2008

Map-based Cloning of an Anthracnose Resistance Gene in

*Medicago truncatula*

Shengming Yang

*University of Kentucky, syang2@uky.edu*

---

**Recommended Citation**


https://uknowledge.uky.edu/gradschool_diss/641

---

This Dissertation is brought to you for free and open access by the Graduate School at UKnowledge. It has been accepted for inclusion in University of Kentucky Doctoral Dissertations by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.
ABSTRACT OF DISSERTATION

Shengming Yang

The Graduate School
University of Kentucky
2008
Map-based Cloning of an Anthracnose Resistance Gene in *Medicago truncatula*

---

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Agriculture at the University of Kentucky

By

Shengming Yang

Lexington, Kentucky

Director: Dr. Hongyan Zhu, Assistant Professor of Crop Science

C0-director: Dr. Todd W. Pfeiffer, Professor of Crop Science

Lexington, Kentucky

2008

Copyright © Shengming Yang 2008
Map-based Cloning of an Anthracnose Resistance Gene in *Medicago truncatula*

Anthracnose, caused by the fungal pathogen *Colletotrichum trifolii*, is one of the most destructive diseases of alfalfa worldwide. Cloning and characterization of the host resistance (*R*) genes against the pathogen will improve our knowledge of molecular mechanisms underlying host resistance and facilitate the development of resistant alfalfa cultivars. However, the intractable genetic system of cultivated alfalfa, owing to its tetrasomic inheritance and outcrossing nature, limits the ability to carry out genetic analysis in alfalfa. Nonetheless, the model legume *Medicago truncatula*, a close relative of alfalfa, provides a surrogate for cloning the counterparts of many agronomically important genes in alfalfa. In this study, we used genetic map-based approach to clone *RCT1*, a host resistance gene against *C. trifolii* race 1, in *M. truncatula*. The *RCT1* locus was delimited within a physical interval spanning ~200 kilo-bases located on the top of *M. truncatula* linkage group 4. Complementation tests of three candidate genes on the susceptible alfalfa clones revealed that *RCT1* is a member of the Toll-interleukin-1 receptor/nucleotide-binding site/leucine-rich repeat (TIR-NBS-LRR) class of plant *R* genes and confers broad spectrum anthracnose resistance. Thus, *RCT1* offers a novel resource to develop anthracnose-resistant alfalfa cultivars. Furthermore, the cloning of *RCT1* also makes a significant contribution to our understanding of host resistance against the fungal genus *Colletotrichum*.

Key words: *Colletotrichum trifolii*, anthracnose, *Medicago truncatula*, TIR-NBS-LRR, alfalfa

Shengming Yang
May 28, 2008
Map-based Cloning of an Anthracnose Resistance Gene in *Medicago truncatula*

By

Shengming Yang

Hongyan Zhu
Director of Dissertation

Todd M. Pfeiffer
Co-director of Dissertation

Charles T. Dougherty
Director of Graduate Studies
RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor’s degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with permission of the author, and with the usual scholarly acknowledgments.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.
Map-based Cloning of an Anthracnose Resistance Gene in *Medicago truncatula*

---

**DISSERTATION**

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Agriculture at the University of Kentucky

By

Shengming Yang

Lexington, Kentucky

Director: Dr. Hongyan Zhu, Assistant Professor of Crop Science

C0-director: Dr. Todd W. Pfeiffer, Professor of Crop Science

Lexington, Kentucky

2008

Copyright © Shengming Yang 2008
Acknowledgement

I thank my advisor Dr. Hongyan Zhu for invoking my interest in molecular biology, for proposing interesting projects and for continually inspiring and encouraging me while I was a graduate student in his laboratory; Dr. Todd Pfeiffer for devoting his time and expertise, and comments in polishing my thesis; and Dr. Lisa Vaillancourt for precious guidance in fungus culture and pathogen inoculation, Dr. Arthur Hunt for serving as member of my committee and being a good teacher. I am indebted to my friends and colleagues form the Zhu lab. Dr. Muqiang Gao, for teaching me the basic technique and important participating in plant transformation. Dr. Caiyan Chen, for being a good colleague, many fruitful discussions and technical advice, and easy to talk to both about work and spare time activities. Dr. Chenwu Xu, for enthusiastically contributing in mapping resistance gene to Colletotrichum trifolii race 2. Jianchang Gao, for invaluable and instant help offered in RACE and sequencing, so make it possible to complete this program on time. Cui Fan, for always available assistance in pathogen inoculation. Jared Criswell, for planting seed and taking care of the experiment material. Thanks to Dr. David Zaitlin and Dr. Shouan Zhang from KTRDC for generous and skillful help in AFLP analysis and sequencing.

I am so grateful to my parents and my in-laws for their love and supporting me in any detail of my life. Big thanks to Dandan Li. Your love and trust have been encouraging me to perform anything with sufficient confidence. I am so fortunate to have you as my wife. And thank you, Dannie, my sweet baby girl, for bringing me incomparable happiness from God every day. Finally, all the honors and praise belong to God, our savior!
TABLE OF CONTENTS

ACKNOWLEDGMENTS.........................................................................................iii

LIST OF TABLES.................................................................................................vi

LIST OF FIGURES...............................................................................................vii

LIST OF FILES.....................................................................................................viii

CHAPTER I

INTRODUCTION....................................................................................................1

Disease resistance in plants................................................................................1
Plant disease resistance (R) genes......................................................................4
Anthracnose in alfalfa.........................................................................................12
Model legume *Medicago truncatula*.................................................................17

CHAPTER II

Genetic and Physical Localization of *RCT1*......................................................24

Introduction.........................................................................................................24
Materials and Methods.......................................................................................26

*Mapping population and DNA isolation*..........................................................26
*Disease reaction assay*.....................................................................................27
*Bulked segregant analysis using AFLPs*..........................................................28
*Sequencing of linked AFLP marker*.................................................................30
*PCR amplification, marker development and genetic mapping*.......................30
*Physical mapping and sequence analysis*........................................................33

Results................................................................................................................33

*Disease reaction assay and segregation analysis*..............................................33
*Identification of markers linked with the *RCT1* locus*....................................36
*Fine mapping and physical localization of the *RCT1* locus*............................39
*RCT1 region is rich in NBS-LRR genes*............................................................41

Discussion..........................................................................................................44

CHAPTER III

*RCT1* cloned from *M. truncatula* confers Broad Spectrum Resistance to *C. trifolii* in Alfalfa.................................................................48

Introduction.......................................................................................................48

Materials and Methods.....................................................................................50

*Plant materials*................................................................................................50
*Disease resistance assay*.................................................................................50
*DNA sequencing and sequence analysis*........................................................51
*Transformation vector construction*...............................................................51
*Alfalfa transformation*.....................................................................................53
*Rapid Amplification of cDNA Ends (RACE)*......................................................54
LIST OF TABLES

Table 2.1. Molecular markers described in this study .................................................32
Table 2.2. Predicted genes in the RCT1-region ..........................................................42
Table 3.1. Enzymes and recognition sites used in vector construction for transformation .................................................................53
Table 3.2. Gene specific primers used in RACE for every cycle of PCR amplification...55
LIST OF FIGURES

Figure 1.1. Domain structure of NBS–LRR proteins..................................................8

Figure 1.2. Anthracnose symptoms of alfalfa caused by *C. trifolii*......................12

Figure 1.3. A *M. sativa* R-like protein sequence (AAN62760, CC-NBS-LRR) is globally conserved with members of a CC-NBS-LRR cluster located on the *M. truncatula* BAC clone MtH2-07M14 (AC135229)........................................21

Figure 2.1. Symptoms of *Medicago truncatula* Jemalong A17 (resistant) and F83005.5 (susceptible) under the infection of *C. trifolii* race1..........................36

Figure 2.2. BSA based on AFLP. The red peaks indicate size markers.................37

Figure 2.3. Part of BAC overlaps in LG4 which cut from genome assembly brower......38

Figure 2.4. Genetic and physical mapping of the RCT1 locus................................40

Figure 2.5. Distribution and phylogeny of NBS-LRR genes in the RCT1 region........44

Figure 3.1. Map-based cloning of RCT1 ..................................................................57

Figure 3.2. Complementation test of the RCT1 candidate genes..........................60

Figure 3.3. Expression analysis of RCT1 in *M. truncatula* and transgenic alfalfa by RT-PCR..........................................................63

Figure 3.4. Schematic representation of the At and Rt involved in heteroduplex formation and of the two types of heteroduplexes, adapted from Eckhart et al. (1999).................................................65

Figure 3.5. 5’ UTR sequence of resistant RCT1 allele in Jemalong A17...............66

Figure 3.6. Alignment of transcript variants (tv) from the 3’ UTR region of the RCT1 allele in Jemalong A17.......................................................67

Figure 3.7. Structure of the RCT1 protein(s).........................................................68

Figure 3.8. Expression analysis of additional resistant and susceptible alleles in *M. truncatula*........................................................................69

Figure 3.9. Alignment of coding sequences of resistant and susceptible alleles in Jemalong A17 and F83005.5, respectively..........................73
Figure 3.10. Part of alignment of *RCT1* (*rct1*) alleles from 12 ecotypes. 74
Figure A.1. Alignment of 5’ UTR region of the resistant and susceptible alleles in Jemalong A17 and F83005.5, respectively. 81

Figure A.2. Alignment of transcript variants (tv) from the 3’ UTR region of the *rct1* allele in F83.005.5. 82

Figure A.3. Sequence polymorphisms between resistant and susceptible alleles in Jemalong A17 and F83005.5, respectively. 83

Figure A.4. Size polymorphism in 3’ UTR region of genomic DNA between resistant parental line A17 and susceptible parental line F83005.5. 84
CHAPTER I

Introduction

Disease resistance in plants

Despite a lack of a sophisticated immune system like animals, plants have evolved to recognize and response to invading pathogens and to induce disease resistance with adaptive singling pathways and mechanisms. There are two categories of plant resistance: host resistance and nonhost resistance. An old distinction of different forms of host resistance is the division into horizontal and vertical resistance (Parleviet and Zadoks, 1977). Horizontal resistance limits the disease progression of a wide range of pathogen genotypes, which is often inherited as quantitative trait loci (QTLs). This type of resistance can be controlled by multiple factors, and is in some cases referred to as basal resistance (Hammond-Kosack and Parker, 2003). The basal resistance is induced from the detection of general pathogen-associated molecular patterns (PAMPs) by membrane-resident pattern recognition receptors (PRRs). PAMPs include variant substances such as plant cell wall degradation products, lipopolysaccharides (LPS), flagellin and chitin (Gomez-Gomez and Boller, 2000; Ramonell et al., 2005; Shen et al., 2007). The horizontal/basal resistances also can be conferred through non-induced component such as physical characteristics of the plant like physical barrier of waxy cuticle, toxin resistance conferred by antimicrobial secondary metabolites (such as gulcosinolates, phytoalexins, oxylipins etc.). The horizontal resistance primarily function in resistance to necrotrophs (feeding on dead tissue).

Vertical resistance is commonly defined as the ability of the plant to completely block growth of a pathogen. In contrast to the basal resistance that is induced by general
elicitors, the vertical resistance relies on specific recognition. Vertical resistance is subdivided into race-specific and race-nonspecific resistance. Race-specific resistance is denoted as the ability that is active against some races of the pathogen, whereas others remain virulent. Race non-specific resistance is the ability to block all known isolates of a pathogen, but some plant genotypes show susceptible phenotype (Hammond-Kosack and Parker, 2003). Vertical resistance can be due to the presence of a resistance (R) gene, that recognizes a pathogen avirulence (Avr) gene, leading to gene for gene type resistance (Flor, 1947). Such defense response is known as the hypersensitive response (HR), normally associated with a rapid, localized program cell death (PCD) to suppress spread of pathogen (Goodman and Novacky, 1994).

During interactions with attacking pathogens, basal resistance just confers weak immune responses through PRR and slow down pathogen growth (Jones and Dangl, 2006). As a secondary immune receptor, resistance (R) proteins can identify directly or indirectly specific pathogen effectors, encoded by Avr genes (Jones and Dangl, 2006), so pathogen colonization is completely blocked in plants. PRR-triggered responses are found to be linked with mitogen activated protein kinase (MAPK) signaling, the accumulation of reactive oxygen species (ROS), and the activation of defense-related genes involving WRKY transcription factors (Asai et al., 2002; Shen et al., 2007). Basal defense does not prohibit pathogen colonization, it just limits the extent of its spread (Glazebrook et al., 1997). Even gene for gene defenses differ quantitatively and kinetically from basal defense, one very interesting fact is that ROS accumulation is also one of consequences of R protein–triggered immune responses (Tao et al., 2003). It implicates a cross-talk between PRR- and R protein–triggered signaling, but the
mechanisms remain unknown. Recent research on barley powdery mildew resistance identified new WRKY transcription factors, which interact with plant \( R \) protein to recognize fungal avirulence effector (Shen et al., 2007) and introduce resistance. However, these WRKY proteins were also identified to repress and compromise basal defense triggered by PAMP.

Finally, not all pathogens are able to attack all plants. The cases where all interactions between all genotypes of a pathogen and all genotypes of a plant are incompatible (no disease develops) are defined as non-host resistance. Individual genes have been identified by mutational analysis that contribute to non-host resistance, such as Non-Host1 (\( NHO1 \)) against \( Pseudomonas syringae pv. Phaseolicola \) in \( Arabidopsis \) (Lu et al., 2001) and the PENETRATION1 (\( PEN1 \)), \( PEN2 \) and \( PEN3 \) genes that prevent haustorial penetration of the barley powdery mildew fungus (Collins et al., 2003; Lipka et al., 2005; Stein et al., 2006). Two models of non-host resistance are currently in the focus of interest. The first one is that because the pathogen is lack of specific effectors, and it is unable to recognize the plant as a potential host. Thereby PAMP-triggered defense response (basal defense) is not compromised and is invincible in nature (Holub and Cooper, 2004). The second model postulates that plant is equipped with stacks of multiple resistance (\( R \))-genes, which can work together to detect all the avirulence (\( Avr \)) proteins secreted by races of a pathogen. Thus the resistance is durable because it is extremely difficult to defeat such redundant plant \( R \) genes (Hammond-Kosack and Parker, 2003; Holub and Cooper, 2004). Different pathosystem provides strong physiological evidence for each model. Effector activity of compatible barley powdery mildew \( Blumeria graminis \) f.sp. \( hordei \) (\( Bgh \)) suppresses defense against a second challenge
inoculation on barley (Kunoh et al., 1986). This local defense suppression also induced the breakdown of nonhost resistance against the fungi that normally grow on wheat or oat (Olesen et al., 2003), which demonstrated non-compromised basal defense in nonhost interactions. On the other hand, there exist genetic data from crosses of powdery mildew fungi that support the model of single or stacked “classical” $R$ genes as basis of nonhost resistance in wheat (Tosa, 1989; Matsumura and Tosa, 1995). Nevertheless, all these proposed mechanisms of non-host resistance need to be further tested.

**Plant disease resistance ($R$) genes**

$R$ genes have been cloned from numerous plant species, and these genes confer resistance to a wide range of plant pathogens including bacteria, fungi, oomycetes, viruses, and nematodes (Dangl and Jones, 2001). To date, more than 40 plant $R$ genes have been cloned and characterized (Martin et al., 2003). Only four of these 40 cloned genes are recessive (Song et al., 1995; Yoshimura et al., 1998; Iyer and McCouch, 2004; Sun et al., 2004; Gu et al., 2005; Chu et al., 2006). Each of the cloned recessive $R$ genes has very different structure, suggesting that they function differently and are involved in various defense mechanisms. Most dominant $R$ genes appear to encode a limited set of products with several common protein motifs, indicating that plants have evolved similar mechanisms for the expression of resistance to a wide range of unrelated pathogens. The protein motifs encoded by $R$ genes include a nucleotide-binding site (NBS), leucine-rich repeats (LRR), a transmembrane domain (TM), and a serine/threonine protein kinase domain. Based on the various combinations of these motifs, $R$ genes can be grouped into

The largest class of $R$ genes encodes a nucleotide-binding site (NBS) and a leucine-rich repeat (LRR) region. Leucine-rich repeats (LRRs) are 20–29-residue sequence motifs present in a number of proteins with diverse functions. The primary function of these motifs appears to provide a versatile structural framework for the formation of protein–protein interactions. The LRRs generally contain a conserved 11-residue segment with the consensus sequence LxxLxLxxN/CxL (x can be any amino acid and L positions can also be occupied by valine, isoleucine and phenylalanine) (Kobe and Kajava, 2001). Various studies indicate that the LRR motif is responsible for recognition specificity in plant defense response (Kobe and Deisenhofer, 1994; Ellis et al., 1999; Jia et al., 2000; Leister and Katagiri, 2000). The NBS domain of plant $R$ genes is characterized by several sequence motifs found in many ATP- and GTP binding proteins (Traut, 1994), including the Ras superfamily (signal transduction cascades and motility) and some animal genes like Ced-4 (cell death abnormal) and Apaf-1 (apoptotic protease activating factor) (Li et al., 1997). The latter genes regulate the activity of proteases that can initiate apoptotic cell death. As defense mechanisms in plants also include apoptotic-like hypersensitive responses, the appearance of the plant/animal homologies is particularly intriguing. The NBS region may function as an effector domain that initiates signaling cascades leading to resistance responses.

According to the N-terminal structural domains, the NBS-LRR family of $R$ genes can be further classified into two subfamilies. One subfamily codes for a domain with homology to the intracellular signaling domains of the Drosophila Toll and mammalian
Interleukin (IL)-1 receptor (TIR-NBS-LRR), whereas another subfamily codes for a putative coiled-coil domain in the N-terminal region (CC-NBS-LRR). In addition to their structural divergence, TIR-NBS-LRR and CC-NBS-LRR genes have also been found to operate through somewhat distinct signaling pathways requiring either \textit{EDS1} or \textit{NDR1}, respectively (Aarts et al., 1998), suggesting divergent evolution of R genes in the major classifications of land plants. In \textit{Drosophila}, the Toll receptor is essential for establishing dorsoventral pattern in embryos and inducing the immune response in the adult fly. Several human homologues of the Toll protein have been isolated and shown to signal adaptive immunity via NF-kB (transcription factor of nuclear factor-kappa B) and mediate lipopolysaccharide-induced cellular signaling (Medzhitov et al., 1997; Yang et al., 1998). Thus, Toll homologues play a role in pathogen pattern recognition receptor (PRR) triggered signaling in different multicellular organisms. In plants, the TIR motif is implicated in pathogen recognition along with the LRRs (Luck et al., 2000). Members of this subfamily include \textit{N} from tobacco, \textit{L6} and \textit{M} from flax, and \textit{RPP1}, \textit{RPP4}, \textit{RPP5}, and \textit{RPS4} from \textit{Arabidopsis}, and so on (Anderson et al., 1997; Botella et al., 1998; Gassmann et al., 1999; Van der Biezen et al., 2002). Most non-TIR-NBS-LRR subfamily encode a putative coiled-coil (CC) domain in their N-terminus (Pan et al., 2000). CCs are bundles of two to five helices with a distinctive packing of amino-acid side chains at the helix-helix interface, of which the leucine zipper motif is one example (Lupas, 1996). The CC motif is implicated to be involved in protein-protein interaction (Tao et al., 2000). This subfamily includes \textit{RPS2}, \textit{RPS5}, \textit{RPS8}, \textit{RPM1} from Arabidopsis, \textit{I2}, \textit{Mi}, \textit{R3a} from tomato, \textit{Xa1} and \textit{Pib} from rice (Bent et al., 1994; Grant et al., 1995; Ori et al., 1997; McDowell et al., 1998; Milligan et al., 1998; Warren et al., 1998; Wang et al., 1999;
Huang et al., 2005). Figure 1.1 shows the domain structure and putative function of NBS-LRR proteins.

Another intriguing feature of TIR-NBS-LRR genes from different plant species is their capacity to generate alternative transcripts with truncated open reading frames (ORFs), which encode putative TIR-NBS proteins that lack the LRR and C-terminal domains (Jordan et al., 2002). Analysis of >2 million cloned human mRNAs also revealed that alternative splicing of Toll proteins was particularly prevalent (Modrek et al., 2001), indicating that splice variants represent important genetic modifiers of the intricate animal immune system. The functional relevance of alternative variants in plants is unclear. Recent studies of the tobacco N and the Arabidopsis RPS4 gene, both encoding TIR-NBS-LRR proteins, showed that intron-deprived genes (genomic construct with all introns removed) have no or reduced activity (Dinesh-Kumar and Baker, 2000; Zhang and Gassmann, 2003, 2007). These findings suggest that alternative splicing is crucial to defense responses mediated by TIR-NBS-LRR proteins.
Figure 1.1. Domain structure of NBS–LRR proteins (Belkhadir et al., 2004). “A schematic representation of NBS–LRR proteins shows a domain-based platform for the assembly of various putative regulatory factors necessary for controlled signaling. These domains also link to a possible intramolecular regulatory region on the carboxy-terminal (C-terminal) LRR. The cartoons in yellow represent putative interactors assembled on and carboxyl to the CC/TIR domains. The blue square represents ATP, but could also be GTP. The gray cartoons that are associated with the amino-terminal part of the LRR domain represent another set of putative interactors that might be positive regulators (Warren et al., 1998).”(Belkhadir et al., 2004)

NBS-LRR genes are widely distributed in plant genomes. Analysis of the complete genome sequence of Arabidopsis (Col-0) revealed the presence of 149 NBS-LRR-encoding genes plus 58 related genes lacking LRRs (Meyers et al., 2003). R genes are unevenly distributed in plant genomes, and many reside in local multigene families. The clustering is a well-known phenomenon observed at many R gene loci and plays an important role in evolutionary arms race involved between host and pathogen (Dawkins and Krebs, 1979; Clay and Kover, 1996; Hulbert et al., 2001). A classic arms race is that hosts evolve novel R-gene alleles to recognize Avr factors that previously avoided detection in a plant population; likewise, pathogens evolve to overcome host defenses
and to survive on plants. The clustered distribution of \( R \) genes provides a reservoir of genetic variation from which new specificities can evolve. Several mechanisms such as duplication, unequal crossing-over, ectopic recombination, gene conversion, and diversifying selection have been proposed to contribute to the evolution of novel resistance specificities and the structure of \( R \) gene clusters (Michelmore and Meyers, 1998; Ellis et al., 2000; Young, 2000; Hulbert et al., 2001). The LRR regions of clustered \( R \) genes have regions with high levels of sequence homology, which increases the likelihood of unequal crossover events between LRR repeats responsible for encoding recognition specificities. This would reschedule LRRs, mixing existing variability to generate new genes encoding novel combinations of LRRs and potentially new recognition capabilities (Dixon et al., 1998).

The genetic interaction between \( R \) and \( Avr \) genes was simply explained by receptor-elicitor model, which is also called receptor-ligand model. In this model, \( Avr \) genes encode ‘specific elicitors (ligands)’ that interact directly with a ‘receptor’ encoded by the corresponding plant \( R \) genes. However, research on numerous sets of \( R \) and \( Avr \) proteins have revealed only three interactions that support this model (Tang et al., 1996; Jia et al., 2000; Dodds et al., 2006). More reasonable mechanistic explanation of \( R-Avr \) interaction was proposed to be “guard hypothesis” by Van der Biezen and Jones (1998). This model predicts that \( R \) proteins activate resistance when they interact with another plant protein (a guardee) that is targeted and modified by the pathogen. Resistance is triggered when the \( R \) protein detects an attempt to attack its guardee, which might not necessarily involve direct interaction between the \( R \) and \( Avr \) proteins. Compelling evidence for this model was reported, in which an \textit{Arabidopsis} \( R \) protein and other
putative guardee interactions are being investigated (Schneider, 2002; Van der Hoorn et al., 2002). Even more and more evidence was found to support this guard hypothesis, the guardee proteins and how the R proteins are activated by guardees are still remain elusive. A groundbreaking progress was made from the discovery of guardees of RPM1-interacting protein (RIN4) and PBS1 (avrPphB susceptible 1) (Mackey et al., 2002; Axtell and Staskawicz, 2003; Mackey et al., 2003; Shao et al., 2003; Kim et al., 2005), which suggest that the guard hypothesis may no longer be a hypothesis.

In Arabidopsis, RIN4 is the guardee protein for RPM1, and it is also required for RPM1 accumulation. RPM1-dependent resistance is trigged by infection with P. syringae expressing either AvrB or AvrRpm1 which acts as kinases to induce RIN4 phosphorylation (Mackey et al., 2002). These results suggested that RIN4 phosphorylation by AvrB and AvrRpm1 kinase activity result RPM1 activation (Mackey et al., 2002). RIN4 is also required for P. syringae 2 (RPS2)-dependent resistance (Axtell and Staskawicz, 2003; Mackey et al., 2003; Mackie et al., 2003). RPS2 confers resistance against P. syringae expressing the type III effector AvrRpt2 (Bent et al., 1994; Mindrinos et al., 1994). AvrRpt2 is a putative Cys protease (Axtell and Staskawicz, 2003) that induce RIN4 decomposition or cleavage. The activity of RPS2 is suppressed by overexpression of RIN4 thereby postpones RIN4 decomposition with presence of AvrRpt2. So RPS2 activation is achieved with RIN4 decomposition catalyzed by AvrRpt2 (Axtell and Staskawicz, 2003; Mackey et al., 2003; Kim et al., 2005; Desveaux et al., 2007). Another Arabidopsis-Pseudomonas syringae plant-pathosystem also exemplified that R proteins detect indirectly Avr proteins activity (Shao et al., 2003). Two genes, RPS5 (encodes NBS-LRR protein) and PBS1 (encodes a protein kinase) are required for
resistance to *Pseudomonas syringae* strains expressing the *avrPphB* in Arabidopsis. *AvrPphB* was found to proteolytically cleave PBS1, and this cleavage was required for RPS5-mediated resistance, which indicates that *AvrPphB* is also recognized indirectly via its enzymatic activity (Shao et al., 2003).

Other *R* proteins and their pathogen effectors also showed the indirect recognition mechanism. The LRR receptor-like *Cf-2* protein recognizes its pathogen effector by monitoring a host cysteine protease (Rooney et al., 2005). *Pto*, which was originally identified as an resistance gene (Martin et al., 1993), may actually be the guardee factor of the NBS-LRR protein, *Prf* (Mucyn et al., 2006). Interestingly, TIR-NBS-LRRs comprise approximately 60% NBS-LRRs in the *Arabidopsis* genome (Meyers et al., 2003), however, most guardee factors have been described only for CC-NBS-LRRs and the LRR-kinase *Cf-2* protein. Recently, an interactor of *N* receptor-interaction protein 1 (*NRIP1*) for tobacco *N*, a TIR-NBS-TIR gene, was reported (Caplan et al., 2008). This protein that normally localizes to the chloroplasts is recruited to the cytoplasm and nucleus by the 50 kDa helicase (*p50*) domain of *Tobacco Mosaic Virus* (TMV) *Avr* protein, and directly interacts with both *N*’s TIR domain and *p50* (Caplan et al., 2008).

**Anthracnose in alfalfa**

Alfalfa (*Medicago sativa* L.) is the most important and widely grown forage legume worldwide. One of the most desirable characteristics of alfalfa is its high nutritional quality as animal feed. Alfalfa is rich in proteins, vitamins, and minerals, making it highly favorable for hay production and pasture for livestock, especially dairy
cows. In the United States, alfalfa ties with wheat as the third most important crop after corn and soybean. The annual value of alfalfa hay is approximately $7.2 billion (Crop Values 2003 Summary, USDA). Alfalfa is also an integral component of crop rotations because of its capacity for symbiotic nitrogen fixation, reducing the need for chemical fertilizers and thereby reducing ground water pollution. In the US alone, alfalfa is grown on over 23 million acres, where it can fix approximately 5 million metric tons of nitrogen worth an estimated $2 billion per year. However, alfalfa is susceptible to numerous plant pathogens that can limit forage production. On an annual basis, approximately 25% of the U.S. alfalfa hay crop is lost to disease, which amounts to losses exceeding $1 billion. Among the numerous pathogens that are responsible for severe economic losses on alfalfa, the *Colletotrichum trifolii*, causing the anthracnose disease, is particularly damaging.

*Colletotrichum* is one of the most widespread and important disease-causing fungi of plants worldwide. The genus contains over 35 morphological species which cause anthracnose or blight on a wide range of temperate and tropical plants, including grain and pasture legumes, cereals, and fruits (Bailey and Jeger, 1992). Species of *Colletotrichum* have been used as model systems for many years to study fungal differentiation and fungal-plant interactions because of the haploid genome, the ease of *in vitro* culture, and the availability of a reproducible and efficient transformation system (Perfect et al., 1999). During the colonization of plant hosts, most fungal pathogens exhibit either biotrophy, where nutrients are derived from living host cells, or necrotrophy, where nutrients are obtained from dead host cells which have previously been killed by the fungus. However, many species in the genus *Colletotrichum* utilize an
interesting hemibiotrophic infection strategy, in which the pathogen initially develops inside living host cells before switching to a destructive necrotrophic mode of infection (O’Connell et al., 1993). Pathogenicity of *Colletotrichum* spp. depends on a precisely orchestrated sequence of developmental transitions including conidial attachment, germination of the conidium to form a germ tube, differentiation of the germ tube into a specialized infection structure called appressorium, penetration of the plant cell by a penetration peg, biotrophic hyphal growth and nutrient assimilation within plant tissue, and eventual differentiation of hyphal tips into asexual conidia which rupture through the plant via acervuli (Perfect et al., 1999; Dickman, 2000). Therefore, *Colletotrichum* species provide excellent models for studying the molecular and cellular bases of fungal pathogenicity (Yang and Dickman, 1999; Perfect et al., 2000).

Anthracnose of alfalfa, induced by *Colletotrichum trifolii*, causes significant losses on alfalfa in the United States and many other regions of the world (Elgin and Ostazeski, 1982), especially when alfalfa is grown under humid and warm (20-25°C) conditions. The same pathogen also causes anthracnose on a number of closely related legume hosts such as annual medic (*Medicago* spp.), red clover (*Trifolium pratense*), sweet clover (*Melilotus alba*), and vetch (*Vicia* spp.) (Welty, 1982; Stuteville and Erwin, 1990). Symptoms of infected plants are manifest on stems as straw-colored, brown-bordered, and diamond-shaped lesions in which black acervuli develop (Figure 1.2A) (Stuteville and Erwin, 1990). Under favorable conditions, these lesions enlarge, coalesce, girdle, and kill one or more stems. The fungus then spreads internally into crown tissues from lesions on stem bases. A bluish-black discoloration of invaded tissue characterizes the crown rot phase of the disease (Figure 1.2B). The rotting can then further extend
through the crown and into the taproot, killing the entire plant. Symptoms also include blackening and killing of petioles and formation of a shepherd’s crook when the stem wilts and dies suddenly (Figure 1.2C). As alfalfa is a perennial, the fungus can persist in stems and crowns of alfalfa grown in warmer areas and re-infect the surrounding plants when conditions become favorable again.

Anthracnose limits alfalfa production by affecting plant growth, forage yield and quality, and plant vigor. Severe infection in susceptible alfalfa varieties can cause 25-30% losses in forage yield as well as losses in plant stand and vigor (Barnes et al., 1969). In fact the full extent of its influence on stand reduction and subsequent yield loss was not realized until the early 1970s when resistant cultivars were developed and yields were compared to those of susceptible cultivars. Average annual forage yields were 10% higher for resistant cultivars compared with the susceptible cultivars (Elgin, 1981). Despite a lack of experimental data, anthracnose may also lower forage quality by reducing the protein and amino acid content, decreasing the concentration of water-soluble carbohydrates and *in vitro* dry matter digestibility, and thus causing reduction in animal production. In addition, estrogenic compounds, produced as a result of fungal infection, may reduce the reproductive capacity of female animals (Sherwood et al., 1970). Historically, anthracnose has been more of a problem in hay production in the
eastern and southeastern United States, but in the last decade the incidence and severity of anthracnose has increased dramatically in the north central states.

In the United States, two races of *C. trifolii* were previously identified (Ostazeski et al., 1979). Race 1 has been detected in all areas where alfalfa is grown, while race 2 seems to be confined to the eastern states. Resistances against these races in alfalfa are controlled by two single dominant independent segregating loci, *An1* and *An2*. *An1* conditions resistance against race 1, whereas *An2* confers resistance to both race 1 and race 2 (Elgin and Ostazeski, 1985). Race 3 of *C. trifolii*, which was less virulent on race 1 susceptible cultivars, was reported in 1982 (Allen et al., 1982); however, this fungus has subsequently been reported to most likely be *C. destructivum* (O’Neill, 1996b). Race 4 of *C. trifolii* which is virulent on *An2* was first reported in Australia (Mackie et al., 2003). Also Race 4 has recently been reported from Ohio, USA (Ariss and Rhodes, 2006). *An1* may also confer resistance to race 4. Therefore, it is hypothesized that only plants carrying genes *An1* and *An2* are resistant to races 1, 2 and 4. However, few follow-up studies were performed on mapping and characterization of *An1* and *An2* in alfalfa.

Defense responses in legume-*Colletotrichum* pathosystems include hypersensitive reactions (incompatible interaction) and induced resistance mechanisms (compatible interaction) (Esquerré-Tugayé et al., 1992; O’Neill, 1996a). In the case of *C. trifolii* and alfalfa, resistance has been associated with the production of pterocarpan and isoflavonoid phytoalexins following fungal infection in both compatible and incompatible interactions (O’Neill, 1996a; Salles et al., 2002). Several genes required for fungal pathogenicity have been cloned from *C. trifolii* (Dickman, 2000; Dickman et al.,
2003), but little is known about how host resistance genes recognize the pathogen and trigger resistance responses. Cloning and characterization of the host R genes will help to gain a better understanding of the process of host recognition and to develop novel mechanisms for disease control. Unfortunately, cultivated alfalfa has an intractable genetic system because of its autotetraploid (2n = 4x =32) and out-crossing nature. Alfalfa cultivars are usually developed as synthetics by intercrossing a large number of parental lines from diverse germplasm sources. As a consequence, alfalfa cultivars are composed of heterozygous individuals that are genetically heterogeneous. Such attributes of cultivated alfalfa (i.e., tetrasomic inheritance, cross-pollination, and population heterogeneity) have severely limited the ability to carry out genetic analysis of agronomically important traits including disease resistance. So, overcoming such roadblocks is crucial for successful mapping and cloning anthracnose resistance gene for alfalfa improvement.

**Model legume *Medicago truncatula***

The genus *Medicago* contains more than 54 characterized species, including both diploid annuals and tetraploid perennials (Lesins and Lesins, 1979). The most important species of *Medicago* is the tetraploid perennial alfalfa (*M. sativa*), although several annual medics are grown on a limited scale as forage crops or for intercropping as a means to enhance soil nitrogen. *M. truncatula*, also known by the common name “barrel medic”, is native to the Mediterranean basin, and has long been cultivated as winter forage in Australia (Davidson and Davidson, 1993). In recent years, investigators have adopted *M. truncatula* as a model system to study legume genomics and to address
biological issues that are either unique to, or best studied in, the legume family (Cook, 1999). The natural attributes of *M. truncatula* that make it desirable as an experimental system include its annual habit, diploid (2n=2x=16) and self-fertile nature, short lifecycle, relatively small genome (~500Mb), abundant natural variation, and close phylogenetic relationships to the major crop legumes such as alfalfa (Cook, 1999).

Over the past decade researchers have developed the tools and infrastructure for basic research in *M. truncatula*, including efficient transformation systems (Chabaud et al., 2003; Zhou et al., 2004), collections of induced variation (Penmetsa and Cook, 2000), well-characterized cytogenetics (Kulikova et al., 2001), and a collaborative research network ([http://www.medicago.org](http://www.medicago.org)). Research efforts on *M. truncatula* encompass a broad range of issues in plant biology, from studies of population biology (Bonnin et al., 1996b; Bonnin et al., 1996a) and resistance gene evolution (Cannon et al., 2002; Zhu et al., 2002) to the molecular biology of symbiotic interactions (e.g., (Endre et al., 2002a; Limpens et al., 2003; Ane et al., 2004; Levy et al., 2004), plant natural products (Dixon and Sumner, 2003), and micronutrient homeostasis (Ellis et al., 2003). Of importance to these hypothesis-driven investigations is the parallel development of tools for genome analysis, including in excess of 190K ESTs in the public domain, public microarray resources including a 6K cDNA array and a 16K N-linked 70mer oligonucleotide set, TILLING (targeting induced local lesions in genomes) and RNAi (RNA interference) reverse genetics tools for high throughput study of gene function, detailed genetic maps with comparative map connections to the major clades of crop legumes (Choi et al., 2004a), a physical map ~of 20X coverage, an ongoing whole genome sequencing effort, and corresponding activities on metabolic profiling and proteomics (reviewed in
VandenBosch and Stacey, 2003). These genomic resources have been developed under an international collaboration, with funding derived from public and private sources in the United States, Australia, and multiple countries in Europe.

It is anticipated that detailed information about genome structure and function gained from the model species can be readily applied to other closely related plant species. Even though alfalfa and *M. truncatula* differ significantly in genome size, both species have the same basic number of chromosomes (x=8), and the global syntenic relationships between *Medicago truncatula* and diploid alfalfa have been well established (Choi et al., 2004b). Through comparative genetics, conserved genome structure will allow using model species as a surrogate genome for map-based cloning of agronomically important genes in other crops with complex genomes. Moreover, detailed knowledge of the molecular basis of conserved phenotypes in model species can be translated efficiently and potentially to great advantage for gene discovery in related species. The validity of this approach has been illustrated by the simultaneous cloning of genes involved in legume-specific phenotypes (e.g., nodulation) from several legumes including *M. truncatula*, alfalfa, and pea (Endre et al., 2002a; Limpens et al., 2003; Ane et al., 2004; Levy et al., 2004). For example, Nodulation Receptor Kinase (NORK) gene, also called Does Not make Infections 2, (DMI2), was first mapped in tetraploid alfalfa (Endre et al., 2002b). Due to the scarcity of genome sequence information and molecular markers in alfalfa, it is extremely difficult to clone the NORK gene from alfalfa. Taking advantage of high level of synteny between the *M. sativa* and *M. truncatula*, the researchers isolate the NORK gene in *M. truncatula* (Endre et al., 2002a). Rely on the reliable *Agrobacterium rhizogenes*-transformed root system of *M. truncatula* (Boisson-
Dernier et al., 2001), the ability of the wild type NORK gene to complement the non-
nodulation mutant was analyzed in *M. truncatula* too. As can be seen from this case, just
because of almost interchangeable genomic sequence gene order between these two
*Medicago* species, with sufficient molecular biological tools in *M. truncatula*, NORK
gene could be successfully identified from alfalfa. It is expected that a similar strategy
will be also applicable for cloning disease resistance genes.

Genes from *M. truncatula* share extremely high sequence identity to their
counterparts from alfalfa, even for the fast-evolving *R* genes. A combination of genetic
and physical mapping was used to assign the genetic position of a minimum of 150
distinct NBS-LRR homologs in *M. truncatula*. In many cases, the mapped RGHs are also
organized into clusters, and, few, if any, of these clusters contain both TIR- and CC-
NBS-LRR sequences (Zhu et al., 2002). Phylogenetic analysis of these *R*-like protein
sequences indicates a high level of diversity in *M. truncatula*. Members of *R* gene
clusters within the same BAC clone are generally closely related to each other, indicating
recent duplication from a common ancestor. *M. truncatula* resistance gene homologs
(RGHs) also share both conserved gene location and close phylogenetic relatedness to
those genes in other legume species (Zhu et al., 2002). RGH sequences from *M.
truncatula* are closest to sequences from *Medicago sativa* (with sequence identity as high
as 95%). The frequent presence of paired *Medicago* RGH sequences in the phylogenetic
analysis between the two species suggests a high frequency of recent orthologous genes,
which we anticipate will likely keep similar biological functions. An example of
conservation of *R* gene sequences between *M. sativa* (Ms) and *M. truncatula* (Mt) is
illustrated in Figure 1.3. Such high level of sequence conservation between the two
species allowed the direct application of non-optimized genetic markers in either direction (Julier et al., 2003; Choi et al., 2004a). The marker alignment between the two *Medicago* maps reveals an extremely high level of conserved gene order. The conserved genome structure between the two species provides a tool for map-based cloning of alfalfa genes using *M. truncatula* as a surrogate genome. As many of the pathogens of *M. truncatula* including *Colletotrichum trifolii*, are also pathogens of closely related alfalfa. It should be possible to clone resistance genes that are active against pathogens of crop legume species in *M. truncatula*. In addition, due to the close relationship of resistance gene sequences between these species, it is likely that functional resistance genes can be moved across species boundaries by transgenic approaches.

**Figure 1.3.** A *M. sativa* R-like protein sequence (AAN62760, CC-NBS-LRR) is globally conserved with members of a CC-NBS-LRR cluster located on the *M. truncatula* BAC clone MtH2-07M14 (AC135229). AAN62760 is the only full-length *R*-like protein sequence available from alfalfa in the GenBank. Only part of the alignment is shown here.

More and more interest in using annual *Medicago* species as a potential source of resistance genes for alfalfa improvement has led to the evaluation of a *Medicago* core collection (Diwan et al., 1994) against several economically important alfalfa pathogens,
including *Peronospora trifoliorum* (Yaege and Stuteville, 2000), *Erysiphe pisi* (Yaege and Stuteville, 2002), *Colletotrichum trifolii* (O’Neill and Bauchan, 2000), *Phoma medicaginis* (O’Neill et al., 2003), and *Aphanomyces euteiches* (Vandemark and Grünwald, 2004). These studies have revealed a wide range of genetic variation against these pathogens within each *Medicago* species, indicating the possibility of performing genetic analysis of resistances in these annual medics. Further efforts in this subject have been focused on *M. truncatula* because of the well-developed genomic tools available for this model system. In fact, a common theme of several of the *M. truncatula* genome projects is a strong focus on plant-microbe interactions, including both pathogenic and symbiotic microorganisms (Cook, 1999). *M. truncatula* EST libraries have been developed from tissues challenged with pathogenic microorganisms ([http://www.tigr.org/tdb/mtg](http://www.tigr.org/tdb/mtg)). Moreover, there is a large increase in the numbers of researchers focusing on pathogen and insect pests of *M. truncatula*, with the specific intent of identifying resistance phenotypes.

Detailed characterization and validation of the *M. truncatula-C. trifolii* pathosystem has been reported by Torregrosa et al. (2004). The cv. Jemalong A17 is resistant and genotype F83005.5 is susceptible to *C. trifolii* race 1. Historical examination of pathogen development revealed that the infection process and resistant reactions were similar to those observed in alfalfa and other annual *Medicago* species (Mould et al., 1991a, 1991b; O’Neill and Bauchan, 2000). There was no significant difference in fungal development from spore germination to appresorium formation between compatible and incompatible interactions, however, fungal spores in resistant tissues of Jemalong failed to penetrate
and produce the primary and secondary hyphae characteristic of susceptible interactions. Interestingly, the resistant phenotype of Jemalong was associated with a hypersensitive response (HR) typical of ‘gene for gene’ resistance. Such HR is associated with the production of reactive oxygen species (ROS) at the pathogen penetration sites of A17. Examination of resistant/susceptible phenotypes of an F2 population suggests that a single dominant gene confers the resistant phenotype (Torregrosa et al., 2004). Molecular components of the resistance were also analyzed through a small scale cDNA macroarray experiment consisting of 92 genes which were selected for their putative functions in plant defense or signal transduction. Differential profiling was observed between resistant genotype Jamalong and susceptible genotype F83005.5 (Torregrosa et al., 2004). The result fits well with the widely accept concept that defense responses are delayed and less intensive in susceptible plants than in resistant ones.

In the present research, we use *M. truncatula* as a surrogate genome to identify and clone the host resistance gene against *C. trifolii* race 1 (*RCT1*), the causal agent of anthracnose on alfalfa. The *RCT1* will provide new tools for the improvement of cultivated alfalfa, either by means of transgenic approaches or by providing comparative molecular markers that can help with the cloning of the orthologs in alfalfa and enable marker-assisted selection for disease resistance in alfalfa breeding. The specific objectives of this research are: 1) genetic mapping of *RCT1* in *M. truncatula*; 2) map-based cloning of *RCT1*; and 3) complementation test in susceptible *M. sativa*.

Copyright © Shengming Yang 2008
Chapter II

Genetic and Physical Localization of RCT1

(Most content of this chapter was published in Theoretical and Applied Genetics (2007), 116:45-52. The full text of that paper was adopted in this chapter with permission of license number 1990400574233)

Introduction

Alfalfa (Medicago sativa L.) is the most important and widely grown forage legume worldwide. In the United States, alfalfa ranks third in dollar value after corn and soybeans (USDA Crop Values, 2005 Summary). In addition to providing highly nutritious hay and pasture for animal and dairy production, alfalfa is also an integral component of crop rotations because of its capacity for symbiotic nitrogen fixation, underlying its importance as a source of nitrogen in natural and agricultural ecosystems.

Alfalfa is susceptible to numerous damaging pests and pathogens, causing significant losses in forage production (Nutter et al. 2002). An improved understanding of genetic and molecular mechanisms underlying host resistance will facilitate the development of resistant cultivars, thus providing the most efficient and environmentally sound strategy to control alfalfa diseases. Unfortunately, cultivated alfalfa has an intractable genetic system because of its autotetraploid (2n=4x=32) and out-crossing nature. The model legume Medicago truncatula, a close relative of alfalfa, has the potential to serve as a surrogate for genetic analysis of disease resistance in alfalfa and to provide new sources of host resistance (Zhu et al. 2002).

M. truncatula (Barrel Medic) is native to the Mediterranean basin and has long been cultivated as winter forage in Australia. It was chosen as a model legume because
of its annual habit, diploid (2n=2x=16) and self-fertile nature, short lifecycle (2-3 months), ample seed production, relatively small genome (~500Mb), abundant natural variation, and close phylogenetic relationships to the major crop legumes (Cook 1999). In the past decade, abundant genetic and genomic tools and resources have been developed for this model legume (Vandenbosch and Stacey 2003; Young et al. 2005). Since *M. truncatula* and alfalfa share many common pathogens (Yaege and Stuteville 2000, 2002; O’Neill and Bauchan 2000; O’Neill et al. 2003; Vandemark and Grünwald 2004; Tivoli et al. 2006), it is potentially feasible to clone resistance genes in *M. truncatula* that are active against alfalfa pathogens. Due to the close phylogenetic relationship between the two species, it is likely that functional resistance genes can be moved across species boundaries by transgenic approaches (Zhu et al. 2002).

Anthracnose of alfalfa, caused by the fungus *Colletotrichum trifolii*, is one of the most destructive diseases of alfalfa worldwide. The disease causes lesions on stems and leaves, and in advanced stages results in crown and root rot which eventually kills the plant (Stuteville and Erwin 1990). Severe infection in susceptible alfalfa varieties can cause up to 25-30% losses in forage yield as well as losses in plant stand and vigor (Barnes et al. 1969). Two races of *C. trifolii*, races 1 and 2, were identified in North America (Ostazeski et al. 1979). Resistances to the two races in alfalfa were reported to be controlled by two dominant genes, *An1* and *An2* (Elgin and Ostazeski 1985). *An1* conditions resistance to race 1, whereas *An2* confers resistance to both race 1 and race 2. Race 3 of *C. trifolii* was reported in 1982 (Allen et al. 1982), but this fungus was subsequently reported to be likely *C. destructivum* (O’Neill 1996b). Most recently, a new *C. trifolii* race, named race 4, was identified in Australia and in the US (Mackie et al. 2002).
An1 may also confer resistance to race 4 (Mackie et al. 2003; 2007). Few follow-up studies were performed on mapping and characterization of An1 and An2 in alfalfa. Mackie et al. (2007) recently described mapping of quantitative trait loci (QTLs) that condition resistance to the C. trifolii races 1, 2, and 4 in autotetraploid alfalfa.

Defense responses in the alfalfa-Colletotrichum pathosystem include both hypersensitive reactions (incompatible interactions) and induced resistance mechanisms (compatible interactions) (Esquerré-Tugayé et al. 1992; O’Neill 1996a). Resistance was also associated with the production of pterocarpan and isoflavonoid phytoalexins following fungal infection in both compatible and incompatible interactions (O’Neill 1996a; Salles et al. 2002). Several genes required for fungal pathogenicity were isolated from C. trifolii (Dickman 2000; Dickman et al. 2003), but little is known about how host resistance genes recognize the pathogen and trigger resistance responses. Detailed characterization of the M. truncatula-C. trifolii pathosystem revealed that the infection process and resistant reactions were similar to those observed in alfalfa and other annual Medicago species (Mould and Robb, 1992; O’Neill and Bauchan 2000; Torregrosa et al. 2004). The resistance of Jemalong, a M. truncatula genotype, to C. trifolii race 1 was associated with a hypersensitive response (HR) and likely controlled by a single dominant gene (Torregrosa et al. 2004). Cloning and characterization of the host resistance genes will help to gain a better understanding of the process of host recognition and to develop novel mechanisms for disease control.
Materials and Methods

Mapping Population and DNA Isolation

The *Medicago truncatula* F2 mapping population was derived from the cross between Jemalong A17 (resistant) and F83005.5 (susceptible) (Torregrosa et al., 2004). For phenotyping, seedlings of parents and the segregating population were grown in a growth chamber with a 16h light, 23°C / 8h dark, 20°C regime for about 4 weeks before inoculation.

Leaf DNA was extracted from 100mg fresh leaf tissue with 2x CTAB buffer (2% CTAB; 1.4 M NaCl; 100 mM pH 8.0 Tris-HCl; 20 mM pH 8.0 EDTA) (Stewart and Via, 1993).

Disease Reaction Assay

*C. trifolii* Bain and Essary race 1 (isolate 2sp2), as determined with alfalfa cultivars, was kindly provided by Dr. Nichole O’Neill (USDA-ARS, Beltsville, MD). Mycelium was routinely grown on ANM plates (malt extract 2%, bactopeptone 0.1%, glucose 2%, and agar 2%) in the dark at 23°C in Petri dishes. Conidia were produced after a week at 23°C on YPSS medium (yeast extract 0.4%; solutable starch 2%; KH$_2$PO$_4$ 0.1%; MgSO$_4$ 0.05% and agar 1.2%). Spores were collected and washed three times in sterile water with the final concentration being adjusted to 2 x 10$^6$ spores per ml. A stem inoculation method was performed for living plants. The six-week-old plants were inoculated by injection of spores into the stems of living plants using a latex free syringe with a thin needle (0.4mm x 13 mm) (1ml 27G1/2, Becton Dickinson & Co) (Ostazeski and Elgin, 1982; Mackie et al., 2003; Mackie et al., 2007). At least two stems of each
plant were inoculated. Inoculated plants were then transferred to a growth chamber with a 16h light, 23°C/8h dark, 20°C regime with >90% humidity. Symptoms were recorded 7 days post inoculation. The plants were scored as either resistant (no symptom) or susceptible (stem collapse).

**Bulked Segregant Analysis Using AFLPs**

Bulked segregant analysis (BSA) was used to identify AFLP markers linked to *RCT1*. The essence of the BSA is to create a bulk sample of DNA for analysis by pooling DNA from individuals with similar phenotypes (Quarrie et al., 1999). AFLP is a DNA fingerprinting technique which detects multiple DNA restriction fragments by means of Polymerase Chain Reaction (PCR) amplification. Producing high density markers and requiring no prior sequence information makes AFLP very attractive for finding markers linked with disease resistant genes (Michelmore et al., 1991; Tabor et al., 2000; Ouedraogo et al., 2002; Asnaghi et al., 2004; Mackie et al., 2007). In the present study, two parental line bulks, one resistant and four susceptible bulks were prepared for AFLP analysis. The resistant and susceptible bulks were obtained by pooling equivalent amounts of DNA from each of 10 resistant and 12 susceptible F2 individuals.

AFLP procedure was performed essentially as described by Vos et al. (Vos et al., 1995), adapted for automated fluorescent detection of the amplified DNA fragments (Zaitlin, personal communication; the AFLP protocol was included in this dissertation with permission). AFLP was performed with an AFLP Core Reagent Kit (catalog #10482-016) from Invitrogen. Fluorescent Eco+3 primers were labeled at the 5’ end with either Well Red D2 or D4 dyes (Beckman Coulter), and were synthesized by Proligo LLC (Boulder, CO) and Synthegen (Houston, TX), respectively. Genomic DNA samples
(0.1-0.25µg) were digested to completion with EcoRI and MseI at 37°C in a volume of 25µl. Following ligation of the E0 and M0 specific adaptor sequences overnight at 16°C (E0: 5’-GACTGCGTACCAATTC-3’; M0: 5’-GAT GAG TCC TGA GTA A-3’), the reactions were diluted 10-fold into 10mM Tris-HCl, 0.1mM EDTA (pH 8.0). Pre-selective amplifications (PSAs) were performed in 25µl of 1x FailSafe ‘A’ premix (Epicentre, Madison, WI), with E01 and M02 primers (E01: E0+A; M02: M0+C) at 0.5µM, Taq DNA polymerase (New England Biolabs) at 40U/ml, and 3µl of diluted digested/ligated DNA for 22 cycles of 94°C for 30 s, 56°C for 60 s, and 72°C for 60s. Twenty percent of each PSA was examined by electrophoresis (1.5% w/v agarose gel in TBE) for the presence of a predictable banding pattern before proceeding. For selective amplification (SA), PSA reactions were diluted 20-fold into deionized water and amplified with E+3 and M+3 primers (3 selective bases at the 3’ end of E0/M0) using the ‘touchdown’ cycling profile of Vos et al. (1995); an initial denaturation step of 94°C for 2 min, followed by 10 cycles of 94°C for 20s, 66°C for 30s, and 72°C for 2 min with the annealing temperature decreased by 1°C/cycle, and then 20 cycles of 94°C for 20s, 56°C for 30s, and 72°C for 2 min with a final 30min at 60°C. Each SA reaction contained 1X FailSafe ‘A’, 1.5 pmol dye-labeled E+3 primer, 6.25 pmol unlabeled M+3 primer, 50U/ml Taq DNA polymerase, and 4µl of diluted PSA DNA in a total volume of 20µl. All DNA amplifications were performed in a Peltier Thermal Cycler (Bio-Rad, Model: PTC-220). For analysis, SAs were diluted 1/30 into Sample Loading Solution (SLS; Beckman Coulter, Fullerton, CA) containing a 1/100 dilution of DNA Size Standard-600 (Beckman Coulter). DNA fragments were separated by capillary electrophoresis on an automated DNA sequencing instrument (Beckman Coulter).
CEQ8000 Genetic Analysis System) using the Frag-4 method (capillary temperature = 50°C, denaturation 90°C for 2 min, sample injection 30 sec at 2kV, electrophoretic separation 65 min at 4.8kV).

**Sequencing of linked AFLP marker**

A ~490bp fragment was found to be linked with RCT1 locus when E39 (E0+ AGA) and M60 (M0+CTC) were used for SA amplification. After the target band was excised with a sharp blade from polyacrylamide gel, DNA fragment was eluted with 2x pK buffer (Haley et al., 2003) (200mM Tris–Cl, pH 7.5; 25mM EDTA, pH 8.0; 300mM NaCl; 2% sodium dodecyl sulfate). The eluted DNA sequence was cloned into pGEM®-T easy vector (Promega) and transformed into competent cells DH10B (GIBCO/BRL) according to the supplier’s protocols. Plasmid DNA was extracted using the QIAprep Spin Miniprep Kit (QIAGEN). DNA inserts were sequenced using primers complementary to the polylinker M13 site. Sequencing was performed by Dye Terminator Cycle Sequencing (DTCS) Quick Start Kit (Beckman Coulter). The total volume of sequencing PCR is 20µl, including 50-100fmol DNA template; 0.5µM sequencing primer and 8µl DTCS Quick Start Master Mix. The thermal cycling program was set as 30 cycles of 96°C for 20s, 50°C for 20s, and 60°C for 4 min. After ethanol precipitation and purification, the sequencing PCR product was resuspended in 40µl Sample Loading Solution (SLS, Beckman Coulter) for loading into the instrument (Beckman Coulter CEQ8000 Genetic Analysis System).

**PCR Amplification, Marker development and genetic mapping**

SSR (simple sequence repeat) markers surrounding the AFLP marker were mapped to localize the approximate position of RCT1, according to the procedures described by
Mun et al. (2006). Additional markers were then developed from ESTs (expressed sequence tags) and BAC (bacterial artificial chromosome) sequences that were mapped close to the RCT1 locus (Zhu et al., 2002; Choi et al., 2004b; Mun et al., 2006). The primers were designed with online software Primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Markers were based on SNPs (single nucleotide polymorphisms) identified between the two parents, which were converted to CAPS (cleaved amplified polymorphic sequences) markers, as described elsewhere (Zhu et al. 2002; Choi et al. 2004). PCR reactions of 10µl volume contained 20ng DNA template, 1x PCR reaction buffer, 2.0 mM MgCl₂, 0.25mM of each dNTP, 2.5pmole of each primer, and 0.5 unit of DNA polymerase (New England Biolabs). PCR was performed with a 4-min initial denaturation step at 94°C, followed by 35 cycles of 94°C for 1 min, 55°C (adjusted with different primers) for 30 seconds, and 72°C for 1 min, followed by single final extension of 7 min at 72°C. For CAPS markers, 2µl PCR product was applied for digestion with restriction enzymes. Each digestion reaction contained 1-2 units of the corresponding restriction enzyme and 1x compatible buffer in a total volume of 10µl. Enzyme digestions were incubated at the suitable temperature for at least two hours. Digestion products were resolved on agarose gels of appropriate percentage and scored for the respective homozygous parental and heterozygous genotypes. Only susceptible plants (homozygous recessive for the susceptible alleles) were used for genetic mapping.

All markers described in this research are listed in Table 2.1. The genetic map was constructed using the software MAPMAKER version 1.0 (Lander et al., 1987).
Table 2.1. Molecular markers described in this study.

<table>
<thead>
<tr>
<th>Marker name</th>
<th>Template sequence accessions no.</th>
<th>Marker type</th>
<th>A17 restriction enzyme</th>
<th>F3005.5 restriction fragment of pattern of CAPS</th>
<th>Forward primer</th>
<th>Reverse primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>MtB331</td>
<td>AC144503 SSR</td>
<td>SSR</td>
<td>N/A</td>
<td>N/A</td>
<td>GGCCTCTCTGATGCTG</td>
<td>ACAAGCAGGGTGAC</td>
</tr>
<tr>
<td>MtB99</td>
<td>AC127674 SSR</td>
<td>SSR</td>
<td>N/A</td>
<td>N/A</td>
<td>GTTAG</td>
<td>ACACA</td>
</tr>
<tr>
<td>AW257289</td>
<td>AW257289CAPS</td>
<td>Bsmal</td>
<td>341+137</td>
<td>478</td>
<td>GGTCGTTTTTTTCTT</td>
<td>CGGGTGACAGATTAT</td>
</tr>
<tr>
<td>CAP20</td>
<td>AC124959 DominantN/A</td>
<td>Dominant</td>
<td>N/A</td>
<td>N/A</td>
<td>GAAGGAGTGT</td>
<td>ACTTAGTGTG</td>
</tr>
<tr>
<td>CAP25</td>
<td>AC140914 CAPS</td>
<td>HinfI</td>
<td>518+46</td>
<td>394+170+46</td>
<td>AAATTCACCTCAAAACAA</td>
<td>CCGGTATACCCATAC</td>
</tr>
<tr>
<td>CAP29</td>
<td>AC165943 DominantN/A</td>
<td>N/A</td>
<td>518</td>
<td>0</td>
<td>AGTGTGGTTGGCATCGTGTTTGAAACCTGCACACT</td>
<td>AACAATTTAAG</td>
</tr>
<tr>
<td>CAP30</td>
<td>AC138016 CAPS</td>
<td>DraI</td>
<td>156+400</td>
<td>556</td>
<td>AAATTTGAGGA</td>
<td>AGTCAATTTTCCCTG</td>
</tr>
<tr>
<td>61P8-1</td>
<td>CG952991 DominantN/A</td>
<td>NlaIII</td>
<td>463</td>
<td>0</td>
<td>TCTTTGATTTG</td>
<td>ACTTACCATAG</td>
</tr>
<tr>
<td>71O16-2</td>
<td>CG959738 CAPS</td>
<td>StuI</td>
<td>249+78</td>
<td>327</td>
<td>TTTACGAGCA</td>
<td>GCCCTCTATGAGCC</td>
</tr>
<tr>
<td>71O16-1</td>
<td>CG959746 CAPS</td>
<td>BbvI</td>
<td>312+140</td>
<td>452</td>
<td>ATTCATATCCCGTA</td>
<td>CCCCCGGATGGCTTT</td>
</tr>
<tr>
<td>61P8-2</td>
<td>CG928897 CAPS</td>
<td>MaelI</td>
<td>307+180 190+190+117</td>
<td>+32+26</td>
<td>GAATTCTTTT</td>
<td>CAAATACCTCATCA</td>
</tr>
<tr>
<td>81B21-1</td>
<td>CG929447 CAPS</td>
<td>DraI</td>
<td>436+190</td>
<td>626</td>
<td>CCACTAAGCCCTATTGCAGAAGGTCAGTAGGAGGTAAG</td>
<td></td>
</tr>
<tr>
<td>h2_119b6a</td>
<td>AC149473 SSR</td>
<td>SSR</td>
<td>N/A</td>
<td>N/A</td>
<td>CGCAAGGTGTTAGTGATGTCGCGACAGAGGGTCTGAT</td>
<td>CAAGGTCAG</td>
</tr>
<tr>
<td>h2_13m22a</td>
<td>AC164520 SSR</td>
<td>SSR</td>
<td>N/A</td>
<td>N/A</td>
<td>TCAACTCTAAGCGCCACCAAGGTCGAGTCATGGAGGTAAG</td>
<td>CAAGGTCAG</td>
</tr>
</tbody>
</table>
Physical mapping and sequence analysis

In this present study, we took advantage of the availability of the integrated genetic and physical map of the *M. truncatula* genome (Mun et al. 2006; http://www.medicago.org). The genomic and BAC-end sequences allowed anchoring the mapped markers onto the BAC contigs by means of BLAST analysis. Sequencing of BACs H2-144L3 and H2-152N14 were carried out at the Advanced Center for Genome Technology, Department of Chemistry and Biochemistry, University of Oklahoma. Gene prediction was performed using the FGENESH program (Solovyev and Salamov, 1997). Domains were predicted using Pfam 21.0 (Bateman et al., 2004) with an initial E-value cutoff of 0.1. Sequence alignments and phylogenetic analysis were performed using ClustalX (Thompson et al. 1997). Phylogenetic trees were constructed using the neighbor-joining method as implemented in ClustalX with 1,000 bootstrap sampling steps.

Results

Disease reaction assay and segregation analysis The ability to unambiguously distinguish between resistant and susceptible phenotypes is crucial for accurately mapping and subsequent positional cloning of a disease resistance gene. This is even a challenge for many gene-for-gene-type resistance traits governed by a single dominant gene.

To assay for disease resistance and susceptibility, we used an inoculation method based on the injection of inoculum into the stems of living plants (Ostazeski and Elgin 1982; Mackie et al. 2003; Mackie et al. 2007). This inoculation technique allowed for
unequivocal differentiation between resistant and susceptible phenotypes. Seven days post inoculation, the inoculated stems of the susceptible genotype (F83005.5) formed large lesions at the inoculation site and collapsed with severe anthracnose symptoms, while the inoculated stems of the resistant genotype (Jemalong A17) grew normally and were completely symptomless (Figure 2.1A, B). The resistance was clearly associated with a hypersensitive response (HR) at the inoculation site in which infected host cells underwent rapid cell death and further fungal colonization was arrested. In contrast, the fungus can successfully colonize on susceptible stems with well-developed dark acervuli (Figure 2.1C, D, E, F). Abundant spores could be collected from the inoculated sites of the susceptible plants, while sporulation never occurred from the inoculation sites of resistant plants (Figure 2.1G, H). Consistent with Koch’s rules, the spores collected from susceptible plants successfully re-infected and colonized susceptible parental lines and exhibited similar anthracnose phenotypes. The advantage of this inoculation technique was its consistency to cause the breaking and subsequent death of the inoculated stems of susceptible plants, but not causing such symptoms for the inoculated stems of resistant plants. Through this assay, we were able to score each of the F2 individuals as either resistant or susceptible.

The resistant genotype Jemalong A17 and the susceptible genotype F83005.5 were crossed to produce an F2 mapping population. Initial analysis of 231 F2 plants identified 51 susceptible and 180 resistant individuals. The segregation of resistance and susceptibility fits 3:1 ratio ($\chi^2 = 1.12, df = 1, P = 0.29$), suggesting that a single dominant gene controls the anthracnose resistance in Jemalong A17. The resistance gene in Jemalong A17 was named as $RCT1$ (for resistance to *Colletotrichum trifolii* race 1).
Figure 2.1. Symptoms of Medicago truncatula Jemalong A17 (resistant) and F83005.5 (susceptible) under the infection of *C. trifolii* race1.  

A. The inoculated stems of Jemalong A17 grew normally and were completely symptomless.  

B. The inoculated stems of F83005.5 became dry seven days after inoculation.  

C. A scar like wound appeared at A17 stem inoculation site.  

D. F83005.5 stem formed large lesions at the inoculation site and collapsed with severe anthracnose symptoms.  

E. A hypersensitive response (HR) was detected at A17 stem inoculation sites.  

F. F83005.5 stem was disrupted by colonization of the fungus.  

G. No spores were extracted from inoculated sites of A17 stems.  

H. Abundant spores were collected from the inoculated sites of the susceptible plants.
Identification of markers linked with the RCT1 locus AFLP-based bulked segregant analysis (BSA) identified a ~490bp fragment (E39M60-490) that was associated with the RCT1 locus(Figure 2.2). The E39M60-490 fragment was present in the resistant parent A17 and bulk R1. Though there was a week peak in the susceptible bulk S3 (Figure 2.2), it was absent in the susceptible parent and the other 3 susceptible bulks (S1, S2, S4), indicating that E39M60-490 was linked with RCT1.

**Figure 2.2.** BSA based on AFLP. The red peaks indicate size markers. E39M60-490 was indicated as blue peak. E39M60-R indicated the resistant pool. E39M60-S1,-S2, -S3 and –S4 were susceptible pools.
E39M60-490 was excised on polyacrylamide gel and cloned into pGEM®-T easy vector for sequencing. The complete sequence was obtained as:

```
AGTCTCTTGTAGTTGTGCTTGAGCAATAGAAGTCTCTTGATTGTGTTCAGGAGCATAGGA
AGTCTCTTGCAGTTGTGCTTGAGCAATTGAAGACTCTTACTAAGTTTAGTAGTGAGCAT
TTGTAATCAGATATTACATAGTGAACCTCTCTTTGGAAGTGCAAGGGGGACATGACC
TCCGGTTTGTGGAAGAAACCTGTATAAATTGCTTGTGTCTTTCTTCTCCCTCTCTCTTA
TCTGTTTTTACTGCCGTGATCTAGTTCTGAACATCTCTCAAGATAGACTCTATCTGCT
TCTGAATTGCAATTTCAGTGAGGAAAGAGAGAAAAACCTAACACAAATTCAACCC
CTTCTTGTGTTTTTCTCACCCTCATTATATGTTATGCTCTTCTGCGTAGTGTGTTTAC
TCGTTCTGTCAAATAGCCTAGCGT
```

BLAST analysis allowed us to anchor the sequence of onto two overlapping BAC sequences AC174309 and AC119416 on the linkage group 4 (LG4) of *Medicago truncatula* (Figure 2.3). Therefore, we inferred that the *RCT1* locus locate on LG4.

![Overlaps for chromosome chr4](image)

**Figure 2.3.** Part of BAC overlaps in LG4 which cut from genome assembly browser (http://www.medicago.org/genome/). The red arrows indicated BAC AC174309 and AC119416. Unsequenced regions are shown on chromosome as hollow lines. The solid line means sequenced region on chromosome.

To confirm the genetic location of *RCT1*, selected SSR markers (Mun et al., 2006) surrounding the AFLP marker were mapped. These SSRs were originally developed
from the sequenced BAC clones and have been used to integrate genetic and physical map of Jemalong A17, the reference genotype of the *M. truncatula* genome project and the resistant parent of this study (Mun et al., 2006). During the mapping process, we gave preferences to the SSR markers that were linked to the clusters of resistance gene homologs on LG4 (Zhu et al. 2002; Mun et al. 2006).

Of the F2 mapping population, only individuals that were susceptible to pathogen infection were selected for genetic mapping of the *RCT1* locus. One advantage of this strategy was that the susceptible plants were homozygous recessive for the susceptible alleles (*rct1/rct1*) and thus were more informative to detect recombination events, while the resistant plants can be either homozygous (*RCT1/RCT1*) or heterozygous (*RCT1/rct1*). Furthermore, selection of the susceptible plants for genetic mapping avoided the possible experimental errors that might occur during the phenotyping process, because a susceptible plant was surely susceptible, while it was possible, though unlikely, that a plant scored as resistant was indeed susceptible due to escape of infection.

Initial mapping of a base population of 93 susceptible individuals in a 96-well PCR plate (including three DNA samples from the two parents and an F1 plant) identified four SSR markers, MtB99, H2-119H6a, H2-13M22a, and MtB331, on the top of linkage group 4 that were linked to the *RCT1* locus (Figure 2.4A), confirming the *RCT1* location determined with AFLP marker E39M60-490. Additional molecular makers that were mapped in this region on LG4 (Zhu et al. 2002; Choi et al. 2004; Mun et al. 2006) in the F2 population from the cross of Jemalong A17 x A20 was investigated. Through this process, three SNP-based markers, CAP25, CAP30, and AW257289 were identified to be closely linked with the *RCT1* locus (Figure 2.4A.). In particular, AW257289, an EST-
based marker, co-segregated with the RCT1 locus, while a single recombinant event was detected between CAP25 and RCT1 and between CAP30 and RCT1. Based on this initial mapping experiment, it was concluded that RCT1 is located between the markers CAP30 and CAP25.

**Figure 2.4.** Genetic and physical mapping of the RCT1 locus. A. Genetic map of the RCT1 region. RCT1 is located on one end of the M. truncatula molecular linkage group 4. The position of the RCT1 gene was delimited to an ~0.4 cM region between markers CAP29 and 71O16R (as indicated by the solid box). B. Physical map of the RCT1 locus in contig 1357. The BAC contig covers the genetically defined interval containing RCT1 (as indicated by the open box). Numbers indicate the number of recombination breakpoints separating the marker from RCT1. The maps are drawn to scale.

**Fine mapping and physical localization of the RCT1 locus** For the purpose of accurately delimiting the RCT1 locus relative to the closely linked flanking markers, the
mapping population was increased to include 466 susceptible F2 plants. Despite the use of a larger mapping population, there were still no recombination events between AW257289 and RCT1, while nine and five recombinants were identified between RCT1 and CAP25 and between RCT1 and CAP30, respectively (Figure 2.4A). This observation indicated that AW257289 is tightly linked to the RCT1 locus. Therefore AW257289 was used as a query to electronically search for M. truncatula BAC clones that harbor AW257289 and to initiate physical mapping of the RCT1 locus.

The availability of a high-throughput (~20X) physical map and abundant genomic, BAC-end and EST sequence information in M. truncatula offered an in silico approach to physically localize the RCT1 locus. After multiple-step of “sequence walking” through BLAST searching of the M. truncatula Gene-index (http://compbio.dfci.harvard.edu/tgi/) and the NCBI BAC-end sequence (Genome Survey Sequence or GSS) database, the contiguous sequence of AW257289 was anchored onto one end of the M. truncatula BAC clone H2-152N14. Searching the M. truncatula physical map database (http://www.medicago.org) using the BAC ID H2-152N14 as a query enabled H2-152N14 to be assigned onto a single BAC contig of ~700 kb (contig 1357; Figure 2.4B).

The BAC-end sequences as well as high-throughput sequences from several BAC clones of this contig (e.g., H2-61P21, H2-8D13, and H2-34D21) served as templates to develop new SNP markers for fine mapping of this physically defined region. Genotyping the 466 susceptible plants did not reveal any recombination events between RCT1 and markers developed from the BAC ends H2-71O16L, H2-152N14L (AW257289), H2-61P8R, and H2-144L3R. Nonetheless, one recombinant event was
observed between \textit{RCT1} and the markers developed from the BAC end H2-61P8L and the BAC sequence of H2-34D21 (CAP29). Furthermore, two independent recombination events were also detected between H2-71O16R and \textit{RCT1}. Therefore, \textit{RCT1} was determined to locate between 71O16R and CAP29 that span \textasciitilde200 kb. Complete sequencing of the BACs H2-144L3 (AC203223) and H2-152N14 (AC203224) that cover the 200-kb interval identified sixteen genes (Table 2.2). Five of the predicted genes are members of NBS-LRR gene family as described below.

\textbf{Table 2.2.} Predicted genes in the \textit{RCT1}-region.

<table>
<thead>
<tr>
<th>Homology</th>
<th>Copy no.</th>
<th>(E) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Translation elongation factor (\textit{Arabidopsis})</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>Aberrant lateral root formation 4 (\textit{Arabidopsis})</td>
<td>1</td>
<td>5e-90</td>
</tr>
<tr>
<td>AT hook motif-containing protein-related (\textit{Arabidopsis})</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>Nucleic acid-binding, OB-fold (\textit{M. truncatula})</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>Protein kinase family protein (\textit{Arabidopsis})</td>
<td>1</td>
<td>e-142</td>
</tr>
<tr>
<td>mRNA capping enzyme family protein (\textit{Arabidopsis})</td>
<td>2</td>
<td>2e-89</td>
</tr>
<tr>
<td>Defective in exine formation 1 (\textit{Arabidopsis})</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>NBS-LRR homologs (\textit{M. truncatula})</td>
<td>5</td>
<td>0.0</td>
</tr>
<tr>
<td>Hypothetical protein (\textit{Arabidopsis})</td>
<td>1</td>
<td>3e-54</td>
</tr>
<tr>
<td>Proton-dependent oligopeptide transport (POT) family protein (\textit{Arabidopsis})</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>Copper ion binding/oxidoreductase (\textit{Arabidopsis})</td>
<td>1</td>
<td>0.0</td>
</tr>
</tbody>
</table>

\textit{RCT1} region is rich in NBS-LRR genes The majority of plant disease resistance (\textit{R}) genes identified to date belong to the nucleotide binding site (NBS)-leucine rich repeat (LRR) family (Hulbert et al. 2001). NBS-LRR genes can be further divided into two subfamilies based on their N-terminal structural domains, namely, TIR-NBS-LRR and CC-NBS-LRR (Meyers et al. 1999). It is noteworthy that the association between NBS-LRR genes and QTLs conferring resistance to Colletotrichum species was also reported in other plant species (Ferrier-Cana et al., 2003; Abad et al., 2006). In \textit{Phaseolus
vulgaris, QTLs associated with anthracnose resistance to \textit{C. lindemuthianum} were mapped to a cluster of CC-NBS-LRR genes on one end of the common bean LG B4 (Ferrier-Cana et al., 2003). In maize, a candidate gene conferring resistance to anthracnose stalk rot, caused by \textit{C. graminicola}, also encodes a CC-NBS-LRR protein (Abad et al., 2006). Taken together, there is strong evidence that NBS-LRR genes confer gene-for-gene type resistance to \textit{Colletotrichum} species in diverse plant hosts.

Annotation of the ~570-kb contig assembled from the BAC sequences of H2-61P21, H2-8D13, H2-34D21, H2-144L3, and H2-152N14 identified 16 NBS-LRR genes of the TIR-type. Ten of the 16 predicted genes contain complete open reading frames (ORFs), while the remaining are truncated genes lacking either a TIR or LRR domain. The distribution of the predicted NBS-LRR genes is indicated in Figure 2.5A. This set of NBS-LRR genes formed a monophyletic clade in the phylogenetic tree consisting of ~200 \textit{M. truncatula} TIR-NBS-LRR genes, indicating a recent common ancestor of this gene cluster. Phylogenetic analysis of the ten complete TIR-NBS-LRR (TNL) genes revealed two minor clades (Figure 2.5B). Overall, the phylogenetic distances are correlated with physical proximity, suggesting that independent gene duplication has played a role in radiation of this gene cluster. In particular, the 200-kb interval, where the \textit{RCT1} was predicted to be located, contains five tandem duplicated NBS-LRR genes (l to p, Figure 2.5A), three of which (l, m, and p) contain complete ORFs and share ~80% identity with each other at the amino acid level, whereas the other two are truncated genes lacking either a TIR or an LRR domain. These three genes, hereafter referred to as TNL-1, TNL-2 and TNL-3 respectively, were considered as the candidate genes of \textit{RCT1}. 

42
**Figure 2.5.** Distribution and phylogeny of NBS-LRR genes in the *RCT1* region.  
**A.** Organization of the NBS-LRR genes around the *RCT1* locus. Only 16 predicted NBS-LRR genes (represented by letters a through p) are shown. The orientations of the genes are indicated by arrows. TNL, TIR-NBS-LRR; NL, NBS-LRR lacking a TIR domain; TN, TIR-NBS lacking a LRR domain. **B.** Phylogeny of TIR-NBS-LRR genes in the *RCT1* region. Phylogenetic analyses of protein sequences were performed using the ClustalX program (Thompson et al., 1997), and trees were constructed using the neighbor-joining method. Numbers are the percentage of 1000 bootstrap replications supporting the particular nodes.
Discussion

The *RCT1* locus in *M. truncatula* that confers resistance to *C. trifolii* race 1 was finely mapped in the present study. The ability to accurately delimit the *RCT1* locus within a small physical interval was attributed to the use of the stem injection inoculation method (Ostazeski and Elgin 1982; Mackie et al. 2003; 2007). This inoculation technique resulted in qualitative disease reactions and thus allowed reliable discrimination between resistant and susceptible phenotypes in the F2 mapping population. However, the use of spray inoculation or detached leaf assay on the same genotypes caused quantitative disease responses (Torregrosa et al. 2004; Mackie et al. 2007). This difference could be due to differential resistance mechanisms of the plants when inoculated by different inoculation methods (Mackie et al. 2007). Stem injection inoculation by-passed the pre-penetration and penetration process, thus likely resulting in only gene-for-gene type responses (Dickman et al. 2003; Mackie et al. 2007). In contrast, the disease reactions from spray inoculation and detached leaf assay might also involve genes and/or environmental and physiological factors associated with penetration events and pathogenicity. Two interesting phenomena revealed distinctive defense responses against hemibiotrophic *Colletotrichum* spp. in detached and attached *Arabidopsis* leaf assay (Liu et al., 2007). A near-adapted isolate *Colletotrichum linicola A1* could launch a typical infection only on detached, but not attached, *Arabidopsis* leaves. Remarkably, resistance gene-like locus *RCH1*-mediated resistance in intact plants also was compromised in detached leaves during the attacks with the virulent reference isolate *C. higginsianum*. Further validation identified that both the salicylic acid- and ethylene-dependent pathways were required for resistance to *C. higginsianum* and were associated
with induced expression of pathogenesis-related genes *PRI* et al. in intact *Arabidopsis* plants. In contrast, disease symptom development in detached leaves appeared to be uncoupled from these defense pathways and more closely associated with senescence (Liu et al., 2007). The research performed by Liu et al. (2007) highlighted the significance in setting up an appropriate plant-pathogen system during resistance gene cloning, because contrasting phenotypes could be derived from different assays with the same individual plant when attacked by same pathogen.

It was reported that the resistance response to *C. trifolii* occurred at the time of penetration of the cuticle and epidermal cell by the penetration peg (Churchill et al., 1988). Dickman et al. (2003) isolated a lipid-induced protein kinase (LIPK) from *C. trifolii*, which was specially induced by plant cutin. The LIPK was required for appressorium formation, and the mutants of LIPK were unable to infect intact host tissue, but able to colonize host tissue following artificial wounding (Dickman et al. 2003).

Despite the quantitative reactions observed by spray inoculation and detached leaf assay, it was evident that major genes (or QTLs with major effect) conditioning resistance response to the three *C. trifolii* races (i.e., races 1, 2 and 4) exist (Torregrosa et al. 2004; Mackie et al. 2007). Detailed genetic and physical mapping of the *RCT1* locus described here supports the hypothesis that resistance to *C. trifolii* race 1, as determined by the stem injection inoculation, is controlled by a single dominant gene in *M. truncatula*.

The *RCT1* locus mapped to a region on the top of the *M. truncatula* linkage group 4 that contains numerous genes related to previously characterized TIR-NBS-LRR type *R* genes. The *RCT1* region is apparently gene rich. Based on annotation of ~570-kb contig, the gene density is about one gene per 5.2 kb, which is much higher than the estimated
overall gene density for the genespaces in *M. truncatula* (~ 7.9 kb/gene) (Cannon et al. 2006). In particular, the 200-kb interval spanning H2-144L3 and H2-152N14, where the *RCT1* gene was predicted to be located, contains three complete TIR-NBS-LRR genes. This observation suggests that the *RCT1* might also encode an NBS-LRR type resistance protein.

The association between NBS-LRR genes and QTLs conferring resistance to *Colletotrichum* species was also reported in other plant species (Ferrier-Cana et al. 2003; Abad et al. 2006). In *Phaseolus vulgaris*, QTLs associated with anthracnose resistance against *C. lindemuthianum*, a closely related species of *C. trifolii*, were mapped to a cluster of CC-NBS-LRR genes on one end of the common bean linkage group B4 (Ferrier-Cana et al. 2003). In maize, a gene conferring resistance to anthracnose stalk rot, caused by *C. graminicola*, also encodes a CC-NBS-LRR protein (Abad et al. 2006). Taken together, there is strong evidence that NBS-LRR genes confer gene-for-gene type resistance to *Colletotrichum* species in diverse plant hosts.

Mackie et al. (2007) reported the mapping of QTLs responsible for resistance to the three races (races 1, 2 and 4) of *C. trifolii* in alfalfa. Disease reaction was assayed by both spray and stem injection inoculation. The plants were scored as either resistant or susceptible for the stem injection inoculation, as described in this study, while a disease index of 1 to 5 was used for spray inoculation. Despite the use of different scoring systems, a single locus for the injection assay and the strongest QTL for the spray assay were co-incident on a *M. sativa* group linkage that appeared to be homologous to *M. truncatula* linkage group 8, based on mapping two *M. truncatula* SSR markers 36b12e and 115m15b in alfalfa (Mackie et al. 2007). It is uncertain whether the *RCT1* in *M.
truncatula is orthologous to the locus mapped in M. sativa. It was reported that a chromosomal translocation event occurred between chromosomes 4 and 8 in the M. truncatula ecotype Jemalong A17 (Kamphuis et al. 2007). Further work is needed to determine the functional and evolutionary relationship between the RCT1 locus in M. truncatula and the locus described by Mackie et al. (2007). This can be done by mapping the candidate RCT1 gene of M. truncatula in alfalfa as well as by mapping more M. truncatula markers in alfalfa that are closely linked to the mapped resistance locus.
Chapter III

*RCT1* cloned from *M. truncatula* confers Broad Spectrum Resistance to *C. trifolii* in Alfalfa

Introduction

Alfalfa (*Medicago sativa* L.), known as the “Queen of Forages”, is the world’s most important and widely grown forage legume. Alfalfa is rich in proteins, vitamins and minerals, providing highly nutritious hay and pasture for animal and dairy production. In the United States, alfalfa ranks with wheat as the third most important crop after corn and soybeans (USDA Crop Values, 2005, 2006; http://www.nass.usda.gov/). Like other legume species, alfalfa contributes to the sustainability of agricultural ecosystems because of its capacity for symbiotic nitrogen fixation. Moreover, the combination of its high biomass production, perennial growth habit, and ability to fix atmospheric nitrogen, have led to an increased interest in using alfalfa as a biofuel feedstock for production of ethanol and other industrial materials.

Alfalfa production has been negatively impacted by damaging pests and pathogens. On an annual basis, ~20% of the U.S. alfalfa hay crop is lost to disease, amounting to losses exceeding $1 billion (Nutter et al., 2002). An improved understanding of genetic and molecular mechanisms underlying host defense will offer novel tools to develop resistant alfalfa cultivars, thus providing an efficient and environmentally sound strategy to control alfalfa diseases. Cultivated alfalfa is autotetraploid (*2n=4x=32*) and out-crossing, making it recalcitrant to genetic analysis, while its diploid relative *Medicago truncatula* is a comparatively simple genetic and genomic system, and has emerged as a reference species for the study of legume biology.
(Cook, 1999). The two species share conserved genome structure and content (Choi et al., 2004b), and thus it is anticipated that *M. truncatula* can serve as a surrogate for cloning the counterparts of many economically important genes in alfalfa. In the case of disease resistance, the family of NBS-LRR disease resistance (*R*) genes has been extensively characterized at the sequence and phylogenetic levels in *M. truncatula* (Zhu et al., 2002; Ameline-Torregrosa et al., 2008a). In parallel, the long history of cultivation of alfalfa provides numerous examples of disease phenotypes that could be mitigated, if an *R* gene(s) with appropriate specificities were identified. In such cases, discovery of *R* genes with novel specificities in *M. truncatula* could have direct applicability to cultivated alfalfa.

Anthrascnose of alfalfa, caused by the fungal pathogen *Colletotrichum trifolii*, is one of the most destructive diseases of alfalfa worldwide. The same pathogen also causes anthracnose on closely related forage legumes, including annual medic species (*Medicago* spp.) and clovers (*Trifolium* spp.). Three races of *C. trifolii* (i.e., races 1, 2 and 4) have been described based on differential responses of alfalfa cultivars (Ostazeski et al., 1979; Mackie et al., 2003; Ariss and Rhodes, 2006), with strain specificity in alfalfa conferred by two independent dominant resistance genes, *An1* and *An2* (Elgin and Ostazeski, 1985; Mackie et al., 2003). *An1* confers resistance to race 1 and likely, race 4, whereas *An2* confers resistance to races 1 and 2. It is noteworthy that the race 3 of *C. trifolii* was reported in 1982 (Allen et al., 1982), but this fungus was subsequently reclassified as *C. destructivum* (O’Neill, 1996b).

Defense responses of *M. truncatula* against *C. trifolii* are similar to those observed in alfalfa and other annual *Medicago* species, including hypersensitive reactions
in incompatible interactions and delayed induction of resistance mechanisms in compatible interactions (Mould and Robb, 1992; O’Neill, 1996a; O’Neill and Bauchan, 2000; Torregrosa et al., 2004; Yang et al., 2007). Alfalfa responses to \textit{C. trifolii} infection also involve the production of pterocarpan and isoflavonoid phytoalexins (O’Neill, 1996a; Salles et al., 2002). In Chapter II, we described the genetic and physical localization of the \textit{RCT1} (for resistance to \textit{C. trifolii} race 1) locus in \textit{M. truncatula}. Here we report the map-based cloning of \textit{RCT1}. \textit{RCT1} encodes a TIR-NBS-LRR type \textit{R} protein that confers broad-spectrum anthracnose resistance when transferred into the susceptible alfalfa plants. Thus, \textit{RCT1} provides a new resource to develop anthracnose-resistant alfalfa cultivars and contributes to our understanding of disease resistance mechanisms against the fungal genus \textit{Colletotrichum}. This study also highlights the potential of ‘translational’ research from \textit{M. truncatula} to the forage legume alfalfa.

### Materials and Methods

#### Plant Materials

The F2 mapping populations were derived from the cross between \textit{M. truncatula} genotypes Jemalong A17 (resistant) and F83005.5 (susceptible). Seedlings were grown in growth chambers programmed for 16h light at 23°C and 8h dark at 20°C.

#### Disease Resistance Assay

\textit{C. trifolii} race 1 (isolate 2sp2), race 2 (isolate H4-2) and race 4 (isolate OH-WA-520) were used for inoculation as described by Yang et al. (2007). Briefly, conidia were produced after a week at 23°C on YPSS medium. Spores were collected and washed
three times in sterile water with the final concentration being adjusted to \(2 \times 10^6\) spores per ml. The six-week-old plants were inoculated by injection of spores into the stems of living plants using a latex free syringe with a thin needle (0.4mm x 13 mm) (1ml 27G1/2, Becton Dickinson & Co). Inoculated plants were then transferred to a growth chamber with a 16h light, 23°C/8h dark, 20°C regime with 100% humidity. Symptoms were recorded 7 days post inoculation. The plants were scored as either resistant (no symptom) or susceptible (stem collapse).

**DNA Sequencing and Sequence Analysis**

Sequencing of BACs H2-144L3 (AC203223) and H2-152N14 (AC203224) were carried out at the Advanced Center for Genome Technology, Department of Chemistry and Biochemistry, University of Oklahoma. Gene prediction was performed using the FGENESH program (Solovyev and Salamov, 1997). Domains were predicted using Pfam 21.0 (Bateman et al., 2004). Sequence alignments were performed using ClustalX (Thompson et al., 1997).

**Transformation Vector Construction**

The \(RCT1\) locus was delimited within a physical interval spanning \(~200\) kb located on the top of \(M.\ truncatula\) LG 4, which is assembled from the BAC sequences of H2-144L03 and H2-152N14. The 200kb interval contains five tandem duplicated NBS-LRR genes, three of which contain complete open reading frames (ORFs) and share 80% identity with each other at the amino acid level. These three genes serve as strong candidates of \(RCT1\), hereafter referred to as TNL-1, TNL-2, TNL-3 respectively. DNA
constructs that contain the individual candidate genes under the control of their native promoters were used for complementation test.

The genomic DNA of BAC H2-144L03 was digested with SacI and KpnI to obtain a 12.9 kb genomic fragment that contained the ~5.0 kb TNL-1 coding region plus ~3.6 kb upstream of the start codon and ~4.7 kb downstream of the stop codon. The same BAC also was digested with StuI and BglII to obtain a 10.3 kb genomic fragment that covered the TNL-2 coding region (~5.0 kb) and ~3.0 kb and ~2.2 kb up- and downstream sequence respectively. The TNL-3 genomic fragment was obtained by digestion of the BAC H2-152N14 with SpeI and SgrAI. This digestion produced ~10.0 kb fragment that contained TNL-3 coding region plus ~3 kb promoter region and ~300bp 3’ untranslated region (UTR). The transformation vector used was pCAMBIA 2300. Conditioned with availability of restriction endonuclease recognition sites on pCAMBIA 2300, isocaudarners of enzymes used in BAC digestion were selected to cut vector to acquire same blunt/sticky ends with plant genomic fragment. Isocaudarner, a special kind of DNA restriction endonucleases, produces the same sticky end in DNA fragment when digested different sequences. All enzymes used in vector constructions were listed in table 3.1 in detail.

**Table 3.1.** Enzymes and recognition sites used in vector construction for transformation.

<table>
<thead>
<tr>
<th>Plant BAC (Enzymes/recognition site)</th>
<th>Vector (Enzymes/recognition site)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TNL-1 (H2-144L03)</strong></td>
<td>SacI GAGCT/C</td>
</tr>
<tr>
<td></td>
<td>KpnI GGTA/C</td>
</tr>
<tr>
<td></td>
<td>SacI GAGCT/C</td>
</tr>
<tr>
<td></td>
<td>KpnI GGTA/C</td>
</tr>
<tr>
<td><strong>TNL-2 (H2-144L03)</strong></td>
<td>StuI AGG/CCT (blunt)</td>
</tr>
<tr>
<td></td>
<td>BglII A/GATCT</td>
</tr>
<tr>
<td></td>
<td>Smal CCC/GGG (blunt)</td>
</tr>
<tr>
<td></td>
<td>BamHI G/GATCC</td>
</tr>
<tr>
<td><strong>TNL-3 (H2-152N14)</strong></td>
<td>SpeI A/CTAGT</td>
</tr>
<tr>
<td></td>
<td>SgrAI CA/CCGGTG</td>
</tr>
<tr>
<td></td>
<td>Xbal T/CTAGA</td>
</tr>
<tr>
<td></td>
<td>XmaI C/CCGGG</td>
</tr>
</tbody>
</table>
Alfalfa Transformation

The transformation method was adapted from Austin et. al. (Austin et al., 1995). The Agrobacterium strain used was LBA4404 (rifampicin and strepclin resistant). New growth leaves were cut from Regen SY plants susceptible to C. trifolii 1 and sterilized with 70% alcohol 10s and 20% bleach 1.5m with 0.05% Tween 20, followed by three rinses in sterile distilled water. Leaf edges were removed and tissue dropped into SHO liquid medium (Schenk and Hildebrand salt s, Schenk and Hildebrand vitamins, 30 g/L sucrose, 0.5 g/L MES, pH 5.7 with KOH). When sufficient explants had been taken, Agrobacterium cells from an overnight culture grown in liquid YEP selection medium (10 g/L protease peptone, 10 g/L yeast extract, 5 g/L NaCl) was added to SHO medium (1ml Agrobacterium/4 ml SHO). Cell density was adjusted to fall between 0.6-0.8 at A_{660}. After 30 minutes inoculation, the explants were gently blotted on filter paper and placed on B5H medium (3.1 g/L Gamborg's B5 salts, 1.0 ml/L 1000x Gamborg's B5 vitamins, 0.5 g/L KNO3, 0.25 g/L MgSO4(7H2O), 0.5 g/L proline, 30 g/L sucrose, pH to 5.7 with KOH, 8 g/L Phytagar) for 3 days of co-culturation. At the end of this period, they were rinsed three times and transferred to B5hKTc selection medium (B5H with stock amino acids and hormones plus 50 mg/L kanamycin and 500 mg/L ticarcillin). B5h stock amino acids contain 6.65 g L-glutamine, 0.83 g serine, 0.004 g adenine. Hormones for B5h are 1 mg/L 2,4-D and 0.1 mg/L kinetin. Plates were maintained at 24° C, 16 h photoperiod and light intensity of 60-80 μE/m²s. Explant-derived calli (and occasionally embryoids) which formed within 3 weeks on this medium were moved to B5hOKTc regeneration medium (similar with B5hKTc but without hormones). After 3-4 weeks, embryos were transferred to MMSKTc medium (4.3 g/L Murashige and Skoog salts, 1
ml/L 1000x Nitsch and Nitsch vitamin stock, 0.1 g/L myo-inositol, 30 g/L sucrose, pH to 5.7 with KOH, 7.0 g/L phytagar. Before pouring add 500 mg/L ticarcillin, 50 mg/L kanamycin.). Over the next 1-3 weeks embryos will form a shoot and sometimes a root. Green plantlets were moved to MMSTc medium (similar with MMSKTc but lacking kanamycin) for further shoot and root development until a good root system formed.

**Rapid Amplification of cDNA Ends (RACE)**

After RCT1 was defined among candidate genes by complementation test, the full-length cDNA was determined with RACE. Two-week-old Jemalong A17 and F83005.5 plants were inoculated with pathogen by spraying spore suspension (1x10^6/ml) to the seedlings. ~100 mg young leave sample was collected for RNA extraction. RNA was isolated using RNeasy miniprep kit (Qiagen). Reverse transcription was completed with Superscript™ II reverse transcriptase (Invitrogen). The 5’ and 3’ ends of the cDNAs were amplified with SMART RACE cDNA kit (Clontech) (Zhu et al., 2001). Two cycles of nested PCR were performed followed by primary PCR reaction to get the final amplification product. After the three cycles of PCR, the product was cloned into pGEM®-T easy vector (Promega) for sequencing. All the gene specific primers (GSPs) used in RACE were listed in Table 3.2.

**Table 3.2. Gene specific primers used in RACE for every cycle of PCR amplification.**

<table>
<thead>
<tr>
<th>Cycle</th>
<th>5’ RACE</th>
<th>3’ RACE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>CAACAAATCCAGCAAGGCCAGCCGC AAC</td>
<td>TGCCCAAGGCTGTCTCAGGTTTCCCATA</td>
</tr>
<tr>
<td>2nd</td>
<td>AGCCGCAACACGAAGCTCATTCTCCAC</td>
<td>AAGCCTTTGGCTGACATGTGGATCAGA</td>
</tr>
<tr>
<td>3rd</td>
<td>TGCGCGAGTGCTTCTCCTCGGAAACTC</td>
<td>TGCTGCCCATGAGGATCTCTCAGGCAAG</td>
</tr>
</tbody>
</table>
Analysis of Gene Expression by RT-PCR

For gene expression analysis, plants were inoculated with *C. trifolii* race 1 by spraying spore suspension (2x10^6/ml) to the seedlings maintained in a growth chamber. Leaves at 0, 1, 2 and 3 dpi were collected for RNA isolation. Total RNA was isolated by the Qiagen Plant RNeasy. Two micrograms of RNA was used to perform RT reactions using M-MLV reverse transcriptase (Invitrogen) in a 20-µL reaction mixture. Two microliters of the RT reaction was used as a template in a 20-µL PCR reaction solution. The PCR primers were as follows: *MtActin*, 5'-GGAGAAGCTTGCATATGTTG-3' and 5'-TTAGAAGCACTTCCTGTGGA-3'; *RCT1*, F1: 5'-AAATGGTTTGCTCCAGGTAG-3', F2: 5'-CAAAGGTGAGGGACTG-3', F3: 5'-CATAGATCTTCTCTTCTTCTCTTCC-3', F4: CCAAAAAGACTACATAAAGCCTGTGA, R1: 5'-TTGCCACACAAGITTTAGCATTG-3', R2: 5'-ATTTCGACGACTGGTTCATC-3', R3: GCCACCAATGTAAGCATAAAATCTGCAA, R4: TTGGCCTTAACGTAACACTTG.

Results

Map-based Cloning of *RCT1* The *RCT1* locus was previously mapped to *M. truncatula* chromosome 4, based on an F2 mapping population derived from the cross between the resistant genotype Jemalong A17 and the susceptible genotype F83005.5 (Yang et al., 2007). Fine mapping using 466 susceptible individuals (*rect1/rect1*) selected from the F2 population identified an EST (Expressed Sequence Tag)-based CAPS (Cleaved Amplified Polymorphic Sequence) marker, AW257289, which co-segregated with the *RCT1* locus (Figure 3.1). AW257289 anchors one end of the *M. truncatula* BAC
(Bacterial Artificial Chromosome) clone H2-152N14, which is located on the physical map of *M. truncatula* within the ~700 kb contig #1357 ([http://www.medicago.org](http://www.medicago.org)).

**Figure 3.1.** Map-based cloning of *RCT1*. The position of *RCT1* was delimited to a genomic region between markers CAP29 and 71O16R. Numbers indicate the number of recombination breakpoints separating the marker from *RCT1*. Candidate genes of *RCT1* are indicated. Arrows point to the transcriptional direction of each candidate gene. TNL, TIR-NBS-LRR; NL, NBS-LRR lacking a TIR domain; TN, TIR-NBS lacking a LRR domain. The map is drawn to scale.

To more precisely delimit the *RCT1* locus within a physical interval, DNA sequence information from contig 1357 was utilized to develop new CAPS markers that flank AW257289 (Figure 3.1). Through this process, a total of three flanking recombination events were identified: one between AW257289 and CAP29, and two between AW257289 and H2-71O16R. No recombination events were detected between AW257289 and markers H2-71O16L (CG959746), H2-61P8R (CG928897), and H2-
144L3R (CR501753). It was therefore determined that the RCT1 locus resides within an ~200 kb window between 71O16R and CAP29. Sequencing and annotation of the BACs H2-144L3 (AC203223) and H2-152N14 (AC203224) identified 5 tandemly arrayed TIR-NBS-LRR (TNL) type R gene homologs. Three of the five NBS-LRR genes contain complete open reading frames (ORFs) and share ~80% identity at the amino acid level, whereas the other two R gene homologs are truncated genes lacking either a TIR or an LRR domain. The three TNL genes, hereafter referred to as TNL-1, TNL-2, and TNL-3, respectively, were considered as candidate genes of RCT1.

RCT1 locus co-segregated with resistance to C. trifolii races 2 and 4 in M. truncatula

M. truncatula genotype Jemalong A17 was resistant to all three known races of C. trifolii, whereas F83005.5 was susceptible to the same three races. Parallel to mapping and cloning of RCT1, the A17 X F83005.5 F2 mapping population was also phenotyped for resistance to C. trifolii races 2 and 4. Segregation data suggested that the resistance to C. trifolii race 2 is likely controlled by two independent dominant genes, as only 76 susceptible individuals were identified from a total of 1,166 F2 plants, which fits the 15:1 (resistant versus susceptible) ratio ($\chi^2=0.13$, $df=1$, $p=0.72$). Genotyping the 76 susceptible plants using the marker AW257289 did not detect any recombination events; that is, all susceptible plants have the allele coming from the susceptible parent. A similar experiment was also performed for the C. trifolii race 4 in a mapping population consisting of 262 F2 individuals. Of the 262 F2 individuals, the ratio of resistant versus susceptible (206:56) statistically fits 3:1 ($\chi^2=1.84$, $df=1$, $p=0.18$), suggesting that the resistance to C. trifolii race 4 is controlled by a single dominant gene. Strikingly, the
resistance to *C. trifolii* race 4 also co-segregated with the RCT1 locus based on mapping with the AW257289 marker. These data suggest that the resistance to the three *C. trifolii* races is either tightly linked or controlled by a single RCT1 gene in *M. truncatula*.

**RCT1 confers broad spectrum resistance to anthracnose disease when transferred into susceptible alfalfa clones** To validate candidate genes for the RCT1 locus, genomic constructs (i.e., introns included) of TNL-1, TNL-2 and TNL-3 were cloned under the control of their native promoters. Since the *M. truncatula* genotype F83005.5 was recalcitrant to transformation and regeneration, two independent clones from the susceptible Regen SY cultivar of alfalfa were selected as a study system. The two selected clones, designated as Regen SY-6 and Regen SY-11, were susceptible to all three races of *C. trifolii*, a feature that enabled us to test whether RCT1 confers broad spectrum resistance, as suggested by linkage mapping in *M. truncatula*.

Transgenic alfalfa plants developed from the three constructs were first inoculated with *C. trifolii* race 1. Independent transgenic plants containing TNL-2 (n=55), TNL-3 (n=15), the empty vector pCAMBIA2300 (n=26), and untransformed wild-type plants (n=10) vegetatively propagated from the original clones were all susceptible to *C. trifolii* race 1 (Figure 3.2). Three to four days post inoculation (dpi), the inoculated stems of the susceptible plants formed a large lesion at the inoculation site and subsequently collapsed at 7 dpi. In contrast, independent transformants containing the TNL-1 transgene (n=42) were completely resistant to the pathogen. Thus, TNL-1 was defined as the RCT1 gene.
**Figure 3.2.** Complementation test of the *RCT1* candidate genes. Transgenic plants containing individual candidate genes and the empty vector (pCAMBIA2300) as well as wild-type plants were inoculated with the races 1, 2, and 4 of *C. trifolii*. Only TNL-1 transgenic plants showed resistance to *C. trifolii*. Arrows indicate inoculated stems. S = susceptible, R = resistant.
For purposes of evaluating resistance to *C. trifolii* races 2 and 4, vegetative clones were propagated from all transgenic lines and rated for disease phenotypes following pathogen inoculation. Strikingly, all the transgenic plants containing the TNL-1 transgene were resistant to races 2 and 4, whereas all transgenic plants containing either TNL-2 or TNL-3 transgenes, as well as control vector only and non-transgenic plants, were susceptible. These data, along with the fact that the resistance to the three *C. trifolii* races co-segregated with the *RCT1* locus in *M. truncatula*, strongly indicated that the *RCT1* gene confers broad spectrum resistance to the three *C. trifolii* races.

**RCT1 is constitutively expressed and alternatively spliced**  

*RCT1* transcription unit was deduced based on a combination of *ab initio* predictions using FGENSEH (Solovyev and Salamov, 1997) and alignment of genomic and cDNA sequences (Figure 3.3A). These analyses revealed gene coding sequence composed of five exons, (557; 1,105; 276; 819; and 540 bp respectively), with inferred intron positions typical of many TIR-NBS-LRR type R gene homologs described in *Arabidopsis* (Meyers et al., 1998) and *M. truncatula* (Ameline-Torregrosa et al., 2008b).

Semi-quantitative reverse transcriptase (RT)-PCR using the *RCT1*-specific primers (F1 and R1 as indicated in Figure 3.3A) was performed to analyze the expression profile of *RCT1*, following inoculation with *C. trifolii* race 1. Leaf tissue of resistant and susceptible parents was collected at four different time points (0, 1, 2, and 3 dpi). The RT-PCR result (Figure 3.3B) indicated that *RCT1* was constitutively expressed in the resistant parent Jemalong A17, and the level of expression was not regulated by fungal infection. This conclusion was further supported by analysis of the *M. truncatula* gene
index (MtGI) database (http://compbio.dfci.harvard.edu), from which all of the cognate expressed sequences of RCT1 (i.e., TC96909, TC97262, and BF643292) were from EST libraries of non-infected tissues.
Figure 3.3. Expression analysis of \textit{RCT1} in \textit{M. truncatula} and transgenic alfalfa by RT-PCR. \textbf{A.} Gene structure of \textit{RCT1}. The exons and introns are indicated by boxes and lines, respectively. Numbers indicate length of individual exons and introns. Arrows indicate the position of the primers used for RT-PCR analysis. \textbf{B.} Constitutive expression of the resistant (\textit{RCT1}) and susceptible (\textit{rct1}) alleles in Jemalong A17 and F83005.5, respectively. Primers used were F1 and R1 that span the intron 2. \textbf{C.} Alternative splicing of intron 4 of \textit{RCT1} alleles in Jemalong A17 and F83005.5. The \textit{M. truncatula Actin} gene was used as a control. Primers used were F2 and R2 that span the intron 3 and intron 4. At=alternative transcript that retained intron 4; Rt=regular transcript with intron 4 spliced out; Het=heteroduplex resulting from RT-PCR of alternatively spliced mRNAs of \textit{RCT1}. \textbf{D.} Expression and alternative splicing of the transgene \textit{RCT1} in alfalfa. The primers used were F3 and R3 from the 5’- and 3’-UTR regions, respectively. This primer pair only amplified the transgene \textit{RCT1} but not homologs of alfalfa.
Alignment of the \textit{RCT1} genomic sequence with its expressed sequences, TC97262 and TC96909, from the MtGI database revealed that TC97262 and TC96909 can be assembled into a single sequence contig. Interestingly, this contig contains the entire fourth intron of 448 bp. This observation was unlikely due to DNA contamination, because part of sequences within TC96909 and TC97262 were from the same cDNA clones in which the second and third introns were spliced out. RT-PCR using exonic primers spanning the third and fourth introns (primers F2 and R2 as indicated in Figure 3.3A) confirmed the presence of two transcripts, corresponding to the splicing out (\~1.5 kb) and retention of intron 4 (\~2 kb), respectively, based on sequence analysis of the RT-PCR products (Figure 3.3C). The regular (Rt, intron 4 spliced out) and alternative (At, intron 4 retained) transcript appeared to be equally present in the RNA profile of Jemalong A17, based on the semi-quantitative RT-PCR analysis (Figure 3.3C). The expression of the alternatively spliced transcript of \textit{RCT1} was not obviously regulated by the pathogen infection. The similar expression pattern was also observed for the \textit{RCT1} transgene in transgenic alfalfa plants (Figure 3.3D). Two amplification products (\~3.9 kb and \~3.5kb) were derived in transgenic alfalfa with primers F3 and R3 as indicated in Figure 3.3A. It is noteworthy that a weak band of \~1.8 kb was detected at high cycle numbers (Figure 3.3C), and sequence analysis indicated that this product was a heteroduplex formed by RT-PCR of alternatively spliced mRNAs of \textit{RCT1}. The heteroduplex DNA strand adopts an \textit{Ω}-like conformation (Eckhart et al., 1999) in which the sequence corresponding to intron 4 forms a single-stranded loop (Figure 3.4). Overall, though 3 bands appeared on the gel (Figure 3.3C), only 2 kinds of single strand
mRNA were transcribed. DNA heteroduplexes and heteroduplexes complexes were also previously reported following RT-PCR of spliced mRNAs (Eckhart et al., 1999).

![Heteroduplex Diagram](image)

**Figure 3.4.** Schematic representation of the At and Rt involved in heteroduplex formation and of the two types of heteroduplexes, adapted from Eckhart et al. (1999). DNA strands are represented by lines with arrowheads at their 3’ end to indicate their orientation. The segment corresponding to intron 4 is shown as a bold line. Interactions between heteroduplexes that mediated the formation of a heteroduplex-duplex are indicated by double-headed arrows.

cDNA sequences from 5’- and 3’-untranslated regions (UTR) of *RCT1* were obtained through 5’ and 3’ RACE (rapid amplification of cDNA ends) experiments. The 5’-UTR was identified to be 188-bp long, and no intron was detected based on alignment with the genomic sequence (Figure 3.5). In contrast, we obtained three transcript variants from the 3’UTR of 721, 734, and 801 bp, respectively (Figure 3.6). Alignment of the 721-bp fragment with genomic sequence revealed three additional introns of 203, 95, and 80-bp, respectively. The 801-bp fragment resulted from the retention of the 80-bp intron, whereas the 734-bp fragment was due to the splicing out only 67 bp of the 80-bp intron.
but retaining the 13-bp at the 3’-direction. These results document multiple transcript variants present in the *RCT1* transcript profile, with added complexity possible if alternative splicing events in the coding and non-coding regions occur independently.

```
GATTGAATAAATTTACAAATTATTCAATAATCTGACTGATCAGTGCAG
ACTGCTGCATTATTTACTACTTTTCAAGTAACAGGAGCATCTCTTC
CTTCTTTTCTTTTTCTTCTTGCAGTGGTTTCGAAACCACAATCATTCATCC
```

**Figure 3.5.** 5’ UTR sequence of resistant *RCT1* allele in Jemalong *A17*. No intron was found in this region.
Alignment of transcript variants (tv) from the 3’ UTR region of the RCT1 allele in Jemalong A17. tv1=the regular transcript with an 80-bp intron spliced out; tv2=the transcript with only 67 bp of the 80-bp intron spliced out; tv3=the transcript with retention of the 80-bp intron.

**Figure 3.6.**
Structure of the RCT1 protein(s)  The fully processed *RCT1* (no intron) is predicted to encode a protein of 1098 amino acids with a molecular weight of ~125kDa, consisting of an N-terminal TIR domain, a centrally located NBS domain with typical conserved motifs (Meyers et al., 1999), 7 degenerate LRRs at C-terminal to the NBS domain (Figure 3.7). The extreme C-terminus of RCT1 is highly conserved with members of TIR-NBS-LRR genes in *M. truncatula* but less conserved between species. The alternatively spliced transcript results in a shift in the reading frame and is predicted to encode a truncated protein of 936 amino acids with a molecular weight of ~106kDa. The first 920 amino acids of the truncated protein are identical to those of the full-length protein. Nevertheless, the truncated protein consists of the entire portion of the TIR, NBS, and LRR domains but lacks the C-terminal domain of the full-length RCT1 protein (Figure 3.7).

Figure 3.7. Structure of the RCT1 protein(s). The conserved motifs within the TIR and NBS domains are underlined. The seven predicted LRRs are highlighted in red color. The alternatively spliced transcript is predicted to encode a truncated protein lacking 178 amino acids in the N-terminal domain (green color).
Expression-level Polymorphisms between Resistant and Susceptible Alleles To explore the molecular nature of resistance and susceptible alleles, sequence analysis of the rct1 allele from the susceptible genotype F83005.5 was carried out, and the expression profile was characterized (Figure 3.3B). RT-PCR experiment using RCT1-specific primers revealed that the fully spliced rct1 allele was constitutively transcribed in the susceptible parent F83005.5 (Figure 3.3B). By contrast, the expression of the alternatively spliced transcript that retains intron 4 was very low and undetectable at 25 cycles of RT-PCR. The correlation between an absence of alternative splicing and disease susceptibility was further examined by sequencing of RCT1 alleles from 12 additional genotypes of M. truncatula (9 resistant and 3 susceptible). As shown in Figure 3.8, the alternative transcript isoform was common to all resistant genotypes, but undetectable or very low in susceptible genotypes. Thus, alternative splicing of RCT1 is correlated with disease resistance to C. trifolii.

Figure 3.8. Expression analysis of additional resistant and susceptible alleles in M. truncatula (the same as the panel C). Primers used were F2 and R2 that span the intron 3 and intron 4 as indicated in Figure 3.3A. At=alternative transcript that retained intron 4; Rt=regular transcript with intron 4 spliced out; Het=heteroduplex resulting from RT-PCR of alternatively spliced mRNAs of RCT1.
Sequence-level Polymorphisms between Resistant and Susceptible Alleles

Sequence polymorphisms occurred in both the coding sequence and the UTR region between parental lines. cDNA coding sequences of the rct1 allele from F83005.5 was also obtained in this study. A total of 27 single nucleotide polymorphisms (SNPs) were detected when aligned with the 3294-bp coding sequence of the RCT1 allele from Jemalong A17. These included a two-base-pair deletion in the first exon (Figure 3.9). The 2-bp deletion leads to a shift of the open reading frame (ORF) and an immediate stop codon. If we assume that the translation of the rct1 allele uses the same ORF and starts with the next available start codon, the rct1 allele would encode an NBS-LRR protein lacking the first 115 amino acids of the TIR domain. However, sequencing additional M. truncatula genotypes revealed this deletion appears to be F83005.5 allele specific and does not represent a universal mechanism to generate susceptible alleles (Figure 3.10).
JEMALONG A17    ATGGGAAGGGAAATCATTCGTGAGAAATCACCAATGGAGCCTGAGGAACGTAGTAGGTTG
F83005.5      ATGGGAAGGGAAATCATTCGTGAGAAATCACCAATGGAGCCTGAGGAACGTAGTAGGTTG
*********************************************************
JEMALONG A17    TGGTTTCATGATGATGTGCTTGATGTATTGTCAGAACATACTGGAACAAAAGCTGTTGAG
F83005.5      TGGTTTCATGATGATGTGCTTGATGTATTGTCAGAACATACTGGAACAAAAGCTGTTGAG
*********************************************************
JEMALONG A17    GGACTGACTTTGAAGATGCCATGTCATAGTGCACAACGATTTAGTACTAAAACATTTGAG
F83005.5      GGACTGACTTTGAAGATGCCATGTCATAGTGCACAACGATTTAGTACTAAAACATTTGAG
*********************************************************
JEMALONG A17    AACATGAAGAAACTCAGATT
GCTGCAACTTTCTGGTGTACAACTTGATGGAGATTTTAAA
F83005.5      AACATGAAGAAACTCAGATT
GCTGCAACTTTCTGGTGTACAACTTGATGGAGATTTTAAA
*********************************************************
JEMALONG A17    TATATTTCAAGAAATTTAAAATGGCTGCACTGGAATGGATTTCCTTTAAGATGCATACCT
F83005.5      TATATTTCAAGAAATTTAAAATGGCTGCACTGGAATGGATTTCCTTTAAGATGCATACCT
*********************************************************
JEMALONG A17    TCAAACTTCTTATCACAAGAATAATAGTATCTTCCATGTAGTTAGAAACAGCATGCTAAAATCT
F83005.5      TCAAACTTCTTATCACAAGAATAATAGTATCTTCCATGTAGTTAGAAACAGCATGCTAAAATCT
*********************************************************
JEMALONG A17    GTGTGGAAAGAGATTCAGAGGATGGAGCAGCTGAAGATTCTAAATCTTAGTCATTCTCAT
F83005.5      GTGTGGAAAGAGATTCAGAGGATGGAGCAGCTGAAGATTCTAAATCTTAGTCATTCTCAT
*********************************************************
JEMALONG A17    CATTTGACACAGACCCCTGACTTTTCATACTTGCCTAATCTTGAAAAGCTAGTGCTCGAA
F83005.5      CATTTGACACAGACCCCTGACTTTTCATACTTGCCTAATCTTGAAAAGCTAGTGCTCGAA
*********************************************************
JEMALONG A17    GATTGCCCAAGGCTGTCT
CAGGTTTCCCATAGCATTGGACATCTCAAAAAAGTTGTTTTG
F83005.5      GATTGCCCAAGGCTGTCT
GAGGTTTCCCATAGCATTGGACATCTCAAAAAAGTTGTTTTG
*********************************************************
JEMALONG A17    ATAAATTTGAAAGATTGTATTAGCCTTTGTAGCCTTCCAAGAAACATCTATACGTTGAAA
F83005.5      ATAAATTTGAAAGATTGTATTAGCCTTTGTAGCCTTCCAAGAAACATCTATACGTTGAAA
*********************************************************
JEMALONG A17    ACTCTGAA
TACTCTCATTCTATCGG
GGATGTTTAATGATTGACAAGTTGGAAGAGGACTTG
F83005.5      ACTCTGAA
TACTCTCATTCTATCGG
GGATGTTTAATGATTGACAAGTTGGAAGAGGACTTG
*********************************************************
JEMALONG A17    GAACAAATGGAATCTTTAACCACCCTGATTGCAAATAA
TACTGGTATAACAAAAGTTCCC
F83005.5      GAACAAATGGAATCTTTAACCACCCTGATTGCAAATAA
TACTGGTATAACAAAAGTTCCC
*********************************************************
JEMALONG A17    TTTTCA
TTTAGTAAGGTCAAAAAGCATTGGATTTATTTCTCTGTGTGGATATGAAGGATTC
F83005.5      TTTTCA
TTTAGTAAGGTCAAAAAGCATTGGATTTATTTCTCTGTGTGGATATGAAGGATTC
*********************************************************
JEMALONG A17    TCACGTGATGTGTTTCCTTCTATCATTT
GCTTTTCAAACAGCTTCTCACATGTCATCCCTTGTGTCTTTAGAGGCATCAACTTGTATT
F83005.5      TCACGTGATGTGTTTCCTTCTATCATTT
GCTTTTCAAACAGCTTCTCACATGTCATCCCTTGTGTCTTTAGAGGCATCAACTTGTATT
*********************************************************
JEMALONG A17    GCAAGTTCTATGGAATTGGAATCAACTGCAACTAC
ATCACAAGTACCAGATGTGAATTCA
F83005.5      GCAAGTTCTATGGAATTGGAATCAACTGCAACTAC
ATCACAAGTACCAGATGTGAATTCA
*********************************************************
JEMALONG A17    TTCTCAATTAGTGAAAGCTTGGATTATTATTCTCTGTGTTGGAATATTGAGGATTC
TTTCAAGTAGTGAAGTGCAAAAGCATGTTATTATTTCTCTGTGTTGGAATATTGAGGATTC
F83005.5      TTCTCAATTAGTGAAAGCTTGGATTATTATTCTCTGTGTTGGAATATTGAGGATTC
TTTCAAGTAGTGAAGTGCAAAAGCATGTTATTATTTCTCTGTGTTGGAATATTGAGGATTC
*********************************************************
JEMALONG A17    TCACGTGATGTGTTTCCTTCTATCATTT
GCTTTTCAAACAGCTTCTCACATGTCATCCCTTGTGTCTTTAGAGGCATCAACTTGTATT
F83005.5      TCACGTGATGTGTTTCCTTCTATCATTT
GCTTTTCAAACAGCTTCTCACATGTCATCCCTTGTGTCTTTAGAGGCATCAACTTGTATT
*********************************************************
JEMALONG A17    GCTTTTCAAACAGCTTCTCACATGTCATCCCTTGTGTCTTTAGAGGCATCAACTTGTATT
F83005.5      GCTTTTCAAACAGCTTCTCACATGTCATCCCTTGTGTCTTTAGAGGCATCAACTTGTATT
*********************************************************
JEMALONG A17    TCCCAGATCATCTACATCTATTATCTGATTGCTCTTCCAAAGCTTGGCTTGAGCA
A
TCCCAGATCATCTACATCTATTATCTGATTGCTCTTCCAAAGCTTGGCTTGAGCA
G
F83005.5      TCCCAGATCATCTACATCTATTATCTGATTGCTCTTCCAAAGCTTGGCTTGAGCA
A
TCCCAGATCATCTACATCTATTATCTGATTGCTCTTCCAAAGCTTGGCTTGAGCA
G
*********************************************************
JEMALONG A17    TGTTGATCAGAACATTCAACTTACAACAAAGTGCAACTGGAATTTAGGTA
TGTTGATCAGAACATTCAACTTACAACAAAGTGCAACTGGAATTTAGGTA
F83005.5      TGTTGATCAGAACATTCAACTTACAACAAAGTGCAACTGGAATTTAGGTA
TGTTGATCAGAACATTCAACTTACAACAAAGTGCAACTGGAATTTAGGTA
*********************************************************
JEMALONG A17    GCAAGTTCTATGGAATTGGAATCAACTGCAACTAC
ATCACAAGTACCAGATGTGAATTCA
F83005.5      GCAAGTTCTATGGAATTGGAATCAACTGCAACTAC
ATCACAAGTACCAGATGTGAATTCA
*********************************************************
Figure 3.9. Alignment of coding sequences of resistant and susceptible alleles in Jemalong A17 and F83005.5, respectively. Polymorphisms were indicated in red.

Sequence polymorphisms were also detected in the 5’- and 3’-UTRs (in Appendix Figure A.1-1.3). Although only one SNP was identified, the 5’-UTR of the \( \text{rct1} \) allele contains a 48-bp fragment deletion resulting from the deletion of genomic sequence (Appendix Figure A.1). The 3’-UTR region of the \( \text{rct1} \) allele shares the same structure as that of the \( \text{RCT1} \) allele which carries three introns of 203, 95, and 80 bp, respectively. Alternative splicing of the 80-bp intron was also detected for the \( \text{rct1} \) allele (Appendix Figure A.2). However, the 299-bp in 3’ end of \( \text{RCT1} \) cDNAs and the 119-bp in 3’ end of
the *rct1* cDNAs are not shared (Appendix Figure A.3). The 299-bp sequence is present in the genomic sequence of *RCT1*, but the 119-bp fragment is absent in the *RCT1* region in Jemalong A17. Long range PCR amplification and sequencing F83005.5 genomic DNA containing the 119-bp fragment revealed a ~10-kb insertion has occurred in 3’ UTR of *rct1* in F83005.5 (Appendix Figure A.4). The F83005.5-specific 199-bp in 3’ UTR is exactly included in the start region of this 10-kb insertion.

<table>
<thead>
<tr>
<th>Allele</th>
<th>Sequence 1</th>
<th>Sequence 2</th>
<th>Sequence 3</th>
<th>Sequence 4</th>
<th>Sequence 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>A17 (R)</td>
<td>AACAAAAAAG TATGACGTGT TTTTGAGTTT CCGAGGAGAA GACACTCGCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEPHI (R)</td>
<td>AACAAAAAAG TATGACGTGT TTTTGAGTTT CCGAGGAGAA GACACTCGCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYPRUS (R)</td>
<td>AACAAAAAAG TATGACGTGT TTTTGAGTTT CCGAGGAGAA GACACTCGCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DZA105 (R)</td>
<td>AACAAAAAAG TATGACGTGT TTTTGAGTTT CCGAGGAGAA GACACTCGCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRC064 (R)</td>
<td>AACAAAAAAG TATGACGTGT TTTTGAGTTT CCGAGGAGAA GACACTCGCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BOUNG (R)</td>
<td>AACAAAAAAG TATGACGTGT TTTTGAGTTT CCGAGGAGAA GACACTCGCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DZA315 (R)</td>
<td>AACAAAAAAG TATGACGTGT TTTTGAGTTT CCGAGGAGAA GACACTCGCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DZA220 (R)</td>
<td>AACAAAAAAG TATGACGTGT TTTTGAGTTT CCGAGGAGAA GACACTCGCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HARBINGER (S)</td>
<td>AACAAAAAAG TATGACGTGT TTTTGAGTTT CCGAGGAGAA GACACTCGCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DZA220 (S)</td>
<td>AACAAAAAAG TATGACGTGT TTTTGAGTTT CCGAGGAGAA GACACTCGCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F20061 (S)</td>
<td>AACAAAAAAG TATGACGTGT TTTTGAGTTT CCGAGGAGAA GACACTCGCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F83005.5 (S)</td>
<td>AACAAAAAAG TATGACGTGT TTTTGAGTTT CCGAGGAGAA GACACTCGCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3.10.** Part of alignment of *RCT1* (*rct1*) alleles from 12 ecotypes. 2bp deletion of *rctl* allele of F3005.5, as indicated in red, was not present in other lines. R means resistant to *C. trifolii* 1; S means susceptible to *C. trifolii* 1.

**Discussion**

*Colletotrichum* spp. are one of the most widespread and important disease causing fungi of plants worldwide. The genus contains over 35 species which cause anthracnose or blight on a wide range of temperate and tropical plants, including grain and pasture legumes, cereals, and fruits (Bailey and Jeger, 1992). During colonization of plant hosts, many species of *Colletotrichum*, including *C. trifolii*, utilize a hemibiotrophic infection strategy, in which the pathogen initially develops inside living host cells before switching to a destructive necrotrophic mode of infection (O’Connell et al., 1993). To date, no other resistance gene against the 13 genus *Colletotrichum* has been published in any plant.
hosts. Thus, our work presented here will contribute significantly to our understanding of molecular mechanisms underlying host resistance against the hemibiotrophic fungal pathogens in the genus *Colletotrichum*.

The model legume *M. truncatula* is native to the Mediterranean basin and has long been cultivated as winter forage in Australia. The past decade has seen the development of abundant genetic and genomic tools for this model species, which has greatly facilitated our understanding of legume genomics and biology (Zhu et al., 2005). The value of this model system has been enhanced by its close relationship with crop legumes, which is reflected in similar genome structures and conserved phenotypes such as legume-rhizobial symbiosis (Zhu et al., 2005). Of crop legumes, alfalfa has become an immediate beneficiary from the study of the *M. truncatula* genomics, not only because alfalfa is a close relative of *M. truncatula*, but also because alfalfa itself is not amenable to genetic analysis. In addition to a focus on symbiotic plant-microbe interactions, significant efforts have taken advantage of *M. truncatula* as a model system to characterize legume-pathogen interactions (Tivoli et al., 2006). Importantly, most alfalfa pathogens also are pathogens of *M. truncatula*, leading to two key predictions: (1) that *M. truncatula* can serve as a tool to clone disease resistance genes for common pathogens of alfalfa, and (2) that functional disease resistance will be maintained when genes are moved across species boundaries by transgenic approaches.

Here we validate these predictions by isolating and characterizing the *M. truncatula R* gene *RCT1*. The *RCT1* locus in *M. truncatula* that confers resistance to *C. trifolii* was finely mapped and isolated. Genetic linkage analysis in *M. truncatula* and transgenic tests performed in alfalfa indicated that *RCT1* confers broad-spectrum
resistance to the three known races of *C. trifolii*. Broad-spectrum disease resistance conferred by NBS-LRR type *R* genes has been reported from other plant hosts. For example, the *RB* and *RPI* genes from wild potato species confer broad-spectrum resistance to nearly all known races of the late blight pathogen *Phytophthora infestans* in cultivated potato (Song et al., 2003; van der Vossen et al., 2003; van der Vossen et al., 2005). In alfalfa, resistance to the three races of *C. trifolii* was reported to be controlled by two independent dominant genes, namely *An1* and *An2* (Elgin and Ostazeski, 1985; Mackie et al., 2003). *An1* confers resistance to race 1 and likely race 4, whereas *An2* confers resistance to races 1 and 2. Thus, only plants carrying genes *An1* and *An2* are resistant to all three races (Mackie et al., 2003). By contrast, we demonstrate that *M. truncatula RCT1* confers broad-spectrum anthracnose resistance in cultivated alfalfa. These results highlight a fundamental difference between these two species and demonstrate the potential of using *M. truncatula* genes for genetic improvement of alfalfa.

Based on extensive conserved synteny and highly similar NBS-LRR sequences between *M. truncatula* and alfalfa (Zhu et al., 2002; Choi et al., 2004b), one might predict that many disease resistance genes identified in *M. truncatula* will be conserved and located in syntenic regions of *M. sativa*. In the case of anthracnose, *Medicago* (*Medicago* spp.) and clovers (*Trifolium* spp.) share the same races of *C. trifolii* as pathogens, suggesting that anthracnose resistance may have originated before speciation within the *Trifolieae* tribe. Under such a scenario, with pressure from a common pathogen gene pool, *RCT1* might represent a slow-evolving *R* gene (Kuang et al., 2004). It is interesting, therefore, that the genetic basis of resistance to *C. trifolii* differs between *M. truncatula* and *M. sativa*. In particular, resistance to races 1, 2 and 4 of *C. trifolii* is
determined by two unlinked genes in tetraploid alfalfa, whereas only a single gene
confers resistance to all three races in diploid *M. truncatula*. To the extent that broad
spectrum resistance of *RTC1* is ancestral to *Medicago* spp, then *RTC1* function may have
been partitioned between homologous genes during the evolution of the tetraploid
genome. Further work is needed to address the evolutionary relationship between *RCT1*
in *M. truncatula* and the *An1* and *An2* genes in cultivated alfalfa, and the possible impact
of polyploidy.

It was demonstrated that *RCT1* was constitutively expressed and alternatively
spliced. Constitutive expression indicates that *RTC1* is expressed in the absence of the
corresponding *Avr*-expressing pathogen, similar to other *R* genes that function in
pathogen surveillance (Hammond-Kosack and Jones, 1997). Alternative splicing has
been frequently detected for TIR-NBS-LRR type *R* genes, such as flax *L6* and *M*,
tobacco *N* and the *Arabidopsis RPS4* genes (Lawrence et al., 1995; Dinesh-Kumar and
Baker, 2000; Zhang and Gassmann, 2003; Schmidt et al., 2007; Tan et al., 2007; Zhang
and Gassmann, 2007). Interestingly, alternative transcripts of the tobacco *N* and the
*Arabidopsis RPS4* genes are both required for complete disease resistance (Dinesh-
Kumar and Baker, 2000; Zhang and Gassmann, 2007). Furthermore, the expression of
alternatively spliced transcripts of the *N* and *RPS4* genes was both upregulated by
pathogen infection (Dinesh-Kumar and Baker, 2000; Zhang and Gassmann, 2007). In
contrast to *N* and *RPS4*, an intronless flax rust resistance *L6* gene that fails to produce
alternative transcripts expresses resistance indistinguishable from that of the wild-type
gene (Ayliffe et al., 1999). However, because another flax rust resistance gene *M* which
is homologous to *L* is also alternatively spliced, and no flax line is available that lacks
other L alleles or genes at the M locus, those authors could not rule out possible functions provided by alternatively spliced transcripts of other L alleles or genes at the M locus in the transgenic plants (Dinesh-Kumar and Baker, 2000). All PCR-based screens aimed at detecting splice variants of known TIR-NBS-LRR gene were restricted to intron 2 and intron3. The alternative transcripts generally possess premature termination codons and thus encode putative truncated proteins lacking the LRR and/or C-terminal domains (Jordan et al., 2002). Evaluation of splice variants derived from N and RPS4, reveals that splice derivatives exert a positive rather than a negative regulatory function. Thus, LRR domain of plant R protein is speculated to have a negative regulatory function in the absence of the Avr elicitor (Jordan et al., 2002).

In terms of RCT1, which shares similar gene structure with those of the N and RPS4 genes, alternative splicing was detected at both coding and 3’-UTR regions. Thus, there are likely multiple transcript variants present in the RCT1 expression profiles. In contrast to the tobacco N and the Arabidopsis RPS4 genes for which alternative splicing involves intron 2 and/or intron 3, alternative splicing of RCT1 in the coding region affects the retention of intron 4. The alternatively spliced transcript is predicted to encode a truncated protein consisting of the entire portion of the TIR, NBS, and LRR domains but lacks the C-terminal domain of the full-length RCT1 protein. It is unknown if the alternative splicing events in the coding and non-coding regions are correlated. It is also unclear whether the alternatively spliced transcripts are required for the functionality of RCT1. Nevertheless, expression-level polymorphisms were detected for the alternatively spliced transcript involving intron 4 between the resistant and susceptible alleles. This
observation suggests that alternative splicing of *RCT1* may be required for its functionality.

Sequence comparison between the coding regions of resistant (Jemalong A17) and susceptible (F83005.5) alleles identified 27 single nucleotide polymorphisms (SNPs), including a 2-bp deletion in the first exon. The 2-bp deletion changes the open reading frame and leads to an immediate stop codon. Thus, this deletion presumably abolishes the *RCT1* function, resulting in the susceptible allele in F83005.5. However, this deletion appears to be unique for the F83005.5 allele and does not represent a conserved mechanism to generate susceptible alleles in *M. truncatula*, because sequencing additional susceptible alleles at this site did not detect such a deletion. In fact, sequence polymorphisms are more significant in the 5’- and 3’-UTRs. For example, the 5’-UTR of the *rct1* allele in F83005.5 contains a 48-bp fragment deletion resulting from the deletion of genomic sequence. The 3’-UTR regions are even more diversified at the poly-A site because of a ~10-kb insertion in *rct1* allele of genomic DNA. Previous experimental data (Dinesh-Kumar and Baker, 2000) suggested that 3’ genomic sequence (GS) plays a crucial role in the regulation of the tobacco *N* gene alternative splicing. After the 3’ GS of *N* was substituted by another sequence, no or significantly reduced alternative transcripts were produced. The 3’ GS represents a distant regulatory element which is necessary for the generation or stability of the alternative splicing product (Dinesh-Kumar and Baker, 2000). The influence 5’ GS on alternative splicing of pre-mRNA has not been reported to date. Taken together, our data suggest that both alternative splicing- and sequence-level polymorphisms may explain the molecular mechanisms underlying the evolution of resistant and susceptible alleles of *RCT1*. 
Development of disease-resistant cultivars is effective to control diseases if sufficient genetic variation for resistance is available. When sources of resistance are limited, breeders will refer to the secondary gene pool for species to hybridize with the cultivated species. Molecular technique allows them to transfer resistance between much more distantly related species. In the Solanaceae, several R genes been investigated to confer resistance reactions to pathogens carrying the appropriate Avr gene when transferred to other Solanaceous species. Transferring tomato Cf-9 to tobacco and potato, pepper Bs2 to tomato, tomato Pto to tobacco, tobacco N to tomato, demonstrated that Avr-dependent R protein-triggered signaling cascades are conserved among Solanaceous species (Thilmony et al., 1995; Whitham et al., 1996; Hammond-Kosack et al., 1998; Tai et al., 1999). RCT1 cloned in this study confers broad spectrum resistance to C. trifolii races. Its biological activity is still retained after introduction to alfalfa. Interspecific transfer of an R gene cloned from M. truncatula will provide a novel resistance resource and enrich the crop genetic background. The direct application of beneficial genes to crop legumes will definitely broaden the genetic utility of M. truncatula and highlight its potential for crop improvement as model plant in the future.

Copyright © Shengming Yang 2008
Appendix

**Figure A.1.** Alignment of 5’ UTR region of the resistant and susceptible alleles in Jemalong A17 and F83005.5, respectively.
Figure A.2. Alignment of transcript variants (tv) from the 3’ UTR region of the rct1 allele in F83.005.5.

tv1=the regular transcript with an 80-bp intron spliced out; tv2=the transcript with retention of the 80-bp intron.
Figure A.3. Sequence polymorphisms between resistant and susceptible alleles in Jemalong A17 and F83005.5, respectively. The 299-bp 3’ end of RCT1 cDNAs and the 199-bp 3’ end of the rct1 cDNAs are not shared and thus not included in the alignment.
**Figure A.4.** Size polymorphism in 3’ UTR region of genomic DNA between resistant parental line A17 and susceptible paternal line F83005.5. With same primer pair of F4 and R4 within 3’ UTR indicated in Fig 3.3, 0.47-kb product was amplified in A17 (lane 1); however, the amplification product is ~10-kb in F83005.5 (lane 2). Size marker was indicated on the left.
REFERENCES


truncatula for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. Mol Plant Microbe Interact 14, 695-700.


Colletotrichum trifolii and further analysis of inheritance of resistance in autotetraploid lucerne. Theor Appl Genet 114, 1417-1426.


VITA

SHENGMING YANG

Personal Information
Date of birth 09/06/1975, Shandong, China

Education
M.S.,
Department of Plant Physiology
Shandong Agricultural University, China
July, 2001

B.S.,
College of Agriculture
Shandong Agricultural University, China
July, 1998

Publications

