events (Conrath 2011). There have been several examples of induced resistance being carried over to the next generation, thus giving rise to a transgenerational response (transgenerational response is reviewed by Hauser et al. 2011; Holeski et al. 2012). To some extent, primed plants, whether in the same or next generation, have an elevated level of basal resistance and this prepared state allows for the plant to defend itself from subsequent stress and possibly offering a broad-spectrum resistance (Kathiria et al. 2010; Conrath 2011).

Several pathogens such as single-stranded positive sense ss(+) RNA viruses, Gram-negative bacteria or synthetic chemicals resembling a pathogen elicitor have demonstrated the ability to generate resistance in a transgenerational manner (Kathiria et al. 2010; Slaughter et al. 2012; Luna et al. 2012). *Nicotiana tabacum* was primed by *Tobacco mosaic virus* (TMV), a ss(+)RNA virus, and the progeny of the treated had lower TMV titer, up-regulation of SA pathway marker pathogenesis related 1 (PR1) and more abundant callose deposition than the mock-treated control group (Kathiria et al. 2010). In a second study, transgenerational pathogen resistance to virulent *Pseudomonas syringae DC3000* pv *tomato* (Pst) and up-regulation of pathogen defense genes were

observed in the progeny of *Arabidopsis thaliana* primed with β-aminobutyric acid or an avirulent isolate of Pst. (Slaughter et al. 2012). In a third study, *A. thaliana* was primed with Pst and transgenerational pathogen resistance was measured as fewer colonies of bioluminescent Pst and altered regulation of pathogen defense genes in primed plants in comparison to the non-treated control plants (Luna et al. 2012). These studies clearly demonstrate that these pathogens trigger a limited transgenerational effect; however, to explore transgenerational diversity and specificity, more types of plant pathogen groups need to be used as stressors in the parent generation. One of the broad-based plant pathogen types that remains to be explored includes dsDNA viruses. Cauliflower mosaic virus (CaMV) is a dsDNA virus that uses reverse transcriptase and a RNA intermediate during replication (Scholthof et al. 2011). CaMV infects a host plant, which most often belongs to family Brassicaceae, following transmission in a non-circulative, semi-persistent manner by an aphid vector such as *Myzus persicae* (Haas et al. 2002). The virus systemically infects young host plants and produces severe symptoms including leaf mottling and mosaic, reduced growth, developmental abnormalities and stunting.

Transgenerational effects have been mainly demonstrated to occur in model laboratory plants (i.e., tobacco and *Arabidopsis thaliana*) (Kovalchuk et al. 2003; Boyko et al. 2007; Boyko et al. 2010). To characterize transgenerational effects in economically important plant species, disease responses in *Brassica rapa* was evaluated as the next step in exploring economically important members of the

Brassicaceae family. *B. ra*pa (AA, n=10) is a diploid species and hybridizes with *Brassica oleracea* (CC, n=9) to give rise to the allotetraploid *Brassica napus* (AACC, n=19), also known as canola. Together, *B. napus* and *B. rapa* are major crops in Canada and they are grown for the production of seed oil, high grade animal feed and biofuel (Rempel et al. 2014). This study examined the transgenerational response of *B. rapa* following exposure to CaMV, producing a compatible pathogen interaction that elicits a disease response. Since host response to a pathogen is often dosage-dependent and influenced by the developmental stage of the host plant (Gutiérrez et al. 2012), these variables were examined experimentally for the onset of a transgenerational response in the form of physiological attributes and pathogen resistance. In addition, small RNA transcriptome sequencing was used to identify candidate genes in biochemical pathways or signaling transduction influencing the transgenerational responses. Evidence is presented to test the hypothesis that transgenerational disease resistance is inducible in economically important plant species, resistance persists for extended periods and critical physical and biochemical characteristics of the plant can be improved.

3.2. Material and Methods

3.2.1. Plant material and experimental design

To evaluate transgenerational inheritance, seed was collected from one *B. rapa cv* R018 parent plant and used to generate the first self-fertilized generation (S1). All plants were grown at 20 °C in controlled greenhouse conditions with 16

h photoperiod with light levels of 100 $\mu E.m^2 s^{-1}$. The parental generation was exposed to either 50, 100 or 200 ng of purified CaMV at host plant age of two, three or four weeks following germination. Each treatment included exposure of five plants to the virus and the entire experiment was replicated two times (N = 65 for each experiment). Purification of CaMV virions was carried out according to Hull and Shepherd (1976) and the concentration of particles was determined using spectrophotometry using an $OD_{260} = 7$ equivalent to 1 mg mL^{-1} while adjusting for light scattering. Individual plants were inoculated with one 10-ul suspension containing either 50, 100 or 200 ng of virus and abrasive 250-400 mesh carborundom (Sigma, Canada). Leaves containing the inoculation sites were removed from the plants within 24 h following pathogen exposure to explore signalling rather than pathogen movement throughout the plant. Plants were grown to set seed and the resulting self-fertilized progeny of plants treated with the pathogen were called P0pS1. Control plants consisted of healthy (P0cS1) or plants that were treated with the inoculation buffer consisting of 0.01 M sodium phosphate, pH 7.2 (P0bS1) **(Figure 3.1)**.

3.2.2. Examination of stable complex traits and virus resistance

Progeny were screened for seed size, stable complex traits and CaMV resistance. Seed size was estimated using image analysis software and a transmitted light flatbed scanner as described by Herridge et al. (2011). Briefly, 50-300 seeds per plant were spread onto the scanner bed ensuring that no seeds were touching one another. An image was taken for each plant at a

resolution of 1200 dpi with transmitted light. ImageJ particle analysis software was used to measure seed area using the threshold feature (Abramoff et al. 2004). The greyscale value was 162 and the lower limit of particle analysis was 30,000 μm^2. Other stable complex traits that were measured in the progeny were rate of germination, number of days until first flower, number of days until first 10 flowers, foliage dry weight, total height, root collar diameter, total number of leaves and average crown radius with the latter four measurements being completed at four and eight weeks following germination.

To examine CaMV resistance, progeny were challenged with CaMV and virus titer was measured at 14 days post-inoculation (dpi) using a double antibody enzyme linked immunosorbent assay (DAS-ELISA). Polyclonal and alkaline phosphatase conjugated goat anti-rabbit (Sigma, Canada) were used as the primary and the secondary antibody, respectively. CaMV titer was measured for three to nine progeny for each treatment during three separate experiments. To remove the influence of wounding, the measured variables for pathogen treated plants (P0pS1) were normalized by the average of buffer treated plants (P0bS1) and the pathogen treatments were compared to the healthy treatments. Data were compiled in Microsoft Excel and statistics completed using SAS version 9 (SAS Institute, USA).

3.2.3. cDNA library preparation and sequencing for transcriptome analysis

Fresh leaf tissue was homogenized in liquid nitrogen and total RNA extracted using a Plant/Fungi purification kit (Norgen Biotek Corp., Canada). The quality of

RNA was assessed with agarose gel electrophoresis and spectrophotometrically before generating the mRNA-Seq library and sequences previously described by Kalischuk et al. (2013).

The mRNA-Seq library was generated following Illumina's sample preparation recommendations. Briefly, the poly[A]+ RNA was enriched from 20 μg of total RNA using Oligo(dT) magnetic beads. This RNA was fragmented into small (200-400 bp) fragments and the short fragments were used as templates for random-hexamers to prime first strand followed by second strand cDNA synthesis. Short fragments were purified with a QiaQuick PCR Extraction Kit (Qiagen) and used in cluster generation on Illumina's Cluster Station. Sequencing was performed as paired-end of 101 nt read length on Illumina HiSeqTM 2000. Raw sequencing intensities were extracted and the bases were called using Illumina's real-time analysis software, followed by sequence quality filtering.

3.2.4. Sequence Analysis

All raw reads generated from the sequencer were *de novo* assembled into contigs using the Trinity program (Hass et al. 2013). Assembled contigs were aligned to sequences of 2,487 proteins of *B. rapa* from the NCBI database (http://www.ncbi.nlm.nih.gov/protein/?term=txid3711[Organism:exp]) using BLASTx and homologous genes with the e-value $<10^{-5}$ were identified. The Blast2GO program was used to obtain alignments to the Gene Ontology (GO) database and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway

42

database (http://www.blast2go.org). Trancript abundance evaluated as

Fragments per kilobase of transcript per million mapped reads (FPKM) was

determined by mapping raw reads back to the assembled contigs using the

Tophat and Cufflinks suite (Trapnell et al. 2012).

3.3. Results

To determine whether prior exposure to a pararetrovirus produces a

transgenerational effect conferring disease resistance and other physiological

changes, *Brassica rapa* cv. 018 was inoculated with cauliflower mosaic virus

isolate LRC2010 (CaMV). The parent plants were inoculated with one of three

concentrations of CaMV (50, 100 or 200 ng purified virions) and virus exposure

for the hosts was either one, two or three weeks old following germination.

Control plants included healthy uninoculated or inoculated with virus suspension

buffer (i.e., without virus). Parental plants were grown to set seed and the

resulting self-fertilized progeny were called P0pS1 for pathogen treated, P0cS1

for healthy and P0bS1 for the buffer treated **(Figure 3.1.).**

3.3.1. Low dose of cauliflower mosaic virus applied just before bolting

To explore if phenotype could be used to describe the primed state of *B. rapa*

after exposure to CaMV, the progeny of the CaMV infected plants (P0pS1) were

compared with the control plants (P0cS1 and P0bS1) to evaluate differences in

physiological attributes, progeny development and pathogen resistance.

Agronomical traits of *B. rapa* that included germination rate, flowering rate,

foliage dry weight, total height, root collar diameter, number of leaves and crown

radius were observed while treatment effects of pathogen dose and timing of host plant pathogen exposure were not detected using one-way or two-way analysis of variance (ANOVA). However, when the parent generation was exposed to 50 ng purified CaMV particles per plant, plants that produced larger seeds generated progeny that were more resistant to CaMV (**Figure 3.2**).
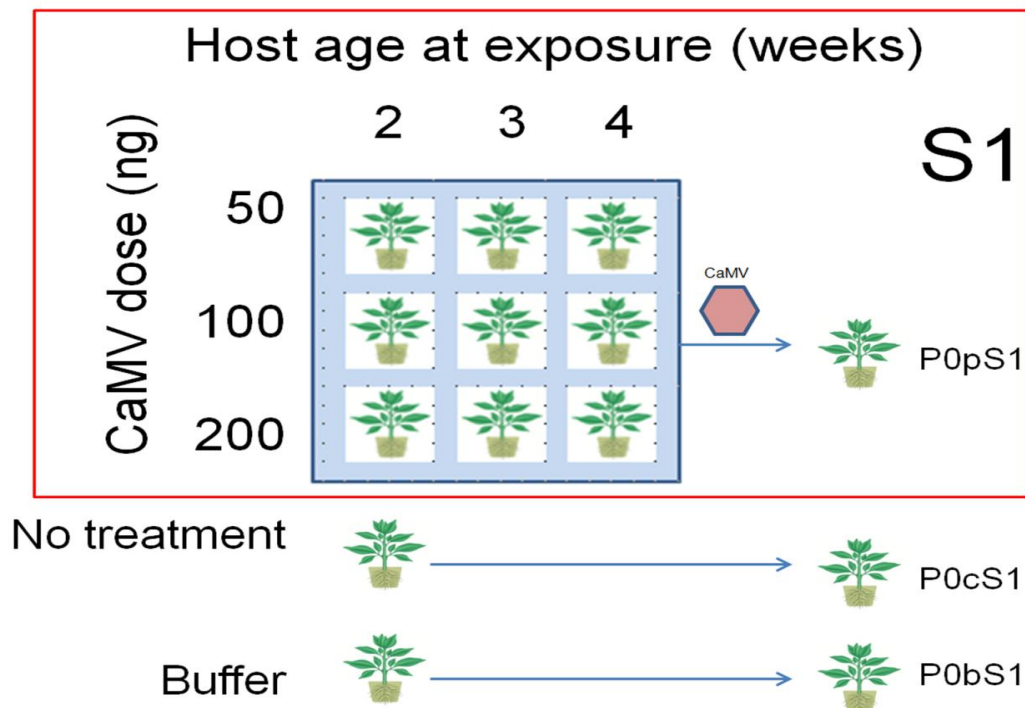
Figure 3.1. Study design for evaluating transgenerational response in *Brassica rapa* following exposure to cauliflower mosaic virus. Seed collected from one *B. rapa cv* R018 parent plant was used to generate the first self-fertilized generation (S1). The parental generation was exposed to either 50, 100 or 200 ng of purified CaMV at host plant age of two, three or four weeks following germination. Inoculation sites were removed at 24 h following pathogen exposure. Parent plants were grown to set seed and the resulting self-fertilized progeny treated with the pathogen were called P0pS1. Control plants consisted of healthy (P0cS1) or plants that were treated with the inoculation buffer which was absent of infectious material (P0bS1). Progeny were screened for stable complex traits and CaMV resistance. Measured variables for pathogen treated plants (P0pS1) were adjusted for by the buffer-treated plants (P0bS1).
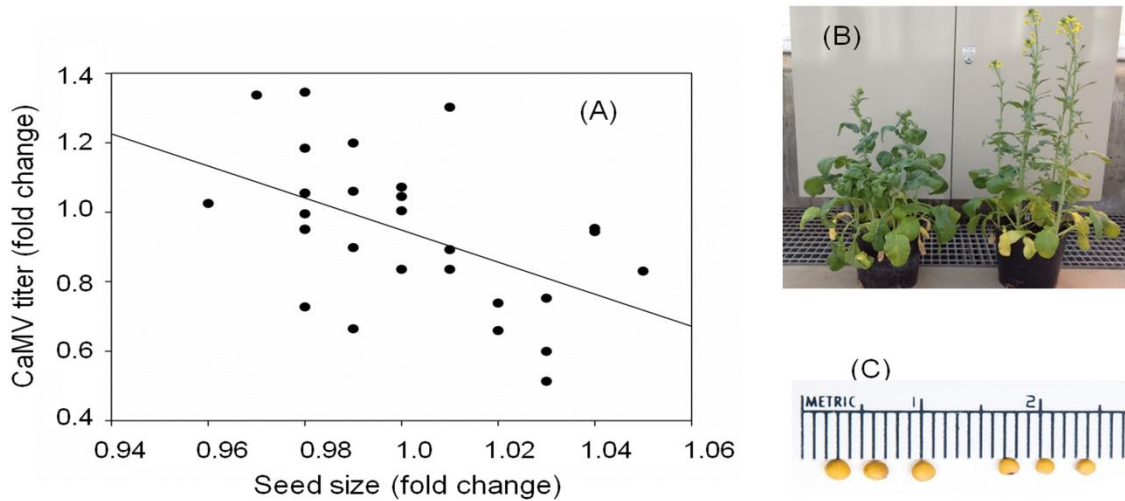
Figure 3.2. Relationship between S1 *B. rapa* seed size and cauliflower mosaic virus (CaMV) titer.

The parents of the S1 were exposed to 50 ng of purified CaMV at either 2, 3 or 4 weeks following germination. (A) The line is defined by the equation CaMV titer (fold change) = 5.62- 4.7 x seed size (fold change). All variables in the equation were significant $p < 0.01$ and the overall model was significant at $p < 0.001$, R^2 = 0.282, N=29. Each data point represents one plant. CaMV titer was determined using a double antibody enzyme linked immunosorbent assay (DAS-ELISA) and sampling included measuring titer for three to nine seeds per plant during three separate experiments. The number of plants measured were 12, 10 and 7 for 2, 3 and 4 week pathogen exposure times, respectively (N=29). Seed size was estimated for 50-300 seeds per plant. The denominator for calculating fold change was based on 1-6 plants that did not receive treatment in the parent generation. **(B)** Symptoms of leaf mosaic, branch stunting, leaf distortion and delayed flowering of host plant *B. rapa* cv 018 infected with CaMV (left) compared to healthy (right). Plants were the same ages and grown under identical conditions. **(C)** Differences in seed size of S1 showing resistance (left) and susceptible (right) to CaMV. Seeds were picked randomly for each treatment and each bar on the scale represents 1 mm (10 bars = 1 cm).
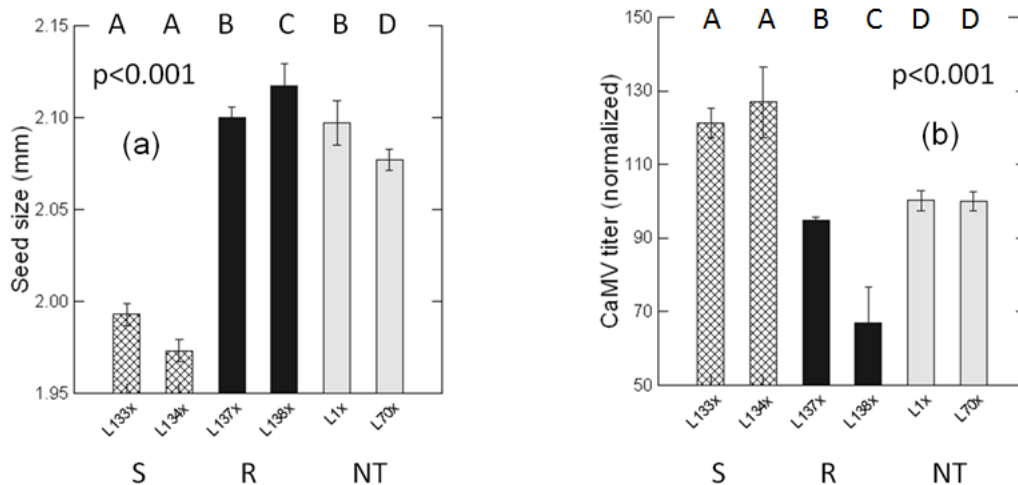
Figure 3.3. Seed size and cauliflower mosaic virus (CaMV) resistance of _B. rapa_ progeny with the parental generation exposed to 50 ng purified CaMV at four weeks following germination.
Seed size (a) and cauliflower mosaic virus (CaMV) titer (b) of B. rapa progeny with the parental generation exposed to 50 ng purified CaMV at four weeks following germination. Hatched bars represent lines of plants with small seeds and rated as susceptible (S) to CaMV. Dark bars represent lines of plants with large seeds and rated as resistant (R) to CaMV. Bars with light shading represent the line of plants that were not challenged during the parent generation (NT). Different letters above the bars indicate significance at p<0.001.

The timing of pathogen exposure was also important in this relationship. For example, later exposure to virus (i.e., at four versus two weeks) increased the number of plants producing larger seeds and the abundance of CaMV resistant progeny. From these analyses and for further comparisons involving P0pS1, only plants exposed to 50 ng purified CaMV at four weeks following germination were characterized and the pathogen treated plants (P0pS1) were separated into two groups designated as P0pS1R for large seed and CaMV resistant phenotype and P0pS1S for small seed and susceptible to CaMV phenotype.

Significant differences in seed size and CaMV titer were detected for healthy (P0cS1), susceptible (P0pS1S) and resistant phenotypes (P0pS1R) **(Figure 3.3)**. The average seed sizes were 2.12 +/-0.08 mm, 1.99 +/-0.08 mm and 2.09 +/-0.09 mm for the resistant, susceptible and healthy treatments, respectively **(Figure 3.3)**. Evaluation of CaMV resistance in the progeny involved double-antibody sandwich enzyme linked immunosorbant assay (DAS-ELISA) and showed that virion titer was significantly lowest in the resistant plants from large seed (P0pS1R) plants, and significantly highest in the susceptible, small seed (P0pS1S) plants **(Figure 3.3)**. Symptom expression following the CaMV challenge also correlated well with CaMV titer measurements obtained by DAS-ELISA. Following CaMV inoculation, plants grown from small seed (P0pS1S) and both control plants (P0cS1 and P0bS1) displayed similar uniform and severe symptoms of CaMV infection including stunting, reduced growth and flowering, leaf deformities, mosaic and mottling within two weeks after being exposed to the virus. Surprisingly, when challenged with the same CaMV pathogen, plants grown from large seeds (P0pS1R) showed a continuum of symptoms ranging from healthy to mild which consisted of a minor intensity rating of stunting and leaf mottling and mosaic. The onset of symptoms that appeared in the P0pS1R plants after being challenged by CaMV were delayed 10 to 14 days in comparison to similarly challenged P0pS1S, P0cS1 or P0bS1.

47

3.3.2. Small RNA sequencing

High throughput small RNA transcriptome sequencing was used to identify differentially expressed loci involved in biochemical pathways that showed a relationship to seed size and/or CaMV resistance. Deep sequenced samples included CaMV challenged P0pS1R, CaMV challenged P0bS1 and healthy P0cS1. Sequences were obtained from a pooled sample of tissue 21 days after pathogen challenge for treatments and time of sampling corresponded to 100% of P0bS1, 20% of P0pS1R and 0% of P0cS1 plants showing CaMV symptoms (N=1). The RNA sequencing produced 55 551 366, 54 658 348 and 54 233 142 raw reads from P0pS1R, P0bS1S and P0cS1, respectively. *De novo* assembly of all sequences generated 39 183 contigs with a mean size of 724 bp that ranged between 201-16 032 bp.

3.3.3. Resistant and susceptible phenotypes exposed to CaMV have contrasting profiles of differentially expressed loci

Genes displaying significant changes in expression were identified in CaMV challenged resistant and susceptible phenotypes. A total of 644 (365 up-regulated and 299 down-regulated) and 3193 (1250 up-regulated and 1943 down-regulated) differentially expressed genes (DEGs) were detected after exposure to CaMV in resistant (P0pS1R) and susceptible (P0bS1S) phenotypes, respectively **(Figure 3.4)**.

3.3.4. Functional annotation of differentially regulated genes

To understand the functions of DEGs, the transcripts yielding a two-fold increase or decrease relative to the healthy control group were mapped to terms in the Gene Ontology (GO) database (Gene Ontology Consortium 2004; **Figure 3.5 and 3.6**). Fisher exact testing was used to determine enrichment of sequences mapping to GO term annotations between the resistant and susceptible phenotypes and was performed using a false discovery rate (FDR) adjusted p-value of <0.01 as the cut-off.

For DEGs up-regulated in resistant and susceptible plants relative to the healthy control, cellular component GO annotations for the up-regulated transcripts that were enriched in the resistant treatment were "cell wall", "intracellular organelle", "chloroplast stroma", "ribosome" and "intra-cellular non-membrane-bound organelle" and represented 27, 82, 11, 29 and 44% of the DEGs, respectively. "Integral to membrane" was the only cellular component annotation enriched for the susceptible treatment whereby representation was 8% and 1% of the total DEGs for the susceptible and resistant phenotypes, respectively.
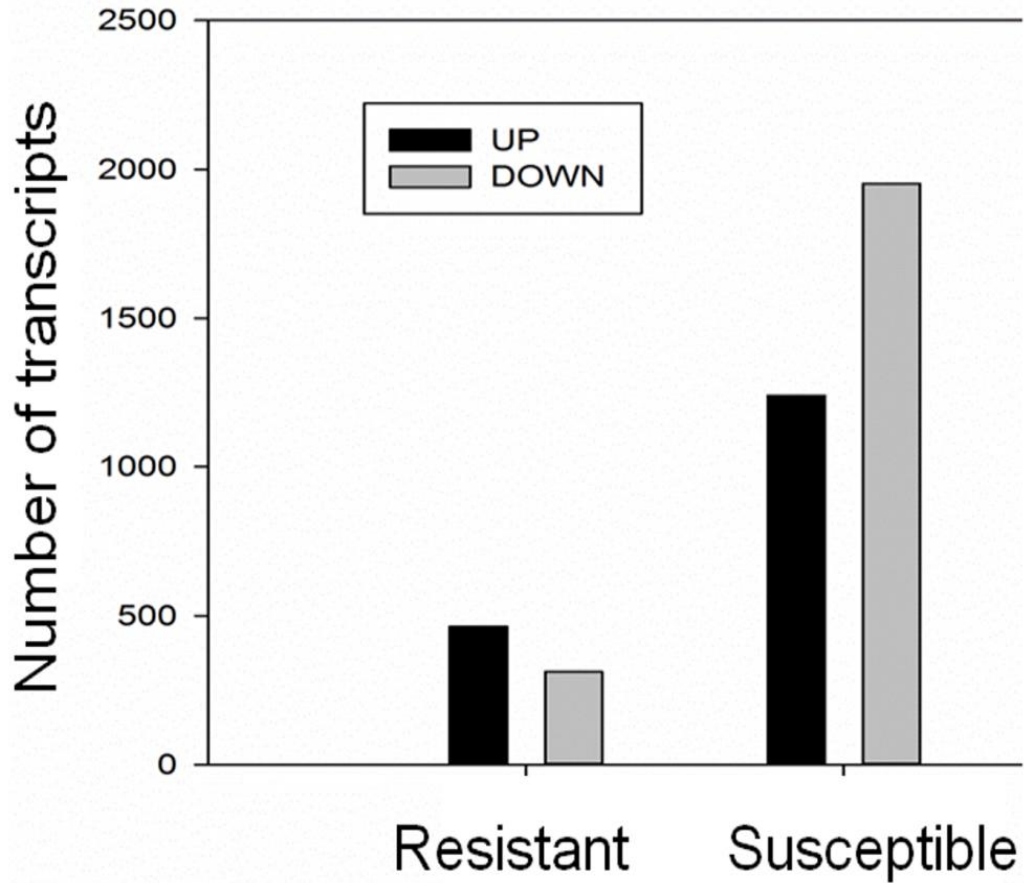
Figure 3.4. Number of differentially expressed genes (DEGs) in S1 *Brassica rapa*, 14 days after being challenged with CaMV.
The cut-off for assessing the regulation of gene expression was based on a two-fold difference. The denominator for calculating resistant and susceptible phenotype was abundance of transcript of the healthy S1 *B. rapa*. Abundance was based on fragments per kilobase of transcript per million mapped reads (FPKM) estimated using Cufflinks RNA-Seq analysis tools.
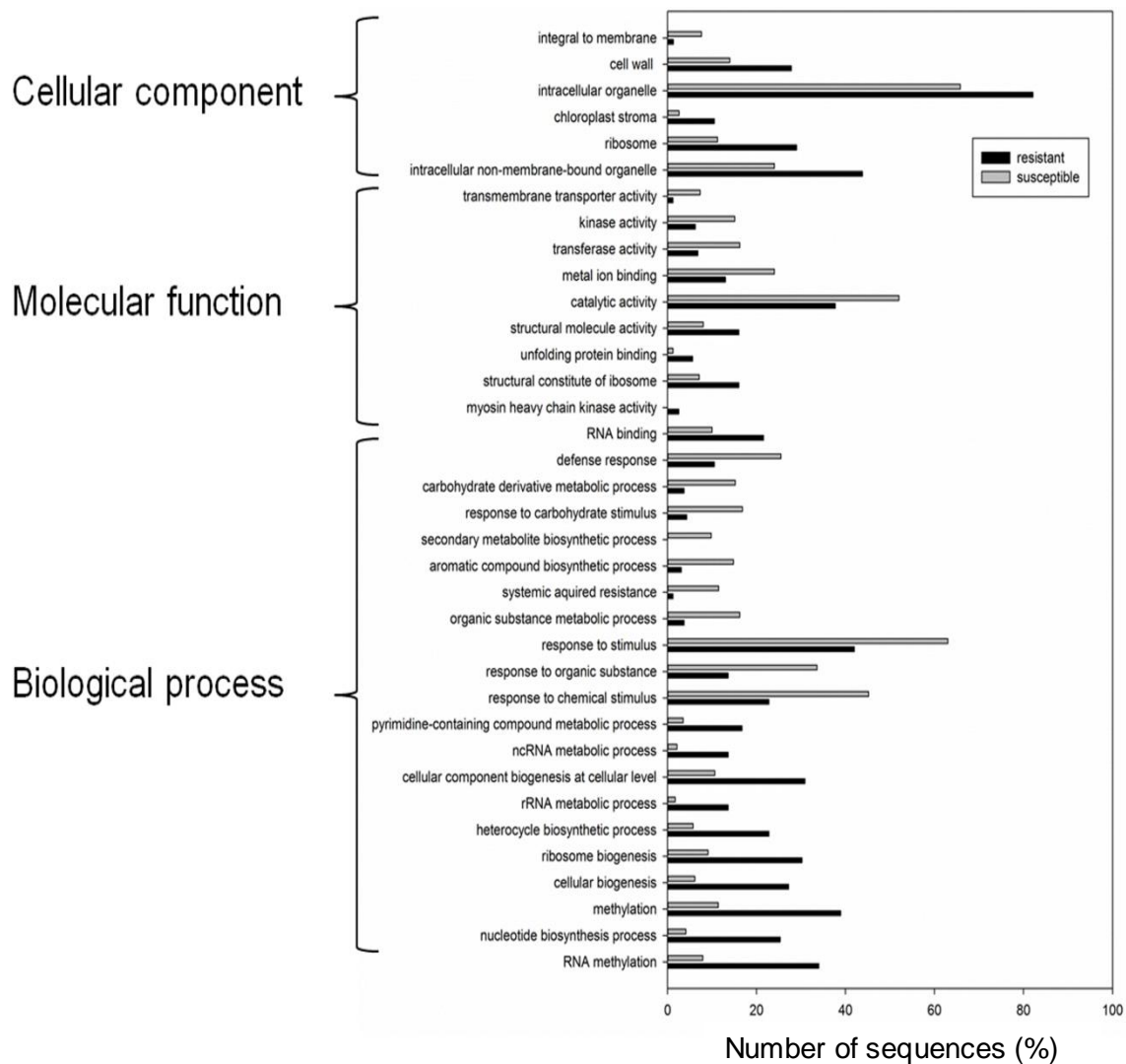
Figure 3.5. GO annotations for up-regulated sequences of resistant and susceptible plants relative to healthy plants.
Significant differences were detected between resistant and susceptible plants for all annotations using Fisher Exact Testing (p<0.01). Only the top 5-10 over and under represented genes for cellular component, molecular function and biological process are shown.
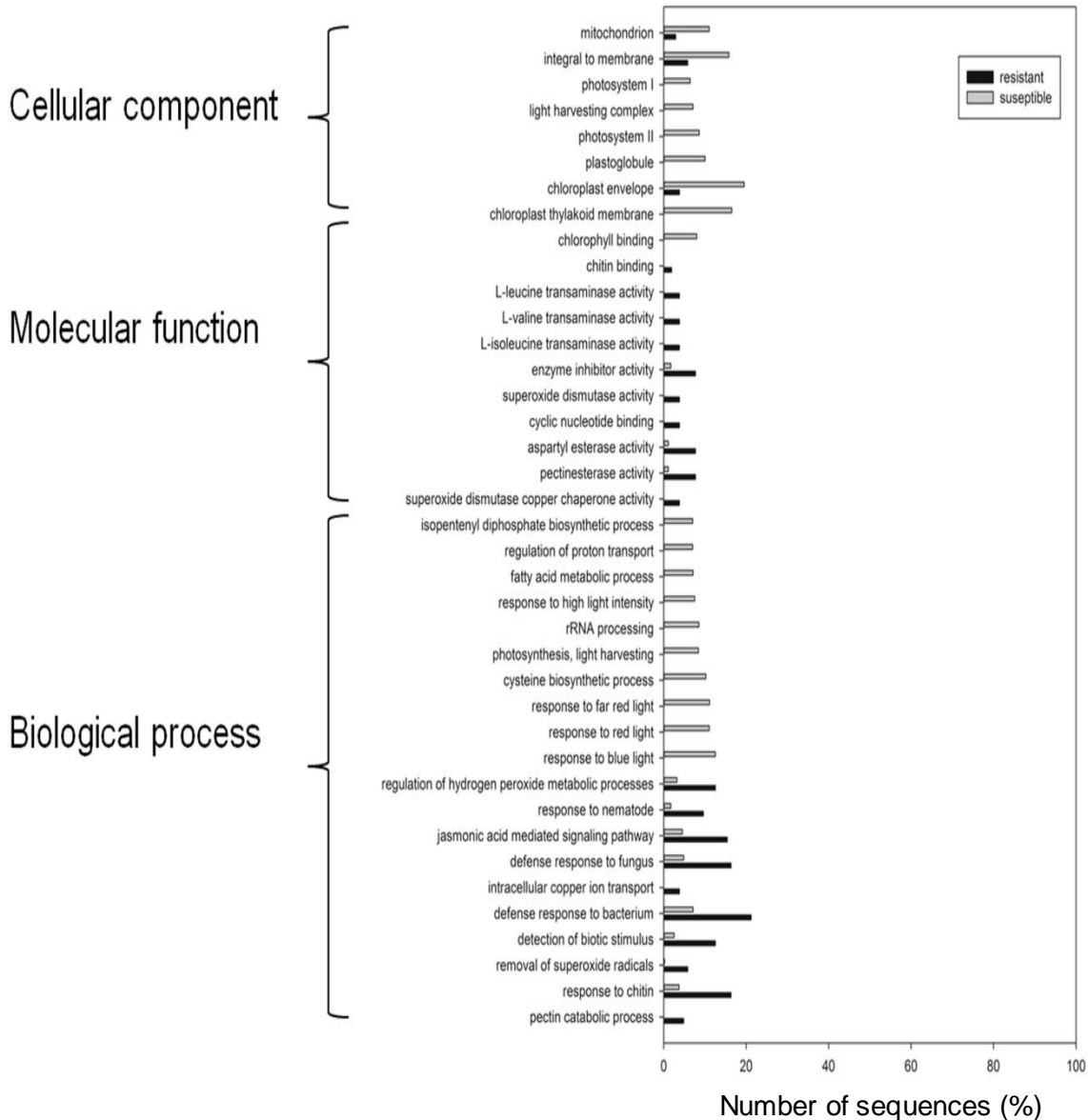
Figure 3.6. GO annotations for down-regulated sequences of resistant and susceptible plants relative to healthy plants.
Significant differences were detected between resistant and susceptible plants for all annotations using Fisher Exact Testing (p<0.01). Only the top 5-10 over and under represented genes for cellular component, molecular function and biological process are shown.

According to molecular function for the up-regulated transcripts relative to the healthy control, the DEGs that mapped to "transmembrane transporter activity", "kinase activity", "transferase activity", "metal ion binding" and "catalytic activity" were enriched in the susceptible plants and represented 7-52% of the DEGs. The resistant plants were enriched in different annotations including "RNA binding", "myosin heavy chain kinase activity", "structural constitute of ribosome", "unfolding protein binding" and "structural molecule activity" representing 3-22% of the total DEGs.

Functional annotation using GO terms categorized into biological process for the up-regulated transcripts suggested that the resistant plants were enriched in "RNA methylation", "nucleotide biosynthetic process", "methylation", "cellular nitrogen compound biosynthetic process", "ribosome biogenesis", "heterocycle biosynthetic process", "rRNA metabolic process", "cellular component biogenesis", "ncRNA metabolic process" and "pyrimidine-containing compound metabolic process". Mapping of these annotations were represented by 13-34% of the total DEGs for the resistant phenotype and 2-11% of the total DEGs for the susceptible phenotype. "Response to chemical stimulus", "response to organic substance", "response to stimulus", "organic substance metabolic process", "systemic acquired resistance", "aromatic compound biosynthetic process", "secondary metabolite biosynthetic process", "response to carbohydrate stimulus", "carbohydrate derivative metabolic process" and "defense response" were enriched in the susceptible treatment, in comparison to the resistant

treatment. These annotations were represented by 10-63% of the total DEGs of the susceptible treatment and 0.1-42% of the total DEGs of the resistant treatment.

For DEGs down-regulated in resistant and susceptible plants relative to the healthy control, cellular component GO annotations for the up-regulated transcripts enriched in the susceptible plants were "mitochondrion", "integral to membrane", "photosystem I", "light harvesting complexes", "photosystem II", "plastoglobule", "chloroplast envelope" and "chloroplast thylakoid membrane". These transcripts represented between 6-20% of the total DEGs. Cellular component annotations down-regulated relative to the healthy control but enriched for the resistant treatment were absent.

According to molecular function for the down-regulated transcripts relative to the healthy control the DEGs that mapped to "L-leucine transaminase activity", L-valine transaminase activity", "L-isoleucine transaminase activity", "enzyme inhibitor activity", "superoxide dimutase activity", "cyclic nucleotide binding", "aspartyl esterase activity", "pectinesterase activity" and "superoxide dismutase copper chaperone activity" were enriched in the resistant plants and represented 2-8% of the total DEGs. "Chlorophyll binding" was the only molecular function annotation enriched for in the susceptible treatment and represented 8% of the total DEGs.

Functional annotation using GO terms categorized into biological process for the down-regulated transcripts suggested that the resistant plants were enriched

in "pectin catabolic process", "response to chitin", "removal of superoxide radicals", "detection of biotic stimulus", "defense response to bacterium", "intracellular copper ion transport", "defense response to fungus", "jasmonic acid mediated signalling pathway", "response to nematode", and "regulation of hydrogen peroxide metabolic processes". These annotations represented 4-21% of the total DEGs for the resistant phenotype and 0-5% of the total DEGs for the susceptible phenotype. "Response to blue light", "response to red light", "response to far red light", "cysteine biosynthetic process", "photosynthesis and light harvesting", rRNA processing", "response to high light intensity", "fatty acid metabolic process", "regulation of protein transport" and "isopentenyl diphosphate biosynthetic process" were enriched in the susceptible treatment in comparison to the resistant treatment. These annotations represented 7-13% of the total DEGs of the susceptible treatment and 0-0.1% of the total DEGs of the resistant treatment.

3.3.5. Mapping differentially expressed genes to KEGG pathways

Mapping the DEGs to KEGG pathways shows that defence pathways, such as flavonoid biosynthesis, fatty acid biosynthesis and glucosinolate biosynthesis were more active in the resistant plants than the susceptible plants **(Figure 3.7)**. Pathways associated with growth and development such as starch and sucrose metabolism, carbon fixation and glycolysis were more active in the susceptible plants, and less active in the resistant.

3.3.6. Common and unique genes among phenotypes

Of the 3857 total number of DEGs analyzed, 222 and 2751 were unique to the resistant and susceptible phenotypes, respectively (**Figure 3.8**). Among the 442 common DEGs, hierarchical clustering was used to detect any differences in the abundance of transcripts between the two treatments **(Figure 3.8)**. Most notable, transcripts of the susceptible and healthy responded more similarly than in comparison to the resistant phenotype.
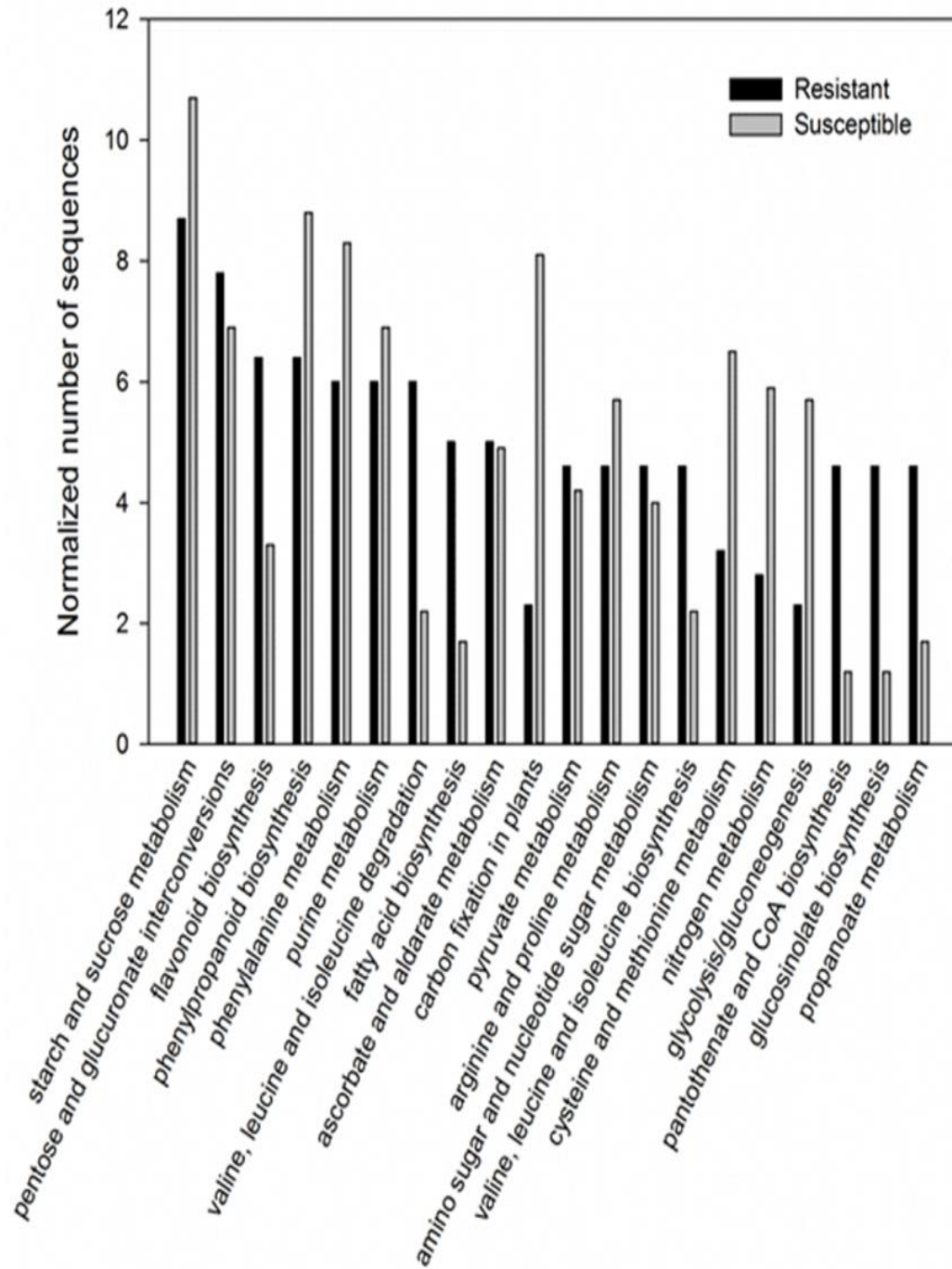
Figure 3.7. KEGG pathways enriched in the DEGs between resistant and susceptible phenotypes.
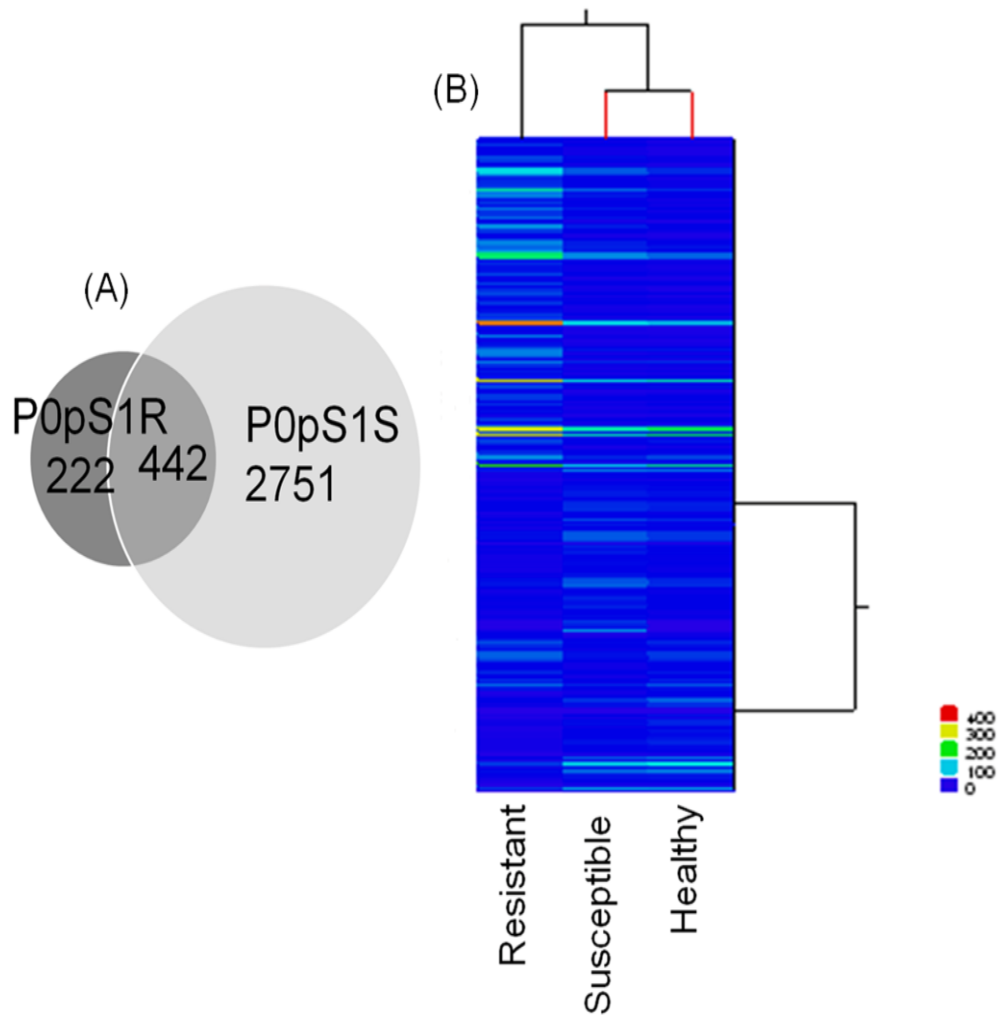
Figure 3.8. Differentially expressed genes of resistant and susceptible phenotypes and hierarchical clustering of the 442 common DEGs among the resistant and susceptible phenotype. (A) Venn diagram illustrating the number of differentially expressed genes of resistant and susceptible phenotypes. Dark shade represents resistant phenotype and light shade represents susceptible phenotype. **(B)** Hierarchical clustering of the 442 common DEGs among the resistant and susceptible phenotype.

3.4. Discussion

This study shows the changes in transcriptome associated innate immunity and confirms that in addition to ss(+) RNA viruses, Gram negative bacteria and chemical elicitors (Kathiria et al. 2010; Slaughter et al. 2012; Luna et al. 2012), dsDNA viruses can also generate a transgenerational response in the form of pathogen resistance. Resistant and susceptible plants displayed differences in seed size, pathogen titer and expression profiles. Pathogen response measured as prolonged resistance was exhibited in the progeny and appeared as a 10 to 14 day delay in symptom response and in some lines, symptoms and/or titers of CaMV were absent. The data support temporal differences in defense responses between primed plants showing pathogen resistance and plants that did not receive a priming treatment, similar to that observed in model hosts (Kathiria et al. 2010). Following pathogen exposure at the same developmental stage, the plants that were not primed exhibited elevated levels of infection whereas, symptom development progressed much more slowly in the primed plants. Symptom development, pathogen titer and small RNA sequence transcriptome profiling was used to characterise these differences between primed first generation stress resistant and susceptible phenotypes.

This study reports pathogen resistance as a trangenerational response in *B. rapa* and this finding is similar to many other recent studies with other pathogens and non-agricultural plants (Kathiria et al. 2010; Luna et al. 2012; Slaughter et al. 2012). However, this study differs from previous studies in that pathogen titer

and symptom development evaluations were observed at 21 days post

inoculation (dpi) rather than at early phases (2–48 h) of local pathogen infection.

At 21 dpi CaMV is systemic in the host enabling the evaluation of advanced

stages in pathogen infection, including quantification of the pathogen using

immunological assays (Farzadfar and Pourrahim 2013). Transgenerational

responses in the form of pathogen resistance that have been reported only last

6-48 h (Kathiria et al. 2010). Others have made similar suggestions in that the

expression of transgenerational effects may not be confined to the seedling

stage, but that they may also be observed in mature plants (Galloway & Etterson

2009; Case et al. 1996). In this study, a prolonged period of 10 to 14 days of

delayed symptom and titer was obtained by measuring the time-point of systemic

rather than local infection.

Susceptible and resistant first generation progeny were challenged with

CaMV and during systemic viral infection, small RNAs were extracted and

sequenced to obtain the transcriptomes. Comparisons among transcriptomes

revealed distinct differences between the resistant and susceptible phenotypes.

Progeny exhibiting strong resistance to CaMV had an increased abundance of

transcripts associated with pathogen defense including glucosinolate

biosynthesis, flavonoid biosynthesis and fatty acid biosynthesis, while the

susceptible phenotype did not share this profile. The glucosinolate, flavonoid

and fatty acid biosynthesis pathways were previously identified as being involved

with pathogen defense (Montillet et al. 2013; Ferreyra et al. 2012; Wang et al.

2011). Although the transcriptome profile of the susceptible plants also displayed

pathogen defense activity the profile was different from the resistant phenotype.

The susceptible phenotype had activity in phenylpropanoid biosynthesis,

phenylalanine metabolism and there was an abundance of transcripts involved in

systemic acquired resistance (SAR). Love et al. (2005) previously demonstrated

that glutathione S-transferase (GST1), pathogenesis-related protein 1 (PRs1),

PRs2 and PRs5 had increased expression at 20 dpi in the compatible reaction

between *Arabidopsis thaliana* and CaMV. GST and PRs are markers for reactive

oxygen species (ROS) and SA signaling, respectively. This current study

provides evidence that SAR appears to be involved in pathogen defense during

the later stages of compatible CaMV infected *B. rapa* in a transgenerational

mechanism.

Differences in the activity of pathways involved in primary metabolism (i.e.,

growth and development) between the two phenotypes were also detected. The

susceptible plants had an increased abundance of transcripts associated with

growth and development processes including photosynthesis and glycolysis,

relative to the resistant plants. These findings coupled with observations of early

bolting and flowering indicates faster flowering as a developmental response of

B. rapa infected with CaMV. Alteration in reproductive development after

exposure to pathogens was previously described (Peters 1999; Agnew et al.

2000). For example, accelerated reproductive development was observed in

Arabidopsis infected with *Pseudomonas syringae*, *Xanthomonas campestris* or

Personospora parasitica (Korves & Bergelson 2003). Transcription factors and phytohormones such as SA, JA, ethylene or auxins link pathogen defense and development responses because they are involved in both processes with cross-talk between gene networks (Wang et al. 2002; Liu et al. 2014). This study further shows strong evidence of differences between pathogen defence and development response supported by an increased abundance of transcripts involved with cellular components associated with structure including the cell wall, cell membrane constituents and intracellular organelles displayed in the resistant plants, relative to the susceptible. Plants coordinate their cellular structures with the production of secondary metabolites to inhibit the movement of viruses and other pathogens during infection (Zavalieu et al. 2011). Additionally some pathogens including CaMV alter the host's cellular structure to facilitate movement throughout the plant (Carluccio et a. 2014). These defense mechanisms alter gene expression of structural host cell machinery and reduce CaMV titer in resistant plants leading to a high level of resistance.

The mechanism behind pathogen resistance as a transgenerational response remains elusive but the meiotic inheritance of epigenetic signatures, such as DNA methylation, acetylation of histone tails, chromatin remodelling and small RNAs have been suggested to give rise to transgenerational responses (Jablonka 2013; Jablonka & Raz 2009). Although in this study, the mechanism behind this resistance appears to involve several pathways and possibly mechanisms, enhanced methylation and non-coding RNA metabolic processes

were detected in the transcriptome of the resistant plants, suggesting an epigenetic mechanism behind the resistance. Further support for an epigenetic mechanism is the nuclear localization of CaMV nucleic acid throughout it's lifecycle. Replication of CaMV requires the viral translational transactivator protein P6 that is present in the nucleus and is essential for CaMV infectivity (Haas et al. 2008). One nuclear function for P6 is to supress RNA silencing, a gene regulation mechanism that provides antiviral capablilties by inactivating the nuclear protein DRB4 that is required by the major plant antiviral silencing factor DCL4. Besides a regulatory role, RNA silencing confers a sequence-specific antiviral immunity to plants through virus-derived short interfering RNA (reviewed in Ding & Voinnet 2007). Localization of CaMV components such as P6 in the nucleus of the plant cell provides an opportunity for genome and transcriptome modifications. These modifications may provide transgenerational protection to other pathogens (Kathiria et al. 2010) and influence other phenotypic characteristics such as seed size as observed in this study.

A relationship was uncovered whereby parent plants that produced larger seeds produced progeny more resistant to CaMV. The phenotypic transgenerational effect of an increase in seed size was previously observed with exposure to low temperature, herbivory or shaded conditions (Case et al. 1996; Agrawal 2001; Galloway & Etterson 2007). Although a linkage between these two traits remains to be determined, as single variables, both of these traits were demonstrated to be inherited in a transgenerational manner and have shown to

be correlated to DNA methylation (Kathiria et al. 2010; Amoah et al. 2012; Luna et al. 2012). Although there is uncertainty over stable transgenerational inheritance (Pecinka et al. 2009; Boyko & Kovalchuk 2011; Iwasaki & Paszkowski 2014) because a mechanism has not yet been determined, this study provides evidence that under a specified treatment regime, seed size and CaMV resistance may be transferred to first generation progeny in a durable transgenerational fashion. The results also indicate that disease resistance was not acquired equally across all progeny, but that there was variability between progeny produced from the same plant, limiting the probable role of maternal inheritance in influencing the traits observed.

Further examination of the inheritance of stable epigenetic signatures and metabolic profiling would advance our understanding of the mechanism behind the resistance and determine if there is a linkage between seed size and resistance. This work provides insight into the use a non-transgenic means for introducing stress resistance into *B. rapa* germplasm which is an economically important crop. It would be beneficial to further explore the type of resistance that was demonstrated in this study and continue to investigate if it could be applied to other economically important crops.

4. COMPLETE GENOMIC SEQUENCE OF *RUBUS YELLOW NET VIRUS* AND DETECTION OF GENOME-WIDE PARARETROVIRUS-DERIVED SMALL RNAS

4.1. Introduction

Red raspberry (*Rubus idaeus* L.) is grown in temperate regions worldwide for fresh market, jam, juice, purée or individually quick-frozen fruit. Raspberry is a host to many different viruses and mixed virus infections appear to be common (Quito-Avila & Martin 2012). In North America and Europe, *Rubus yellow net virus* (RYNV) causes raspberry mosaic disease (RMD) while in association with *Black raspberry necrosis virus* (BRNV) and *Raspberry leaf mottle virus* (RLMV) (McGavin & MacFarlane 2010). Although complete genomic sequence is available for many of the viruses involved in RMD, the complete genomic sequence of RYNV is currently unavailable, limiting molecular characterization of this disease that reduces berry growth and yield. Based on a bacillifom particle morphology and partial sequence from a highly conserved portion of the virus genome, RYNV was previously identified as a putative member of the family *Caulimoviridae*, genus *Badnavirus* (Jones et al. 2002).

Badnaviruses include some of the most destructive plant viruses, often causing devastating losses to infected crops (Harper et al. 2004; Huang & Hartung 2001; Thresh et al. 1986). Initially, it was reported that the badnaviruses were found infecting hosts growing exclusively in tropical and sub-tropical climates. However, badnaviruses have also been found infecting host plants occupying North American temperate zones. These include host shrub-like species including red raspberry, gooseberry and ornamental spiraea and the infectious agents RYNV, *Gooseberry vein banding virus* (GVBV) and *Spiraea yellow leaf spot virus* (SYLSV), respectively.

Badnaviruses have a 120-150 x 30-nm non-enveloped bacilliform capsid containing a singular circular dsDNA genome that is 7.3-8.0 kb in size (Lockhart 1990). Virions enter a host in a non-circulative semi-persistent manner by a mealybug or an aphid vector. In the case of raspberry, RYNV is vectored by *Amphorophora agathonica* Hottes, also known as the large raspberry aphid. The typical genomic structure for badnaviruses consists of three open reading frames (ORFs), and all of the genes are encoded on the same discontinuous strand (Xu et al. 2011). The first two ORFs encode proteins that are 17 and 15 kDa, respectively, and little is known about their function. Most information regarding the badnaviruses is derived from the characterization, organization and highly homologous nature of ORF 3, which encodes a 216 kDa polyprotein that is cleaved post-translationally by the viral-encoded aspartic protease to produce a movement protein, a coat protein and a replicase comprised of a reverse transcriptase and ribonuclease H (Laney et al. 2012; Sether et al. 2012).

High throughput small RNA sequencing has increasingly been used for plant virus identification, virus genome characterization and viral genome assembly (Hwang et al. 2013; Kreuze et al. 2009). RNA silencing is a universal system among most eukaryotes to developmentally and temporally regulate gene expression through the production of small RNAs (Huntzinger & Izaurralde 2011). RNA silencing also plays an important role in antiviral defense whereby, small RNAs (sRNAs) accumulate to abundant levels in infected organisms (Ratcliff et al. 1997; Waterhouse et al. 2001). In plants, RNA silencing is orchestrated by a

diverse family of endonucleases, known as dicer-like proteins (DCLs). In *Arabidopsis thaliana*, four Dicers (DCL1-4) have been identified (Margis et al. 2006). The dicer-like proteins involved in antiviral defense include DCL2 and DCL4 that generate 22-24-nt small interfering RNAs, from double-stranded replicative intermediates of plant RNA viruses (Voinnet 2001).

To examine the molecular structure and characteristics of RYNV, nucleic acid from the virus was cloned, sequenced and compared to other members of the genus *Badnavirus*, family *Caulimoviridae*. The RYNV genomic sequence most closely resembled that of GVBV, the only other member of the genus *Badnavirus* from North America with a fully sequenced genome. Small RNA sequence profiling of RYNV infected red raspberry leaf tissue yielded a highly uneven genome-wide distribution of virus-derived RNAs (vsRNAs) with strong clustering to small defined regions distributed over both strands of the RYNV genome. This suggests that the raspberry interfering RNA pathway targets specific segments of a plant dsDNA virus sequence, possibly as an antiviral defense mechanism.

4.2. Material and methods

4.2.1. Source of infected plant material

Leaf tissue to obtain the genomic sequence of RYNV was collected from one red raspberry (*Rubus idaeus* L.) plant located near Lethbridge, Alberta, Canada expressing symptoms characteristic of RYNV infection (Kalischuk et al. 2008). Tissue was ground in liquid nitrogen and total DNA extracted using hexadecyltrimethylammonium bromide, also known as CTAB procedure (Doyle &

Doyle 1990). Total RNA for the small RNA sequencing was extracted using the

miRNeasy spin column kit (Qiagen, Canada) from three RYNV-infected plants

showing the symptoms similar to those described above and located at the same

location. RNA sequencing was performed with a pooled sample from the three

independent plants. The quantity and quality of the RNA were checked using

spectrophotometry and agarose gel electrophoresis.

4.2.2. Genomic sequencing of RYNV

The complete genomic sequence of RYNV was determined by sequencing

two DNA fragments generated using PCR. Oligonucleotides

5'ATATAGGAGCTAGCCGCAGCTGTGGA3' and

5'AGCGAATTCGCCTGATTCCGAGCTGCTTGTTG3' were used to PCR amplify

a 4501-bp product that was TA cloned into pCR-XL-TOPO (Invitrogen, Canada).

Oligonucleotides 5'AGCGTCGACCTGGTGAACAAGGAGGTACAGAGC3' and

5'TCCACAGCTGCGGCTAGCTCCTATAT3' were used to amplify a 5871-bp

fragment; however, due to instability in *Escherichia coli*, the TA cloning of this

fragment was unsuccessful. Instead, the 5871-bp fragment was digested with

BamHI and each of the two DNA fragments cloned separately. The SalI/BamHI

fragment was cloned into the corresponding sites of pBluescript II SK(-)

(Stratagene, Santa Clara, USA) and the NheI/BamHI fragment was cloned into

the corresponding sites of pLitmus 38i (New England Biolabs, USA). Standard

procedures were used for the remaining cloning and Sanger sequencing

reactions. Primer walking was used to obtain sequence from the entire length of

each insert, and sequencing was performed on both DNA strands to produce at

least three copies of overlapping sequence for the entire genome **(Table 4.1 &**

4.2).

Table 4.1. Primers for sequencing the sense strand of *Rubus yellow net virus*

Primer name	Primer sequence for sense strand (5' to 3')	Genome location (bp)
Rb20	TCCTTTAGTGTCGGCAGCATCTC	7421-7443
Fc20f	GCAGCAACTAGGCAAGAAACC	309-329
Rb11	GGAAATCAAGGCCACACAGT	808-827
Rb103f	AGAAGGACCTTCAGTAGGGACTT	1580-1602
Rb105Ff	GATCGTGGAAGAAGGAGGACCATC	2371-2394
Fa13mf	GCTGCATGCTAGTCCTTCCAC	3285-3305
Rb107f	CAAGTACAAGCACACATCTGG	3997-4018
Rb1f1	TGGAGGACGTGTCCATTC	4713-4729
Rb2f2	CAAGTGTAAGCTTGACATCAAG	6183-6217
Rb55	CAGACTCAGCTGGCAAGGAAG	6597-6617

Table 4.2. Primers for sequencing the antisense strand of *Rubus yellow net virus*

Primer name	Primer sequence for antisense strand (5' to 3')	Genome location (bp)
Rb34r	GGTAAATCCTTACTGTTCTCTTG	7321-7343
Fa6r	CACCTGCACCACCTGCAACC	143-163
Rb53	TGATGTTTTGTGCAAGATCGTCG	747-769
Fd5r	CTCACTGGTCTTGCTGTATAG	1509-1529
Rg4r	GGTATTGTGCCTTCGTCTC	2276-2294
Ff1r	GGATGTGGTCGATGGGAA	3128-3145
Ff3r	GCATTTGCACTTCTTCCCA	3900-3918
Rb4r	CCTTGCATGTCTGCTGGT	4741-4758
Rb5r	CCTTCTGGCCTTCCCTC	5515-5531
Ff7r	CTGAAGGTCCTTCTCATTGC	6340-6359

4.2.3. RYNV genomic sequence assembly and analysis

Assembly of the RYNV genome was completed using Sequencher V4.7 (Gene Codes Inc., Ann Arbor MI, USA) and the alignments used only high quality value bases as determined by Sequence Scanner Software v1.0 (Applied Biosystems, USA). Nucleotide and the deduced amino acid sequence were compared to existing sequences in GenBank and in other databases such as conserved protein domain (CDD) (http://www.ncbi.nlm.nih.gov/structure/cdd.shtml), Pfam (http://pfam.sanger.ac.uk/), SMART (http://smart.embl-heidelberg.de/) and Prosite (http://prosite.expasy.org/) using the basic local alignment search tool (BLAST). Putative ORFs were identified using ORF Finder

(http://www.ncbi.nlm.nih.gov/projects/gorf/) and theoretical molecular weights

were predicted using Compute pI/Mw (http://www.expasy.org/tools/pi_tool.html).

Potential promoter elements and non-coding RNA's were identified using PLACE

(http://www.dna.affrc.go.jp/PLACE/), PlantCARE

(http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) and Signal Scan

(http://www.dna.affrc.go.jp/sigscan/signal.html/). Secondary structure predictions

were made using Jpred (www.compbio.dundee.ac.uk/www-jpred/) and Paircoil

(http://groups.csail.mit.edu/cb/paircoil/cgi-bin/). Patterns and profiles were

predicted using ELM (http://elm.eu.org/search/). Direct and inverted symmetry

repeats were identified using Radar

(http://www.ebi.ac.uk/Tools/Radar/index.html).

4.2.4. Phylogenetic analysis

Phylogenetic relationships among the badnaviruses were estimated using

CLUSTAL X v2 multiple alignment with ORF 3 amino acid sequence. The

neighbor-joining clustering algorithm was used to estimate the phylogeny and a

rooted tree constructed using CLC Main Workbench software (CLC Biosciences,

USA). Confidence estimates were based on bootstrap sets of re-sampling

alignments with 1000 replicates and these values were added to the nodes of the

trees. GenBank accessions used in the analysis are listed in **Table 4.6**.

4.2.5 Small RNA sequencing and mapping of RYNV-infected leaf tissue

The RNA was sequenced on an Illumina Genome Analyzer IIx platform (Illumina, Inc., USA). Briefly, the processing by Illumina consisted of the following successive steps: purification of 20-30 nt RNAs by filtering through a polyacrylamide gel, ligation of 3' and 5' adapters to the 20-30nt RNAs, c-DNA synthesis by reverse transcription followed by acrylamide gel purification and final steps of bridge amplification and paired-end sequencing.

RYNV-derived sRNA sequences were identified by mapping the reads to the RYNV genome. Raw reads were filtered for quality and adaptor sequences trimmed using FastX Test Kit (http://hannonlab.cshl.edu/fastx_toolkit). A mappable read had a Phred score greater than 30, a 3'ADT and greater than 15 bases after the 3'ADT cut. Reads that were 18-25 nt and that met the filtering criteria were mapped to the RYNV genome using Bowtie (http://bowtie-bio.sourceforge.net/index.shtml). The seed length was set to 16 nt and one mismatch was allowed in the seed region during mapping. The command line for Bowtie on a Mac Unix system was as follows: bowtie -q -l 16 -n 1 -e 40 -k 1 raspberry_virus Rubida.fq. Output data were exported to a Microsoft Excel spreadsheet and summarized using standard graphics software (http://www.corel.com).

4.3 Results

The complete genomic sequence of RYNV was determined and consisted of 7932 bp (Genbank Accession: KF241951; **Appendix A**).

4.3.1. Non-coding regions of RYNV genomic DNA

The intergenic region (IR) of RYNV was 969 bp and contained many of the conserved nucleic acid sequences previously described for plant dsDNA viruses (Benfey & Chua 1990; Medberry & Olszewski 1993). A plant tRNA^{Met} sequence was predicted with 67 percent nucleotide identity to the complementary sequence of 3'ACCAUAGUCUCGGUCCAA5', and this region has been previously described as one of the priming site for reverse transcription (Medberry et al. 1990). The 5' end of the tRNA ^{Met} was designated as the first nucleotide for numbering the RYNV genome as done with previously characterized badnaviruses. Nucleic acid motifs previously identified and located in the intergenic region of other members of the family *Caulimoviridae* and identified along the RYNV intergenic region included a TATA box (nt 7523-7536), a cap signal associated with the TATA box (nt 7564-7561), a hexamer motif (nt 7633-7638), an as-1-like sequence (nt 7363-7379) and a GATA box motif (nt 7486-7490) (**Table 4.3**).

4.3.2 Coding regions of RYNV genomic DNA

Rubus yellow net virus ORF 1 potentially encodes a 210-amino acid protein of 24-kDa **(Table 4.4)** with two nuclear export signals (NES), two coiled-coil regions and a proline rich C-terminus. Homology was detected between RYNV ORF 1 and both P24 of *Rice tungro bacilliform virus* (RTBV) and ORF 1 of GVBV.

Rubus yellow net virus ORF 2 encodes a predicted 17-kDa protein (153 aa) with homology to ORF 2 from other badnaviruses including *Citrus mosaic virus*

(CMBV), *Dracaena mottle virus* (DMV), *Dioscorea bacilliform virus* (DBV),
Kalanchoe top-spotting virus (KTSV) and *Sugarcane bacilliform virus* (SCBV).
RYNV ORF 2 contained one NES, one 14-amino acid direct repeat, one coiled-
coil domain and a proline rich C-terminus.

   Rubus yellow net virus ORF 3 was predicted to produce a 1971-amino acid
protein with a molecular weight of 210-kDa **(Table 4.4)**. RYNV ORF 3 resembled
that of other badnaviruses encoding a polyprotein containing the movement
protein (MP), coat protein (CP), aspartic protease (PR), reverse transcriptase
(RT) and ribonuclease H (RNaseH) (Medberry et al. 1990; Laco et al. 1995;
Marmey et al. 1999; Tzafrir et al. 1997). The active site
$HXGX_9HRX_3GX_8D...EXDX_3G$ (where X represents any amino acid) for the MP
was located at RYNV ORF 3 amino acid position 137-182. In addition to
similarities to the badnaviruses (Bouhida et al. 1993; Tzafrir et al. 1997), the
active site for the MP region of the RYNV genome also showed homology to the
30K superfamily movement protein from members of the *Caulimoviridae*,
Flexiviridae and the *Tobamoviruses* (Melcher 2000). A second motif
$GX_2SXRFXNYX_7P$ (where X represents any amino acid) recognized to be
involved in systemic infection was also found within the RYNV genome at amino
acids 299-315 (Tzafrir et al. 1997). A total of five proline, glutamic acid, serine
and threonine PEST sequences found in proteins with short cell half-lives, were
located within the putative movement protein region of RYNV. PEST sequences
have been described previously (Kawchuk et al. 2001).

The badnavirus coat proteins were previously identified in ORF 3 as containing two cysteine-histidine rich motifs (Cys-His motif), with sequences $CXCX_2CX_4HX_4C$ and downstream $CX_2CX_5HX_5CX_2CX_4CX_2C$ (where X represents any amino acid) (Medberry et al. 1990). These two Cys-His motifs within ORF 3 of RYNV were located at amino acids 864-879 and 998-1024, respectively. At amino acids 1209-1216, 3' to the RYNV putative CP, was an aspartic protease (PR) active site motif AX_2DXGXT (where X represents any amino acid). The consensus sequences for RT and RNaseH were identified at the 3' end of ORF 3 at amino acids 1553-1556 and 1553-1823, respectively.

Open reading frame 4 (136 aa) encodes a predicted protein of 15-kDa **(Table 4.4)**. RYNV ORF 4 overlapped the 3' region of ORF 3 but in a different reading frame. Homology was absent between RYNV ORF 4 and other known proteins. Comparative genomics between ORF 4 and a partial sequence for another RYNV isolate (Jones et al. 2002) showed 76% amino acid similarity. Interestingly, ORF 4 appeared to be more conserved than the corresponding region of ORF 3 that is 3' to the RNaseH.

Rubus yellow net virus ORF 5 was located 305-bp downstream of ORF 3, in the same reading frame as ORF 3 and potentially encoded a 17-kDa protein of 152-amino acids **(Table 4.4).** Multi-functional roles in protein coding and non-coding sequences were predicted in this region of the RYNV genome because it overlapped the putative consensus sequence complementary to the plant tRNA^{Met} and other predicted promoter elements. Features predicted for RYNV

ORF 5 included a signal peptide for localization within the secretory pathway near the N-terminus (1-21 aa), a transmembrane helix near the central portion of the sequence (66-85 aa) and an endocytosis signal (YxxF) towards the C-terminus, with all of these features being previously described (Kawchuk et al. 2001).

4.3.3. Open reading frames along the antisense strand

There were two ORFs predicted at greater than 10-kDa along the antisense strand of the RYNV genome and they were designated as ORFs 6 and 7 **(Figure 4.1)**. ORF 6 would produce a 145-amino acid 16.2-kDa protein. Similarly, ORF 7 would encode a 143-amino acid 16.7-kDa protein that contained a zinc finger-like motif with 86% amino acid similarity to the second Cys-His motif of the RYNV putative coat protein.

Table 4.3. Rubus yellow net virus genomic features and motifs

Start	End	Feature	Description
1	18	complementary sequence to tRNA^{Met}	5'TGGTATCAGAGCTTTAGC3'
7349	385	intergenic region	
7523	7536	TATA box	5'TTTGTTTAAAGGTA3'
7564	7561	cap signal associated with TATA box, antisense	5'GCAG3'
7362	7378	as-1-like element	TGACGcaaggaaTGACT, 160nt downstream of TATA
7265	7260	hexamer motif	ACGTCA (antisense)
7486	7493	GATA motif	tcaaGATA
386	1018	ORF 1	210 aa protein of 24 kDa
608	653	nuclear export signal (NES)	located as portion of ORF 1
779	785	nuclear export signal (NES)	located as portion of ORF 1
779	860	coiled-coil region	located as portion of ORF 1
920	989	coiled-coil region	located as portion of ORF 1
1015	1476	ORF 2	153 aa protein of 17 kDa
1201	1222	nuclear export signal (NES)	located as portion of ORF 2
1147	1185	14 amino acid direct repeat - arm 1	located as portion of ORF 2
1204	1243	14 amino acid direct repeat - arm 2	located as portion of ORF 2
1433	7348	ORF 3	1971 aa of 210 kDa
1844	1979	binding motif putative movement protein	Identified in all members of *Caulimoviridae* HXGX9HRX3GX8D...EXDX3G
1988	2006	hairpin	LOCATED AS PORTION OF orf 3 (mp), TTGGTATACATAATACCAA
2330	2378	systemic infection MP motif	badnavirus specific, GX2SXRFNYX7P
4025	4070	zinc finger-like motif	highly conserved in pararetroviruses and reverse transcribing elements
4427	4505	second zinc finger-like motif	highly conserved, badnavirus specific and close similarity to zinc finger found in retroviruses
5060	5081	active site for aspartate protease	AX2DXGXT
6092	6101	reverse transcriptase motif	YXDD
6092	6902	ribonuclease H motif	
6940	7356	ORF 4	136 aa producing 15 kDa
7654	180	ORF 5	
7654	7674	signal peptide	located in portion of ORF 5
7852	7909	transmembrane helix domain	located in portion of ORF 5
7830	7840	endocytosis signal	located in portion of ORF 5, YxxF
3330	3767	ORF 6	145 aa encoding a 16.2 kDa protein
7906	405	ORF 7	143 aa encoding a 16.7 kDa protein

Table 4.4. Coding capacity of Rubus yellow net virus open reading frames

ORF	1st nucleotide	Last nucleotide	Size (bp)	No. amino acids	M_r (kDa)
1	386	1018 (TGA)	630	210	24
2	1015	1476 (TGA)	462	153	17
3	1433	7345 (TGA)	2181	1971	210
4	6937	7353 (TAG)	417	136	15
5	7651	180 (TGA)	456	152	17
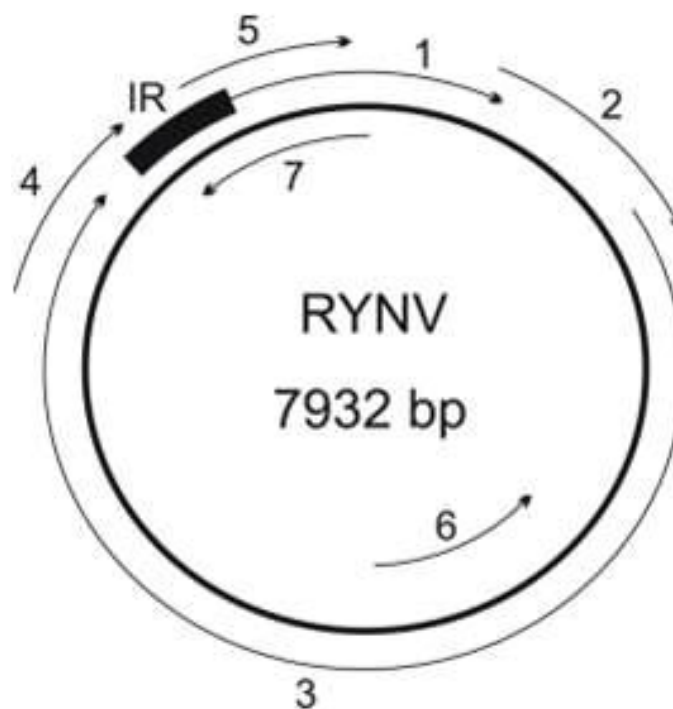
Figure 4.1. *Rubus yellow net virus* (RYNV) genomic organization.
A schematic showing the dsDNA genomic organization of RYNV. The rectangle represents the intergenic region containing the tRNA^{Met} consensus sequence. Lines with arrows represent the open reading frames 1-7. ORF 3 encodes a polyprotein containing the putative movement protein, coat protein, protease, reverse transcriptase and ribonuclease H.

4.3.4. Phylogenetic Analysis

The phylogeny derived from the badnavirus ORF 3 amino acid sequence

(Figure 4.2, Table 4.5 & 4.6) showed that RYNV was most closely related to

GVBV but they remain as two distinct species as they have differences in host

range and differences in polymerase nucleotide sequences greater than 20%.
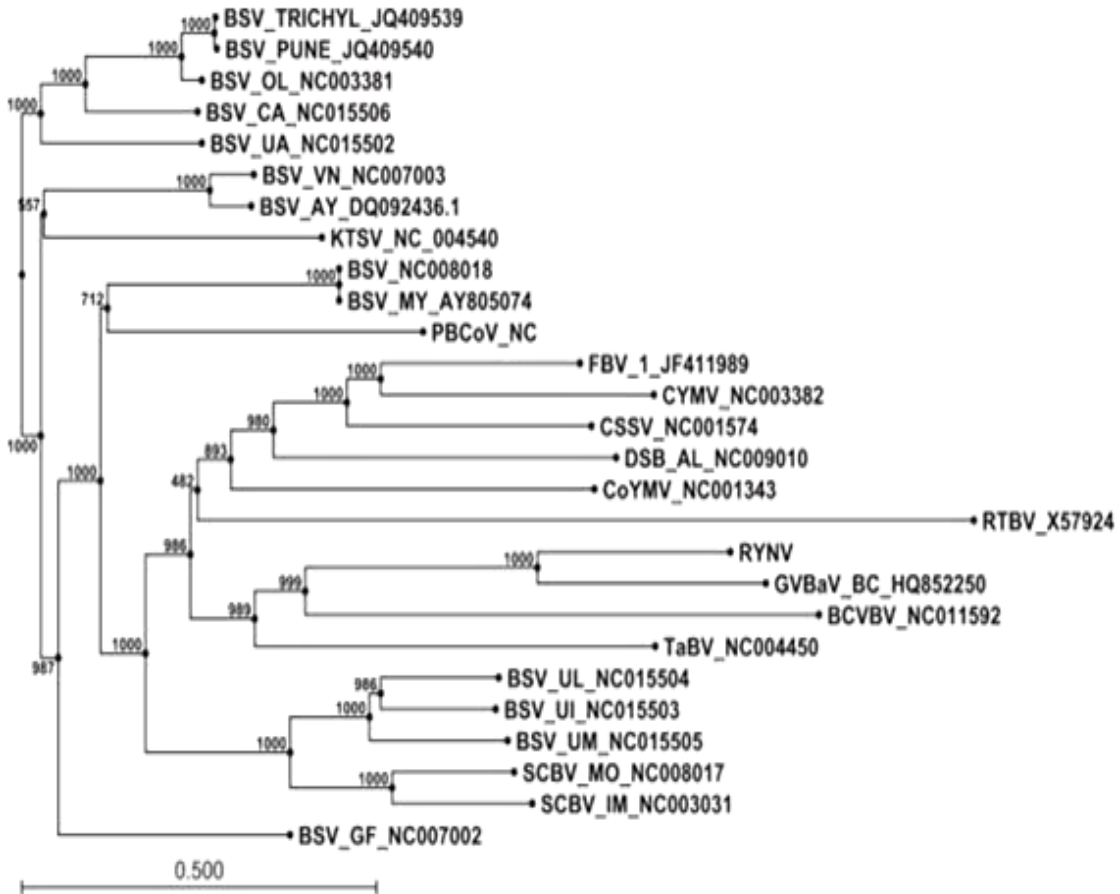
Figure 4.2. Neighbor-joining dendrogram of sequence relationships determined using the amino acid sequence alignment of the reverse transcriptase among species within genus *Badnavirus*.
The tree was rooted to rice trugro bacilliform virus (RTBV). Details of the accesions used in the analysis are in Table 4.6. Alignments were produced by a CLUSTAL algorithm and the dendrogram was produced by CDC Main Workbench software. Horizontal distances were proportional to sequence distances and vertical distances were arbitrary. The dendrograms was bootstrapped 1000 times (shown at nodes).

Table 4.5. Pairwise sequence alignments for members of *Badnavirus*

Virus	Genome size (bp)	Nucleotide identity[a] (%)	Amino acid identity (similarity) (%)[a]				
			ORF 1	ORF 2	ORF 3	ORF 4[b]	ORF5[c]
BSV_OLV	7389	53	16 (36)	17 (40)	37 (51)		
BSV_Trichyl	6950	51	19 (39)	17 (41)	36 (50)		
BSV_Pune	6950	51	15 (35)	17 (40)	36 (50)		
BSV_CA	7408	54	19 (36)	22 (43)	37 (51)		
BSV_UA	7519	53	20 (39)	20 (39)	36 (51)		
BSV_VN	7801	54	15 (29)	19 (40)	36 (51)		
BSV_AY	7722	54	19 (39)	22 (42)	36 (50)		
SCBV_MV	7568	53	19 (34)	24 (38)	34 (50)		
SCBV_IMV	7687	53	17 (34)	22 (39)	33 (49)		
KTSV	7591	53	20 (40)	27 (45)	35 (50)		
CSSV	7161	54	18 (33)	23 (36)	40 (54)	21 (35)	
CMBV	7559	53	18 (32)	31 (48)	39 (53)	22 (37)	
ComYMV	7489	52	18 (34)	20 (38)	36 (51)		
FBV_1	7140	54	18 (34)	29 (46)	40 (54)		
DBV	7261	52	15 (30)	26 (45)	36 (52)		
DMV	7531	54	19 (34)	21 (47)	39 (55)	19 (33)	18 (26)
BCVBV	8759	53	17 (28)	20 (35)	35 (50)		
GVBAV	7649	61	49 (70)	38 (64)	56 (68)		
TaBV	7458	52	18 (35)	23 (40)	38 (53)		
PBV_Co	7451	54	17(35)	24(41)	35(51)		
RTBV	8002	50	19 (43)	18 (29)	24 (39)		9 (16)

[a]Needleman-Wunsch algorithm was used for alignments. Identity and similarity were calculated as # matches divided by longest total sequence length of either query or subject multiplied by 100. [b] Based on same position along the genome, RYNV ORF 4 was compared to CSSV ORF Y, CYMV ORF 6, and DMV ORF 6. [c] RYNV ORF 5 was compared to DMV ORF 7 and RTBV P46.

Table 4.6. Accession numbers for open reading frame 3 sequence used in badnavirus phylogenetic analysis

Virus	Abbreviation	Accession Numbers ORF 3 protein	Reference
Banana streak OL virus	BSV_OLV	NP569150.1	Harper and Hull, 1998
Banana streak virus Trichyl isolate	BSV_Trichyl	AFH88829.1	Khurana and Baranwal, 2011*
Banana streak virus Pune isolate	BSV_Pune	AFH88829.1	Khurana and Baranwal, 2011*
Banana streak CA virus	BSV_CA	YP004442836.1	James et al., 2011
Banana streak UA virus	BSV_UA	YP004442824.1	James et al., 2011
Banana streak virus strain Acuminata Vietnam	BSV_VN	YP233110.1	Lheureux et al. 2007
Banana streak virus Acuminata Yunnan	BSV_AY	AAY99427.1	Zhuang and Liu, 2005*
Banana streak UL virus	BSV_UL	YP004442830.1	James et al., 2011
Banana streak UI virus	BSV_UI	YP004442827.1	James et al., 2011
Banana streak UM virus	BSV_UM	YP004442833.1	James et al., 2011
Banana streak Mysore virus	BSV_MYV	AAW80648.1	Geering et al. 2005
Sugarcane bacilliform Mor virus	SCBV_MV	YP595725.1	Bouhida et al. 1993
Sugarcane bacilliform IM virus	SCBV_IMV	NP149413.1	Geijskes et al. 2002
Cacao swollen shoot virus	CSSV	NP041734.1	Hagen et al. 1993
Kalanchoe top-spotting virus	KTSV	NP777317.1	Yang et al. 2003
Pineapple bacilliform comosus virus isolate HI1	PBV_CO	AEV42076.1	Sether et al. 2012
Fig badnavirus 1 isolate Arkansas	FBV_1	AEF56561.1	Laney at al. 2012
Citrus mosaic bacilliform virus	CMBV	NP569153.1	Huang and Hartung, 2001
Commelina yellow mottle virus	ComYMV	NP039820.1	Medberry et al., 1990
Dioscorea bacilliform virus	DBV	ABI47983.1	Seal and Muller, 2007
Dracaena mottle virus	DMV	ABE77344.1	Su et al., 2007
Gooseberry vein banding virus BC isolate	GVBAV	AEE39276.1	Xu et al., 2011
Bougainvillea spectabilis chlorotic vein-banding virus	BCVBV	YP002321513.1	Wang et al., 2008
Taro bacilliform virus	TaBV	ANN75640.1	Yang et al., 2003
Rice tungro bacilliform virus	RTBV	CAA40997.1	Hay et al., 1991
Grapevine vein clearing virus	GVCV	YP004732983	Zhang et al. 2011

4.3.5. Virus-derived small RNA profiling

A total of 14 million sRNA reads between 15-and 30-nt in length were obtained from high throughput sequencing of RYNV-infected raspberry. After filtering the data and using a greater than 99% accuracy cut-off value for base calls, 6.7 million reads were mapped against the RYNV genomic sequence. Of the mappable reads, a total of 0.4% showed sequence homology with RYNV, with a greater number of sRNAs represented by the sense strand than by the antisense strand. The genomic coverage of RYNV by sRNAs was 84% and the size classes of the RNAs were mainly 21-24 nt with 22-nt being the most abundant size class **(Figure 4.3).** Mapping the viral small RNAs (vsRNA) to the RYNV genome indicated several prominent areas targeted for RNA silencing. ORF 2 exhibited concentrated high levels of RYNV vsRNAs and these regions corresponded to predicted secondary structures **(Figure 4.3)**. There were forty-four regions dispersed across the RYNV genome that were devoid of vsRNAs. Although 82% of these regions were less than 100 nt in length, one identified sRNA desert was remarkable in that it comprised 6.5% of the RYNV genome or 514-nt **(Figure 4.3)**. This 514-nt region corresponded to the 5' portion of the large polyprotein sequence and contained the sequence for the active site for the MP **(Figure 4.3)**.
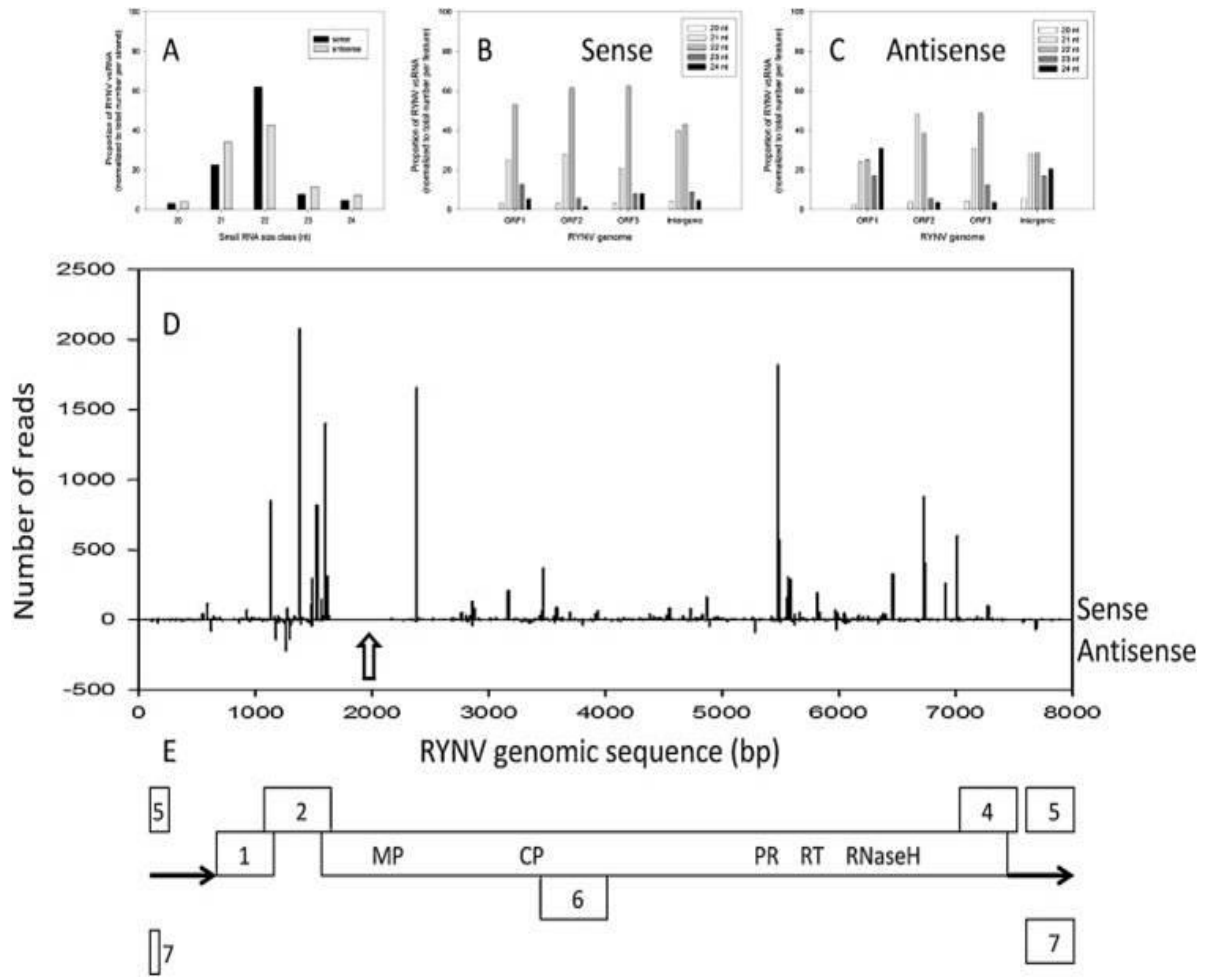
Figure 4.3. Illumina deep-sequencing analysis of RYNV vsRNA from infected *Rubus idaeus* leaf tissue.

(A) Bar graph showing the proportion of mapped sense and antisense RYNV vsRNA classified into 20-24 nt distribution. Bar graphs showing the proportions of 20-24 nt RYNV vsRNA mapped to ORF 1-3 and the intergenic region on the sense (B) and antisense (C) strand of the RYNV genomic sequence. (D) Genome-wide map of RYNV vsRNA at single-nucleotide resolution. The diagram plots the number of 20-24 nt vsRNA at each nucleotide position of the 7932 bp RYNV genome. Base numbering for the RYNV genome begins at the 5' nucleotide position of the tRNA^{Met} consensus sequence, on the sense strand. Vertical lines above the axis represent sense reads starting at each respective position and those below the axis represent the antisense reads. The arrow indicates the 514 bp region devoid of RYNV-derived sRNAs. (E) In scale, linear genomic map of RYNV sense and antisense strands. Lines with arrows represent the intergenic region and rectangles represent ORFs. ORF 3 is the polyprotein consisting of the movement protein (MP), coat protein (CP), aspartic protease (PR), reverse transcriptase (RT) and the ribonuclease H (RNaseH).

4.4. Discussion

The genome of a pararetrovirus that infects red raspberry and is transmitted by an aphid vector was sequenced and found to be 7932 base pairs. All badnaviruses, including RYNV, share ORFs 1-3 which have approximately the same size and location within the genome (Lockhart 1990; Medberry et al. 1990). RYNV ORF 3 encodes a large polycistronic transcript critical to the dsDNA viral lifecycle, as it produces the movement protein, coat protein and reverse transcriptase. Phylogenetic analysis of the amino acid sequence for ORF 3 suggests that RYNV is a member of genus *Badnavirus* and it is most related to GVBV, another member belonging to genus *Badnavirus* found infecting a temperate climate host. *Cauliflower mosaic virus* (CaMV), like RYNV, is another pararetrovirus transmitted by aphids that infects hosts growing in temperate climates but is classified in the genus *Caulimovirus*. As shown in this study, the evolutionary relationship among the two pararetroviruses RYNV and CaMV remain distinct at the genus level (i.e., *Badnavirus* and *Caulimovirus*) even though the viruses are both aphid-transmitted temperate climate pararetroviruses.

Analysis of ORF 3 indicated that RYNV closely resembled other members of genus *Badnaviruses* irregardless of many biological and molecular differences. Unlike ORF 3, ORF 1 and 2 had low homology with the badnaviruses and other known proteins. RYNV ORF 1 and 2 contained a proline rich C-terminus indicating possible non-sequence specific nucleic acid binding potential. Non-specific nucleic acid binding was demonstrated at the proline rich terminus region

for ORF 2 of *Cacao swollen shoot virus* (CSSV) and *Rice tungro bacilliform virus* (RTBV) (Jacquot et al. 1996; Jacquot et al. 1997). RYNV ORF 1 and 2 also contained coiled-coil regions that are associated with protein-protein interactions. Coiled-coil domains located in these ORFs are widespread across plant pararetroviruses (Leclerc et al. 1998) and protein-protein interactions were detected through tetramerization (Stavolone et al. 2001). RYNV ORF 2 was especially interesting because in addition to possible protein-protein interactions and nucleic acid binding potential, the DXG motif necessary for aphid transmission of CaMV (Schmidt et al. 1994) was present at the C-terminus of the sequence and may function in aphid transmission. Evidence supporting additional function includes the prediction of nuclear export signals (NES) contained in ORFs 1 and 2, suggesting a nucleocytoplasmic function.

Unique to the RYNV genome were two small ORFs located on the sense strand. CSSV, *Citrus mosaic virus* (CMBV) and *Dracaena mottle virus* (DMV) also have an ORF located in relatively the same position as RYNV ORF 4 (Su et al. 2007; Huang & Hartung 2001). Although homology was not detected between RYNV ORF 5 and any other proteins, DMV, CaMV and RTBV have an ORF of a similar size near the same vicinity.

Another feature unique to the RYNV genome was that ORF 6 and ORF 7 occur in the antisense sequence and may encode several zinc finger-like motifs. ORF 7 encoded a zinc finger-like motif with 86% amino acid similarity to the second Cys-His motif identified in the RYNV putative coat protein. Zinc finger

motifs are often involved in forming finger-like protrusions involved in binding DNA, RNA, protein and/or lipid substrates. There are many occurrences of zinc finger signatures in viral genomes and functional studies have suggested diverse roles such as in structure and regulation (Tanchou et al. 1998). Since RYNV ORFs 6 and 7 are located on the antisense strand and are not preceded by sequences associated with promoter activity, the transcription of RYNV is likely asymmetric like that of the other members of genera *Badnavirus* and *Caulimovirus* (Medberry et al. 1990).

Genomes of RNA viruses produce a double-stranded RNA intermediate during replication (Weber et al. 2006). These dsRNA products are recognized by the RNAi mechanism and cleaved into virus-derived siRNAs that target homologous ssRNA for further degradation. Although DNA viruses do not require dsRNA intermediates, they are known to produce dsRNA during infection (Weber et al. 2006). Interestingly, it was recently shown that infection of *Drosophila melanogaster* with *Invertebrate iridescent virus 6* (IIV-6) elicited the RNAi pathway involving *Dcr-2* and *Argonaute-2* (Bronkhorst et al. 2012). The highly uneven genome wide distribution with clustering of vsRNAs to defined regions designated hotspots were also observed with RYNV. Results with IIV-6 showed that antisense transcipts were produced during infection that potentially base-paired to form dsRNA with overlapping sense transcripts and that RNAi provided an antiviral defense against dsDNA viruses (Bronkhorst et al. 2012). The method

87

that plant pararetroviruses use to produce dsRNA during replication remains to be determined.

In this study, approximately 0.4% of total sRNAs generated by raspberry were RYNV-derived vsRNAs which spanned 84% of the RYNV genome. Mapping the vsRNAs to the viral genome revealed a 514 nt region that appears to have escaped the RNA silencing machinery. This sequence lacking sRNAs spanned the region of the RYNV genome that contained the binding domain indicative of the movement protein in badnaviruses. Another study that used high throughput sequencing to explore novel viruses in sweet potato reported partial vsRNA profiles for two badnaviruses (Kreuze et al. 2009) that also showed a lack of vsRNA in a similar location.

Involvement of all known DCLs in the production of a diverse pool of 21-24 nt vsRNAs has been demonstrated as a general host plant response to pararetrovirus infection (Blevins et al. 2006; Moissiard & Voinnet 2006). Sequencing of RYNV-infected raspberry leaf tissue sRNA revealed a diverse pool of 21-24-nt vsRNAs, supporting the aforementioned interaction between pararetrovirus and host. The 22-nt vsRNAs were the predominant size class of vsRNAs spanning the RYNV coding regions and concentrated in areas such as ORF 2 that produced considerable secondary structure. Another study that partially examined size and distribution of vsRNAs in badnavirus-infected tissue, where it was also found that the majority of vsRNAs were 22-nt in size and the sRNAs were distributed along the genome (Kreuze et al. 2009). Mapping the

vsRNAs to the RYNV intergenic region revealed that they were predominately 21 and 22-nts in size **(Figure 4.3)**. Previously, it was demonstrated that in *Arabidopsis thaliana* and *Brassica rapa* exposed to CaMV, the 24-nt size class of vsRNAs was predominant and that massive amounts of vsRNAs clustered within the long pararetrovirus leader sequence (Moissiard & Voinnet 2006; Blenins et al. 2011), rather than being dispersed throughout the genome, as with RYNV. Interestingly, the evidence provided by this study suggests that plant pararetroviruses do indeed induce the host plants siRNA antiviral defense pathway and that the trigger of the siRNA pathway is represented throughout the genome of both sense and antisense strands of RYNV. These results contrast with the 24-nt siRNAs being clustered at the intergenic region of CaMV infecting *A. thaliana* (Blenins et al. 2011). It is possible that these differences may result because of the relative contribution of distinct DCLs in vsRNA biogenesis in different host plants (Akbergenov et al. 2006). Nevertheless, the findings in this study clearly demonstrated that red raspberry induces the small interfering pathway against RYNV, a member of genus *Badnavirus,* but the exact trigger within the replication cycle remains to be determined. Small RNA sequencing of other host plants with badnavirus infection will give insight into whether or not the small interfering pathway is broadly used against members of the genus *Badnaviruses*.

4.5. Conclusion

Advances in technologies such as small RNA sequencing and the use of model organisms will continue to be valuable sources of knowledge that enable us to increase the strategies available for managing diseases such as those associated with the badnaviruses in economically important plants. Sequence from RYNV indicated that it is a distinctly related member of genus *Badnavirus* and several previously unreported nucleic acid and amino acid sequences were reported, increasing the complexity of the pararetroviruses and provided evidence of genome-wide distribution of vsRNA in a badnavirus.

5.0 Conclusion and Future Direction

Abiotic and biotic stresses alter genome stability and physiology of plants. Under some stressful situations, a state of stress tolerance can be passed on to the offspring rendering them more suitable to stressful events than their parents. In plants, the exploration of transgenerational response has remained exclusive to model species, such as *Arabidopsis thaliana*. This study expands transgenerational research to include *Brassica rapa*, a close relative to economically important plant canola (*B. napus*), as it is exposed to the biotic stress of a double-stranded DNA virus Cauliflower mosaic virus (CaMV).

Parent plants exposed to a low dose of 50 ng purified CaMV virions just prior to the bolting stage produced significantly larger seeds than mock inoculated and healthy treatments. The progeny from these large seeds displayed resistance to the pathogen stress applied in the parental generation. Differences in defense pathways involving fatty acids, and primary and secondary metabolites were detected by *de novo* transcriptome sequencing of CaMV challenged progeny exhibiting different levels of resistance.

This study highlights the biological and cellular processes that may be linked to the growth and yield of economically important *B. rapa*, in a transgenerational manner. Although much remains unknown as to the mechanism behind transgenerational inheritance, this work shows a disease resistance response that persists for several weeks and is associated with an increase in seed size. Evidence suggests that a number of changes involved in the persistent stress

adaptation are reflected in the transcriptome. Future work should be directed towards characterizing the biochemistry of the seeds that provide resistance and examining other economically important crops for a similar response. The results from this study demonstrate that treating *B. ra*pa with dsDNA virus within a critical time frame and with a specified amount of infectious pathogen produces economically important agricultural plants with superior coping strategies for growing in unfavorable conditions.

A dsDNA virus from a red raspberry (*Rubus idaeus* L.) plant exhibiting symptoms of mosaic and motling in the leaves was cloned, sequenced and characterised. The genomic sequence indicates that the virus was a distinct member of the genus *Badnavirus*, with 7932-bp and seven ORFs, the first three corresponding in size and location to the ORFs found in the type member Commelina yellow mottle virus. Analysis of the genomic sequences detected several features including nucleic acid binding motifs, multiple zinc finger-like sequences and domains associated with cellular signaling. Subsequent sequencing of the small RNAs from RYNV-infected *R. idaeus* leaf tissue was used to determine any RNA sequences targeted by RNA silencing and identifies abundant virus-derived small RNAs (vsRNAs). The majority of the vsRNAs were 22-nt in length. A highly uneven genome-wide distribution of vsRNAs with strong clustering to small defined regions was observed. Future work should focus on locating the silencing suppressor for the members of genus *Badnavirus* and characterization of the interfering RNA pathway in red raspberry more fully.

The results from this study show that sequences of the aphid-transmitted pararetrovirus RYNV are targeted in red raspberry by the interfering RNA pathway, a predominant antiviral defense mechanism in plants. Together these studies on pararetroviruses advance our understanding of host-pathogen interactions and transgenerational memory in a commercially valuable crop.

Since completion of this work, there have been sevrnal new advances in plant biology directly related to this work. A novel DNA virus with grapevine vein-clearing and vine decline was identified and sequenced (Zhang et al. 2011). Phylogentic analysis (**Appendix B**) indicated that *Grapevine vein-clearing virus* (GVCV) is most closely related to *Rubus yellow net virus* and *Gooseberrry vein banding virus.* Advances in high throughput seqencing technology is expected to continue to facilitate the identification and characterization of novel viruses from plant tissue. Similarily, the availability of seqences from high-throughput platforms continues at an accelerated rate. For example, at the time of this study a database of 2500 proteins were publically avaiable and used in the analysis. Within one year after publication of these data form this thesis, the pubically available database increased to include 140 000 proteins. The advances in high throughput sequencing capabilities and accelerated rates of sequence data acquisition continues to provide us with a better understanding of the biological world.

6. Bibliography

Abramoff MD, Magelhaes PJ, Ram SJ. 2004. Image processing with ImageJ. Biophotonics International 11: 36-42.

Agerbirk N, Olsen CE. 2012. Glucosinolate structure and evolution. Phytochemistry 77: 16-45.

Agnew P, Koella JC, Michalakis Y. 2000. Host life history responses to parasitism. Microbes and Infection 2: 891-896.

Agrawal A. 2001. Transgenerational consequences of plant responses to herbivory: an adaptive maternal effect? American Naturalist 157: 555-569.

Akbergenov R, Si-Ammour A, Blevins T, Amin I, Kutter C, et al. 2006. Molecular characterization of geminivirus-derived small RNAs in different plant species. Nucleic Acids Research 34: 462-471.

Amoah S, Kurup S, Lopez CMR, Welham SJ, Powers SJ, et al. 2012. A hypomethylated population of *Brassica rapa* for forward and reverse epi-genetics. BMC Plant Biology 12: 193-210

Anandalakshmi R, Pruss GJ, Ge X, Marathe R, Mallory AC, et al. 1998. A viral suppressor of gene silencing in plants. Preceedings of the National Academy of Sciences 95: 13079-13084.

Angell SM, Davies C, Baulcombe DC. 1996. Cell-to-cell movement of potato virus X is associated with a change in the size-exclusion limit of plasmadesmata in trichome cells of *Nicotiana clevelandii*. Virology 216: 197-201.

Aravin AA, Hannon GJ, Brennecke J. 2007. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science 318: 761-764.

Baulcombe D. 2004. RNA Silencing in plants. Nature 431: 356-363.

Baulcombe DC. 1996. Mechanisms of pathogen-derived resistance to viruses in transgenic plants. Plant Cell 8: 1833-1844.

Baulcombe DC. 1996. RNA as a target and an initiator of post-transcriptional gene silencing in transgenic plants. Plant Molecular Biology 32: 79-88.

Baumberger N, Tsai C, Lie M, Havecker E, Baulcombe DC. 2007. The polerovirus silencing suppressor P0 targets ARGONAUTE proteins for degradation. Current Biology 17: 1609-1614.

Bayne EH, Rakitina DV, Morozov SY, Baulcombe DC. 2005. Cell-to-cell movement of Potato Potexvirus X is dependent on suppression of RNA silencing. Plant Journal 44: 471-482.

Beachy RN. 1997. Mechanisms and applications of pathogen-derived resistance in transgenic plants. Current Opinion in Biotechnology. 8: 215-220.

Benfey PN, Chua N. 1990. The *cauliflower mosaic virus* 35S promoter: combinational regulation of transcription in plants. Science 250: 959-966.

Bennetzen JL, Wang H. 2014. The contributions of transposable elements to the structure, function and evolution of plant genomes. Annual Review of Plant Biology 65: 505-530.

Bernstein E, Caudy AA, Hammond SM, Hannon GJ. 2001. Role of a bidentate ribonuclease in the initiation step of RNA interference. Nature 409: 363-66.

Blevins T, Rajeswaran R, Shivaprasad PV, Beknazariants D, Si-Ammour A, et al. 2006. Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. Nucleic Acids Research 34: 6233-46.

Blevins T, Rajeswaran R, Aregger M, Borah BK, Schepetilnikov M, et al. 2011. Massive production of small RNAs from a non-coding region of *Cauliflower mosaic virus* in plant defense and viral counter-defense. Nucleic Acids Research 39: 5003-5014.

Boller T, Felix G. 2009. A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. Annual Review of Plant Biology 60: 379-406.

Bollman KM, Aukerman MJ, Park MY, Hunter C, Berardini TZ, et al. 2003. HASTY, the *Arabidopsis* ortholog of exportin 5/MSN5, regulates phase change and morphogenesis. Development 130: 1493-1504.

Bouche N, Lauressergues D, Gasciolli V, Vaucheret H. 2006. An antagonistic function for *Arabidopsis* DCL2 in development and a new function for DCL4 in generating viral siRNAs. EMBO Journal 25: 3347-56.

Boutet N, Vazquez F, Liu J, Beclin C, Fagard M, et al. 2003. Arabidopsis HEN1: a genetic link between endogenous miRNA controlling development and siRNA controlling transgene silencing and virus resistance. Current Biology 13: 843-48.

Boyko A, Blevins T, Yao Y, Golubov A, Bilichak A, et al. 2010. Transgenerational adaption of Arabidopsis to stress requires DNA methylation and the function of dicer-like proteins. Plos One 5: e9514.

Boyko A, Kathiria P, Zemp FJ, Yao Y, Pogribny I, et al. 2007. Transgenerational changes in the genome stability and methylation in pathogen-infected plants (Virus-induced plant genome instability). Nucleic Acids Research 35: 1714-1725.

Boyko A, Kovalchuk I. 2011. Genome instability and epigenetic modification - heritable responses to environmental stress? Current Opinion in Plant Biology 14: 260-266.

Brigneti G, Voinnet O, Li WX, Ji LH, Ding SW, et al. 1998. Viral pathogenicity determinants are suppressors of transgene silencing in *Nicotiana bethamiana*. EMBO Journal 17: 6739-46.

Bronkhorst, A., van Cleef K.W.R., Vodovar, N., Ince I.A., Blanc, H., et al. 2012. The DNA virus invertebrate iridescent virus 6 is a target of the Drosophila RNAi machinery. Proceedings of the National Academy of Sciences 109: E3604-E3613.

Brosnan CA, Mitter N, Christie M, Smith Na, Waterhouse PM, et al. 2007. Nuclear gene silencing directs reception of long-distance mRNA silencing in *Arabidopsis*. Proceedings of the National Academy of Sciences 104: 14741-14746.

Bouhida M, Lockhart BEL, Olszewski NE. 1993. An analysis of the complete sequence of a *sugarcane bacilliform virus* genome infectious to banana and rice. Journal of General Virology 74: 15-22.

Buxdorf K, Yaffe H, Barda O, Levy M. 2013. The effects of glucosinolates and their breakdown products on necrotrophic fungi. Plos One 8(8): e70771.

Carluccio AV, Zicca S, Stavolone L. 2014. Hitching a ride on vesicles: Cauliflower mosaic virus movement protein trafficking in the endomembrane system. Plant Physiology DOI:10.1104/pp.113.234534.
Carrillo-Tripp J, Shimada-Beltrán H, Rivera-Bustamante R. 2006. Use of geminiviral vectors for functional genomics. Current Opinion in Plant Biology 9: 209-215.

Case A, Lacey E, Hopkins R. 1996. Parental effects in *Plantago lanceolata* L. II. manipulation of grandparental temperature and parental flowering time. Heredity 76: 287-295.

Chapman EJ, Prokhnevsky AI, Gopinath K, Dolja VV, Carrington JC. 2004. Viral RNA silencing suppressors inhibit the microRNA pathway at an intermediate step. Genes & Deveopment 18: 1179-1186.

Chellappan P, Vanitharani R, Fauquet CM. 2005. MicroRNA-binding viral protein interfers with *Arabidopsis* development. Proceedings of the National Academy of Sciences 102: 10381-10386.

Cogoni C, Irelan JT, Schumacher M, Schmidhauser TJ, Selker Eu, et al. 1996. Transgene silencing of the al-1 gene in vegetative cells of *Neurospora* is mediated by a cytoplasmic effector and does not depend on DNA-DNA interactions or DNA methylation. EMBO Journal 15: 3153-3163.

Conrath U. 2011. Molecular aspects of defence priming. Trends in Plant Science 16: 524-531.

Conti E, Zadra C, Salerno G, Leombruni B, Volpe D, et al. 2008. Changes in the volatile profile of *Brassica oleracea* due to feeding and oviposition by *Murgantia histrionica* (Heteroptera: Pentatomidae). European Journal of Entomology 105: 839-847.

Curtin SJ, Watson JM, Smith NA, Eamens AI, Blanchard CL, et al. 2008. The roles of plant dsRNA-binding proteins in RNAi-like pathways. FEBS Letters 582: 2753-2760.

Dalmay T, Hamilton AJ, Rudd S, Angell S, Baulcombe DC. 2000. An RNA-dependent RNA polymerase gene in *Arabidopsis* is required for posttranscriptional gene silencing mediated by a transgene but not a virus. Cell 101: 543-553.

Dalmay T, Horsefeld R, Braunstein TH, Baulcombe DC. 2001. SDE3 encodes an RNA helicase required for post-transcriptional gene silencing in Arabidopsis. EMBO Journal 20: 2069-2078.

Dangl JL, Jones JD. 2001. Plant pathogens and integrated defence responses to infection. Nature 411: 826-833.
Deleris A, Gallego-Bartolome J, Bao J, Kasschau KD, Carrington JC, et al. 2006. Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. Science 313: 68-71.

Diaz-Pendon JA, Li F, Li WX, Ding SW. 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. Plant Cell 19: 2053-2063.

Ding SW, Voinnet O. 2007. Antiviral immunity directed by small RNAs. Cell 130: 413-426.

Diwan N, Fluhe R, Eshed Y, Zamir D, Tanksley SD. 1999. Mapping Ve in tomato: A gene conferring resistance to the broad spectrum pathogen *Verticillium dahlia* race 1. Theoretical Applied Genetics 98: 315-319.

Dodds PN, Rathjen JP. 2010. Plant immunity: towards an integrated view of plant-pathogen interactions. Nature Reviews Genetics 11: 539-548.

Doyle JJ, Doyle JL. 1990. Isolation of plant DNA from fresh tissue. Focus 12: 13-15.

Dunoyer P, Schott G, Himber C, Meyer D, Takeda A, et al. 2010. Small RNA duplexes function as mobile silencing signals between plant cells. Science 328: 912-915.

Durrant WE, Dong X. 2004. Systemic acquired resistance. Annual Reviews in Phytopathology 42: 185-209.

Eamens AL, Smith NA, Curtin SJ, Wang MB, Waterhouse PM. 2009. The *Arabidopsis thaliana* double-stranded RNA binding protein DRB1 directs guide strand selection from microRNA duplexes. RNA 15: 2219-2235.

Feinberg EH, Hunter CP. 2003. Transport of dsRNA into cells by the transmembrane protein SID-1. Science 301: 1545-1547.

Feng F, Zhou JM. 2012. Plant-bacterial pathogen interactions mediated by type III effectors. Current Opinion in Plant Biology 15: 469-476.

Fordyce JA, Agrawal AA. 2001. The role of plant trichomes and caterpillar group size on growth and defence of the pipevine swallowtail *Buttus philenar*. Journal of Animal Ecology 70: 997-1005.

Foyer CH, Noctor G. 2013. Redox signaling in plants. Antioxidants & Redox Signaling 18: 2087-2090.

Fu ZQ, Dong X. 2013. Systemic acquired resistance: turning local infection into global defense. Annual Review of Plant Biology 64: 839-863.

Fusaro AF, Matthew L, Smith NA, Curtin SJ, Dedic-Hagan J, et al. 2006. RNA interference-inducing hairpin RNAs in plants act through the viral defence pathway. EMBO Reports 7: 1168-1175.

Farzadfar S, Pourrahim R. 2013. Biological and molecular variation of Iranian Cauliflower mosaic virus (CaMV) isolates. Virus Genes 47: 347-356.

Ferreyra MLF, Ruis SP, Casati P. 2012. Flavonoids: biosynthesis, biological functions, and biotechnology applications. Frontiers in Plant Science 3: 222-235.

Galloway LF, Etterson JR. 2009. Plasticity to canopy shade in a monocarpic herb: within-and between- generation effects. New Phytologist 182: 1003-1012.

Gao Z, Wang M, Blair D, Zheng Y, Dou Y. 2014. Phylogenetic analysis of the endoribonuclease Dicer family. Plos One DOI: 10.1371/journal.pone .0095350.

Geering, A.D., Pooggin, M.M., Olszewski, N.E., Lockhart, B.E., Thomas, J.E., 2005. Characterization of Banana streak Mysore virus and evidence that its DNA is integrated into the B genome of cultivated *Musa*. Archives of Virology 150, 787-796.

Geijskes RJ, Braithwaite KS, Dale JL, Harding RM, Smith GR. 2002. Sequence analysis of an Australian isolate of sugarcane bacilliform badnavirus. Archives of Virology 147: 2393-2404.

Gene Ontology Consortium. 2004. The gene ontology (GO) database and informatics resource. Nucleic Acids Research 32(suppl1): D258-D261.

Glazebrook J. 2005. Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. Annual Reviews of Phytopathology 43: 205-227.

Gómez-Gómez L, Boller T. 2000. An LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. Molecular Cell 5: 1003-1011.

Gutiérrez S, Yvon M, Pirolles E, Garzo E, Fereres A, et al. 2012. Circulating virus load determines the size of bottlenecks in viral populations progressing within a host. Plos Pathogens 8:e1003009.

Greenberg JT, Yao N. 2004. The role and regulation of programmed cell death in plant-pathogen interactions. Cellular Microbiology 6: 201-211.

Gruber MY, Wang S, Ethier S, Holowachuk KJ, Bonham-Smith PC, et al. 2006. Hairy Canola - Arabidopsis *GL3* induces a defense covering of trichomes on *Brassica napus* seedlings. Plant Molecular Biology 60: 679-698.

Gutierrez JR, Balmuth AL, Ntoukakis V, Mucyn TS, Gimenez-Ibanez S, et al. 2010. Prf immune complexes of tomato are oligomeric and contain multiple Pto-like kinases that diversify effector recognition. Plant Journal 61: 507-518.

Haas BJ, Kamoun S, Zody MC, Jiang RHY, Handsaker RE, et al. 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature 461: 393-398.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, et al. 2013. *De novo* transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. Nature Protocols 8: 1494-1512.

Haas G, Azevedo J, Moissiard G, Geldreich A, Himber C, et al. 2008. Nuclear import of CaMV P6 is required for infection and suppression of the RNA silencing factor DRB4. EMBO Journal 27: 2102-2112.

Haas M, Bureau M, Geldreich A, Yot P, Keller M. 2002. Cauliflower mosaic virus: still in the news. Molecular Plant Pathathology 3: 419-429.

Hagen LS, Jacquemond M, Lepingle A, Lot H, Tepfer M. 1993. Nucleotide sequence and genomic organization of cacao swollen shoot virus. Virology 196: 619-628.

Hamilton A, Voinnet O, Chappell L, Baulcombe D. 2002. Two classes of short interfering RNA in RNA silencing. EMBO Journal 21: 4671-4679.

Hannon GJ. 2002. RNA interference. Nature 418: 244-251.

Harper G, Hart D, Moult S, Hull R. 2004. Banana streak virus is very diverse in Uganda. Virus Research 100: 51-56.

Harper G, Hull R. 1998. Cloning and sequence analysis of banana streak virus DNA. Virus Genes 17: 271-278.

Hartig JV, Tomari Y, Forstemann K. 2007. piRNAs the ancient hunters of genomic invaders. Genes & Development 21: 1707-1713.

Hauser MT, Aufsatz W, Jonak C, Luschnig C. 2011. Transgenerational epigenetic inheritance in plants. Biochimica Biophysica Acta (BBA) - Gene Regulatory Mechanisms 1809: 459-468

Hay JM, Jones MC, Blakebrough ML, Dasgupta I, Davies JW, Hull R. 1991. An analysis of the sequence of an infectious clone of rice tungro bacilliform virus, a plant pararetrovirus. Nucleic Acids Research 19: 2615-2621.

Herridge RP, Day RC, Baldwin S, Macknight RC. 2011. Rapid analysis of seed size in Arabidopsis for mutant and QTL discovery. Plant Methods 7: 3-13.

Hiragui A, Itoh R, Kondo N, Nomura Y, Aizawa D, et al. 2005. Specific interactions between Dicer-like proteins and HYL1/DRB-family dsRNA-binding proteins in *Arabidopsis thaliana*. Plant Molecular Biology 57: 173-188.

Holeski LM, Jander G, Agrawal AA. 2012. Transgenerational defense induction and epigenetic inheritance in plants. Trends in Ecology and Evolution 27: 618-626.

Huang Q, Hartung JS. 2001. Cloning and sequence analysis of an infectious clone of *Citrus yellow mosaic virus* that can infect sweet orange via *Agrobacterium*-mediated inoculation. Journal of General Virology 82: 2549-2558.

Hull R, Shepherd RJ. 1976. Cauliflower mosaic virus: An improved purification procedure and some properties of the virus particles. Journal of General Virology 31: 93-100.

Huntzinger E, Izaurralde E. 2011. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nature Reviews Genetics 12: 99-110.

Huot B, Yao J, Montgomery BL, He SY. 2014. Growth-defense tradeoffs in plants: A balancing act to optimize fitness. Molecular Plant 7: 1267-1287.

Hwang YT, Kalischuk M, Fusaro A, Waterhouse PM, Kawchuk L. 2013. Small RNA sequencing of *Potato leafroll virus* infected plants reveals an additional subgenomic RNA encoding a sequence-specific RNA binding protein. Virology 438: 61-69.

Ito H, Kakutani T. 2014. Control of transposable elements in *Arabidopsis thaliana*. Chromosome Research 22: 217-223.
Iwasaki M, Paszkowski J. 2014. Identification of genes preventing transgenerational transmission of stress-induced epigenetic states. Proceedings of the National Academy of Sciences doi/10.1073/pnas.1402275111.

Jablonka E. 2013. Epigenetic inheritance and plasticity: The responsive germline. Progress in Biophysical and Molecular Biology 11: 99-107.

Jablonka E, Raz G. 2009. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. Quarterly Reviews of Biology 84: 131-176.

Jacquot E, Hagen LS, Jacquemond M, Yot P, 1996. The open reading frame 2 product of *cacao swollen shoot badavirus* is a nucleic acid-binding protein. Virology 225: 191-195.

Jacquot E, Keller M, Yot P. 1997. A short domain supports a nucleic acid-binding activity in the *rice tungro bacilliform virus* open reading frame 2 product. Virology 239: 352-359.

James AP, Geijskes RJ, Dale JL, Harding RM. 2011. Molecular characterisation of six badnavirus species associated with leaf streak disease of banana in East Africa. Annuals of Applied Biology 158: 346-353.

Jones AT, McGavin WJ, Geering ADW, Lockhart BEL. 2002. Identification of *Rubus yellow net virus* as a distinct badnavirus and its detection by PCR in *Rubus* species and in aphids. Annuals of Applied Biology 141: 1-10.

Jones JD, Dangl JL. 2006. The plant immune system. Nature 444: 323-329.

Kalischuk ML, Johnson D, Kawchuk LM. 2014. Priming with a double-stranded DNA virus alters *Brassica rapa* seed architecture and facilitates a defense response. Gene (in press).

Kalischuk ML, Fusaro AF, Waterhouse PW, Pappu HR, Kawchuk LM. 2013. Complete genomic sequence of *a Rubus yellow net virus* isolate and detection of genome-wide pararetrovirus-derived small RNAs. Virus Research 178: 306-313.

Kalischuk ML, Kawchuk LM, Leggett F. 2008. First report of *Rubus yellow net virus* on *Rubus idaeus* in Alberta, Canada. Plant Disease 92: 974.

Kathiria P, Sidler C, Golubov A, Kalischuk M, Kawchuk LM, et al. 2010. Tobacco mosaic virus infection results in an increase in recombination frequency and resistance to viral, bacterial and fungal pathogens in the progeny of infected tobacco plants. Plant Physiology 153: 1859-1870.
Kasschau KD, Carrington JC. 1998. A counterdefensive strategy of plant viruses: suppression of posttranscriptional gene silencing. Cell 95: 461-70.

Kawchuk LM, Hachey J, Lynch DR, Kulcsar F, van Rooijen G, et al. 2001. Tomato *Ve* disease resistance genes encode cell surface-like receptors. Proceedings of the National Academy of Sciences 98: 6511-6515.

Kawchuk LM, Kalischuk ML, Bizimungu B, Leggett F. 2009. Advances in molecular assisted breeding for disease resistance in Solanaceous plants. Recent Research Developments in Genetics & Breeding 3: 49-60.

Khurana D, Baranwal VK. 2011. Full genome characterization of Banana streak virus from India. unpublished.

Khvorova A, Reynolds A, Jayasena SD. 2003. Functional siRNA and miRNAs exhibit strand bias. Cell 115: 209-216.

Klattenhoff C, Theurkauf W. 2008. Biogenesis and germline functions of piRNAs. Development 135: 3-9.

Korves T, Bergelson J. 2003. A developmental response to pathogen infection in Arabidopsis. Plant Physiology 133: 339-347.

Kovalchuk I, Kovalchuk O, Kalck V, Boyko V, Filkowski J, et al. 2003. Pathogen-induced systemic plant signal triggers DNA rearrangements. Nature 12: 760-762.

Kreuze JF, Perez A, Untiveros M, Quispe D, Fuentes S, et al. 2009. Complete viral genomic sequence and discovery of novel viruses by deep sequencing of small RNAs: A generic method for diagnostics, discovery and sequencing of viruses. Virology 388: 1-7.

Laco GS, Kent SBH, Beachy R. 1995. Analysis of the proteolytic processing and activation of the *rice tungro bacilliform virus* reverse transcriptase. Virology 208: 207-214.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody M, et al. 2001. Initial sequencing and analysis of the human genome. Nature 409: 860-921.

Laney AG, Hassan M, Tzanetakin IE. 2012. An intergrated badnavirus is prevalent in fig germplasm. Phytopathology 102: 1182-1189.

Lau P, Guiley K, De N, Potter CS, Carragher B, et al. 2012. The molecular architecture of human Dicer. Nature Structural and Molecular Biology 19: 436-440.

Leclerc D, Burri L, Kajava AV, Mougeot J, Hess D, et al. 1998. The open reading frame III product of *cauliflower mosaic virus* forms a tetramer through a N-terminal coiled-coil. Journal of Biological Chemistry 273: 29015-29021.

Lee SW, Han SW, Sririyanum M, Park CJ, Seo YS, et al. 2009. A type I-secreted peptide triggers XA21-mediated innate immunity. Science 326: 850-853.

Lheureux F, Laboureau N, Muller E, Lockhart BEL, Iskra-Caruana ML. 2007. Molecular characterization of banana streak acuminata Vietnam virus isolated

from *Musa acuminata* siameg (banana cultivar). Archives of Virology 152: 1409-1416.

Li HW, Ding SW. 2005. Antiviral silencing in animals. FEBS Letters 579: 5965-5973.

Li H, Li WX, Ding SW. 2002. Induction and suppression of RNA silencing by an animal virus. Science 296: 1319-1321.

Li WX, Li H, Lu R, Li F, Dus M, et al. 2004. Interferon antagonist proteins of influenza and vaccinia viruses are suppressors of RNA silencing. Preecedings of the National Academy of Sciences 101: 1350-1355.

Li HW, Lucy AP, Guo HS, Li WX, Ji LH, et al. 1999. Strong host resistance targeted against a viral suppressor of the plant gene silencing defence mechanism. EMBO Journal 18: 2683-2691.

Li J, Yang Z, Yu B, Liu J, Chen X. 2005. Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in *Arabidopsis*. Current Biology 15: 1501-1507.

Liu L, Zhang J, Adrian J, Gissot L, Coupland G, et al. 2014. Elevated levels of MYB30 in the phloem accelerate flowering in Arabidopsis through the regulation of FLOWERING LOCUS T. PLoS One 9: e89799.

Lindbo JA, Silva-Rosales L, Proebsting WM, Dougherty WG. 1993. Induction of a highly specific antiviral state in transgenic plants: implications for regulation of gene expression and virus resistance. Plant Cell 5: 1749-1759.

Lindeberg M, Cunnac S, Collmer A. 2012. *Pseudomonas syringae* type III effector repertoires: last words in endless argument. Trends in Microbiology 20: 199-208.

Lockhart BEL. 1990. Evidence for a double-stranded circular DNA genome in a second group of plant viruses. Phytopathology 80: 127-131.

Lomonossoff GP. 1995. Pathogen-derived resistance to plant viruses. Annual Review of Phytopathology. 33: 323-343.

Love AJ, Yun BW, Laval V, Loake GJ, Milner JJ. 2005. Cauliflower mosaic virus, a compatible pathogen of Arabidopsis, engages three distinct defense-signaling pathways and activates rapid systemic generation of reactive oxygen species. Plant Physiology 139 : 935-948.

Lu JJ, Boa JL, Wu GS, Xu WS, Huang MQ, et al. 2013. Quinones derived from plant secondary metabolites are anti-cancer agents. Anti-cancer Agents in Medicinal Chemistry 13: 456-463.

Lu R, Maduro M, Li F, Li HW, Broitman-Maduro G, et al. 2005. Animal virus replication and RNAi-mediated antiviral silencing in *Caenorhabditis elegans*. Nature 436: 1040-1043.

Lu R, Martin-Hernandez AAM, Peart JR, Malcuit I, Baulcombe DC. 2003. Virus-induced gene silencing in plants. Methods 30: 296-303.

Luna E, Bruce TJA, Roberts MR, Flors V, Ton J. 2012. Next-generation systemic acquired resistance. Plant Physiology 158: 844-853

Malamy J, Carr JP, Klessig DF, Raskin I. 1990. Salicyclic acid: a likely endogenous signal in the resistance response of tobacco to viral infection. Science 250: 1002-1004.

Mallory AC, Ely L, Smith TH, Marathe R, Anandalakshmir R, et al. 2001. HC-Pro suppression of gene silencing eliminates the small RNAs but not the transgene methylation or the mobile signal. Plant Cell 13: 571-583.

Margis R, Fusaro AF, Smith NA, Curtin SJ, Watson JM, et al. 2006. The evolution and diversification of Dicers in plants. FEBS Letters 580: 2442-2450.

Marmey P, Bothner B, Jacquot E, de Kochko A, Ong CA, et al. 1999. *Rice tungro bacilliform virus* open reading frame 3 encodes a single 37-kDa coat protein. Virology 253: 319-326.
Mathieu J, Schwizer S, Martin GB. 2014. Pto kinase binds two domains of AvrPtoB and its proximity to the effector E3 ligase determines if it evades degradation and activates plant immunity. PLOS Pathogens 10(7): e1004227.

Matzke M, Kanno T, Daxinger L, Huettel B, Matzke AJM. 2009. RNA-mediated chromatin-based silencing in plants. Current Opinion in Cell Biology 21: 367-376.

Matzke MA, Mette MF, Matzke AJM. 2000. Transgene silencing by the host genome defense: implications for the evolution of epigenetic control mechanisms in plants and vertebrates. Plant Molecular Biology 43: 401-415.

McGavin WJ, MacFarlane SA. 2010. Sequences similarities between Raspberry leaf mottle virus, Raspberry leaf spot virus and the closterovirus Raspberry mottle virus. Annals of Applied Biology 156: 439-448.

Medberry SL, Lockhart BE, Olszewski NE. 1990. Properties of *Commelina yellow mottle virus's* complete DNA sequence, genomic discontinuities and transcript suggest that it is a pararetrovirus. Nucleic Acids Research 18: 5505-5513.

Medbery S, Olszewski NE. 1993. Identification of *cis* elements involved in *Commelina yellow mottle virus* promoter activity. Plant Journal 3: 619-626.

Melcher U. 2000. The '30K' superfamily of viral movement proteins. Journal of General Virology 81: 257-266.

Meldau S, Erb M, Baldwin IT. 2012. Defence on demand: mechanisms behind optimal defence patterns. Annals of Botany 110: 1503-1514.

Metraux JP, Signer H, Ryals J, Ward E, Wyss-Benz M, et al. 1990. Increase in salicylic acid at the onset of systemic acquired resistance in cucumber. Science 250: 1004-1006.

Mithöfer A, Boland W. 2012. Plant defense against herbivores: chemical aspects. Annual Review of Plant Biology 63: 431-450.

Moissiard G, Voinnet O. 2006. RNA silencing of host transcripts by cauliflower mosaic virus requires coordinated action of the four *Arabidopsis* Dicer-like proteins. Proceedings of the National Academy of Sciences 103: 19593-19598.

Molnar A, Melnyk CW, Bassett A, Hardcastle TJ, Dunn R. 2010. Small silencing RNAs in plants are mobile and direct epigenetic modification in recipient cells. Science 328: 872-875.

Montillet JL, Leonhardt N, Mondy S, Tranchimands S, Rumeau D, et al. 2013. An abscisic acid-independent oxylipin pathway controls stomatal closure and immune defense in Arabidopsis. PLoS Biol. 11: e1001513.

Mourrain P, Beclin C, Elmayan T, Feuerbach F, Godon C, et al. 2000. *Arabidopsis* SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. Cell 101: 533-542.

Morel JB, Godon C, Mourrain P, Beclin C, Boutet S, et al. 2002. Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. Plant Cell 14: 629-639.

Nagamatsu A, Masuta C, Senda M, Matsuura H, Kasai A, et al. 2007. Functional analysis of soybean genes involved in flavonoid biosynthesis by virus-induced gene silencing. Plant Biotechnology Journal 5: 778-790.

Nakazawa Y, Hiraguri A, Moriyama H, Fukuhara T. 2007. The dsRNA-binding protein DRB4 interacts with the Dicer-like protein DCL4 in vivo and functions in the trans-acting siRNA pathway. Plant Molecular Biology 63: 777-85.

Padmanabhan C, Zhang X, Jin H. 2009. Host small RNAs are big contributors to plant innate immunity. Current Opinion in Plant Biology 12: 465-472.

Palauqui JC, Elmayan T, Pollien JM, Vaucheret H. 1997. Systemic acquired silencing: transgene specific post-transcriptional silencing is transmitted by grafting from silenced stocks to non-silenced scions. EMBO Journal 16: 4738-4745.

Parker JS, Roe SM, Barford D. 2005. Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. Nature 434: 663-666.

Pecinka A, Rosa M, Schikora A, Berlinger M, Hirt H, et al. 2009. Transgenerational stress memory is not a general response in Arabidopsis. PLoS One 4: e5202.

Peele C, Jordan CV, Muanfsan N, Turnage M, Egelkrout E, et al. 2001. Silencing of a meristematic gene using geminivirus-derived vectors. Plant Journal 27: 357-366.

Peters AD. 1999. The effects of pathogen infection and mutation on life history characteristics in *Arabidopsis thaliana*. Journal of Evolutionary Biology 12:460-470.
Petutschnig EK, Jones AM, Serazetdinoval, Lipka U, Lipka V. 2010. The lysin motif receptor-like kinase (LysM-RLK) CERK1 is a major chitin-binding protein in *Arabidopsis thaliana* and subject to chitin-induced phosphorylation. Journal of Biological Chemistry 285: 28902-28911.

Pfeffer S, Zavolan M, Grasser FA, Chien M, Russo JJ et al. 2004. Identification of virus encoded microRNAs. Science 304: 734-736.

Pieterse CMJ, Zamioudis C, Van der Does D, Van Wees SCM. 2014. Signalling networks involved in induced resistance. *In* Induced resistance for plant defense: A sustainable approach to crop production. 2nd ed. John Wiley & Sons Ltd., New York.

Qi Y, He X, Wang XJ, Kohany O, Jurka J, Hannon GJ. 2006. Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. Nature 443: 1008-1112.

Qu F, Ye X, Morris TJ. 2008. *Arabidopsis* DRB4, AGO1, AGO7, and RDR6 participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1. Proceedings of the National Academy of Sciences 105: 14732-14737.

Quito-Avila DF, Martin RR. 2012. Real-time RT-PCR for detection of *Raspberry bushy dwarf virus*, *Raspberry leaf mottle virus* and characterizing synergistic interactions in mixed infections. Journal of Virological Methods 179: 38-44.

Radulovic NS, Blagojevic PD, Stajanovic-Radic ZZ, Stojanovic NM. 2013. Antimicrobial plant metabolites: structural diversity and mechanism of action. Current Medicinal Chemistry 20: 932-952.

Ratcliff F, Harrison BD, Baulcombe DC. 1997. A similarity between viral defense and gene silencing in plants. Science 276: 1558-1560.

Rempel CB, Hutton SN, Jurke CJ. 2014. Clubroot and the importance of canola in Canada. Canandian Journal of Plant Pathology 36: 19-26.

Rojas CM, Senthil-Kumar M, Tzin V, Mysore KS. 2014. Regulation of primary plant metabolism during plant-pathogen interactions and its contribution to plant defense. Frontiers of Plant Science 5: 17.

Rosli HG, Zheng Y, Pombo MA, Zhong S, Bombarely A, et al. 2013. Transcriptomics-based screen for genes induced by flagellin and repressed by pathogen effectors identifies a cell wall-associated kinase involved in plant immunity. Genome Biology 14: R139.
Ruiz-Ferrer V, Voinnet O. 2009. Roles of plant small RNAs in biotic stress responses. Annual Reviews of Plant Biology 60: 485-510.

Ruiz-Medrano R, Xocosnostle-Cazares B, Lucus WJ. 1999. Phloem long-distance transport of CmNACP mRNA: implications for supracellular regulation in plants. Development 126: 4405-4419.

SanMiguel P, Tikhonov A, Jin YK, Motchoulskaia N, Zakharov D, et al. Nested retrotransposon in the intergenic regions of the maize genome. Science 274: 765-768.

Sasaki T, Chino M, Hayashi H, Fujiwawa T. 1998. Detection of mRNA species in phloem sap. Plant Cell Physiology 39: 895-897.

Schmidt I, Blanc S, Esperandieu P, Kuhl G, Devauchelle G, et al. 1994. Interaction between the aphid transmission factor and virus particles is a part of

the molecular mechanism of cauliflower mosaic virus aphid transmission. Proceedings of the National Academy of Sciences 91: 8885-8889.

Scholthof KBG, Adkins S, Czosnek H, Palukaitis P, Jacquot E, et al. 2011. Top 10 plant viruses in molecular plant pathology. Molecular Plant Pathology 12: 938-954.

Seal S, Muller E. 2007. Molecular analysis of a full-length sequence of a new yam badnavirus from *Dioscorea sansibarensis*. Archives of Virology 152: 819-825.

Senthil-Kumar M, Mysore KS. 2013. Non host resistance against bacterial pathogens: retrospectives and prospects. Annual Reviews of Phytopathology 51: 407-427.

Sether DM, Melzer MJ, Borth WB, Hu, JS. 2012. *Pineapple bacilliform CO virus*: Diversity, Detection, Distribution and Transmission. Plant Disease 96: 1798-1804.

Shah J, Zeier J. 2013. Long-distance communication and signal amplification in systemic acquired resistance. Frontiers of Plant Science 4: A30.

Silhavy D, Molnar A, Lucioli A, Szittya G, Hornyik C, et al. 2002. A viral protein suppresses RNA silencing and binds silencing-generated, 21-to 25-nucleotide double-stranded RNAs. EMBO Journal 21: 3070-3080.

Sijen T, Fleenor J, Simmer F, Thijssen KL, Parrish S, et al. 2001. On the role of RNA amplification in dsRNA-triggered gene silencing. Cell 107: 465-476.

Slaughter A, Daniel X, Flors V, Luna E, Hohn B, et al. 2012. Descendants of primed Arabidopsis plants exhibit resistance to biotic stress. Plant Physiology 158: 835-843.

Smith NA, Singh SP, Wang MB, Stoutjesdijk P, Green A, et al. 2000. Total silencing by intron-spliced haripin RNAs. Nature 407: 319-320.

Song JJ, Smith SK, Hannon GJ, Joshua-Tor L. 2004. Crystal structure of Argonaute and its implications for RISC slicer activity. Science 305: 1434-1437.

Soroka JJ, Holowachuk JM, Gruber MY, Grenkow LF. 2011. Feeding by flea beetles (Coleoptera: Chrysomelidae; *Phyllotreta* spp.) is decreased on canola (*Brassica napus*) seedlings with increased trichome density. Journal of Economic Entomology 104: 125-136.

Spitzer B, Bez Zvi MM, Ovadis M, Marhevka E, Barkai O, et al. 2007. Reverse genetics of floral scent: Application of tobacco rattle virus-based gene silencing in petunia. Plant Physiology 145: 1241-1250.

Su L, Gae S, Huang Y, Ji C, Wang D, et al. 2007. Complete genomic sequence of *Dracaena mottle virus*, a distinct badnavirus. Virus Genes 35: 423-429.

Sullivan CS, Grundhoff AT, Tevethia S, Pipas JM, Ganem D. 2005. SV40-encoded microRNAs regulate viral gene expression and reduce susceptibility to cytotoxic T cells. Nature 435: 682-686.

Stavolone L, Etienne H, Leclerc D, Hohn T. 2001. Tetramerization is a conserved feature of the virion-associated protein in plant pararetroviruses. Journal of Virology 75: 7739-7743.

Stein LD, Bao Z, Blasiar D, Blumenthal T, Brent MR, et al. 2003. The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. PLoS Biology 1(2): E45.

Tanchou V, Decimo D, Péchoux C, Lener D, Rogemond V, et al. 1998. Role of the N-terminal zinc finger of human immunodeficiency virus type 1 nucleocapsid protein in virus structure and replication. Jornal of Virology 72: 4442-4447.

Tang G, Reinhart BJ, Bartel DP, Zamore PD. 2003. A biochemical framework for RNA silencing in plants. Gene & Development 17: 49-63.

Thresh JM, Owusu GK. 1986. The control of cocoa swollen shoot disease in Ghana: an evaluation of eradication procedures. Crop Protection 5: 41-52.

Torres MA, Jones JDG, Dangl JL. 2006. Reactive oxygen species signaling in response to pathogens. Plant Physiology 141: 373-378.

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, et al. 2013. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature Protocols 7: 562-578.

Tzafrir I, Ayala-Navarrete L, Lockhart BEL, Olszewski NE. 1997. The N-terminal portion of the 216-kDa polyprotein of Commelina yellow mottle *badnavirus* is required for virus movement but not for replication. Virology 232: 359-368.

van der Biezen EA, Jones JDG. 1998. Plant disease resistance proteins and the gene-for-gene concept. Trends in Biochemical Sciences 23: 454-456.

Vargason JM, Szittya G, Burgyan J, Tanaka Hall TM. 2003. Size selective recognition of siRNA by an RNA silencing suppressor. Cell 115: 799-811.

Vasyukova NI, Ozeretskovskaya OL. 2007. Induced plant resistance and salicyclic acid: A review. Applied Biochemistry and Microbiology 43: 367-373.

Vaucheret H, Vazquez F, Crete P, Bartel DP. 2004. The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. Genes & Development 18: 1187-1197.

Vernooij B, Friedrich L, Morse A, Reist R, Kolditz-Jawhar R, et al. 1994. Salicylic acid is not the translocated signal responsible for inducing systemic aquired resistance. Plant Cell 6: 959-965.

Voinnet O. 2001. RNA silencing as a plant immune system against viruses. Trends in Genetics 17: 449-459.

Voinnet O. 2008. Use, tolerance and avoidance of amplified RNA silencing by plants. Trends in Plant Science 13: 317-328.

Voinnet O, Lederer C, Baulcombe DC. 2000. A viral movement protein prevents spread of the gene silencing signal in *Nicotiana benthamiana*. Cell 103: 157-167.

Wang XH, Aliyari R, Li WX, Li HW, Kim K et al. 2006. RNA interference directs innate immunity against viruses in adult Drosophila. Science 312: 452-454.

Wang HL, Chuang WH, Wang IC. 2008. Sequencing and molecular analysis of the viral genome of Bougainvillea spectabilis chlorotic vein-banding virus from *Bougainvillea spectabilis*. unpublished.

Wang D, Dong X. 2011. A highway for war and peace: the secretory pathway in plant-microbe interactions. Molecular Plant 4: 581-587.

Wang D, Weaver ND, Kesarwani M, Dong X. 2005. Induction of protein secretory pathway is required for systemic aquired resistance. Science 308: 1036-1040. Wang KLC, Li H, Ecker JR. 2002. Ethylene biosynthesis and signaling networks. Plant Cell 14: S131-S151.

Wang H, Wu J, Sun S, Liu B, Cheng F, et al. 2011. Glucosinolate biosynthetic genes in *Brassica rapa*. Gene 487: 135-142.

Waterhouse PM, Wang MB, Lough T. 2001. Gene silencing as an adaptive defence against viruses. Nature 411: 834-842.

Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, et al. 2002. Initial sequencing and comparative analysis of the mouse genome. Nature 2002: 420: 520-562.

Weber F, Wagner V, Rasmussen SB, Hartmann R, Paludan SR. 2006. Double-stranded RNA is produced by positive-strand RNA viruses and DNA viruses but not in detectable amounts by negative-strand RNA viruses. Journal of Virology 80: 5059-5064.

Wi SJ, Ji NR, Park KY. 2012. Synergistic biosynthesis of biphasic ethylene and reactive oxygen species in response to hemibiotrophic *Phytophthora parasitica* in tobacco plants. Plant Physiology 159: 251-265.

Wilkins C, Dishongh R, Moore SC, Whitt MA, Chow M, et al. 2005. RNA interference is an antiviral defence mechanism in *Caenorhabditis elegans*. Nature 436: 1044-1047.

Winston WM, Molodowitch C, Hunter CP. 2002. Systemic RNAi in *C. elegans* requires the putative transmembrane protein SID-1. Science 295: 2456-2459.

Xie Z, Johansen LK, Gustafson Am, Kasschau KD, Lellis AD, et al. 2004. Genetic and functional diversification of small RNA pathways in plants. Plos Biology 2: E104.
Xie Z, Fan B, Chen C, Chen Z. 2001. An important role of an inducible RNA-dependent RNA polymerase in plant antiviral defense. Proceedings of the National Academy of Sciences 98: 6515-6521.

Xing W, Zou Y, Liu Q, Liu J, Luo X, et al. 2007. The structural basis for activation of plant immunity by bacterial effector protein AvrPto. Nature 449: 243-247.
Xu D, Mock R, Kinard G. 2011. Molecular analysis of the complete genomic sequences of four isolates of *Goosebery vein banding associated virus*. Virus Genes 43: 130-137.

Yang Z, Nicolaisen M, Hobbs W, Sattler S, Olszewski NE, et al. 2003. Complete nucleotide sequence of Kalanchoe top-spotting virus (Genus *Badnavirus*) and detection by PCR of two forms of the virus in *Kalanchoe blossfeldiana*. unpublished.

Ye K, Malinina L, Patel DJ. 2003. Recognition of small interfering RNA by a viral suppressor of RNA silencing. Nature 426: 874-878.

Yu D, Fan B, MacFarlane SA, Chen Z. 2003. Analysis of the involvement of an inducible *Arabidopsis* RNA-dependent RNA polymerase in antiviral defence. Molecular Plant and Microbe Interactions 16: 206-216.

Zanoni I, Ostuni R, Marek LR, Barresi S, Barbalat R, et al. 2011. CD14 controls the LPS-induced endocytosis of Toll-like receptor 4. Cell 147: 868-880.

Zavalieu R, Ueki S, Epel BL, Citovsky V. 2011. Biology of callose (β-1,3-glucan) turnover at plasmadesmata. Protoplasma 248: 117-130.

Zhang L, Lockhart BEL, Lheureux F, Geering ADW, Olszewski NE. 2005. Cloning and sequencing of a second banana streak isolate which may also occur as a retroelement, integrated into the host genome. unpublished.

Zhang X, Yuan YR, Pei Y, Lin SS, Tuschl T, et al. 2006. Cucumber mosaic virus-encoded 2b suppressor inhibits *Arabidopsis* Argonaute1 cleavage activity to counter plant defence. Genes & Development 20: 3255-3268.

Zhuang J, Liu ZX. 2005. Cloning the complete genome of a new Banana streak virus isolate and coat protein gene identification. unpublished.

Ziberman D, Cao X, Jacobsen SE. 2003. ARGONAUTE4 control of locus specific siRNA accumulation and DNA histone methylation. Science 299: 716-719.

Zipfel C. 2009. Early molecular events in PAMP-triggered immunity. Current Opinion in Plant Biology 12: 414-420.

Zhang Y, Singh K, Kaur R, Qiu W. 2011. Association of a noval DNA virus with grapevine vein-clearing and vine decline syndrome. Phytopathology 101: 1081-1090.

Appendix A - Fasta of Rubus yellow net virus isolate Canadian 2

>gi|563580448|gb|KF241951.1| Rubus yellow net virus isolate Canadian 2
hypothetical protein genes, partial cds; hypothetical proteins,
polyprotein, ORF 6, and hypothetical protein genes, complete cds; and
hypothetical protein genes, partial cds
TGGTATCAGAGCTTTAGCTCTCACCATGGCAGCTTAAACACTTCCCTTCTTGTCGAGAAACCCAAGTTTC
AGACCAGAACCTTGAGTTTGCTCTCTTTTTCGGAGGGAAGAGGAGTGAGTGTCTGTGTCAAAACCTTGAA
AGATCAAACCCCCATGAAAACTTTCCTCACGGTACCATGAGTTTTCTATCCTTCACTAGTTTGAACCTAC
TGCTCAAACTGCAGGCTTAGGCGTCGAAGCGAAGTACCCTTGTAGCCGTTAGCAGGAGGCGTTAGGCGTT
GATTGGGGAAAACTGACGTAAAGAAGCAGCAGCAACTAGGCAAGAAACCTGACGGGTAGATCACCGGCCG
GAAAGCCAGTAAGCGGCTAGATCTGGGCAGTTTTGATGCAACCTCACGAAATCTCAGCCTTCGAAGAAGA
AAGCAGCTCTTGGGAAAGGTCTGAACGGGCGTATCGACAAGACTTTTTATTCAGAAATCTCAGAACGTAT
CCACGTTGGGAGGCAAATCAGAAAACACCCTCTCTAGACTTTCCTTGCTACCACTTCAACACAACAACCG
GACCACCAGTCCACCGCACTCTCTGCAGACAAGAGAACAGTAAGGATTTACCATTTCTGGTAAACACCCT
GTTCGATCTCAACATCACCGAGATCCACAATCAGGCGATTCTGGACGATAAGATCTCCAGACTCACCCAG
TACCTGACAACAAAGGTTGGTTCACTACCAACAATCCCGGAGGATTCACCCCTCCTGGACCAAGCCACAA
TATCCTTAGATCTTCAAGCCCTCAAGGCAGATCTGAAGGAAATCAAGGCCACACAGTCAGCCCTGAAGCT
AGGCTTCGCACAACTGCAGGAGGCAGTTCAGCTGATCATCACAAGGGAAAACGATCCCAAACCAATCGAA
GCAGCTACTGCACAAGTAGCCGAACAGCTGAGGAAGCAGCTTATTGAGGTCAAGTCCGTCCTCGAGGAGA
CCAAGAAGATCGCGAGATCTCTGTCCCCCGACGGATGAACCCTAGGTGGCAGGATACTGCAACCAAGGAA
ACCTACCTTAAAGCCATACAAGCTACCTCATCTCTCACCTCCAACAACACAGGTCTAGGCTTCATCGAGC
CACATACCTACACCGGAGGACAGCTATCTACCAACCTAGCAAAACAGAACAACACGCTCATCCAGCTGTT
AGTTCAGGTGCTAGAAAAGAACCTCGACCTCGAGCAGGCAATTGTCAACCTCACAGCTCAGGTCACAAGG
CTAGAAAAGACCGTCTCGGAGAAAGACACGGTCAAACTCCCAGAAAGTGTCCTCAACGACCTCACTAAGG
AGTTCGGAAAGGTCAACTTAGGGAAAGGAAAGGGGATAGAAGGAGCAGTCTCATCCAGAGACAAGAACTT
CTACGTCTGGAAGAACCCCTTCAATCAATACAATGAGCAGAAGCCAAACAAGGCTCCAGGCTCCGCCCGC
AACTGAAAGGGCAACAAGCAGCTCGGAATCAGGCACCCCCACCTTGGAGGACCAGATCCGAGGATACAGG
CGCTCCGCAAGGTTACGACACCAGGCGCAGCGAGCAATGAGAAGGACCTTCAGTAGGGACTTCAGAAACA
CCATAGAACGGCAACTAGACCCAGATGCCGAGCTTTCCCTCAGCAGAAGGAGGAGAGCGAACCTAGTACC
CGCGGAGGTACTATATGCACACAATGGCTCTGAGCCAGTAAACCGTGTGTACGAGCACTACAGTGAGCTC
AGCGCTCATGTGGTAGATAGGCAGCAAAACTTCCGGTTCATCGAGGAAGTATCTTACCAGCACCTAGTCA
GAGAAGGCATGCAATTTATACATGTCGGCATGGCGATGGTCAGAATCCAGATGCTGCACAGGACAGATGC
AGGTATATCTGCACTAGTGGTGTTCAGAGACACCAGATGGAGTGATGACAGGCAGGTCATCGGAAGCATG
TCCGTAGACATGACCAGGGGCGCGCAGTTGGTATACATAATACCAAACGCCATGATGTCAATACACGATT
TCTACAATCGTATACAGGTCAGCGTGCAAACCCGAGGCTACGGAACAGGTTGGGAAGGTGGAGACAGTAA
CATGATCATCACGAGATCACTGGTTGGACGTCTCACCAACACCAGTGTGACCAACTTCGAATACCGGATA
GACCAGGTAACAGACTACCTAGCAAGCAACGGTGTGGCTTGCATCCCCGGTCAGAAGTGGAATGTGGCTA
ATAGATCTGGAGAATGGGAGTTACAACCCAGCAGAATCATAGCGCCACTAGTAGTCCCAACTGAAGCAAG

GCTTAGCCAAAAACAGAAATGGCAACATAAGCCTCAGATTCACCGACTTCCGAGATCAGAGGATCGTGGAA
GAAGGAGGACCATCTGAGCCAGAGGGAAGGCCAGAAGGAGAAGACGGTAGCACACACTATGTGCTTATGT
TCAACCACTCAAGGTGGGACACCCTCGGCCAACCAAGCGGAAAATACGACTACATGGTACGCTACGATGC
ACCAGAACCAACCGCATGGCCCACAACCAACATCGGGTGGGACGATGATGAGTCACCAAAACCACCAAGC
TCTGCAAAAGGATCCTTCGAGGTAAACCTCAGAGGAGAAAAGAAGCTGAAGGAGAAAGAACTCGCAGAGT
TCACGCCAGAGACAGATCTGGTGAGCCAGTGGTTAAGTCAGCTGTCAACATCCGCACACAATAGCGGAGC
CTCAAGTTCAGACGAAGAACCAAAGTTCGACGAGGCAGAGGACGAAGACGATGTGTACAACCAGCAAACC
TGGCAAAAGGAAGACAAGGAGAAAAGAGACCTGGAACTACAGGGGTGGAAACCCACTGGGAGACCAGGAA
TCTACGAGATGATCCCCGAAGAAGAAGAAGAGATCTACCTCAGGTACGAGGCAGAAGACGAGGAGGATCA
GGAGGTACAAGTAATAGGTGCCACCACCATTGAGGAGCCGGAAATGGAATACCCAACCAGGCTCGAGGAA
GTTATGGGCAAGCTCAAAAACGTGAGCATGGAAAAGCTGTTCCCAGTGAGCGGGATGGACAGCGAATCCA
GCAGCATTACCGGTGGAGGATTCATCCCACCAAGCCCAGTACCAGGAGCGCAAGAGTACCCACCAGCAAC
TGGAGCATCCGCGTCCACCATTGGACCAGCAGACATGCAAGGATGGGGAGGACGGCTACCTCGGAGCAGG
TCGCCTATAGGCTATGGCAGACCCCAACAACCGTGGTCACTGCCCTCAGCACAGTCTGATAACGGCTGCA
TGCTAGTCCTTCCACAGGACTTCACCCTAGTCCCCGACGTAATCAACAGATGGGAATCCATCACAGTCAA
CCTCATCAACAAGATGATGTTTGATTCCCTACAGGACAAGGCGGACTACGTAGAAAACCTCCTTGGAGAA
AGAGAAAAGGAGACATGGATGACATGGAGAATGCAGTACGAGGAAGAGTACAGGCAACTCCTCACCATGA
GCGGAGACGTAAGGAACCTTACTGCCGCAGTCAAAAGGGTCTTTGGAGTACACGACCCGCACACAGGATC
AGTACACATCCAGAATCAAGCGTACGCAGAGCTGGAGCGCCTCTACTGCAAAAGAACGGATGATGTGATC
CCCTTCCTCTACGACTACTACCAGTTAGCAGCCAAGTCAGGAAGGATGTGGCTCGGACCTGAGCTATCTG
AGAAGCTGTTCAGAAAGCTTCCACCGGAGATAGGCCCAACAATAGAGCAGGCCTATAAAGACAGGTATCC
AGGCCTCACGATTGGAGTTTTGGCAAGGGCCAATTTCATCCTGGAATATCTACAAAACGTCTGCAAGCAA
GCAGCGTTGCAAAGGTCCCTAAAAAGCCTGAGCTTCTGCAGAAACATGCCAGTACCAGGGTACTACGAGA
AGAAGCAATACGGCATCAGAAAGGCTAAAACCTATAAAGGAAAGCCTCACCCTACCCACGTGAAAGTCAT
CAAAAACAAGTACAAGCACACATCTGGGAAGAAGTGCAAATGCTACTTATGTGGGATAGAAGGCCATTAC
GCCAGGGAATGCCCAAAGAAAGTGGTGAAGCCACAAAGAGCGGCATACTTCAATGGCATGGGACTAGACG
ACAACTGGGATGTCGTGTCCGTCGAGCCCGGAGAATCAGATGACGATGAAATCTGTAGCATCTCCGAGGG
AGAAACGCTGGAGGAATGCATGAGCTTATGGCATTCAAGACTCAACTCCCATACCCAGTGGAGTACGAA
GCCAGCGCACCACAGTTCCTGATGCCACGGACACAGGTAACAGTGGAAAGAAGCGAAAAACCTTCCTGGA
GAAGAAGGAAGGAGATACCAAAGGCACAGCAAGAATGCACTCACGCCTGGAGTGATACACAAGAAGTGCC
GATCGAAGGGAGGATATGCAGCATATGCAGCGACGAAACGCCTCATGGGCGAAGAGTCACCTGCACCACC
TGCAACCTCAACCTCTGTCCCCTTTGCGCATACATGGATCATGGGATCAAGCTTATAGCCGCAAAGGACA
CCAAGGACGCAGCTAAGTGGCAATACCACAACAAAGATGAGCTTATACGACATCTCTATGAGCACAACGC
TTTCCTCACTAGAAAAGTCGAAGAGCTCACCAGTCAGCTGCAGGAGTTCCACAACCGCAAACCTGATGAC
CTGATCAGCTTAGCGGATGACTTGGAGGACGTGTCCATTCTGGACAACGCCTCAAAAAGGGGGAAGGAGA
AGGAATCTTTCCAATTCGGAACAACGATTCCCATCGACCACATCCAAAACTTGGAAAACGTGGCAAGGAT
CATCGAGCAATGGAAGGATACCCCCAAGGTGATTATCAAAGAAACAGCTGAAAGTAGCAACAACACTATT
GGAGCCCTCTTAGCAGAAGAAGGAATAGAGGAGCTAGCCGCAGCTGTGGACACGGCATATACCGAGATGC
CAAAAGGAGGACTAAACAAACTCTACAACACCATAGTCGAGTTTGTAATACCACAGGAGAAAGGGGCACC

CGTAAGATTCAGGGTAAGAGCTGTAATAGACACTGGATGCACCTGTACATGCATCAACAGCAAGAAAGTT
CCCAAGGAAGCCCTAGAAGAGGCGAAGTATCAGATGAACTTCGCGGGAGTAAACTCCACTGGAGAAACAA
AGCTGAAGATGAAAAACGGCAAGATGATCGTGTCAGGAAGCGACTTCTACACGCCATACATTGCAGCTTT
CCCAATGGAGCTACCAGACGTAGACATGCTCATCGGCTGCAACTTCTTGCGAGCCATGAAGGGAGGAGTC
AGGCTTGAAGGTACTGAAGTGACGATCTACAAGAAAGTCACCACAATCCAAACAACCCTGGAGCCCCAAA
AGATATCTCTGCTCCGCGCAGAAGCAGAAGTCGGAGAGGAGATCGAGCGTATGTACTACGCAAATGACTA
CTCTGAAGAAGGAGTCAGTCGCCTGAGAAACCACAAACTGCTGCAGGAACTAAAAGAACAAGGCTACATA
GGCGAAGAGCCAATGAAGCACTGGGCGAAAAACGGGATCAAGTGTAAGCTTGACATCAAGAACCCAGACA
TAGTAATCAGCAGTAAACCCCCGGATGCTGTCTCAAAGGAGACGAAGGCACAATACCAGCGGCACATTGA
CGCTCTCCTGAAGATCAAAGTGATCCAGCCAAGCAAGAGCAGGCACAGAACCGCAGCCTTCATCACAAAC
TCGGGCACAACCGTTGACCCGATCACAAAGAAAGAAATCCGAGGAAAAGAAAGGATGGTGTTCGACTACA
GAAGTCTGAACGACAACACCCACAAAGACCAGTATACTTTGCCTGGGATCAACACCATCATATCAGCAAT
CGGCAATGCGAAGATCTTCAGCAAATTTGATCTGAAGTCTGGATTCCACCAAGTATTGATGGACGAAGAA
TCCATCCCGTGGACCGCATTTGTCACACCAGTAGGGTTCTACGAGTGGAAGGTAATGCCTTTCGGACTCG
CAAACGCTCCGGCCGTCTTCCAGAGAAAGATGGACCAGTGTTTTGCAGGAACCTCAGAGTTCATAGCCGT
CTACATCGACGATATCCTGGTCTTCAGCAAGACCTTGAAGGAGCACGAAAAGCACCTGAGCATCATGCTT
GGGATATGTCGAGACAACGGCCTGGTTTTTGTCACCAAGCAAGATGAGGTTAGCAGCAACCGAGATCGACT
TCTTGGGAGCCAGCATTGGTGACGGAAAGATTAAACTCCAGCCTCACATAATCAAGAAGATAGCTGAGGT
GGACGATGAATCTCTGAAGACCCTCAAGGGGCTGAGAAGTTGGTTGGGAGTTCTCAACTATGCCAGGAAC
TACATCCCGAAGTGCGGAACACTCCTAGGCCCACTATACAGCAAGACCAGTGAGCATGGAGACAGAAGGT
GGCATGCTTCGGATTGGGCCTTAGTAAAGAAGATCAAGAGCCTGGTCCAAAATCTCCCAGGCCTCAAACT
GCCCAGTGAGGAGGCCTATATGATCATCGAGACAGATGGTTGTATGGAAGGATGGGGCGGAGTCTGTAAG
TGGAAGCCCAACAAAGCAGACTCAGCTGGCAAGGAAGAAATCTGCGCTTACGCAAGCGGTAAGTTCCCAA
CGGTGAAATCTACCATTGGCGCAGAAATCTTCGCTGTAATGGAGTCCTTAGAAAAATTTAAAATTTTCTA
CATGAACAAGGACGAGATCACCATCAGGACCGACTGCCACGCCATCATCACCTTCTATGAAAAGTTAAAC
GCCAAGAAACCTTCTCGGGTAAGGTGGTTAGCTTTTTGTGATTATATAACAAACTCAGGGGTGAAGATGA
AGTTCGAACACATCAAAGGCAAAGATAATCAGCTCGCTGACAATCTTAGTCGCTTTACCCAACTCATCAC
TGTAGTAAGATGGCTTCCCAAGGAACTAGCGGAGCTCACGGCCGAACTGGTCAAAGGAAGGGACGAAGCC
CTGGTGAACAAGGAGGTACAGAGGAACATCTCATGTTTTCTCGAGACTGCCCTCCTCCAAGCGGAGAAAT
CCGTGACTACTCGCCCATCAGAGCCGCACCATGTACTATGGCGGAGATGGACGAATCCCGAAGAGTGGCC
ATGCAGCGAAGAATCAAGGTCTTCGACGATCTTGCACAAAACATCAGCGACGCCGTATACATCACAGGCA
TCGACCTCGCCGCCGCCAAAGCACGGGCAACCAGGGATAACTGGTACAATGACGTCACCCCGGCATTGGA
AGAACGAGCAGCTGCAGCATGGAGACTCATGGCAGCCTACTCAGACTTCGCCACGTGGAAGGACGTGAAC
GTCTAGTGAAGTGACGCAAGGAATGACTTCACAATTGCCAATGTCGTCACTGCTTACGACTTGGAACTTA
TCCTTTAGTGTCGGCAGCATCTCTTAGCTGTCATAAGTGTGTAAGTGCGCCAGTAGTGCGCTGTGTCAAG
ATAAGGAATCTTATCTCCTTATCTTCTTTCCCTTTGTTTAAAGGTAAAGCTGTAAAGCAGGACTAATTAG
CTGCAGGTCATCAGGTTTGCGGTTGTGGAACTCCTGCAGCTGACTGGTGAGCTCTTCGACTTTTCTAGTG
AGGAAAGCGTTGTGCTCATAGAGATGTCGTATAAGCTCATCTTTGTTGTGGTATTGCCACTTAGCTGCGT
CCTTGGTGTCCTTTGCGGCTATAAGCTTGATCCCATGATCCATGTATGCGCAAAGGGGACAGAGGTTGAG

GTTGCAGGTGGTGCAGGTGACTCTTCGCCCATGAGGCGTTTCGTCGCTGCATATGCTGCATATCCTCCCT
TCGATCGGCACTTCTTGTGTATCACTCCAGGCGTGAGTGCATTCTTGCTGTGCCTTTGGTATCTCCTTCC
TTCTTCTCCAGGAAGGTTTTTC

Appendix B - Phylogenetic Analysis
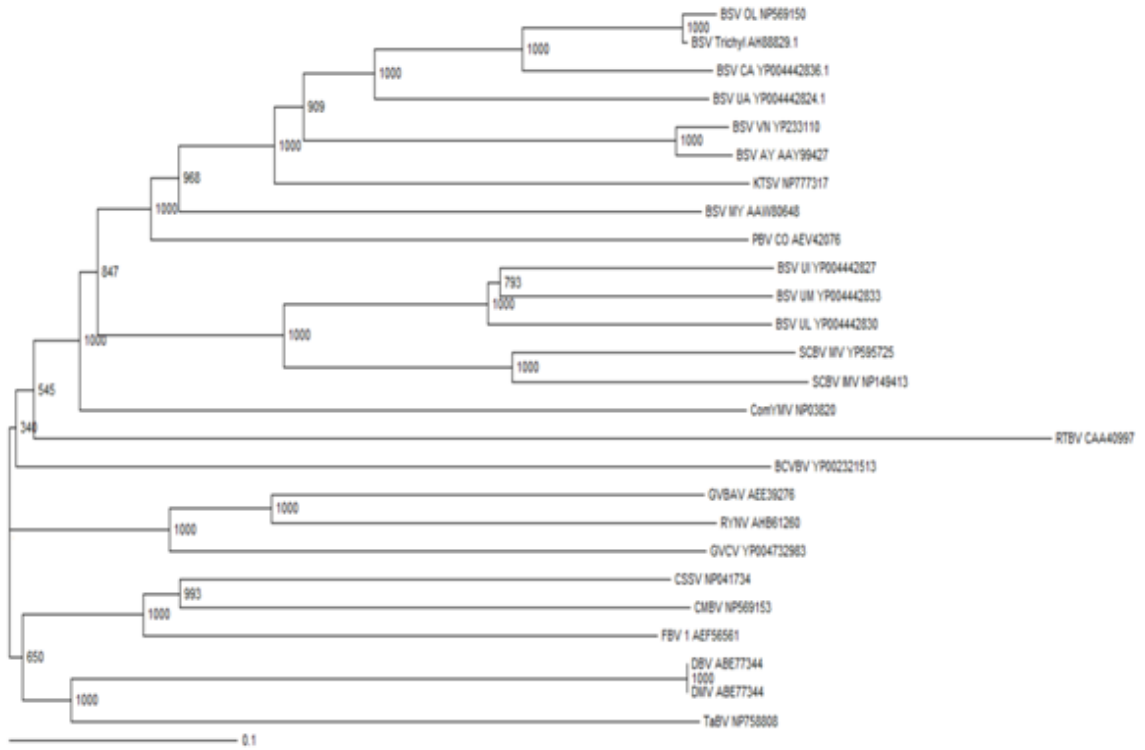
Updated phylogenetic tree: Neighbor-joining dendrogram of sequence relationships determined using the amino acid sequence alignment of the reverse transcriptase among species within genus *Bandavirus*.