Project Description

A. Results from Prior NSF Support. N/A

B. Specific Aims

Our main goal is to understand the role of an *Arabidopsis* three-gene family (RDR3, RDR4, RDR5) thought to be involved in RNA interference (RNAi). This is a novel gene-family that has not been cited in the literature. These genes encode putative RNA-dependent RNA polymerases (referred to as RdRP in general or RDR for *Arabidopsis* specific genes). We will study these genes by using null-mutant T-DNA insertion lines; data mining of microarray experiments; and protein expression studies. We have bioinformatics data suggesting that at least one of the genes is involved in genome regulation in the seed. We have preliminary data showing these genes are involved in a branch of the RNAi pathway important for virus defense. By using the genetic model *Arabidopsis thaliana* we will be able to define functional roles for RDR3, RDR4, and RDR5.

This project has three specific aims focused on determining the diverse roles of RDR3, RDR4, and RDR5 in *Arabidopsis thaliana.*

1. ***To determine the role of RDR3, RDR4, and RDR5 in the virus defense response using a broad spectrum of plant viruses.*** *RDR3*, *RDR4*, and *RDR5* compose a unique three-gene family that has not been studied, yet are believed to be important in fighting virus infections. We have preliminary data suggesting that one of the family members, RDR4, is important in the host defense response to Tobacco Rattle Virus. We will characterize the roles of these three *RDR* genes in the defense response by analyzing differences in susceptibility to a virus infection. These studies will include single and multi-gene *RDR* mutant lines. In addition, transgenic plants overexpressing an *RDR* gene will be generated to analyze for increased resistance to a virus infection.

2. ***To define gene expression profiles for RDR3, RDR4, and RDR5.*** We will use qRT-PCR to determine the specific expression profiles of *RDR3*, *RDR4*, and *RDR5* in multiple plant tissues and during virus infection. We will determine which *RDR* genes are up-regulated in response to a virus infection. We will also analyze *RDR* gene expression in *rdr* mutant lines to determine RDR functional redundancies. In addition, we will continue to mine the publicly available microarray data to further define the diverse roles of RDRs in the plant.

3. ***To define the mechanisms of RDR3, RDR4, and RDR5 activity.*** *RDR3*, *RDR4*, and *RDR5* mRNA sequences will be cloned into Gateway-compatible vectors. These expression vectors will enable us to determine the mechanism of action of the protein products. The site of action will be determined by cellular localization studies using GFP-tagged proteins. HA-tagged proteins will be generated for immunoprecipitation studies to look at protein:protein and protein:nucleic acid interactions.

The research defined in this proposal will contribute significantly to the fields of RNA silencing and virus defense. We will define the functions of three novel genes in *Arabidopsis*. Furthermore, the proposed research will train undergraduate and Masters students in the techniques of molecular biology and bioinformatics. This work will be done at an urban, primarily undergraduate institution with a large minority (21%) and women (54%) student population. Minority and female enrollment in the Biology department is greater than on the campus in general, 32% and 70% respectively.

I have mentored eight students in my research lab in the two years I have been at UCD, six of them women. I have over 10 years experience in the molecular characterization of virus infections, in both human and plant models. As a post-doc I worked in one of the top labs studying RNA silencing and virus-host interactions (Jim Carrington, Oregon State University). I have an ongoing collaboration with the Carrington lab (see letter of support in Supplemental Documents). The three genes we are focusing on have not been described in the literature, nor have they been presented at the most recent plant-RNA biology meeting (May 2006).

C. Background

RNA interference (RNAi) is a pathway involving the generation of small RNA molecules, 21-24 nt in length. Small RNAs have numerous roles in gene regulation at both the RNA and DNA levels, and are involved in virus defense in plants. A simplified model for RNAi is shown in Figure 1. The crux of the RNAi pathway is the effects of these small RNA molecules on genome regulation (Figure 2). In *Arabidopsis*, small RNA molecules can affect gene expression by directing the cleavage of cognate RNA molecules, by directing methylation of DNA or histones, or, less frequently, by translational repression (Figure 2) [4, 9, 10, 20, 23, 24, 27, 30, 35, 41, 46, 50, 64, 69, 73]. These small RNAs are produced by the processing of double-stranded (ds) RNA by one of the four Dicer-like (DCL) enzymes [18, 25, 44, 45]. In plants, RNAi also functions as a viral defense mechanism [38-40, 48]. Small RNAs generated in response to a virus infection result in cleavage of the virus genome, thus diminishing virus replication. In addition, a systemic signal is generated that can move through the plant vasculature and prevent subsequent virus infections.

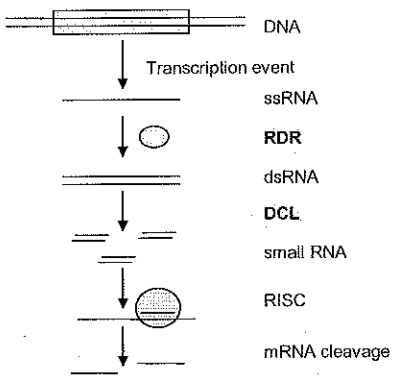

Figure 1. A simplified schematic of RNAi illustrating the general components of the pathway. Double stranded (ds) RNA is essential for inducing the RNAi pathway. There are many modes of dsRNA generation; a host RDR can copy the single-stranded (ss) RNA to a dsRNA molecule. A specific DCL recognizes and cleaves the dsRNA into small RNA molecules. These small RNA molecules can then associate with an RNA-induced silencing complex (RISC) and guide RISC to a cognate RNA molecule, such as an mRNA. RISC then cleaves the mRNA molecule, rendering it nonfunctional for translation. For additional functions of small RNAs see Figure 2.

1. RNA-dependent RNA Polymerase. RdRPs are found in plants, fungi (*N. crassa*), and worms (*C. elegans*), but not in insects (*Drosophila*) or mammalian cells [5, 9, 13, 52, 54]. Consequently, studying RdRP functions in *Arabidopsis* provides a unique opportunity to

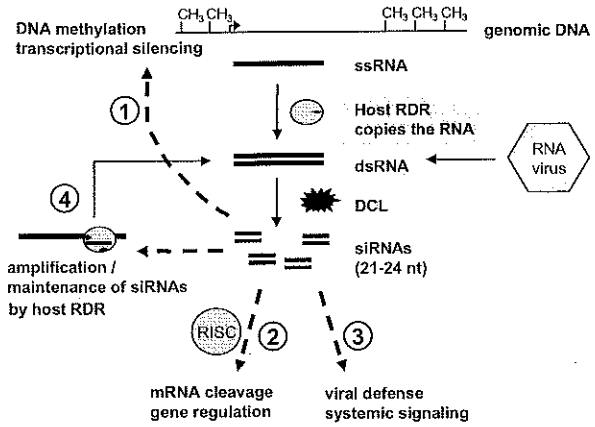

Figure 2. Schematic of some of the roles of siRNAs in a plant cell (dashed lines). siRNAs are generated by DCL cleavage of dsRNA into 21-24 nt RNA molecules. (1) One class of siRNAs are thought to direct DNA methylases in the nucleus, leading to DNA or histone methylation and transcriptional gene silencing. (2) Another class of siRNAs direct RISC to a cognate RNA (mRNA or viral RNA) for cleavage (see Figure 1). (3) Virus-specific siRNAs also provide a defense mechanism for the plant against invading viruses through a local and systemic signal. The systemic signal, most likely nucleic acid, can move through the plant and prevent future infections of homologous viruses by initiating RNAi in distant tissue. (4) A host RDR may amplify the production of siRNAs by using an RNA template and generating more dsRNA. A host RDR may also be required for generation of the systemic signal in virus defense.

study a set of genes important in a plant defense mechanism, which may prove to be important for agricultural biotechnology. The *Arabidopsis* genome encodes for seven potential RdRPs; six are thought to be functional RdRPs and one (*RDR7*) appears to be a truncated protein missing the catalytic domain. Of the six functional RDRs, three (*RDR3*, *RDR4*, and *RDR5*) appear to comprise a gene family, probably as a result of two gene duplication events. As yet, there are no published data on RDR3, RDR4, or RDR5. Diverse roles in virus defense, gene expression, and genome maintenance have been shown previously for RDR1, RDR2, and RDR6 (see below).

2. Known functions of plant RDRs. RdRPs catalyze the synthesis of complementary RNA from single-stranded RNA or DNA molecules [54, 57]. In plants, production of dsRNA triggers the RNAi pathway (Figure 1 and 2). An intense area of research is to resolve how the dsRNA molecules are generated. (We will address this area in Specific Aim 3.) Small RNA Northern blot analysis has shown that RDR2 is required for the generation of the majority of endogenous short interfering RNAs (siRNAs), including those important for heterchromatin formation [4, 10, 21, 69, 73]. RDR6 is important for the production of trans-acting siRNAs, another class of small RNAs [2, 46, 64]. Recently, RDR2 and RDR6 have also been identified as being important for production of some of the antisense RNA products from microRNA (miRNA) targets [51]. miRNAs are another class of small RNAs which also cause gene silencing and are important in organismal development [9].

RNAi is one method of protecting the plant from physical destruction due to a viral infection. Production of small RNAs through the RNAi pathway can halt virus replication by targeting the virus genome for destruction. RDR1 and RDR6 have been shown to be important for the virus defense response to some, but not all viruses [14, 15, 42, 61, 73, 75]. Thus, the roles of individual RDRs in the plant are diverse.

3. Virus-specific siRNAs. The classes of small RNAs we are interested in are those produced during a virus infection. Many plant viruses have RNA genomes and encode their own RdRP. Therefore, a replicating RNA virus generates dsRNA during viral genome replication. Double-stranded viral RNA is a trigger of the RNAi pathway (Figure 2). Cleavage of the double-stranded viral RNA by a DCL results in virus-specific siRNAs that can target the viral genome, through RISC, for destruction. RDRs are thought to play a role in amplification of the

viral targeting signal. The RDRs generate virus-specific dsRNA which can then be processed into siRNAs. These siRNAs direct RISC, resulting in destruction of additional viral genomes. Besides the local defense response, a systemic signal is generated in the plant. This signal moves through the plant vasculature and protects newly emerging tissue from a future infection by the same virus. It has been hypothesized that this systemic signal is a nucleic acid, due to its specificity, but whether it is a small RNA molecule or a larger nucleic acid remains to be determined. RDRs are thought to play a role in the production of this systemic signal by again generating dsRNA which can then be processed into smaller RNAs [23].

For a productive viral infection, many plant viruses have evolved counter-defensive strategies to avoid or block the RNAi defense pathway [3, 6, 7, 11, 12, 22, 28, 29, 31-34, 36, 37, 58, 62, 65-68]. These RNAi suppressors appear to work at many different levels in the general RNAi pathway (Figure 3). The Tombusviral suppressor, p19, interferes with RNAi by binding 21 nt siRNAs, preventing these siRNAs from entering the RISC and guiding RNA degradation [11, 58, 63, 74]. The 2b protein of CMV blocks the systemic signal of RNAi, thereby allowing for subsequent virus infections in newly emerging tissue [7, 36]. These are just two examples of mechanisms of silencing suppression by RNA viruses. Therefore, it is crucial to investigate the infectivity of multiple virus families, encoding distinct viral suppressors, in order to understand the interplay between host and pathogen. This proposal takes the approach of studying representatives from a broad range of plant virus families with the goal of elucidating the functions of the specific *RDR* genes in the RNAi virus defense pathway.

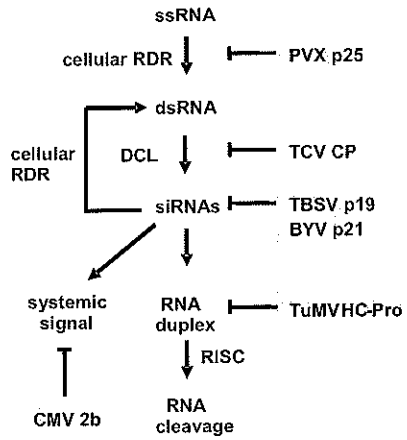

Figure 3. Proposed sites of action of virus suppressors of RNAi. Each virus family appears to have independently evolved an RNAi suppressor, as evidenced by the unique sites of action of the suppressors in wild-type plants.

D. Preliminary Data to support the hypothesis that RDR3, RDR4, and RDR5 have unique and diverse functions in RNAi and the virus defense response.

1. Role of RDRs in Arabidopsis development. To determine if RDR3, RDR4, or RDR5 affect *Arabidopsis* development, developmental timing was monitored (Figure 4). It has been shown that *rdr6* mutants are accelerated in the transition from juvenile to adult vegetative growth [46, 71]. This transition phenotype is characterized by elongated, downward curled rosette leaves (Figure 5), as well as early production of abaxial trichomes. Quantification of developmental stages in the lab for the six functional *RDR* mutant lines has produced comparable phenotypes for *rdr6*. *rdr3, rdr4,* and *rdr5* mutants did not show any quantitative differences in rosette leaf morphology (Figure 4) or growth stage progression (Figure 3) as compared to wild-type Columbia.

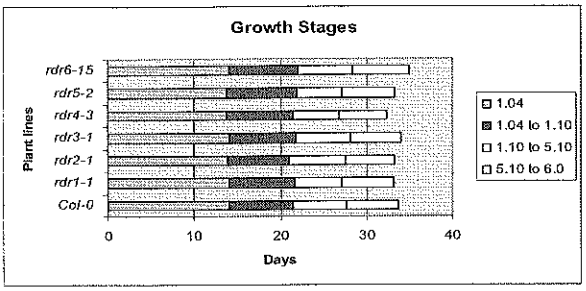

Figure 4. Growth stage progression for mutant plant lines and wild-type (Col-0). Growth stages comprise the appearance of four rosette leaves (1.04), ten rosette leaves (1.10), reproductive bud (5.10), and first flower (6.0). Growth stage progression for mutant lines did not differ from that of wild-type. (n=80-100)

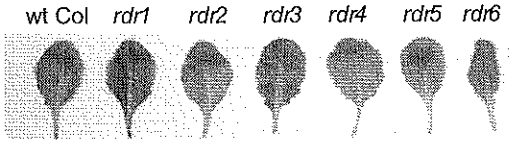

Figure 5. Photographs of rosette leaf 5 from wild-type and mutant lines. Note the downward curling of the *rdr6* leaf. No quantitative differences in leaf morphology for *rdr3*, *rdr4*, or *rdr5* mutant lines were found as compared to wild-type Columbia.

2. Molecular characterization of small RNA production in the various* RDR *mutant lines. Different classes of small RNAs have different functions in the cell. The different small RNA classes are derived through unique pathways involving a specific DCL and sometimes a specific RDR. Northern blot analysis characterizing small RNA accumulation was done to determine if RDR3, RDR4, or RDR5 function in the production of several defined classes of endogenous small RNAs (Figure 6). Our results show no difference in endogenous small RNA production in the *rdr3*, *rdr4*, or *rdr5* mutant lines as compared to wild-type. In contrast, *rdr2* shows a loss of production of most endogenous siRNAs, while *rdr6* shows a loss of trans-acting (ta)-siRNA production, as has been reported previously [2, 46, 64, 73].

3. The rdr4 mutant line is more susceptible to TRV. The viruses listed in Table 1 have been propagated and tested for virus infectivity in *Arabidopsis* Columbia ecotype. Initial susceptibility to virus infection experiments using Tobacco Rattle Virus (TRV) and the six functional *RDR* mutant plant lines showed an increased susceptibility to TRV in the *rdr4* mutant line (Figure 7). Northern blot analysis using a probe specific for the TRV coat protein showed an increased titer of TRV in *rdr4* at **14 days post-infection**, as compared to the wild-type Columbia, *rdr3*, or *rdr5* (Figure 7). These are exciting results, as they indicate a function in the virus defense pathway for RDR4 which is unique from RDR3 and RDR5, (Figure 7), as well as RDR1, RDR2, and RDR6 (data not shown). *RDR3*, *RDR4*, and *RDR5* comprise a three gene family whose DNA sequence and protein products are 79% homologous. We hypothesize that the gene duplication events resulted in genes that have evolved unique functions. This data begins to support this hypothesis.

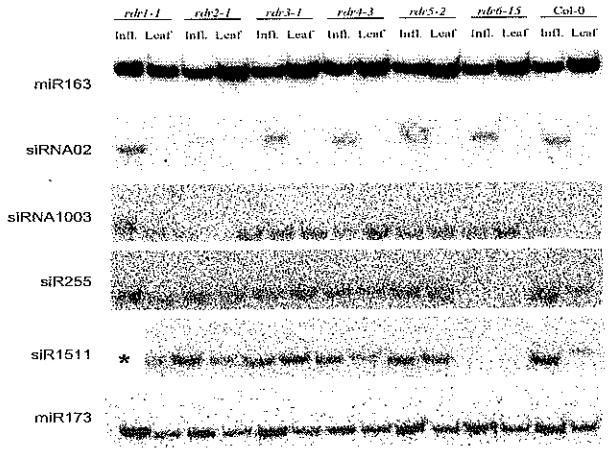

Figure 6. Northern blot analysis of small (21-24 nt) RNAs. miR163 is a miRNA. siRNA02 and siRNA1003 are siRNAs derived from an inverted duplication and 5S rRNA loci, respectively. siR255 and siR1511 are ta-siRNAs. miR173 is a ta-siRNA associated miRNA. (* membrane inadvertently cut, losing *rdr1-1* inflorescence sample)

virus	family	suppressor
TRV **Tobacco rattle virus**	Tobravirus	p16K
PVX **Potato virus X**	Potexvirus	p25
TMV **Tobacco mosaic virus**	Tobamovirus	p30
TCV **Turnip crinkle virus**	Carmovirus	CP
CMV **Cucumber mosaic virus**	Cucumovirus	2b
TEV **Tobacco etch virus**	Potyvirus	HcPro
PVY **Potato virus Y**	Potyvirus	HcPro
TuMV **Turnip mosaic virus**	Potyvirus	HcPro

Table 1. Viruses to be used for studying the role of RDRs in the viral defense response. The multiple virus families were chosen because of the diverse silencing suppressors. TEV, PVY, and TuMV all encode a similar viral RNAi suppressor protein, HCPro, but each virus has varying degrees of infectivity in Arabidopsis ecotype Columbia.

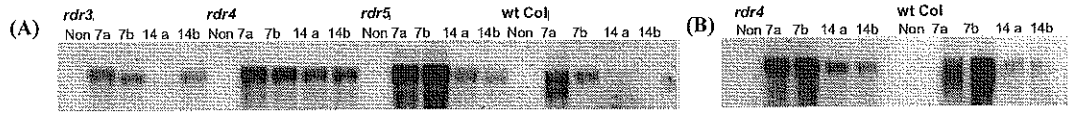

Figure 7. Northern blot analysis from TRV infected tissue. (A) Mutant lines *rdr3*, *rdr4* and *rdr5*, and wild-type Columbia were infected with Tobacco Rattle Virus (TRV) on day 0. Aerial (shoot and inflorescence) tissue was harvested at 7 and 14 days post infection. Four plants were pooled for each replicate (a and b). Non-infected tissue (non) was used as a control. A probe specific to TRV coat protein was used to detect the presence of TRV in the systemic tissue. **(B)** Repeat of *rdr4* susceptibility experiment, again showing an increase in virus susceptibility at d14 in the *rdr4* mutant as compared to wild-type.

4. RDR5 is highly expressed in seed. *Arabidopsis* expression data from 1,436 ATH1 GeneChip microarrays, which have been normalized to enable reliable cross-experiment comparisons, is freely available for download from TAIR. These data have been analyzed for expression profiles of *RDR1*, *RDR2*, *RDR5*, and *RDR6* genes by our bioinformatics student. Preliminary analysis of these data has shown that in multiple experiments *RDR5* is highly expressed in seeds. A subset of 80 samples of these data specifically measures normal *Arabidopsis* development across time and different tissues (AtGenExpress ExpressionSet:1006710873) [55, 59]. Additional work has used this subset to look for correlative expression patterns between *RDRs* and *DCLs* across development. *RDR5* and *DCL2* are coexpressed at many stages in development (Figure 8). *RDR5*, *DCL2* and to a lesser extent, *DCL1*, show increased expression in seed, whereas all other *RDRs* and *DCLs* show decreased expression (Figure 8). DCL1 has been shown to be essential for miRNA production. RDR2 and RDR6 have been identified as being important for production of **some** of the antisense RNA products from miRNA targets, but many of these antisense transcripts are RDR2- and RDR6-independent [51], and thus instead may be synthesized by RDR1, RDR3, RDR4, or RDR5. DCL2 has a role in the virus defense response [73]. Based on these facts, and through our analysis of publicly available microarray data, we have been able to form the following hypotheses: (1) RDR5 functions in the seed with DCL2 or DCL1 and (2) RDR5 and DCL2 form

a functional pair for small RNA generation throughout much of the plants development. These hypotheses are being tested in the lab.

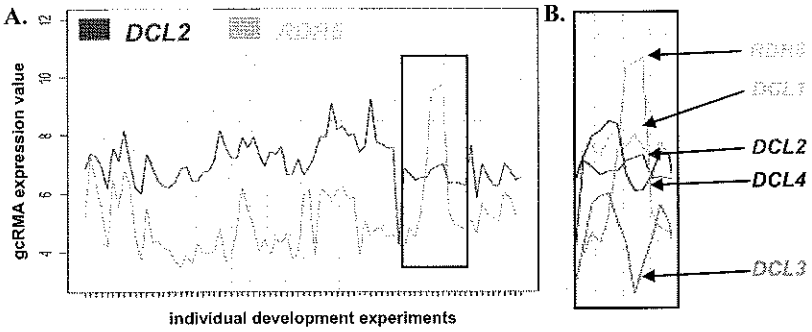

Figure 8. Microarray expression data across *Arabidopsis* development. **A.** Expression data is shown for *RDR5* and *DCL2*. *RDR5* shows a high level of expression in seed. **B.** Enlargement of the boxed region and addition of the DCL1, DCL3, and DCL4 expression profiles. *DCL1* and *DCL2* show a slight increase in expression level in seed, whereas *DCL4* and *DCL3* decrease expression levels in seed. The working hypothesis is that RDR5 and DCL2 and/or DCL1 work in a common gene regulation pathway in seeds.

Since we have not been able to detect an effect on development or endogenous small RNA production in leaf or inflorescence tissue, we hypothesize that RDR3, RDR4, and/or RDR5 (1) have a role in the plant's response to virus infections and (2) have unique functional tissue distribution.

E. Experimental Methods and Analysis.

Specific Aim 1. *To determine the role of* RDR3, RDR4, *and* RDR5 *in virus defense response using a broad spectrum of plant viruses.*

1.a. Background. A significant problem in agriculture is the destruction of crops due to pests. One type of agricultural pest is viruses. *Arabidopsis*, being a model for crop plant stress response, makes it an ideal candidate for studying the virus defense response [49, 59]. The *Arabidopsis* ecotype Columbia is susceptible to a broad range of viruses [70].

The branch of the RNAi pathway we are studying here is specific for targeting the genome of an invading virus. Much like our own immune response, this RNAi pathway is a specific response. RNAi targets RNA molecules, viral or endogenous, with sequence homology to the inducer (dsRNA). Therefore, each infecting virus has the potential to elicit a unique RNAi defense response in the plant. This also means that an RNAi response to one virus will not protect the plant from infection by a non-homologous virus. This is partly due to the fact that many plant viruses have evolved independent suppressors of RNAi. First identified as pathogenicity determinants, these RNAi suppressor genes have been shown to block unique sites in the RNAi pathway (Figure 3) [65]. Thus, virus resistance and susceptibility studies are very complex. A single virus cannot be studied with the hopes of understanding viral pathogenicity as an entire process.

Our goal is to determine if expression of certain genes will render *Arabidopsis* more resistant to virus infections in general. We will explore this hypothesis by first looking at the role of three *RDR* genes (*RDR3*, *RDR4*, and *RDR5*) in susceptibility to virus infection. Our research will analyze the host defense response to numerous virus families (Table 1). The viruses listed in Table 1 have been propagated in the lab and viral stocks are available for the experiments listed below. The purchase of *Arabidopsis* growth chambers, requested in this proposal, will not only provide the lab with a more controlled growth environment, but also allow us to work with more diverse virus families, such as the DNA virus family Geminivirus, and agriculturally important viruses, like Citrus Tristeza virus, both of which have restrictive USDA regulations.

1.b. Rationale. There are data showing involvement of RDR1 and RDR6 in the virus defense response [14, 15, 42, 73, 75]. In summary, *rdr1* mutant lines are more susceptible to TMV and TRV. *rdr6* mutant lines show an increased susceptibility to CMV, PVX, and TRV but not to TuMV, TCV, TMV, or TVCV (Turnip vein clearing virus). Therefore, the response to individual viruses involves different *Arabidopsis* RDRs. We hypothesize that RDR3, RDR4, and RDR5 are also involved in this defense response pathway and each can be independently engaged depending upon the infecting virus. We also hypothesize that the numerous RDRs have compensatory roles in the virus defense response. Therefore, multi-gene mutants will be generated to analyze for genetic redundancy. The ultimate goal is to find a gene, or more likely genes, that comprise a general virus resistance pathway.

1.c. Susceptibility to virus infection studies. Our hypothesis is that RDR3, RDR4, and/or RDR5 are each important for virus defense depending on the invading virus. *rdr3*, *rdr4*, and *rdr5* homozygous mutant plant lines and wild-type Columbia will be infected with a single virus (Table 1). *rdr1*, *rdr2*, and *rdr6* mutants will also be included in these studies as the role of these genes in susceptibility to virus infection is not fully understood. Aerial tissue, the shoot and inflorescence, will be collected at 7, 14, and 21 days post-infection. A minimum of four plants will be pooled into one sample, and duplicates will be collected at each time point. A non-infected, pooled control will also be sampled. In addition, measurements of shoot height, shoot phenotype (such as curling or yellowing), and flower phenotype will be documented through quantitative data and photo documentation. The aerial tissue will be processed for total RNA, which will be subjected to Northern blot analysis. Virus-specific probes will be used to determine the level of virus infectivity in each sample. The production of small RNAs resulting from the RNAi response to the virus infection will also be detected by Northern blot. We will systematically do single-mutant infectivity studies with all of the viruses listed in Table 1.

Preliminary data obtained in our lab suggests a role for RDR4 in TRV susceptibility (Figure 7). TRV does not encode a strong silencing suppressor. In fact, TRV is used as an inducer of silencing and the virus is silenced rapidly in wild-type tissue [8, 15]. Molecular analysis of the *rdr4* line shows an increased titer of TRV at 14 days post-infection as compared to wild-type Columbia or the *rdr3* or *rdr5* mutant lines.

1.d.. Multi-gene mutant studies. Figure 7A also shows an increase in initial virus titer at d7 in the *rdr5* mutant line, followed by a decrease at d14. Similar to what is seen in *dcl* mutant lines [16, 19, 26], the RDRs may have unique functions, but can also have compensatory functions

when there is a loss of an RDR. Therefore, multi-gene knockout plants will be used to further determine RDR redundancies. An *rdr4 rdr5* double mutant may be highly susceptible to TRV.

To determine genetic redundancies, double and triple knockout mutants will be analyzed. The following double and triple mutants are currently being generated by our lab or have been obtained in collaboration with Jim Carrington's laboratory at Oregon State University.

rdr1 x rdr2	*rdr3 x rdr4*	*rdr2 x dcl3*	*rdr1 x rdr2 x rdr6*
rdr1 x rdr6	*rdr4 x rdr5*	*rdr6 x dcl4*	*rdr3 x rdr4 x rdr5*
rdr2 x rdr6	*rdr3 x rdr5*		

Results from the single mutant susceptibility studies will guide the multi-gene knockout studies. Additional multi-gene knockout lines may be required, such as an *rdr1*x *rdr4xrdr6* triple mutant line to study susceptibility to TRV.

1.e.. The requirement to study multiple virus families. When studying susceptibility to Tobacco Mosaic Virus (TMV) we found that there was no difference in susceptibility to virus infection between the *rdr* mutant lines and wild-type Columbia (Figure 9). TMV was equally infectious in all mutant lines. Therefore, it will be very important to study susceptibility to multiple viruses. The viruses we have chosen to initially analyze (Table 1) are representative of multiple virus families, each expressing a different type of silencing suppressor. Three Potyviruses have been chosen due to the varying degrees of infectivity in the Columbia ecotype, from severe (TuMV) to negligible (TEV).

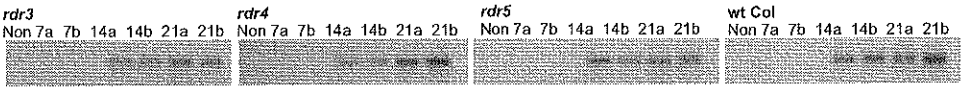

Figure 9. Northern blot analysis from TMV infected tissue. Mutant lines *rdr3*, *rdr4* and *rdr5*, and wild-type Columbia were infected with Tobacco Mosaic Virus (TMV) on day 0. Aerial (shoot and inflorescence) tissue was harvested at 7, 14, and 21 days post infection. Four plants were pooled for each replicate (a and b). Non-infected tissue (non) was used as a control. A probe specific to TMV coat protein was used to detect the presence of TMV in the aerial tissue.

In addition to the viruses listed in Table 1, we would like to begin work with the Geminivirus family. This is a DNA virus that does not encode an RdRP. Geminiviruses do not naturally produce long dsRNA. There is a region of overlap between the leftward and rightward transcripts, but this is small region [56]. Therefore, the requirement for a host RdRP to amplify the RNAi defense pathway may be stronger for a DNA geminivirus than for an RNA virus that encodes its own RdRP. Work on the geminivirus Cabbage leaf curl begomovirus in *Arabidopsis* has suggested a role for RDR6 in VIGS (virus induced gene silencing) but not necessarily in naturally controlling the virus infection [43]. Therefore, RDR3, RDR4, or RDR5 may be important in propagating a silencing signal against members of this virus family. Our long-term goal is to dissect susceptibility to virus infection and endogenous resistance pathways to more agriculturally important viruses, such as Citrus Tristeza virus.

1.f. Generation of transgenic plants overexpressing RDR(s). The experiments described above will lead to information regarding *RDR* genes important in virus defense. Transgenic plants will be made that overexpress each *RDR* gene, to determine if overexpression will result in

increased virus resistance. For instance, since RDR4 appears to be important for TRV defense; an *RDR4*-transgenic line, in the wild-type Columbia background, will be made to assess if overexpression of RDR4 results in greater resistance to TRV, as compared to wild-type plants. Virus infectivity studies will be done as described in Specific Aim 1.c. Though a complication to the generation of transgenic *Arabidopsis* is the silencing of the transgene, identification of a stably expressing transgenic line is feasible. I created several such transgenic lines as a post-doc.

Specific Aim 2. ***To define gene expression profiles for*** **RDR3, RDR4,** ***and*** **RDR5.**

2.a. Background and Rationale. Most experiments characterizing RNAi pathways have used inflorescence tissue; a highly replicating tissue, allowing for isolation of large quantities of RNA. The second most commonly studied tissue is leaf. Differences in small RNA expression levels can be seen between leaf and inflorescence (Figure 6). Given that RNAi is a gene regulation pathway, it is not surprising that there are tissue-specific expression profiles for both the RNAi effector molecules, small RNAs, and the RNAi components, such as RDRs and DCLs. Therefore, a methodical study of tissue expression profiles for *RDR3*, *RDR4*, and *RDR5* is crucial to understanding the roles of these genes in the plant.

2.b. Data mining public databases to determine RDR5 gene function. *RDR3* and *RDR4* are not represented on the ATH1 GeneChip. *RDR3*, *RDR4*, and *RDR5* are represented on the Affymetrix 8K Chip, but there are only a few experiments using the 8K Affymetrix chip and not all of the 8K chip data are in a format (CEL or RMA) that can be used in comparative analysis. Therefore, the majority of our bioinformatics work has been focused on *RDR5*.

Arabidopsis expression data from 80 biologically different samples (AtGenExpress ExpressionSet:1006710873) [55, 59] was analyzed for gene expression of *RDR1*, *RDR2*, *RDR5*, and *RDR6* genes by our bioinformatics student. It was shown that *RDR5* is highly expressed in seed (Figure 8). Expression patterns showed a correlation between *RDR5* and *DCL2* in several tissues (Figure 8). These two findings would not have been possible without the use of microarray data and a skilled bioinformatisist.

For her Masters project, our bioinformatics student is pursuing the hypothesis that *RDR5* and *DCL2* function together in an RNAi gene regulation pathway. She has applied computational methods to the publicly available microarray and small RNA data (http://asrp.cgrb.oregonstate.edu/) to identify differentially expressed gene clusters, predict RNA secondary structures and identify possible binding sites of small RNAs to their potential mRNA targets. From these results she has predicted siRNAs and their target genes which may be regulated via an *RDR5/DCL2* silencing pathway. She is testing her predictions in the lab this summer.

Another interesting finding is that while it has been generally shown that silencing is turned off in pollen [47], *RDR7* shows a dramatic increase in gene expression (Figure 10). In RDR7 appears to be a truncated RDR with two cDNA clones identified but no ESTs reported to date [59]. Interestingly, several of the Argonaute genes, which also form a multi-gene family involved in RNAi, show a dramatic increase in gene expression in pollen (data not shown).

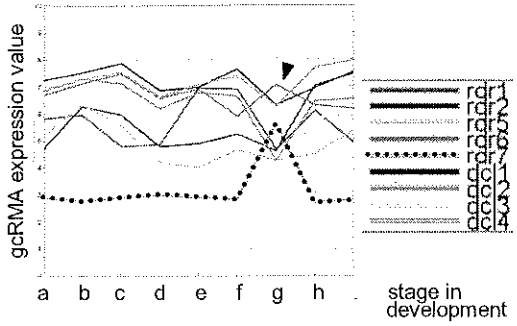

Figure 10. *RDR* and *DCL* gene expression values from wild-type tissue development data. Development stage represented (a) cotyledons, (b) roots, (c) seedling (green parts) (d) leaves 1,2, (e) flowering rosette, (f) flowers stage 15, (g) mature pollen, and (h) siliques. Both RDR7 and RDR6 (▼) show an increase in expression levels in pollen, though the RDR6 level is not as dramatic. The *RDR7* line was retraced (·····) for clarity in the figure (due to lack of color figure output). This high level of expression for *RDR7* has been seen in other pollen microarray data sets.

These data will be validated using Northern blot analysis and may result in a unique role for RDR7 in development.

It is essential for the lab to be able to continue these types of data mining projects. Therefore, funds have been requested to hire a bioinformatics consultant to perform in depth and more complex data analysis than is available from basic gene analysis software. Ideally, this will be the bioinformatics student who is currently working in the lab and will be completing her Thesis this Fall. The data mining projects are relatively inexpensive and can generate hypotheses that would not be formulated without analyzing global gene expression. The hypotheses generated from the bioinformatics work can then be tested in the lab. This will result in projects suitable for undergraduate independent study. Therefore, students will receive training in plant molecular biology, as well as learning some basic bioinformatics. One of the educational objectives of this grant proposal is to train both bioinformatics students and biology students in communication between the two disciplines. By bringing these two disciplines together, a more complete picture of the roles of RDRs in *Arabidopsis* will be formed.

2.c. Using qRT-PCR to determine RDR3 and RDR4 gene expression. *RDR3* and *RDR4* are not represented on the ATH1 GeneChip and there is a limited number of microarray data available from the Affymetrix 8K chip. Though the 8K Chip contains *RDR3* and *RDR4* probes, it does not contain *RDR6*, *DCL2*, *DCL3*, or *DCL4* probes. In addition, the 8K Chip is now considered a custom-array and must be purchased in bulk (40 chips at a time). This would not be a cost-efficient way to determine expression profiles for two genes. Therefore, tissue specificity experiments will be done using real-time RT-PCR (qRT-PCR) to further characterize the expression profiles for *RDR3* and *RDR4*. As seen with the tissue expression findings for *RDR5* and *RDR7*, *RDR3* and *RDR4* may have unique tissue distributions as well. Therefore, total RNA from multiple wild-type tissues will be isolated; including root, seedling, multiple rosette stages, inflorescence, pollen, and seeds. qRT-PCR using *RDR3* and *RDR4* gene specific primers will be done to determine tissue distribution of these two genes. qRT-PCR experiments will be conducted in association with the UCHSC Microarray Core Facility http://microarray.uchsc.edu/.

2.d. Induction of RDR gene expression due to a virus infection. It has been shown previously that *RDR1* gene expression is induced upon infection with TMV and TRV viruses [72, 75]. Our hypothesis is that different viruses may induce expression of individual *RDR* genes. Northern blot analysis will be done using *RDR* gene specific probes. Multiple time points post-infection (5h, 24h, 3d, 5d, 7d, 14d) as well as multiple tissue types (inoculated leaves, non-inoculated leaves, reproductive bud, aerial tissue, cauline leaves) will be sampled to determine the extent of virus infection and effects on *RDR* expression.

*2.e. **Compensatory roles of RDRs and effects on gene expression.*** Though we hypothesize that the members of the *RDR3, RDR4, RDR5* three gene family have independent functions, we also predict that a loss of an RDR can be compensated by a second RDR in the cell. Therefore, *RDR* gene expression levels will be analyzed in the various mutant lines. For instance, we may see an increase in *RDR5* expression in an *rdr4* mutant line; RDR5 may be able to compensate for the loss of RDR4. This result may not be seen under wild-type conditions without induction of the RNAi pathway, thus both non-infected and infected tissue will be analyzed. These sets of data may reveal compensatory gene expression patterns that could not be teased out in wild-type tissue. Gene expression differences will be analyzed by Northern blot or quantitative RT-PCR.

Specific Aim 3. ***To define the mechanisms of* RDR3, RDR4, *and* RDR5 *activity.***

*3.a. **Background and Rationale.*** Most of the characterizations of RDR functions have been through the use of T-DNA insertion lines. Use of these mutant lines determines "what happens if a gene product is absent". In plants there are very few studies that address the question of "how does the protein product normally function". These are obviously much harder questions to address. The experiments described below are designed to initiate such a characterization. The experiments are simple in design, but potentially powerful in data acquisition, as this type of characterization has not been published for any of the RDRs.

*3.b. **Cloning of* RDR3, RDR4, *and* RDR5.** The *RDR3*, *RDR4*, and *RDR5* mRNAs will be amplified from wild-type Columbia tissue by RT-PCR. The coding sequences will be cloned into a Gateway (Invitrogen) entry vector and then shuttled into several pEarleyGate vectors [17]. The expression vectors will allow for either overexpression in plants (pEarleyGate 100), expression of a GFP-His tagged protein (pEarleyGate 103), or expression of a HA-tagged protein for immunoprecipitation experiments (pEarleyGate 205). We have previously cloned *RDR1*, *RDR2*, and *RDR6* into a Gateway vector system and so we are not only proficient with this procedure, but also have three of the *RDRs* already cloned.

*3.c. **To determine the cellular localization of individual* RDRs.** By making GFP-tagged RDR constructs, Agroinfiltration in *Nicotiana benthamiana* leaves can be used to determine cellular localization of the RDRs [73]. The UCD biology department has a fluorescence scope with a digital imaging system available for GFP detection. If necessary, confocal microscopy can be done through the UCHSC DERC Histoloy & Flow Cytometry Core.

*3.d. **To determine RDR:protein and RDR:nucleic acid interactions.*** To define the molecules that the RDRs interact with, immunoprecipitation experiments will be done The pEarleyGate HA-tag vector will be used to make transgenic plants [11]. Using an antibody specific to the HA-tag, RDR protein complexes and RDR nucleic acid complexes will be pulled down from tissue lysates. The precipitates will be assayed either for protein:protein interactions using SDS-PAGE and silver staining, or for RDR:nucleic acid interactions. For the RDR:nucleic acid interactions, bound nucleic acid will be recovered by proteinase K treatment and ethanol precipitation. The recovered nucleic acids will then be cloned, via adapter ligation, and the individual clones sequenced.

The immunoprecipitations will be done with both non-infected and virus infected tissue. These experiments will reveal the types of RNA molecules that a given RDR binds to, leading to an understanding of the molecular intermediates involved in dsRNA production. These experiments are unique to this proposal. RdRP activity has been shown in tomato and wheat germ extracts, but mechanistic studies for the *Arabidopsis* RDRs have not been published [53, 60].

Future Directions

Though most of the proposed methods use conventional approaches for the initial characterization of RDR3, RDR4, and RDR5, they are methods that are well established in the research lab and amenable to undergraduate and Masters-level researchers. In addition, they are techniques that can begin to quickly define the roles of the *RDR* genes during a virus infection. This is important in a field that is moving so quickly. We want to answer some basic questions first and generate publications that will help to build the reputation of the lab. Once we understand the basic gene functions we can carefully dissect the specific modes of action by a more detailed examination of the protein:protein interactions, protein:nucleic acid interactions, and molecular intermediates formed during the processing of RNA by the RDRs. We will be set up for these types of experiments as we will have already cloned the *RDR* genes into a Gateway entry vector that allows for shuttling of the cloned gene into numerous expression vectors. We will also have worked out protocols for isolation of RDR:protein and RDR:nucleic acid complexes. Furthermore, the continuation of data mining is essential for producing more complex hypotheses regarding the roles of the *RDR* genes in the various RNAi pathways. The advantage of having a bioinformatics consultant is that as new data are made available, we can continue to use this information to study the genes of interest. Our bioinformatics student is able to combine data from numerous sources and analyze it with various sophisticated algorithms, which can quickly and inexpensively yield new insights and hypotheses related to the questions we want to answer. This is one way that our smaller lab can stay on top of the wealth of information generated by the larger labs.

E. Significance of Proposed Research

The goal of this research is to understand the diverse roles of RDRs in susceptibility to virus infection, assign functions to each individual gene, and to train future scientists at both the undergraduate and Masters-level. The uniqueness of the research proposed is to integrate whole organism, molecular, and bioinformatics disciplines into one focus and to have this happening in a single laboratory. By using a multidisciplinary approach, the new scientists will be trained in a broader manner than purely molecular or purely organismal biology. This style also provides opportunities for students with diverse interests to work together. Communication between bench biologists and computational biologists will be a necessary skill for future scientists. This project is designed to prepare both types of scientists.

1. Generation of research tools for the Arabidopsis community. Once published, homozygous T-DNA insertion lines will be available for use by the *Arabidopsis* community. Homozygous seed stocks will be deposited with the ABRC [1] and will be available through the

PI's laboratory. Papers published characterizing the *Arabidopsis* plant lines will be linked to the gene information page on the TAIR website [59]. By doing this, other *Arabidopsis* researchers will be able to identify the mutant alleles of interest to their research and may contact ABRC or the PI for seed requests. The sharing of homozygous mutant lines throughout the community promotes collaboration and accelerates the characterization of the gene functions.

2. Translational biology for crop scientists. Understanding the genes important to the plant defense pathway is essential for applying genetic-based technology to crop plants. As more crop plant genomes are sequenced, knowledge of the roles of the different virus defense pathway components will make translation biology much less time consuming [49, 76]. For instance, *Arabidopsis* AtRDR3 (*RDR3*), AtRDR4, and AtRDR5, cluster with the rice OsRDR3 and OsRDR4 genes. The rice homolog for AtRDR2 is OsRDR1. OsRDR2 is homologous to AtRDR6. AtRDR1 does not have a rice homolog. By understanding the roles of the individual *Arabidopsis RDR* genes in the plant defense response to different families of viruses, crop biologists may be able to target specific members of a crop's multi-gene family without having to dissect the role of each member one at a time.

In addition, this proposal will look at the results of overexpressing genes important for virus defense. If by overexpressing this gene, *Arabidopsis* becomes more resistant to a virus infection, then crop biologists may be directed to endogenous anti-virals. Traditional breeding techniques may allow for the selection of plants that have enhanced expression of these genes. Alternatively, expression of the *Arabidopsis* gene, transiently or transgenicly, may protect the crop plant from viral disease. As sequencing of crop plant genomes continues, we can identify whether or not the plant of interest encodes the required genes for virus protection. As a result, a strategy for viral resistance can be proposed.

F. Broader Impacts of Work: Student involvement and Community Service

The research in this proposal will be carried out primarily by undergraduate and Masters-level students. All the figures in this proposal are based on research conducted by undergraduate and Masters-level students. I have a strong commitment to training student scientists and have already mentored eight students in the research lab since the Fall of 2004. Students not only will be involved in carrying out the experiments, they will also be trained in keeping a proper research notebook, designing experiments, evaluating controls, and troubleshooting. The students read primary literature related to the lab's interests and discuss it in a biweekly Journal Club. We hold lab meetings every 4-6 weeks, at which each student presents their project and progress to date. In addition, the students will be encouraged to submit an abstract for a poster presentation at the International *Arabidopsis* meeting or similar research-based meeting. This will give the students a chance to present their research and interact with other scientists. Currently, the lab participates in the Plant Supergroup meetings held at Colorado State University. This bimonthly meeting allows the students to interact with local plant scientists and provides them an opportunity to discuss their research project in a casual, non-threatening environment.

I have a strong commitment to science education and a goal of reaching as many students as possible. Part of my teaching load includes a Molecular Biology lab. I integrate state-of-the-

field techniques in the classroom, as well as basic molecular biology skills and principles. For instance, throughout the semester, each student works with a unique clone from an *Arabidopsis* cDNA library acquired from the ABRC. The project involves sequencing the clone, designing PCR primers, cloning a region of the cDNA to make a GFP fusion, purifying the GFP:cDNA fusion, performing semi-quantitative RT-PCR, and using the cloned cDNA region as a probe for Northern blot analysis. The students will be isolating RNA from the different *Arabidopsis RDR* mutant lines used in this research project. Each student will be using an individual clone to look for differential gene expression in the various mutant plant lines. Not only will the students gain valuable experience with molecular biology techniques, but they will also be obtaining new data for possible use in my research program. Furthermore, each student has his/her own clone (gene) and so they will be challenged to think independently and write a final paper including data derived from their experiments and a hypothesis as to the function of their unique gene. It is not a cookbook course and the experience is invaluable for instilling confidence as a scientist for these students. The course also enables me to reach more students and expose them to true research than just the few who work in the lab.

A third impact of this research is my involvement in PREP (Partnership for Research and Education in Plants) through the Fralin Biotechnology Center at Virginia Tech. I have donated seed stocks from *Arabidopsis* homozygous mutant lines to PREP. These seeds are used in inquiry-based K-12 science education. The K-12 teachers are trained in *Arabidopsis* plant growth and maintenance. They are given mutant and wild-type seeds and the students carry out research projects of their own design. Since a hypothesis in our lab is that RDRs are involved in the plant stress response, students have designed stress related experiments and gathered data as the plants grow. I have visited several high school and junior highs in the Denver area. I have presented information about the research being done at the Downtown Denver Campus and explained how the students could help the research project by designing their own experiments and using our mutant plant lines as their experimental system. The feedback from the students has been incredible. One high school student approached me at the UCDHSC campus open house and was able to do a summer internship in my lab. Several other students have also approached me at the campus open house and relayed their involvement in the high school project. I believe this is a valuable outreach program for women, for underrepresented students, and for students who might not see the value in their science education experiences. The students showed great enthusiasm to do real research and put a lot of pride in their experimental design and data collection.

In summary, this proposal combines state-of-the-field research with science education. The research will determine gene functions for several novel *Arabidopsis* genes. The research will provide a comprehensive picture of the roles of RDRs in an agriculturally important pathway, virus defense. The research will also do something unique; cloning and characterizing the mechanism of *Arabidopsis* RDR functions. Finally, student scientists at all levels of study will be able to experience the thrill of research.