CREATING RUST RESISTANCE IN WHEAT VIA MODIFICATION OF HOST GENES

by

Bernard Mensah Nyamesorto

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Plant Genetics

MONTANA STATE UNIVERSITY
Bozeman, Montana

May 2020

©COPYRIGHT

by

Bernard Mensah Nyamesorto

2020

All Rights Reserved

DEDICATION

To my true heroes, my late Mom (Mornica Akpanga) and Dad (Amenorpe Kpataku) who through their toil and sacrifice gave me the gift of education that they did not have and to my siblings Beauty (Late), Ema, Yohanes and Clement Nyamesorto for their love and support.

ACKNOWLEDGEMENTS

I would like to express my appreciation to all and sundry who help me throughout my educational life for your council, inspiration support and unrelenting believe in me.

First, my profound gratitude to my principal advisor Dr. Li Huang for her excellent supervision and teaching. I recall when my first attempt to wear a glove failed, I told her I do not have molecular biology skills because I did not like it. She said to me, "do not worry, I will teach you". Whatever I have achieved in the course of this study, I owe it to her guidance and leadership.

My special thanks to Dr. Mike Giroux for his support, recommendations and suggestions for my study. Mike's teaching skills also imparted me after I served as his teaching assistant.

My special appreciation to Dr. Phil Bruckner and Dr. Jennifer Lachowiec serving on my committee and giving very useful suggestions for this study.

Permit me to also acknowledge Dr. Hongtao Zhang (Lab manager) and Mr. Mina Botros (Lab mate) for their help in the lab. Also, thanks to Dr. Agyemang Danquah for recommending me to Li and to my number one cheer leader from Junior high school till date Mr. Gabriel Azangodo.

Finally, I am grateful to NSF BREAD program (Grant No. IOS-0965429) and Montana State University for funding my research and PhD degree.

TABLE OF CONTENTS

1. GENERAL INTRODUCTION....1

Evolution of Cultivated Bread Wheat....1
Wheat Rust Diseases....3
Wheat-Rust Interaction....5
Wheat Rust Resistance Breeding....7
Gene Editing for Rust Resistance....8
References....12

2. WHEAT *MYC4* TRANSCRIPTIONAL FACTOR GENE MODIFICATION ENHANCED HOST RESISTANCE AGAINST RUST PATHOGENS....17

Contribution of Author and Co-Authors....17
Manuscript Information....18
Abstract....19
Introduction....20
Materials and Methods....22
Plant and Pathogen Materials....22
Plant Growth Conditions....22
Rust Pathogen Inoculation and Assessment....23
Sample Collection and Treatment....24
RNA and DNA Extraction....24
qRT-PCR, Conventional PCR and KASP Assay....25
Barley Stripe Mosaic Virus-Induced Gene Silencing Assay....26
Mutant Search and Validation....26
Genetic Analysis....27
Pathogenesis-related (PR) Genes Expression....27
Databases and *in silico* Sequence Analysis....27
Statistical Analysis....28
Results....29
Expression Profiles of *TaMYC4* Gene by Real-Time qPCR....29
Conserved Domains of *TaMYC4* Genes in Wheat....30
Silencing of the *TaMYC4* Genes....32
Expression of *TaMYC4* Homeologs during Stripe Rust Infection....34
Identification of *TaMYC4* Mutants and their Response to Rust Pathogens....36
Genetic Analysis of the Mutations....39
Molecular Mechanism of the new Rust Resistance....40
Discussion....41
Wheat *TaMYC4* Gene Aids Rust Pathogens in Host Colonization....41
Cadenza P12S Mutant Showed Race Specific Resistance....46

TABLE OF CONTENTS CONTINUED

An Approach to Create New Resistant Germplasm 47
Acknowledgement 49
References 50

3. A NEW MODE OF NPR1 ACTION VIA AN NB-ARC-NPR1
FUSION PROTEIN REGULATES DEFENSE RESPONSE
TO STEM RUST PATHOGEN IN WHEAT 56

Contributions of Authors and Co-Authors 56
Manuscript Information 58
Summary 59
Introduction 60
Materials and Methods 63
Plant Materials 63
Sequence Analysis 63
Plant Growth Conditions 64
Pathogens 64
Pathogen Inoculation 64
Stripe Rust Inoculation and Assessment 65
Virus Induced Gene Silencing 66
Gene Expression Analysis by RT-qPCR and RNA-seq 67
Mutant Screen 68
SA/JA Level Analysis with the LC-MS 69
Results 70
Identification of Wheat *NPR1* Homologs 70
Silencing of *TaNPR1* Genes 74
Expression of *TaG7NPR1* Post Rust Inoculations 77
Identification of *Ta7ANPR1* Mutants 79
Levels of SA/JA and PR Genes when *TaG3NPR1* Genes
were Downregulated 82
Transcript Abundance of PR Genes when *Ta7ANPR1* Gene
was Knocked out 83
Discussion 84
A requirement of NPR1 in Wheat Defense Response
to *Puccinia* is Selective 84
NPR1 Proteins with Integrated Domains in the Wheat Genome 86
The Negative Regulation of Ta7ANBS-NPR1 is Specific 87
Alternative Splicing and Innate Immunity 88
Acknowledgements 89
References 90

TABLE OF CONTENTS CONTINUED

4. WHEAT *MYB21* TRANSCRIPTIONAL FACTOR GENE MODIFICATION ENHANCED HOST RESISTANCE AGAINST RUST PATHOGENS 95

Introduction 95
Materials and Methods 97
Results 98
Expression Profiles of *TaMyb21* in a Pair of NIL 98
TaMyb21 Knockdown Enhanced Chinese Spring Resistance to Stem Rust 99
TaMyb21-3B Knockout and Confirmation 100
Infection Type of Cadenza S91L Mutant 102
Discussion 104
Acknowledgements 105
References 106

5. CONCLUSION 109

CUMULATIVE REFERENCES 111

APPENDICES 129

APPENDIX A: Supporting Information for Chapter Two 130
File 2.S1: Alignment of *TaMYC4* homeologs in ClustalW 131
File 2.S2: *TaMYC4* homeologs protein sequences for domain prediction 134
Figure 2.S1: Relative expression of *TaMYC4* 138
Figure 2.S2: FastQC RNA-seq quality check output 139
Figure 2.S3: Basal expressions of PR genes 140
Figure 2.S4: Agronomic Traits of P12S 141
Table 2.S1: Summary of primers 142
APPENDIX B: Supporting Information for Chapter Three 143
File 3.S1: Amino acid sequences of proteins 144
File 3.S2: The alignment of G3NPR1 like homologs with clustalw2 151
Figure 3.S1: Specificity of primers for screening mutations on Ta7ANPR1 165
Figure 3.S2: Transcripts of Ta7ANPR1, Ta4ANPR1 and Ta7DNPR1 in the RNAseq data generated from *Pgt* and *Pst* inoculated wheat 166

APPENDICES CONTINUED

Figure 3.S3: Transcripts of Ta7ANPR1, Ta4ANPR1 and Ta7DNPR1 in the RNAseq data generated from *Pst* inoculated wheat....167
Figure 3.S4: Reverse transcription amplified cDNAs from total RNAs extracted from CS under no-stressed condition, stripe rust inoculated or BSMV inoculated condition....168
Figure 3.S5: Detection of two transcripts of *Ta7ANPR1* in the RNAseq data generated from *Lr47* near-isogenic lines post leaf rust inoculation at 6 time points....169
Figure 3.S6: Transcript of two isoforms of Ta7ANPR1 in Cadenza during the time course of Pgt *TPMKC* infection....170
Figure 3.S7: Transcript abundances....171
Figure 3.S8: Transcript abundances of six copies of NPR1 on group 3 homeologous chromosomes during a time-course study....172
Figure 3.S9: JA concentrations and PR genes expression....173
Figure 3.S10: JA concentrations and PR genes expression....174
Figure 3.S11: Transcript abundances of five *PR* genes in the wild types and mutants without pathogen inoculation....175
Figure 3.S12: Transcript abundances of two PR1 (a) and PR10 (b) genes during a time-course study of post-Pgt inoculation....176
Table 3.S1: Sequences of the Oligos used in silencing....177
Table 3.S2: Sequences of the primers used for quantitative real-time PCR....178
Table 3.S3: Genes silenced by each constructor....179

LIST OF TABLES

Table Page

2.1. Summary of *TaMYC4-1BL* Cadenza mutant identified from wheat TILLING website....37

2.2. Rust evaluations of Cadenza P12S mutant and wildtype cadenza....38

3.1. qRT-PCR analysis of TaNPR1-like gene expression in silenced plants....77

4.1. Summary of *TaMyb21-3B* Cadenza mutant identified from wheat TILLING website....101

LIST OF FIGURES

Figure Page

2.1. Relative expression of *TaMYC4* TF in Alpowa inoculated with Buffer (control) and stem rust (race QFCSC) 30

2.2. Predicted conserved domains of *TaMYC4* homeologs 31

2.3. BSMV-VIGS of *TaMYC4* in Chinese spring (CS) cultivar 34

2.4. Expression of *TaMYC4* homeologs during stripe rust-Avocet Yr5 (resistant) and Vuka (susceptible) interactions 36

2.5. Infection types of *TaMYC4-1BL* mutant 38

2.6. Relative expressions of PR genes responding to stem rust (race TPMKC) in Wildtype and mutant Cadenza 41

3.1. Genomic blot (Southern) hybridization result from genomic DNAs of seven wheat NT lines and W3SNPR1 as a probe 71

3.2. Arrangement and structures of six *NPR1* homologs in wheat group 3 chromosomes 72

3.3. Arrangement and structures of three *NPR1* homologs and three *NB-CC*-like genes in wheat chromosomes 7A, 7D and 4A(7B)...73

3.4. Infection types of *Sr33* near isogenic lines in Chinese Spring background when challenged with stem rust race QFCSC 14 days post inoculation...76

3.5. Genomic DNA and two transcripts of the *Ta7ANPR1* gene..........................78

3.6. Infection types of *Ta7ANPR1* mutants 81

4.1. Relative expression of *TaMyb21* 99

4.2. BSMV-VIGS of MYC4 in Chinese spring (CS) cultivar 100

4.3. Predicted conserved domain of *TaMyb21-3B* homeolog.............................102

4.4. Infection type of S91L mutant 103

LIST OF FIGURES CONTINUED

4.5. Agronomic Traits of S91L 104

ABSTRACT

A major challenge to resistance breeding in bread wheat (*Triticum aestivum* L.) is limited genetic diversity. The traditional approach to combating this problem is introgression of resistant genes from other closely related species into elite but susceptible cultivars. This strategy is often associated with linkage drag. Moreover, pathogens continue to evolve into different and more virulent forms (races) that overcome these resistant genes in a process called resistance breakdown. A typical example is the outbreak of Ug99, a novel African stem rust pathotype that exhibited virulence against numerous stem rust resistance genes. Creating resistance within wheat's own genome is a panacea to the challenges surrounding the traditional method. Biotrophic plant pathogens such as wheat rusts are known to manipulate host genes as a means of overcoming host defense response and acquiring nutrients. Central to wheat-rust interactions is highly sophisticated immune repertoire consisting of diverse signal perception and intracellular signaling pathways which are regulated by transcriptional regulators and co-factors. Unfortunately, pathogen effector proteins also take advantage of host plant genes (so called pathogen susceptible host genes) including transcriptional mechanisms. Hence editing the genes targeted by these pathogens in wheat is a valuable means of creating host resistance that has been neglected. We conducted these studies to identify host genes targeted by rust pathogens through bioinformatics approaches including transcriptome analysis which showed that wheat NPR1 genes (transcriptional regulator) and MYC4 and MY21 (transcription factors) are negatively involved compatible wheat-rust interactions. Subsequently, when these genes were down regulated in susceptible Chinese Spring using Barley Mosaic Virus Induced Gene Silencing (BSMV) assay, the silenced plants became resistant to rust pathogens. Loss-of-function mutations created in these homeologs via Ethyl methanesulfonate mutagenesis conferred resistant to rust pathogens. Consequently, this study led to the development of new rust resistance germplasms.

CHAPTER ONE

GENERAL INTRODUCTION

Evolution of cultivated bread wheat

Globally, wheat is the second most produced grain and ranks third among the United States of America's field crops in 2018/2019. In order to improve and sustain the production of wheat, it is crucial to understand the evolution of the plant. Bread wheat (*Triticum aestivum* L., 2n=6x=42 chromosomes; genomic formula AABBDD) is an allohexaploid that originated from two interspecific hybridizations involving three diploid ancestors (Kihara, 1954; Sax, 1922). The diploid sub-genomes from *T. urartu* (AA), an unknown close relative of *Aegilops speltoides* (BB), and *Ae. tauschii* (DD) underwent two natural hybridization events to produce bread wheat (International Wheat Genome Sequencing, 2014). The first allopolyploidization event, which occurred a few hundred thousand years ago (S. Huang et al., 2002), was between the A and B genome donors to produce the existing tetraploid emmer wheat (*T. turgidum*; AABB). The second hybridization involving a tetraploid emmer wheat and the D genome donor, resulted in the modern hexaploid bread wheat (AABBDD) (Petersen, Seberg, Yde, & Berthelsen, 2006) and originated with modern agriculture ~10,000 years ago (Salamini, Ozkan, Brandolini, Schafer-Pregl, & Martin, 2002).

The progenitor species of the A genome is diploid wild einkorn wheat *T. urartu* (Dvorak, Diterlizzi, Zhang, & Resta, 1993; Mcfadden & Sears, 1946a). It resembles cultivated bread wheat more extensively than *Ae. speltoides* and *Ae. tauschii*, especially

in the morphology and development of spike and seed (Ling et al., 2013). It was observed that the ratio of nucleotide-binding site, leucine-rich repeat (NBS-LRR) type genes in *T. urartu* was considerably higher than that in barley and suggested that there was a specific expansion of resistance (R) genes in the A genome (Ling et al., 2013). The D genome progenitor is *Ae. tauschii* (McFadden & Sears, 1946b; J. R. Wang et al., 2013), and it is highly repetitive in nature. It was estimated that 84.4% of the genome sequence represent transposable elements (TEs). This leads to frequent recombination errors and causes gene duplications and other structural chromosome changes resulting in fast genome evolution, particularly at the ends of the chromosomes (Luo et al., 2017). The origin of the B sub-genome has been an age-old scientific bone of contention. Early studies have controversially considered goat-grass *Ae. speltoides* (2n = 2x = 14, genome SS) as the closest donor of the wheat B genome (Jenkins, 1929; Pathak, 1940; Sarkar & Stebbins, 1956). Despite contrary arguments (Dvorak, 1972; Jenkins, 1929; Riley, Chapman, & Kimber, 1961; Sarkar & Stebbins, 1956) and uncertainties, recent work provides more evidence that *Ae. speltoides* may be the progenitor of the wheat B genome but may not be considered an exclusive donor of this genome (Miki et al., 2019; W. Zhang et al., 2018). This suggests that wheat B genome might have a polyphyletic origin with multiple ancestors involved, including *Ae. speltoides*.

Overall, it was recently estimated that the domesticated bread wheat genome has a very huge size of ~17 Gb (Shrawat & Armstrong, 2018) and a total of 107,891 high-confidence genes with over 85% repetitive DNA sequences, representing a threefold redundancy due to its hexaploid genome (Appels et al., 2018). Also, because bread wheat

evolution and domestication occurred just about 10,000 years ago, there is very low genetic diversity in bread wheat. These present a huge challenge in breeding efforts to improve vital agronomic traits including disease resistance.

Wheat Rust Diseases

Among the different biotic and abiotic stresses of wheat, rust pathogens are among the most destructive. Severe rust disease can cause greater than 90% yield losses (X. M. Chen, 2014). Wheat rusts include stem rust (caused by *Puccinia graminis* Pers.: Pers f. sp. *tritici* Eriks. E. Henn) (*Pgt*), leaf rust (caused by *Puccinia triticina* Eriks.) (*Pt*) and stripe rust (caused by *Puccinia striiformis* Westend f. sp. *tritici* Eriks.) (*Pst*). These fungi are obligate biotrophic parasites that require living host tissue to grow and reproduce and are very specific to their host plants. As obligates, they form close associations with their hosts and as a result, it takes several days for the disease symptoms to develop. Generally, the three *Puccinia* species have up to five life stages. At stage 0 on alternate host, pycnia bearing one-celled pycniospores and receptive hyphae which both have one nucleus (n) fertilize to produce dikaryotic mycelia growing within the host tissue. The dikaryotic mycelia then produce aecia containing one-celled aeciospores with two nuclei (n+n). This aecial stage is known as the stage I at which the aeciospores infect the host plant (wheat). The aeciospores then produce uredinia-bearing urediniospores (dikaryotic) on the host plant which is referred to as stage II. Urediniospores are capable of asexually reproducing several times on the wheat host leading to severe rust epidemic on the plant. When conditions have become unfavorable for urediniospores production and at later stages of wheat host growth, urediniospores

produces telia-bearing teliospores referred to as stage III. Teliospores are two-celled, produced by karyogamy of two haploid nuclei (n+n). Teliospores may or may not go through dormancy before producing basidiospores. *P. graminis* teliospores, have a long dormancy, while those of *P. striiformis* have have a very short or no dormancy. Teliospores undergo meiosis to produce one-celled basidiospores.When a teliospore germinates after dormancy, each cell produces a promycelium (also called basidium) bearing four basidiospores (one-celled and with one or more nuclei); this is referred to as stage IV. These basidiospores lands and germinates on the alternate host and the cycle is repeated (Anikster, 1986; M. N. Wang & Chen, 2015; M. N. Wang, Wan, & Chen, 2015; Zhao, Wang, Chen, & Kang, 2016).

The most common type of rust is leaf rust which has alternate hosts including Meadow rue, Rue-anemone, and Leather flower. While the alternate host of stem rust is barberry plant, stripe has no known alternate host. The diseases obtain their name from their appearance on the host plant. On wheat plants, stem rust has brick-red, elongated pustules that are mostly on the leave sheaths and stem; leaf rust has reddish-orange, circular to oval pustules that are scattered or clustered on the upper leaf surface while stripe rust pustules are orange-brown, circular to oval and appear on leaf stripes. The key notes here are that these fungal pathogens are obligate parasites and therefore always devise means of successful colonization of their host. Their several spore stages and ability to reproduce both sexually and asexually enhances their survival through frequent mutation to produce different pathotypes and rapid rate of proliferation. Also, they can be

transmitted to different places via various means including wind. These together, makes the control and management of wheat rust diseases an unending challenge.

Wheat-Rust Interaction

Plants have evolved to trigger different and complex defense mechanisms to survive against pathogens. This involves activation or repression of various genes involved in different molecular, biochemical and physiological processes acting in defensive processes. Plant response regulatory network against external stresses are mostly controlled at the transcriptional level (W. J. Chen & Zhu, 2004). On the other hand, pathogens have also evolved various means to overcome host barriers and defensive mechanisms. Necrotrophs like genera *Pythium* and *Botrytis* species secrete cell wall degrading enzymes to overcome cell wall barriers. Hemi-biotrophs, such as the rice blast fungus *Magnaporthe grisea* pathogens have developed a special translocation system to deliver effectors into host cells. Biotrophic pathogens such as rust and powdery mildew force open cell walls and produce haustoria which secretes effector proteins to acquire nutrients and manipulate host cellular activities for the pathogens' survival in the host (Mendgen, Hahn, & Deising, 1996).

So far, our knowledge of compatible wheat-rust interactions (successful infection leading to disease) has been divided into perception based and non-perception based. In perception-based interactions, plants use their innate ability to fight pathogens by recognizing pathogen elicitors, in pathogen/microbe-associated molecular pattern triggered immunity (PTI) (Boller & He, 2009). This is the plant's first line of defense against pathogen elicitors such as β-Glucans and chitin by fungi (Bowman & Free, 2006;

Klarzynski et al., 2000). Effector triggered immunity (ETI) is deployed when pathogens elude basal defenses. Effectors are proteins that are released by pathogens into the host cells from the haustoria of fungal pathogens (Garnica, Nemri, Upadhyaya, Rathjen, & Dodds, 2014). Effectors suppress host defense and manipulate host cellular mechanism. Some effectors act as transcriptional activator like proteins (TAL) that activate host cells for the pathogen's benefit (Sanjana et al., 2012). To defense against these proteins, the plant activates the transcription of resistance genes which recognizes effectors in the so-called gene-for-gene resistance model. Resistance evident by hypersensitive response is observed only when a dominant resistance gene (R gene) corresponds with an avirulent effector protein (*Avr* gene) (Flor, 1971; Loegering & Ellingboe, 1987). This led to the identification and cloning of many plant resistance genes. In wheat there are about 187 rust resistance genes; 80 leaf rust, 58 stem rust and 49 stripe rust (Aktar-Uz-Zaman, Tuhina-Khatun, Hanafi, & Sahebi, 2017). Resistance genes are designated as *Lr*, *Sr* and *Yr* genes to mean leaf rust, stem rust and yellow (stripe) rust respectively. Later models include molecular surveillance model (Van Der Biezen & Jones, 1998) that explains the monitoring of host genes modification by effector protein for defense activation and decoy model (van der Hoorn & Kamoun, 2008) in which plants mimic modified 'guardee' proteins (by pathogen effectors) for defense. In non-perception based, pathogens subvert host defense by hijacking pathogen susceptibility (S) genes in the host. S genes are usually recessive unlike R genes and are required by pathogens for host recognition and penetration, proliferation and spread and for negative regulation of

immune signals (van Schie & Takken, 2014). In summary, our knowledge of wheat rust-interaction guides our efforts in breeding for rust resistance in bread wheat.

Wheat Rust Resistance Breeding

The two main methods of controlling rust diseases include the use of chemicals and genetic resistance. Genetic or variety of resistance has been the most economical and environmentally friendly method of control. Traditionally, breeding for rust resistance in wheat involves the use of resistance genes (R) genes and adult plant resistance genes (APR). R genes are effective from seedling to adult stages of wheat but are less durable. Mostly, resistance conferred by R genes are pathogen race-specific and follows the gene-for-gene hypothesis (Flor, 1971). They are only useful when they recognize their corresponding pathogen *Avr* gene. However, some R genes like *Sr24* and *Sr31* have been classified as "broad spectrum" because they confer resistance against several tested races of a rust pathogen species. Several R genes have been cloned; the first was *L6* from flax (Lawrence, Finnegan, Ayliffe, & Ellis, 1995) and in wheat, *Lr21* (L. Huang et al., 2003). Molecular cloning revealed that all R genes (except *Rpg1* a stem rust R gene from barley) encodes the coiled-coil, nucleotide-binding site, leucine-rich repeat (CC-NB-LRR). In contrast, APR genes are mainly useful at the adult stage of the plant and are more durable. For example, *Lr34* (leaf and stripe rust and powdery mildew resistance gene) and *Sr2* (stem rust resistance gene) have been extensively used APR genes in breeding that confer durable but partial resistance in wheat.

Clearly, using one or two resistance genes in breeding is continually becoming inadequate. Pathogens continue to evolve new races that are aggressively virulent to most

resistant genes. For example, the emergence of the broadly virulent stem rust African strain Ug99 and a new virulent Australian stripe rust race in 2002 and their mutational derivatives have overcome several R and APR genes. Ug99 is virulent against *Sr31* which has shown resistance to many *Pgt* races for over 30 years (Pretorius, Singh, Wagoire, & Payne, 2000). This raises huge concerns for breeders. Moreover, the narrow gene pool in bread wheat has prompted the introgression of resistance genes from wild relatives into wheat through interspecific crosses, which is often characterized by linkage drag and labor and cost intensiveness. For example, *Sr24* and *Sr31* genes in wheat are from *Agropyron sp.* and *Secale cereale,* respectively. Though efforts are pointing to using resistance genes from *Aegilops* species (Olivera, Rouse, & Jin, 2018; Rouse, Olson, Gill, Pumphrey, & Jin, 2011) to reduce these aforementioned challenges associated with interspecific crosses, it is still not enough to drastically solve the problems faced in breeding for rust resistance. Notably, we would still have to deal with some level of linkage drag as well as the inability of the genes to fully express in their destined background. It is, therefore, imperative to speed up efforts in considering other ways of creating resistance in bread wheat.

Gene Editing for Rust Resistance

The understanding that plant pathogen effectors target host S genes for their use is driving efforts in the identification and editing of S genes to enhance host resistance to the pathogens. For instance, destruction of tomato *SlMlo1* (Nekrasov et al., 2017) and *DMR6* (de Toledo Thomazella, Brail, Dahlbeck, & Staskawicz, 2016) genes enhanced resistance to Powdery mildew and *Pseudomonas syringae*, *Phytophthora capsici*, and

Xanthomonas spp respectively; apple *DIPM-1*, *DIPM-2* and *DIPM-4* genes enhanced resistance to Fire blight disease and grape *MLO-7* gene gave resistance to Powdery mildew (Malnoy et al., 2016); citrus *CsLOB1* promoter enhanced resistance to Citrus canker (Peng et al., 2017).

The first step to editing host genes for resistance is to identify the possible rust target host genes. To do this, the function of the gene in previous studies could be a clue to hypothesizing a similar or related role in a plant of interest. As you have a candidate gene you could use *in silico*, *in vitro* and/or *in vivo* experiments to test your initial hypothesis under a compatible host-pathogen (wheat-rust) interaction condition. Next, study the candidate gene using functional gene analysis tools like RNA interference (RNAi) and Virus induced gene silencing (VIGS), which are posttranscriptional gene silencing tools. RNAi is used in creating transgenic plants, which comes with transformation challenges in wheat. VIGS is a better choice when you do not need to create a transgenic plant. It is a homology-based post-transcriptional gene silencing (PTGS) technique that exploits plants' natural defense mechanisms against invading viruses. In hexaploid wheat, barley stripe rust mosaic virus (BSMV)-VIGS was first used in silencing phytoene desaturase expression and *Lr21* function (Scofield, Huang, Brandt, & Gill, 2005). If the candidate gene functions as expected, the gene is edited via creation of mutations in the gene. Earlier targeted genome editing technologies rely on engineered endonucleases such as zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs). These nucleases composed of sequence-specific DNA-binding domains fused to a non-specific DNA cleavage module that generate site-specific

double-stranded breaks at target loci which repairs through the endogenous repair systems in cells. TALEN was first exploited in wheat to edit *TaMLO* genes, which resulted in a broad-spectrum resistance to powdery mildew (Y. Wang et al., 2014). However, these techniques have high costs for the construction of the protein domains and have the tendency of off-target cleavage of DNA. Presently, CRISPR/Cas9 system is receiving all the attention for gene editing because of its efficiency over ZFNs and TALENs, multiplexed mutation ability, and simplicity in target design (Kumar, Kaur, Pandey, Mamrutha, & Singh, 2019). Its usage in monocot and dicot plants is growing. In wheat CRISPR/Cas9 was used in editing abiotic stress-responsive transcription factor genes (Kim, Alptekin, & Budak, 2018) and creating low-gluten, transgene-free wheat lines (Sánchez-León et al., 2018). With the release of the fully annotated high-quality reference genome of bread wheat variety ‘Chinese Spring’ (Appels et al., 2018), it is expected that the use of CRISPR/Cas systems will skyrocket. However, the challenge of transformation in wheat and the issue of off-targeting remain limitations of the system.

Chemical mutagenesis has been a well-established system for creating heritable mutations in plants. Ethyl methanesulfonate (EMS) is one of the most used chemical mutagens. Chemical mutagens can be used to create mutant populations with high densities of mutations, making it easier to screen for specific mutations. With the advent of efficient TILLING (Targeting Induced Local Lesions in Genomes) techniques (McCallum, Comai, Greene, & Henikoff, 2000), reverse genetics approaches which were previously not applicable in mutation breeding are now possibilities. Good TILLING populations already exist for wheat and barley (Krasileva et al., 2017; Szurman-

Zubrzycka et al., 2018). Moreover, the availability of new generation sequencing techniques and efficient methods to detect DNA heteroduplexes have made it easier now to screen the populations for mutations in selected target genes. All in all, the development of TILLING, stable mutant populations, now makes the acquisition of mutations in specific genes very efficient and fast for wheat and other crops. In these studies, we identified and edited rust pathogen targetted wheat genes using EMS mutagenized populations and advanced the population to near-isogenic lines via a single seed descent method to create wheat rust resistance germplasm.

REFERENCES

Aktar-Uz-Zaman, M., Tuhina-Khatun, M., Hanafi, M. M., & Sahebi, M. (2017). Genetic analysis of rust resistance genes in global wheat cultivars: an overview. *Biotechnology & Biotechnological Equipment, 31*(3), 431-445. doi:10.1080/13102818.2017.1304180

Anikster, Y. (1986). Teliospore Germination in Some Rust Fungi. *Phytopathology, 76*(10), 1026-1030. doi:DOI 10.1094/Phyto-76-1026

Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., . . . Poland, J. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science, 361*(6403), eaar7191.

Boller, T., & He, S. Y. (2009). Innate immunity in plants: an arms race between pattern recognition receptors in plants and effectors in microbial pathogens. *Science (New York, N.Y.), 324*(5928), 742-744. doi:10.1126/science.1171647

Bowman, S. M., & Free, S. J. (2006). The structure and synthesis of the fungal cell wall. *Bioessays, 28*(8), 799-808.

Chen, W. J., & Zhu, T. (2004). Networks of transcription factors with roles in environmental stress response. *Trends in Plant Science, 9*(12), 591-596. doi:https://doi.org/10.1016/j.tplants.2004.10.007

Chen, X. M. (2014). Integration of cultivar resistance and fungicide application for control of wheat stripe rust. *Canadian Journal of Plant Pathology, 36*(3), 311-326. doi:10.1080/07060661.2014.924560

de Toledo Thomazella, D. P., Brail, Q., Dahlbeck, D., & Staskawicz, B. (2016). CRISPR-Cas9 mediated mutagenesis of a DMR6 ortholog in tomato confers broad-spectrum disease resistance. *bioRxiv*, 064824.

Dvorak, J. (1972). Genetic Variability in Aegilops-Speltoides Affecting Homoeologous Pairing in Wheat. *Canadian Journal of Genetics and Cytology, 14*(2), 371-&.

Dvorak, J., Diterlizzi, P., Zhang, H. B., & Resta, P. (1993). The Evolution of Polyploid Wheats - Identification of the a-Genome Donor Species. *Genome, 36*(1), 21-31. doi:DOI 10.1139/g93-004

Flor, H. H. (1971). Current status of the gene-for-gene concept. *Annual Review of Phytopathology, 9*(1), 275-296.

Garnica, D. P., Nemri, A., Upadhyaya, N. M., Rathjen, J. P., & Dodds, P. N. (2014). The Ins and Outs of Rust Haustoria. *PLOS Pathogens, 10*(9), e1004329. doi:10.1371/journal.ppat.1004329

Huang, L., Brooks, S. A., Li, W., Fellers, J. P., Trick, H. N., & Gill, B. S. (2003). Map-based cloning of leaf rust resistance gene Lr21 from the large and polyploid genome of bread wheat. *Genetics, 164*(2), 655-664.

Huang, S., Sirikhachornkit, A., Su, X., Faris, J., Gill, B., Haselkorn, R., & Gornicki, P. (2002). Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the Triticum/Aegilops complex and the evolutionary history of polyploid wheat. *Proc Natl Acad Sci U S A, 99*(12), 8133-8138. doi:10.1073/pnas.072223799

International Wheat Genome Sequencing, C. (2014). A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. *Science, 345*(6194), 1251788. doi:10.1126/science.1251788

Jenkins, J. A. (1929). Chromosome homologies in wheat and Aegilops. *American Journal of Botany, 16*(4), 238-245. doi:Doi 10.2307/2435839

Kihara, H. (1954). Considerations on the Evolution and Distribution of Aegilops Species Based on the Analyser-Method. *Cytologia, 19*(4), 336-357. doi:DOI 10.1508/cytologia.19.336

Kim, D., Alptekin, B., & Budak, H. (2018). CRISPR/Cas9 genome editing in wheat. *Functional & Integrative Genomics, 18*(1), 31-41. doi:10.1007/s10142-017-0572-x

Klarzynski, O., Plesse, B., Joubert, J. M., Yvin, J. C., Kopp, M., Kloareg, B., & Fritig, B. (2000). Linear beta-1,3 glucans are elicitors of defense responses in tobacco. *Plant physiology, 124*(3), 1027-1038. doi:10.1104/pp.124.3.1027

Krasileva, K. V., Vasquez-Gross, H. A., Howell, T., Bailey, P., Paraiso, F., Clissold, L., . . . Dubcovsky, J. (2017). Uncovering hidden variation in polyploid wheat. *Proceedings of the National Academy of Sciences of the United States of America, 114*(6), E913-E921. doi:10.1073/pnas.1619268114

Kumar, R., Kaur, A., Pandey, A., Mamrutha, H. M., & Singh, G. P. (2019). CRISPR-based genome editing in wheat: a comprehensive review and future prospects. *Molecular Biology Reports, 46*(3), 3557-3569. doi:10.1007/s11033-019-04761-3

Lawrence, G. J., Finnegan, E. J., Ayliffe, M. A., & Ellis, J. G. (1995). The L6 gene for flax rust resistance is related to the Arabidopsis bacterial resistance gene RPS2 and the tobacco viral resistance gene N. *The Plant cell, 7*(8), 1195-1206.

Ling, H. Q., Zhao, S., Liu, D., Wang, J., Sun, H., Zhang, C., . . . Wang, J. (2013). Draft genome of the wheat A-genome progenitor Triticum urartu. *Nature, 496*(7443), 87-90. doi:10.1038/nature11997

Loegering, W. Q., & Ellingboe, A. (1987). HH Flor: pioneer in phytopathology. *Annual Review of Phytopathology, 25*(1), 59-66.

Luo, M. C., Gu, Y. Q., Puiu, D., Wang, H., Twardziok, S. O., Deal, K. R., . . . Dvorak, J. (2017). Genome sequence of the progenitor of the wheat D genome Aegilops tauschii. *Nature, 551*(7681), 498-+. doi:10.1038/nature24486

Malnoy, M., Viola, R., Jung, M.-H., Koo, O.-J., Kim, S., Kim, J.-S., . . . Nagamangala Kanchiswamy, C. (2016). DNA-free genetically edited grapevine and apple protoplast using CRISPR/Cas9 ribonucleoproteins. *Frontiers in Plant Science, 7*, 1904.

McCallum, C. M., Comai, L., Greene, E. A., & Henikoff, S. (2000). (TILLING) for Plant Functional Genomics. *Plant physiology, 123*(2), 439. doi:10.1104/pp.123.2.439

Mcfadden, E. S., & Sears, E. R. (1946a). The Origin of Triticum Spelta and Its Free-Threshing Hexaploid Relatives. *Journal of Heredity, 37*(3), 81-89.

McFadden, E. S., & Sears, E. R. (1946b). The origin of triticum spelta and its free-threshing hexaploid relatives*. *Journal of Heredity, 37*(3), 81-89. doi:10.1093/oxfordjournals.jhered.a105590

Mendgen, K., Hahn, M., & Deising, H. (1996). Morphogenesis and mechanisms of penetration by plant pathogenic fungi. *Annual Review of Phytopathology, 34*(1), 367-386. doi:10.1146/annurev.phyto.34.1.367

Miki, Y., Yoshida, K., Mizuno, N., Nasuda, S., Sato, K., & Takumi, S. (2019). Origin of wheat B-genome chromosomes inferred from RNA sequencing analysis of leaf transcripts from section Sitopsis species of Aegilops. *DNA research : an international journal for rapid publication of reports on genes and genomes, 26*(2), 171-182. doi:10.1093/dnares/dsy047

Nekrasov, V., Wang, C., Win, J., Lanz, C., Weigel, D., & Kamoun, S. (2017). Rapid generation of a transgene-free powdery mildew resistant tomato by genome deletion. *Scientific reports, 7*(1), 482.

Olivera, P. D., Rouse, M. N., & Jin, Y. (2018). Identification of New Sources of Resistance to Wheat Stem Rust in Aegilops spp. in the Tertiary Genepool of Wheat. *Frontiers in Plant Science, 9*(1719). doi:10.3389/fpls.2018.01719

Pathak, G. N. (1940). Studies in the cytology of cereals. *Journal of Genetics, 39*(3), 437-467. doi:Doi 10.1007/Bf02982884

Peng, A., Chen, S., Lei, T., Xu, L., He, Y., Wu, L., . . . Zou, X. (2017). Engineering canker-resistant plants through CRISPR/Cas9-targeted editing of the susceptibility gene Cs LOB 1 promoter in citrus. *Plant biotechnology journal, 15*(12), 1509-1519.

Petersen, G., Seberg, O., Yde, M., & Berthelsen, K. (2006). Phylogenetic relationships of Triticum and Aegilops and evidence for the origin of the A, B, and D genomes of common wheat (Triticum aestivum). *Mol Phylogenet Evol, 39*(1), 70-82. doi:10.1016/j.ympev.2006.01.023

Pretorius, Z., Singh, R., Wagoire, W., & Payne, T. (2000). Detection of virulence to wheat stem rust resistance gene Sr31 in Puccinia graminis. f. sp. tritici in Uganda. *Plant Disease, 84*(2), 203-203.

Riley, R., Chapman, V., & Kimber, G. (1961). Origin of Genetic Control of Diploid-Like Behavior of Polyploid Wheat. *Journal of Heredity, 52*(1), 22-&. doi:DOI 10.1093/oxfordjournals.jhered.a107015

Rouse, M. N., Olson, E. L., Gill, B. S., Pumphrey, M. O., & Jin, Y. (2011). Stem Rust Resistance in Aegilops tauschii Germplasm. *Crop Science, 51*(5), 2074-2078. doi:10.2135/cropsci2010.12.0719

Salamini, F., Ozkan, H., Brandolini, A., Schafer-Pregl, R., & Martin, W. (2002). Genetics and geography of wild cereal domestication in the near east. *Nat Rev Genet, 3*(6), 429-441. doi:10.1038/nrg817

Sánchez-León, S., Gil-Humanes, J., Ozuna, C. V., Giménez, M. J., Sousa, C., Voytas, D. F., & Barro, F. (2018). Low-gluten, nontransgenic wheat engineered with CRISPR/Cas9. *Plant biotechnology journal, 16*(4), 902-910. doi:10.1111/pbi.12837

Sanjana, N. E., Cong, L., Zhou, Y., Cunniff, M. M., Feng, G., & Zhang, F. (2012). A transcription activator-like effector toolbox for genome engineering. *Nature protocols, 7*(1), 171-192. doi:10.1038/nprot.2011.431

Sarkar, P., & Stebbins, G. L. (1956). Morphological Evidence Concerning the Origin of the B-Genome in Wheat. *American Journal of Botany, 43*(4), 297-304. doi:Doi 10.2307/2438947

Sax, K. (1922). Sterility in wheat hybrids. II. Chromosome behavior in partially sterile hybrids. *Genetics, 7*(6), 513-552.

Scofield, S. R., Huang, L., Brandt, A. S., & Gill, B. S. (2005). Development of a Virus-Induced Gene-Silencing System for Hexaploid Wheat and Its Use in Functional Analysis of the Lr21-Mediated Leaf Rust Resistance Pathway. *Plant physiology, 138*(4), 2165. doi:10.1104/pp.105.061861

Shrawat, A. K., & Armstrong, C. L. (2018). Development and application of genetic engineering for wheat improvement. *Critical Reviews in Plant Sciences, 37*(5), 335-421.

Szurman-Zubrzycka, M. E., Zbieszczyk, J., Marzec, M., Jelonek, J., Chmielewska, B., Kurowska, M. M., . . . Szarejko, I. (2018). HorTILLUS-A Rich and Renewable Source of Induced Mutations for Forward/Reverse Genetics and Pre-breeding Programs in Barley (Hordeum vulgare L.). *Frontiers in Plant Science, 9*, 216-216. doi:10.3389/fpls.2018.00216

Van Der Biezen, E. A., & Jones, J. D. (1998). Plant disease-resistance proteins and the gene-for-gene concept. *Trends in biochemical sciences, 23*(12), 454-456.

van der Hoorn, R. A. L., & Kamoun, S. (2008). From Guard to Decoy: A New Model for Perception of Plant Pathogen Effectors. *The Plant cell, 20*(8), 2009. doi:10.1105/tpc.108.060194

van Schie, C. C., & Takken, F. L. (2014). Susceptibility genes 101: how to be a good host. *Annual Review of Phytopathology, 52*, 551-581.

Wang, J. R., Luo, M. C., Chen, Z. X., You, F. M., Wei, Y. M., Zheng, Y. L., & Dvorak, J. (2013). Aegilops tauschii single nucleotide polymorphisms shed light on the origins of wheat D-genome genetic diversity and pinpoint the geographic origin of hexaploid wheat. *New Phytologist, 198*(3), 925-937. doi:10.1111/nph.12164

Wang, M. N., & Chen, X. M. (2015). Barberry Does Not Function as an Alternate Host for Puccinia striiformis f. sp tritici in the U. S. Pacific Northwest Due to Teliospore Degradation and Barberry Phenology. *Plant Disease, 99*(11), 1500-1506. doi:10.1094/Pdis-12-14-1280-Re

Wang, M. N., Wan, A. M., & Chen, X. M. (2015). Barberry as Alternate Host Is Important for Puccinia graminis f. sp tritici But Not for Puccinia striiformis f. sp tritici in the US Pacific Northwest. *Plant Disease, 99*(11), 1507-1516. doi:10.1094/Pdis-12-14-1279-Re

Zhang, W., Zhang, M., Zhu, X., Cao, Y., Sun, Q., Ma, G., . . . Cai, X. (2018). Molecular cytogenetic and genomic analyses reveal new insights into the origin of the wheat B genome. *Theor Appl Genet, 131*(2), 365-375. doi:10.1007/s00122-017-3007-0

Zhao, J., Wang, M. N., Chen, X. M., & Kang, Z. S. (2016). Role of Alternate Hosts in Epidemiology and Pathogen Variation of Cereal Rusts. *Annual Review of Phytopathology, Vol 54, 54*, 207-228. doi:10.1146/annurev-phyto-080615-095851

CHAPTER TWO

WHEAT *MYC4* TRANSCRIPTIONAL FACTOR GENE MODIFICATION ENHANCED HOST RESISTANCE AGAINST RUST PATHOGENS

Contribution of Author and Co-Authors

Manuscript in Chapter 2

Author: Bernard M Nyamesorto

Contributions: Conducted experiments, collected and analyzed data, prepared figures and wrote the manuscript.

Co-Author: Hongtao Zhang

Contributions: conducted RNA-seq experiments, the findings led to this study.

Co-Author: Li Huang

Contributions: designed, supervised the project, and edited the manuscript.

Manuscript Information

Bernard Nyamesorto, Hongtao Zhang and Li Huang

Status of the manuscript:

__X__ Prepared for submission to a peer-reviewed journal

____ Officially submitted to a peer-reviewed journal

____ Accepted by a peer-reviewed journal

____Published in a peer-reviewed journal

ABSTRACT

Transcriptional reprogramming is an essential feature of plant immunity and is governed by transcription factors (TFs) and co-regulatory proteins associated with discrete transcriptional complexes. On the other hand, effector proteins from pathogens have been shown to hijack these vast repertoires of plant TFs. Studies have begun targeting and editing some host genes that benefit pathogens to enhance plants' immunity to pathogens. However, our current knowledge on the role of host genes (including TFs) involved in host colonization is just based on research employing a few model plants such as *Arabidopsis* and rice with minimal efforts in wheat rust interactions. In this study, we identified the wheat *MYC4* transcription factor (TF) located on the chromosome 1A, 1B and 1D (*TaMYC4-1A, 1B, 1D)* was upregulated at 24 hours post- rust inoculation in a susceptible wheat line. Down-regulation of all the *TaMYC4* homeologs using barley stripe mosaic virus-induced gene silencing (BSMV-VIGS) in the susceptible cultivar Chinese Spring enhanced its resistance to stem rust pathogen. Knockout of the *TaMYC4 -1BL* homeolog in Cadenza rendered new resistance to races of stem, leaf, and stripe rust pathogens. From this discovery, we have created new germplasm in wheat via modifications of the wheat *TaMYC4 -1BL* transcription factor.

INTRODUCTION

Plants employ a complex network of signaling pathways to defend themselves against pathogen attacks. Signal integration is dictated by transcription factor (TF) regulatory networks. Transcriptional reprogramming is a major component of plant immunity and is administered by TFs and co-regulatory proteins connected within distinct transcriptional complexes (Moore, Loake, & Spoel, 2011). Over a couple of decades, studies have uncovered numerous TF family members mostly in *Arabidopsis thaliana* and rice that are critical in regulating proper defense responses when plants are confronted by pathogens. Many of these TFs have been categorized into AP2/ERF, bHLH, bZIP, MYB, NAC and WRKY families. *MYC*, a basic-helix-loop-helix (bHLH) family TFs were initially discovered from a homology study between an oncogene carried by the Avian *v*irus, *My*elo*c*ytomatosis (*v-MYC*) and a human gene overexpressed in different cancers, *c*ellular *MYC* (*c-MYC*). They have a DNA binding domain made of 50–60 amino acids, which allows for homo- or heterodimerization to their DNA consensus hexamer sequence CANNTG (Finver et al., 1988). The bHLH TFs have been shown to be key regulators in Jasmonic Acid (JA)-mediated defense responses and in mediating crosstalk with other phytohormones, including salicylic acid (SA), abscisic acid (ABA), gibberellins (GA), and auxin (Kazan & Manners, 2013).

Co-evolved with host defense systems in plants, pathogens are also continually developing counter mechanisms to overcome host defenses. It has become evident that one component of their arsenal is manipulating host cellular processes using effector proteins, including exploiting pathogen susceptible host genes. Efforts in modifying these

pathogen targeted host genes to increase resistance against pathogens have become a go-to approach for disease-resistant breeding in model crops such as rice. For example alteration of rice promoters *OsSWEET14* and *OsSWEET11* and the *OsMPK5* gene enhanced resistance to Bacterial blight, fungal (*Magnaporthe grisea*) and bacterial (*Burkholderia glumae*) pathogens (Jiang et al., 2013; Li, Liu, Spalding, Weeks, & Yang, 2012; Xie & Yang, 2013). Editing of wheat *TaMLO* and *TaEDR1* genes (three homeologs) enhanced resistance to powdery mildew (*Blumeria graminis f. sp. tritici*) (Shan et al., 2013; Wang et al., 2014; Y. Zhang et al., 2017). Nevertheless, the level of application of gene editing is minimal in wheat rust resistance breeding.

Upon the release of the complete genome sequence of hexaploid wheat (*Triticum aestivum* L.) by the International Wheat Genome Sequencing Consortium (IWGSC, 2018), it has become apparent that of the over 107,891 high-confidence genes identified, more than 35,000 are transcriptional factors categorized into 40 families (MYC, MYB, WRKY, etc) and 84 subfamilies (Appels et al., 2018). Limited studies and implications of these TFs in plant-microbe interactions were mostly done in *A. thaliana* but not in *T. aestivum*. In a typical example, it was established in *A. thaliana* that *MYC2* is the best known target of the Jasmonic Acid ZIM domain (JAZ) family of repressor proteins (Dombrecht et al., 2007). Other examples of TFs studies in plant-microbe interaction could be seen in these reviews and research work (J. Chen et al., 2020; Hussain, Sheikh, Haider, Quareshy, & Linthorst, 2018; Tsuda & Somssich, 2015). In a pilot transcriptomic study on wheat-leaf rust interaction, the expression pattern of selected wheat TFs including *MYC4* were found to be upregulated starting at 5 days post-rust inoculation, a

critical rust pathogen development stage when the rust has established infection sites in susceptible wheat cultivars. In this study, we demonstrate how the modification of the wheat TF *MYC4* gene (*TaMYC4*) homeolog in the long arm of chromosome 1B (*TaMYC4 -1BL*) conferred enhanced resistance to wheat against rust pathogens.

MATERIALS AND METHODS

Plant and Pathogen Materials

Alpowa (PI 566596) spring wheat, was obtained from the USDA National Plant Germplasm System (NPGS) and Chinese Spring (CS) was obtained from Dr. Evans Lagudah at Commonwealth Scientific and Industrial Research Organization (CSIRO). Cadenza, a spring wheat, was obtained from the SeedStor via http://www.seedstor.ac.uk.

The *Puccinia graminis* f. sp. *tritici* (*Pgt*) races QFCSC and TPMKC were provided by Dr. Yue Jin from Cereal Disease Laboratory, USDA-ARS, St. Paul, MN. Leaf rust, *P. tritici (Pt)* race PBJJG was provided by Dr. Robert Bowden (USDA-ARS, Manhattan, KS, USA). Stripe rust, *P. striiformis* f. sp. *tritici* (*Pst*) culture (race unknown) was collected from the Bozeman Agricultural Research and Teaching Farm of Montana State University (MSU).

Plant Growth Conditions

For rust screenings, wheat seeds were directly planted into 4-inch small pots (5 seeds/pot) containing only SunGro Horticulture Sunshine mix (HeavyGardens Company, Denver, CO). For seed propagation and crosses, wheat seeds were first germinated in Petri dishes on filter paper at room temperature. At root radical emergence, the seeds

were transferred to 8-inch pots (one seedling/pot) containing a mixture of 1:1 ratio of local soil: Sunshine mix. Growth conditions were set at 22°C/14°C day/night temperatures and a 16 h photoperiod in a greenhouse at the Plant Growth Center, MSU. Plants were watered and fertilized every day with Peters General Purpose Plant Food (Scotts-Miracle-Gro Company, Marysville, OH) at the concentration of 150 ppm.

Rust Pathogen Inoculation and Assessment

All rust screenings were done at the two-leaf seedling stage. Rust inoculations were conducted as described in (Campbell et al., 2012). In brief, plants were inoculated with rust spore in Soltrol 170 isoparaffin oil suspensions. Inoculated plants were then transferred to a Percival I-60D dew chamber (Percival Scientific Inc., Perry, IA) pre-conditioned to an internal air temperature between 15–17 °C for leaf and stripe rust, 19–20°C for stem rust for 24 hours. An additional step of 4 hours exposure to light prior to the removal from the dew chamber was done for stem rust inoculation. Stem rust race QFCSC and TPMKC, leaf rust race PBJJG and stripe rust evaluations were conducted at Montana State University in the greenhouse of the Plant Growth Center. Stem rust races TTKSK and TKTTF GER assay were done in the Cereal Disease Laboratory, USDA-ARS, St. Paul, MN following procedure described by (Jin et al., 2007). The other stripe rust races were assayed at the Wheat Health, Genetics, and Quality Research Unit, USDA-ARS, Pullman, WA, according to procedures described by (Line & Qayoum, 1992) and infections types (IT) recorded 18-20 days after inoculation.

Infection types of seedlings to leaf and stem rust disease responses were assessed using the 0-4 IT scale (McIntosh, Wellings, & Park, 1995) at eight days post leaf rust

inoculation (dpi) and 13-14 dpi for stem rust, respectively when the symptoms of susceptible controls were fully expressed. Stripe rust disease was assessed 14 dpi based 0 (immune)-8 scale (highest susceptible) (Line & Qayoum, 1992).

Sample Collection and Treatment

For RNA extraction, leaf samples were taken for three plants separately per treatment. Each sample was frozen in liquid nitrogen and stored at –80°C. RNA extractions were not done until all the samples at different time points were collected. Sample collection times ranges from 0 to 10 dpi depending on the experiment. The samples at 0 dpi were taken immediately after inoculation prior to placing the inoculated plants in a dew chamber. For DNA extraction, leaf samples were taken and immediately used for extraction.

RNA and DNA Extraction

Total RNAs were isolated and treated with DNase I on columns using the Qiagen RNeasy Plant Mini Kit (Qiagen, Valencia, CA) following the manufacturer's instructions. Genomic DNA extraction for conventional PCR was done using the QIAGEN DNeasy Plant Mini Kit (Qiagen Sciences Inc, Germantown, MD). For KASP assays, DNA was extracted from 96 plants using the 96-well plate extraction procedure modified from (Holleley & Sutcliffe, 2009). The quality and concentration of total RNA/DNA were assessed using 260/280ABS measurements on a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific Inc., Wilmington, DE). The integrity of DNA or RNA was checked via agarose gel electrophoresis with 2 µl of sample, 4 µl of water, 1 µl loading buffer (98 % formamide, 10 mM EDTA, 0.25 % bromophenol blue,

and 0.25 % xylene cyanol) on a 0.8-1% gel stained by GelRed (Bio-Rad, Hercules, CA) at 125 volts for 25 minutes.

qRT-PCR, Conventional PCR and KASP Assay

qRT-PCR was performed using the iScript One-Step RT-PCR Kit with SYBR Green (Bio-Rad, Hercules, CA) on a CFX96 real-time PCR detection system (Bio-Rad, Hercules, CA) following the manufacturer's procedure with 100-150 µg sample RNA at annealing temperatures of 56/57 °C depending on the primers. Actin was used as the house keeping gene for normalization (Appendix A Table 2.S1). qRT-PCR was conducted in triplicate.

PCR amplifications were conducted in 20 µl reactions containing 25 mM $MgCl_2$, 10 mM dNTP, 2 µM of each primer (BN1BL primers, Table 2), 50 ng genomic DNA and 1 unit Go Taq Flexi DNA polymerase (Promega, Madison, Wisconsin). Amplifications were performed at 95 °C for 7 min, followed by 35 cycles at 95 °C for 45 s, 55 °C for 45 s, and 68/72°C for 40 s (depending on primers), with a final extension at 68/72 °C for 10 min.

KASP genotyping was conducted using KASP genotyping trial kit following the manufacturer’s protocol (Biosearch Technologies Genomic analysis by LGC) using manually designed KASP primer (Appendix A Table 2.S1) on a CFX96 real-time PCR detection system. Products from qRT-PCR, PCR and KASP were checked using gel electrophoresis as described previously.

Barley Stripe Mosaic Virus-Induced Gene Silencing Assay

Gene knockdown was conducted via BSMV-VIGS assay. The original BSMV vectors were obtained from Dr. Andrew O. Jackson (UC Berkeley, CA, USA). The target fragment for silencing assay was inserted into the modified γ vector ready for direct PCR cloning as described by (Jackie & Huang, 2010). BSMV RNA transcripts were synthesized in vitro using T7 RNA polymerase (New England Biolabs, Ipswich, MA, USA) from linearized α, β and γ plasmids. The BSMV inoculum was prepared with 3 μL of BSMV RNAs (1:1:1 ratio of α, β, and γ) and 22.5 μL of inoculation buffer. The inoculum was then rub inoculated onto the second leaf of two-leaf-stage plants. Leaf tissues were sampled after successful silencing (at nine days after virus inoculation) to test the silencing efficiency. Stem rust inoculations were done at 14 days post virus inoculation when BSMV-induced target gene silence reached highest level.

Mutant Search and Validation

Cadenza mutants were identified from the wheat-tilling database using the sequence of candidate genes as a query. Wheat-tilling is a resource TILLING population consisting of 2,700 individuals developed via EMS mutagenesis in tetraploid durum cv 'Kronos' and the hexaploid wheat cv 'Cadenza' backgrounds. Each mutant has been full genome sequenced. Mutations of requested mutants were validated via sequencing of the target regions after PCR amplification from the wild-type Cadenza and mutants using gene-specific primers.

Genetic Analysis

Genetic analysis was conducted to test the genotype-phenotype association using 150 seeds from a self-pollinated plant that are heterozygote at the selected locus. Also, 96 F_4 individuals were used for genotyping and phenotyping via KASP assay using KASP designed primers. For mutant P12S, the single nucleotide polymorphism (SNP) was a C-T nucleotide change from wildtype to P12S mutant. The forward oligos were designed as; Allele 1 with wildtype nucleotide (C) and Allele 2 which has the mutant nucleotide (T). A common reverse primer was designed for both allele oligos. A combination of the three oligos was used in assay (Appendix A Table 2.S1).

Pathogenesis-related (*PR*) Genes Expression

PR gene expression was assayed during time courses of the *Pgt* TPMKC infections at 0, 1, 2 dpi in both the wild-type Cadenza and mutant. Leaf samples were collected from three biological replicates per dpi for both wild-type and mutant and pretreated under recommended conditions for RNA extraction. Using corresponding *PR* gene primers (Appendix A Table 2.S1), *PR* genes were quantified using extracted RNA via real-time-qPCR as described earlier.

Databases and *in silico* Sequence Analysis

All BLAST and sequence downloads were completed using International Wheat Genome Sequence Consortium (IWGSC) resources at https://wheat-urgi.versailles.inra.fr/Seq-Repository/BLAST. Multiple sequences alignments were conducted using ClustalW Omega at https://www.ebi.ac.uk/Tools/msa/clustalo/. Gene and conserved domain predictions were performed using Softberry at

http://www.softberry.com/berry.phtml?topic=fgenesh&group=programs&subgroup=gfind, National Center for Biotechnology Information (NCBI) Conserved Domain Architecture Retrieval Tool at https://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi, and the ExPASy bioinformatics portal at https://prosite.expasy.org. Primers were designed either manually or using the PrimerQuest® tool at https://www.idtdna.com/pages/tools/primerquest, and primer specificity was assessed by BLAST search of the IWGSC database. RNA-seq data quality was checked using FastQC Version 0.11.6 (Babraham bioinformatics).

RNA sequence data from the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI) (Dobon, Bunting, Cabrera-Quio, Uauy, & Saunders, 2016) were used to quantify homeologs expression levels at different time points via Kallisto software (Bray, Pimentel, Melsted, & Pachter, 2016). The wheat transcriptome downloaded from IWGSC was indexed and reads of RNA-seq data were also assessed for quality via FastQC in interactive mode. Paired-end sequences were split using Fastq-dump prior to transcript quantification.

Statistical Analysis

Data assessment and analysis were conducted in Microsoft Excel and R-studio software version 1.1.453.0. For real-time PCR results, data were used only if the Ct standard deviation among the triplets was ≤ 0.2 and the mean of the triplet's Ct was used for downstream analysis. Relative expression was calculated using the ΔΔCt method as described in the CFX96 manual (Bio-Rad, Hercules, CA), where fold change = $2^{-\Delta\Delta Ct}$. Expression measurement of genes were conducted in three technical replicates for each of

three biological replicates. Standard deviations were calculated among three biological replicates normally or using delta formula when an estimator depends on a data. Student's t-tests were performed to test whether the expression level at different time points was significantly different. The p-values were calculated based on an unpaired two-tailed distribution. Expression patterns were graphically represented using averages of the three biological replicates.

RESULTS

Expression Profiles of the *TaMYC4* Genes by Real-Time qPCR

To further probe the expression pattern of *TaMYC4* candidate gene observed in the pilot transcriptomics data study (Appendix A Figure 2.S1) (unpublished) in which the expression of *TaMYC4* in the susceptible host was significantly higher than in the resistant near isogenic line, a time-course-study was conducted in Alpowa inoculated with the *Pgt* race QFCSC and buffer Soltrol 170 isoparaffin. The time-course experiment was conducted at 0, 1, 2, 3, 4, 5, 8 and 10 dpi. The expression level of *TaMYC4* was assayed via real-time quantitative PCR using the primers BN4RT that measured all homeologs of the *TaMYC4* gene (Appendix A Table 2.S1). *TaMYC4* expression levels were similar in samples taken at 0 dpi for pathogen and buffer. A significant increase in *TaMYC4* transcript abundance was detected at 1 dpi in *Pgt* infected Alpowa compared to control plants. The *TaMYC4* expression level returned to an undetectable level at 2 dpi and stayed unchanged during the rest of the time points (Figure 2.1). The expression level of *TaMYC4* in buffer inoculated Alpowa showed little and insignificant change along

from 1-3 dpi but showed an increase from 5-8 dpi then declined at 10 dpi. This increase is however, statistically insignificant. This result prompted us to ask whether *Pgt*-induced *TaMYC4* upregulation is beneficial to the pathogen during their colonization of wheat.

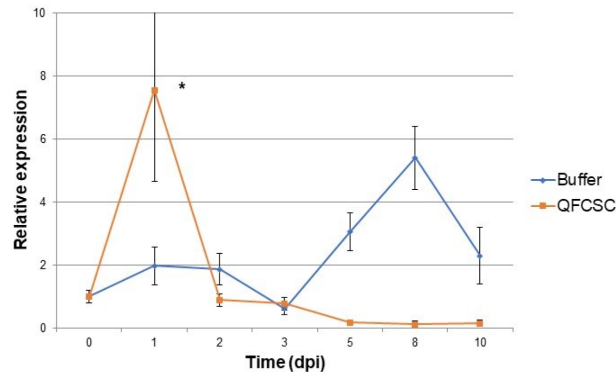


Figure 2.1: Relative expression of *TaMYC4* TF in Alpowa inoculated with Buffer (control) and stem rust (race QFCSC). Alpowa cultivar was inoculated with stem rust at 2-leaf stage. RNA samples were extracted from the leaf samples collected at seven time points. Real-time PCR was used to quantify transcript abundances of *TaMYC4* genes in both buffer and stem rust inoculated samples. The expression of the *TaMYC4* gene at each time point was computed relative to the level at 0-dpi. Error bars represent standard deviation for biological reps computed by delta method (since the relative expression at a time point depends on the comparison between expression in stem rust and buffer inoculated samples) and* denote statistical significance at the $p \leq 0.05$ levels calculated between each time point compared with 0 dpi.

Conserved Domains of *TaMYC4* Genes in Wheat

In order to address our previous question of *TaMYC4* possible negative role in wheat plant defense, we began with further investigation of *TaMYC4* conserved domains to relate its structure to function. We identified three copies of the *TaMYC4* gene in bread wheat by using a cDNA of *TaMYC4* -like sequence named TRIUR3_32014 from

Aegilops tauschii to BLAST search the International Wheat Genomic Sequence Consortium (IWGSC) database. *TaMYC4*-like genes were identified on 1AL, 1BL and 1DL chromosomes with a 97-99% similarity on DNA level or amino acid with a Expect Value of 0.0. Using the protein sequences shown in Appendix A File 2.S2, the three homeologs were predicted to have conserved Proline-rich domains (Pro), Glutamine-rich region (E), basic-helix-loop-helix (bHLH) and Leucine-zipper, each of which is classical *MYC4* conserved domains. Unique domains found on individual homeologs includes the following: Isoleucine-rich region profile (I) in *TaMYC4 -1AL*, Protein kinase C phosphorylation site (P) and Tyrosine kinase phosphorylation site (Y) on *TaMYC4 -1BL*, Arginine rich (Arg), APETALA2/ Ethylene-Responsive Element binding factor domain profile (A) and Bipartite nuclear localization signal profile (B) on *TaMYC4 -1DL* (Figure 2.2).

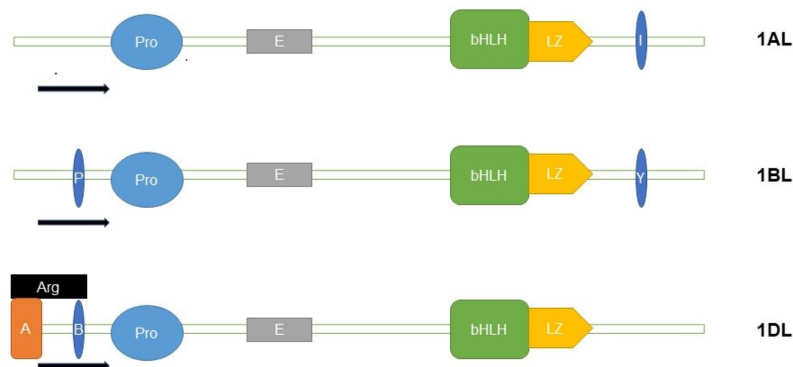


Figure 2.2: Predicted conserved domains of *TaMYC4* homeologs. Gene and conserved domains were predicted using scaffold sequences of wheat *TaMYC4* gene copies (1AL, 1BL and 1DL) from IWGSC via Softberry.com and NCBI's CDART and ExPASy bioinformatics portal, respectively. bHLH= basic helix-loop-helix; LZ= Leucine-zipper; Pro= Proline-rich; E= Glutamine rich; Y= Tyrosine kinase phosphorylation site; I= Isoleucine rich region profile; P= PKC_Phospho_site; Arg= Arginine rich; A= AP2/ERF domain profile; B= Bipartite nuclear localization signal profile. Arrows show the orientation of the gene. The figure is not drawn to scale.

The bHLH and leucine zipper motifs allow binding of *MYC* proteins with DNA and dimerization with other bHLH TFs (including Max), respectively. Proline-rich domains in transcription factors are vital for transcriptional activation presumably through binding to transcription initiation factors like RNA polymerase and probably TFIID (Williamson, 1994). It is known that TFs with a glutamine-rich region influence expression variation of their targets in eukaryotes across different timescales (Gemayel et al., 2015). Only *TaMYC4 -1BL* has phosphorylation sites. From ClustalW Omega multiple sequence alignment, percent identity matrix indicated the three homeologs had over 93%, 90% and 83% identities at the DNA, mRNA and protein levels, respectively. These insights into *TaMYC4* TFs gave additional validity to our question of whether *Pgt*-induced *TaMYC4* upregulation is beneficial to the pathogen during their colonization of wheat.

Silencing of the *TaMYC4* Genes

Next, we downregulated *TaMYC4* genes and examined the effect of the pathogens on the host. We knocked down all the three endogenous copies of *TaMYC4* (to avoid functional redundancy) in Chinese Spring (CS) wheat lines, using a barley stripe mosaic virus-induced gene silencing (BSMV-VIGS) assay. A construct containing a 247-bp fragment conserved among the gene homeologs after multiple sequence alignment (Appendix A File 2.S1) was obtained via PCR amplification using primers VIGS-F/R (Appendix A Table 2.S1) and used to silence all the gene copies on the 3 chromosomes (labeled as BSMV:*MYC4*). A construct carrying only the BSMV genome was used as a

no-target control and labeled as BSMV:00. For short, the BSMV-derived construct with no insert was named as γ00, and each BSMV silencing construct was named after the target gene, for example, γ*MYC4*. The concurrent silencing of BSMV inoculum was made by combining the α: β: γtarget transcripts in an equal ratio with excess inoculation buffer (FES). In each assay, twenty wheat seedlings were inoculated with γ*MYC4* or γ00, as a control. Six days post BSMV inoculations (dpbi), viral symptoms were visualized on the newly emerged leaves of plants inoculated with BSMV. At nine dpbi, plants inoculated with BSMV constructs showed viral-symptom-free leaf segments, indicating BSMV induced gene silencing has occurred. Three viral-symptom-free leaf segments were randomly sampled from plants inoculated with γ00 and γ*MYC4* construct for RT-qPCR. The plants were inoculated with *Pgt* race QFCSC immediately. Infection type (IT) observed 14 dpi showed enhanced disease resistance in plants that had *TaMYC4* silenced. Non-silenced plants were susceptible (Figure 2.3a). Transcript abundances of *TaMYC4* were measured through qRT-PCR using primers BN4RTF/R (Appendix A Table 2.S1), which confirmed a 40% reduction of *TaMYC4* in silenced plants relative to the control (Figure 2.3b). Though the reduction in the expression of the three *TaMYC4* homeologous genes resulted in enhanced resistance to *Pgt* QFCSC, we do not know which homeolog or all three of them are critical for the rust to colonize the wheat host successfully.

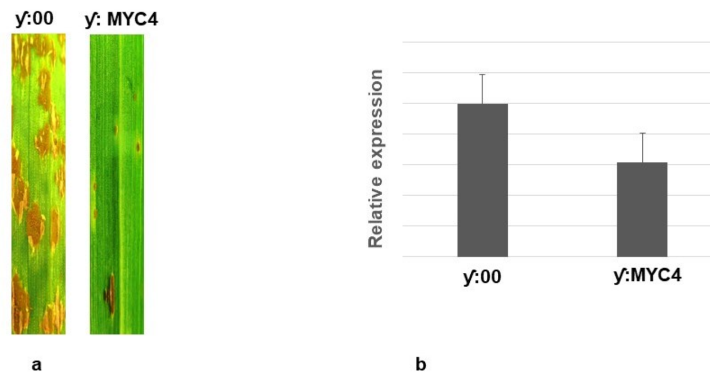


Figure 2.3: BSMV-VIGS of TaMYC4 in Chinese spring (CS) cultivar. a) Infection types of TaMYC4 -silenced plants; b) Relative expression of TaMYC4 in γ:00 and γ:MYC4 silenced plants. The first leaf of each plant was rub inoculated with RNAs of BSMV constructs at the two-leaf stage and then spray inoculated with Pgt race QFCSC at 10 days post BSMV inoculation. The rust infection type was assessed and photographed at 14 days post rust inoculation. CS inoculated with BSMV without wheat target insert was used as a control (labeled as γ:00), CS inoculated with BSMV:MYC4 (labeled as γ:MYC4). RNA was extracted from viral-free leaf segments taken from γ:00 and γ:MYC4 plants prior to rust pathogen inoculation. Transcript abundance was quantified via RT-qPCR. Error bars represent the standard deviation among three biological replicates.

Expression of *TaMYC4* Homeologs during
Stripe Rust Infection

To explore whether all the homeologs or just a specific one are negatively involved in the plant's defense mechanism, we quantified each *TaMYC4* homeolog via Kallisto software using RNA sequence data generated from a study (Dobon et al., 2016) involving Avocet *Yr5* (resistant) and Vuka (susceptible) inoculated with stripe rust pathogen at 0, 1, 2, 3 and 5 dpi and accessed from NCBI. We used just the stripe rust data set since we could not find readily available stem rust generated RNAseq data for the specific time-course study of our interest. Also, the stripe rust study helped us to assess the *TaMYC4* expression pattern in the wheat-*Pst* interaction, which we already examined

beforehand in leaf and stem rusts (Appendix A Figure 2.S1 and Figures 2.1 and 3). FastQC, a quality control tool for high throughput sequence data was used to assess RNA-seq data quality. Overall, FastQC checks on the RNA-seq data indicated good quality features such as Per base sequence quality and Overrepresented sequences. A few measures were red marked poor (Appendix A Figures 2.S2). The RNA-seq was pseudo-aligned and quantified via Kallisto software (Bray et al., 2016). The output of transcript abundance was recorded in transcript per million (tpm). The transcript abundances of the genes of interest (1AL, 1BL, 1DL) were imported using the gene IDs; TRIAE_CS42_1AL_TGACv1_000298_AA0008240.1, TRIAE_CS42_1BL_TGACv1_726352_AA2170300.1, TRIAE_CS42_1DL_TGACv1_061684_AA0201690.1 respectively. Pre-examination of transcript abundance data satisfied normality requirement in R studio. A graph of expression of the gene was plotted using the averages of biological replicates at each time point.

Homeolog *TaMYC4-1BL* was significantly upregulated in the susceptible cultivar at 1 dpi compared with the resistant cultivar. It was also the most expressed of the three homeologs (Figure 2.4). This result led to the search for permanent alteration (mutants) of *TaMYC4-1BL*.

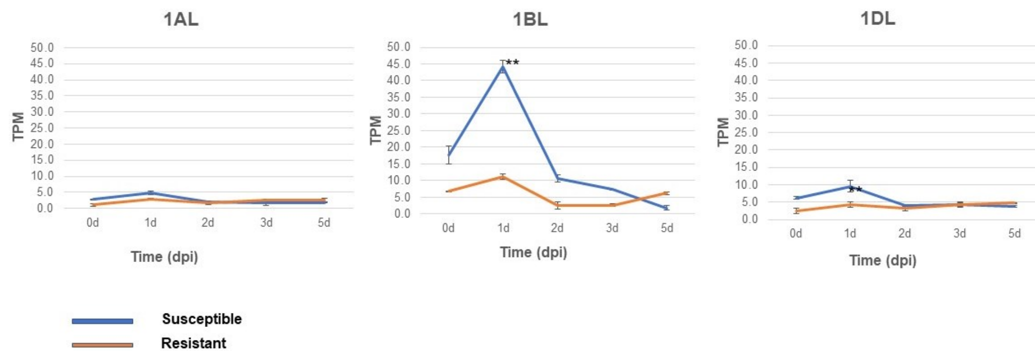


Figure 2.4: Expression of *TaMYC4* (1AL, 1BL and 1DL) homeologs during stripe rust -Avocet *Yr5* (resistant) and Vuka (susceptible) interactions. The RNA-seq data were accessed from the Sequence Read Archive (SRA) of NCBI and quality checks were done on FASTQC. Transcript quantification was carried out via the Kallisto tool and recorded as a transcript per million (TPM) and data analysis was done in R-studio. Error bars represent standard deviation among three biological reps and ** denote statistical significance at the $P \leq 0.01$ levels calculated between each time point and 0 dpi.

Identification of *TaMYC4* Mutants and their Response to Rust Pathogens

Using the *TaMYC4-1BL* cDNA sequence as a query, a BLAST search for matches in the database of wheat-tilling mutant lines revealed more than 40 lines carrying a mutation on the *TaMYC4-1BL* gene. We selected and requested three lines (one heterozygote and two homozygotes) because these mutants have Sorting Intolerant from Tolerant (SIFT) score of 0.0, missense mutation, which a non-conservative amino acid changed and clear location of the mutation for confirmation (Table 2.1). Two of the three mutations were confirmed after genotyping with *TaMYC4 -1BL* specific primers (Appendix A Table 2.S1) for one homozygote, hereafter identified as P12S-MYC4 -1BL or P12S and a heterozygote, identified as M17I-MYC4 -1BL or M17I (Table 2.1).

Table 2.1: Summary of identified *TaMYC4 -1BL* Cadenza mutants

Chromo-some	**Nucleotide change**	**Mutated amino acid position**	**Amino acid changes**	**Type of mutation**	**SIFT score***	**Mutated ID****
1BL	G to A	12	P to S	Homozygote	0	P12S-MYC4 -1BL
1BL	C to T	17	M to I	Heterozygote	0	M17I-MYC4 -1BL

* SIFT score predicts whether an amino acid substitution affects protein function, and ranges from 0 to 1. The amino acid substitution is predicted to be damaging if the score is ≤0.05 and tolerated if the score is >0.05.

**Gene IDs were given based on amino acid change and position of change, gene name and chromosome

The Cadenza P12S mutant showed enhanced resistance to stem rust race TPMKC, leaf rust race PBJJG and one field-collected race of stripe rust (Figure 2.5) to which the wild type background Cadenza was susceptible. Both P12S mutant and wild-type Cadenza were screened by the additional stem and stripe rusts (Table 2.2). Cadenza and both mutants showed resistance to *Pgt* QFCSC and *Pst* race PSTv-4 but the two mutants were as susceptible as Cadenza to *Pgt* races TTKSK and TKTTF, and *Pst* races PSTv-37, PSTv-41 and PSTv-47. The rust screening results indicated the resistance conferred by the mutations in *TaMYC4 -1B* is race specific.

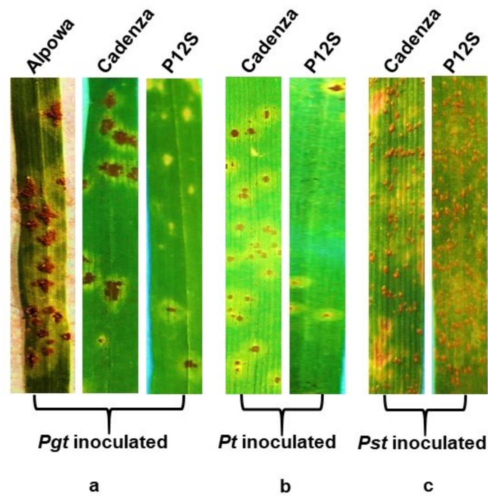


Figure 2.5: Infection types of *TaMYC4 -1BL* mutant. (a) Infection types of wild-types Alpowa and Cadenza and P12S mutant at 14 days post stem rust race TPMKC inoculation. (b) Infection types of wild-type Cadenza and P12S mutant in response to leaf rust race PBJJG at 8 days post inoculation. (c) Infection types of wild-type Cadenza and P12S mutant at 8 days post inoculation of stripe rust.

Table 2.2: Rust evaluations of Cadenza P12S mutant and wild-type cadenza

			Infection Type (IT)			
Pathogen	**Race**	**Isolate**	**Alpowa**	**Avocet**	**Mutant**	**Wild-type**
Pgt	QFCSC		3+	N	1	2
	TMLKC		3+	N	;1-	3+
	TTKSK		N	N	3+	2+3
	TKTTF		N	N	3+	3+
Pt	PBJJG		N	N	;1-	3+
Pst	NK	Field culture	N	N	3	7
	PSTv-4	19WA-200-YrSP	N	8	3	3
	PSTv-37	19ID-11	N	8	8	8
	PSTv-41	19WA-193	N	8	8	8
	PSTv-47	19ID-32	N	8	8	8

Pgt, P. graminis f. sp. Tritici; Pt, Puccinia triticina and Pst, P. striiformis f. sp. Tritici.

For infection type, higher the number, more susceptible plant with a "+" (more than average) or a "–" (less than average) to further quantify the level. A semicolon (;) symbolizes the presence of hypersensitive flecks. N means not tested. Wheat cultivar Alpowa was used as a stem rust susceptible control. The race of stripe rust used is unknown (NK).

Genetic Analysis of the Mutations

To confirm the new rust resistance in the mutant is indeed due to the mutation in *TaMYC4 -1B* and not due to background mutations, we crossed the mutant P12S with a susceptible cultivar Alpowa. F_1 plants were self-pollinated to produce F_2 segregating populations. Meanwhile, the mutant M17I is a heterozygote, self-pollination of the mutant produced a segregating population for genetic analysis. Out of the 150 plants in the first genetics analysis, 148 of them showing clear *Pgt* TPMLKC infection were scored. A resistant to susceptible ratio of 35:113 was observed which fit 1:3 Mendelian phenotypic ratio, χ^2 (1, N = 148) = 0.14, p > .05, indicating resistance was recessive phenotype. Genotyping results from 12 susceptible and 3 resistant plants confirmed the resistant plants had mutant nucleotide (T/A) and the susceptible plants had (C/G) base. KASP assays were designed manually to detect single nucleotide polymorphisms (SNP) for mutant P12S using the *TaMYC4-1BL* scaffold sequence (Appendix A Table 2.S1). The KASP assay result showed 38 monomorphic makers (comprising the susceptible wild-type and resistant mutant alleles), 17 heterozygotes and 41 had none of the markers.

Initial virtual observations of plant height, tiller and seed morphology shows no observable difference between the Cadenza mutants and the wild type (Appendix Figure 2.S4).

Molecular Mechanism of the new Rust Resistance

To understand the genetic mechanism of the new rust resistance, we tested the expression of four *PR* genes, including SA-dependent *PR2* and *PR5* and JA-dependent *PR3* and *PR10* (L. C. Van Loon & Van Strien, 1999) in the wild-type and mutant P12S cadenza with stem rust (race TPMKC) infection using *PR* gene-specific primers (Appendix A Table 2.S1). The basal expression of the *PR* genes at 0 dpi was at a similar level between the wild-type Cadenza and the mutant P12S, and the levels were very minimal, about 0-0.5 relative to the expression of the reference gene actin (Appendix A Figure 2.S3). This means that without rust pathogens attack, PR protein expression in relation to defense pathways is not induced. *PR5* was highly upregulated 12-fold at 1 dpi in P12S relative to its expression in the wild-type Cadenza (Figure 2.6). *PR2, PR3,* and *PR10* had no significant differences in relative expressions across all time points. This suggests that these *PR* proteins are not required in the defense response during this Cadenza P12S mutant-TPMKC interaction.

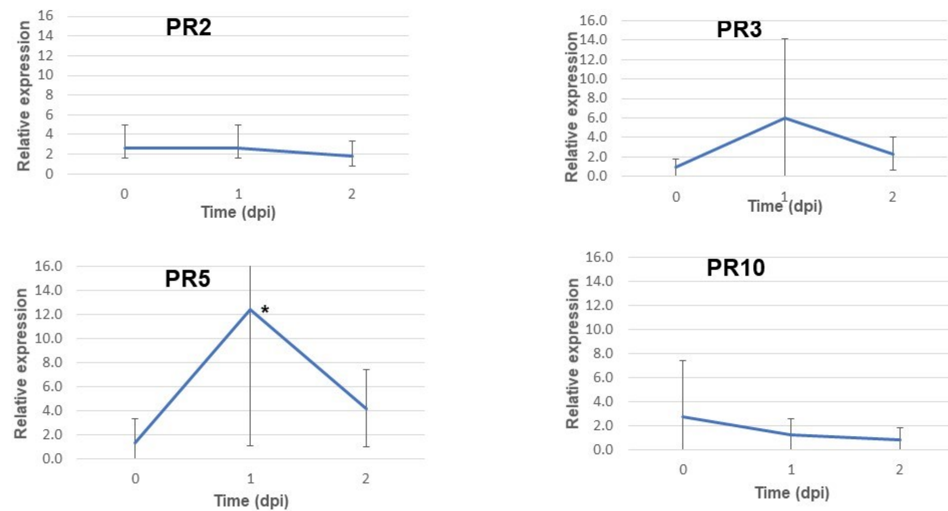


Figure 2.6: Relative expressions of *PR* genes responding to stem rust (race TPMKC) in wild-type and mutant Cadenza. Wild-type Cadenza and mutant P12S were inoculated with stem rust at 2-leaf stage. RNA samples were extracted from the leaf samples collected at three time points. Real-time PCR was used to quantify transcript abundances of the four *PR* genes. The expression of *PR* genes at each time point was expressed in the mutant relative to the level in the wild type. * denote statistical significance at the $P \leq 0.05$ levels compared between wild-type and mutant at each time point. Error bars represent standard deviation computed by the delta method (since the relative expression of a PR gene at a time point depends on the comparison between expression in wild type and cadenza)

DISCUSSION

Wheat *TaMYC4* Gene Aids Rust Pathogens in Host Colonization

The demonstration of the upregulation of *TaMYC4* in stem rust inoculated Alpowa but not in the buffer inoculated Alpowa, provides initial evidence that *TaMYC4* TF was differentially expressed during the pathogen infection (Figure 2.1). It suggests that the effect the *TaMYC4* upregulation was positive to the pathogen and negative to the

susceptible Alpowa wheat host. This led us to hypothesize that *TaMYC4* gene was a target of the pathogen for assistance in colonization. In an example by Chandra et al. (2016), they demonstrated the upregulation of Auxin related proteins, including Auxin-induced protein, 1-aminocyclopropane-1-carboxylate oxidase-like, transcription factor ARF in leaf rust infected susceptible wheat host. They suggested that these genes helped the pathogen through depolymerization of polysaccharides of the host cell wall and an antagonistic effect on the SA defense pathway (Chandra et al., 2016). Host cell wall degradation by pathogens occurs at the initial stages of pathogenesis, suggesting that these host TFs are negatively involved at the preliminary stages of pathogenesis. Like their study, our study showed that the upregulation of *TaMYC4* (*-1B*) occurred at 24 hours post pathogen (hpi) inoculation, as observed in Figure 2.1 and Figure 2.4. This is an indication that *TaMYC4-1B* gene's contribution to the pathogenesis is at the early stage of pathogen colonization. The subsequent drop in the abundance levels of *TaMYC4(-1B)* after 1dpi (from 2dpi) supports this assertion. Fungal hautoria mother cells formation is another vital process at the early stages of pathogenesis. Recent study (Serfling, Templer, Winter, & Ordon, 2016) revealed haustoria mother cells of *P. triticina* formation commences at 24 hpi contrary to previous revelations of 42 hpi (Anker & Niks, 2001; Niks & Dekens, 1991). Haustorium is known to play a vital role in cellular communication between pathogen and host (HEATH, 1997), nutrient acquisition (Hahn & Mendgen, 2001), manipulation of host metabolism, and the suppression of host defenses (Voegele & Mendgen, 2003). This knowledge further strengthens our suspicion of *TaMYC4's* benefit to the pathogen at 1 dpi.

Secondly, we showed that the down-regulation of all the three identical copies of *TaMYC4* reduced Chinese spring's susceptibility to stem rust race QFCSC (Figure 2.3). The relative reduction in *TaMYC4* transcript abundance in the silenced plants (Figure 2.3b) indicates that the observed enhanced resistance is attributable to the *TaMYC4* knockdown. Similar to this result, the silencing of all three wheat homeologs of *COP9* subunit 5-like gene (*TaCSN5*) enhanced resistance to leave rust in Alpowa (Zhang, Wang, Giroux, & Huang, 2017).

Moreover, a recessive loss-of-function mutation in *TaMYC4 -1BL* conferred resistance of the host against the rust pathogens. This provides further support to our claim that the *TaMYC4-1B* gene facilitated infection of rust pathogens in a compatible wheat rust interaction and further explains the significance of *TaMYC4 -1BL* being the most upregulated of three homeologs in Figure 2.4. This is consistent with previous research, which revealed the highest transcription and functionality of a wheat gene homeolog compared with the others (Abdollahi, Kamiya, Kawaura, & Ogihara, 2012; Hu et al., 2013; Nomura, Ishihara, Yanagita, Endo, & Iwamura, 2005).

The three homeologs of *TaMYC4* had bHLH and Leucine-zipper conserved domains characteristic of *MYC* transcription factors. Nonetheless, there were different domains found on individual copies of the gene (Figure 2.2) which may suggest non-redundancy in functions under different conditions. Indeed, several studies (Chaudhary et al., 2009; X. Chen et al., 2011; Hovav et al., 2008; Shitsukawa et al., 2007) have revealed transcriptional divergence among homeologs. Non-redundancy can also arise due to different patterns of expression, not due to the protein sequence. We suggest that

TaMYC4 -1BL gene is acting as a rust pathogen susceptible wheat host gene such that it acts as a factor needed by the pathogens to colonize the host. Pathogen susceptible plant host genes are recessive. The segregation ratio of resistant to susceptible phenotypes confirmed that *TaMYC4 -1BL* confers a recessive phenotype.

Indeed, the upregulation of *PR5* protein in Cadenza P12S mutant at 24 hours post stem rust (race TPMKC) inoculation (Figure 2.6), suggests *PR5* expression is hampered when *TaMYC4 -1BL* transcriptional factor is abundant in the presence of stem rust pathogen. The evidence of *TaMYC4 -1BL* upregulation at the same time point (1 dpi) after rust inoculation further supports this hypothesis. Similar to this result, H. Zhang et al., (2017) established that mutation of *TaCSN5-2A* or *TaCSN5-2D* genes in Alpowa increased PR1 transcription and enhanced resistance to leaf rust (H. Zhang et al., 2017). PR genes are induced by salicylic acid (SA) and jasmonic acid (JA), key signaling molecules of transduction pathways, in induced resistance and plant development (Thomma, Penninckx, Cammue, & Broekaert, 2001). *PR* genes are involved in host defense under different wheat-pathogen race interactions (Zhang, Qiu, Yuan, Chen, & Huang, 2018). PR5 is an SA-dependent molecule that has a thaumatin-like protein function (L. Van Loon, 1982) and has been shown to inhibit the growth of a variety of fungi (Muthukrishnan, Liang, Trick, & Gill, 2001). Strategies of bacteria hijacking plant hormones to manipulate host defense has been well studied. For example, *Agrobacterium tumefaciens* uses its T-DNA to facilitate production host auxin and cytokinin hormones in the formation of crown galls. Various strains of *Pseudomonas syringae* produces coronatine (COR) phytotoxin to manipulate host hormones to enhance bacterial growth

and symptom development (Mittal & Davis, 1995). It was later shown that COR is structurally similar to JA isoleucine hence function in antipathy to SA pathway, which plays a crucial role in defense against this bacterial (Brooks, Bender, & Kunkel, 2005; Browse, 2009). Zheng et al., demonstrated that COR targets host NAC TFs to cause stomata reopening and systemic induced susceptibility (Zheng et al., 2012). We suspect that *PR5* might have been suppressed by rust pathogen effectors using *TaMYC4 -1BL* as a host target gene. The insignificant expression of *PR10* (the other SA-related proteins) in this interaction (Figure 2.6) suggests that for specific signal transduction, separate *PR* proteins may be involved during different plant-pathogen reactions. Also, it has been established that a cross-communication between SA- and JA-dependent defense pathways (Felton & Korth, 2000; Pieterse, Ton, & Van Loon, 2001) which Spoel et al., demonstrated is modulated by *nonexpressor of pathogenesis-related genes 1* (*NPR1*) (Spoel et al., 2003). The relatively lower and insignificant expression levels of the JA-pathway associated *PR* proteins (*PR2* and *PR3*) at 1 dpi (Figure 2.6), alludes to this crosstalk that enables plants to fine-tune their defense reactions depending on the type of stress they encounter.

The bHLH superfamily of transcription factors, including *TaMYC4*, have important regulatory components in transcriptional networks of many developmental pathways (Atchley & Fitch, 1997). In Arabidopsis, *MYC4* TFs are known to bind to G-box of promoters and are involved in JA gene regulation (Niu, Figueroa, & Browse, 2011). Collectively, *MYC4*, *MYC2* and *MYC3* were shown to control JA-dependent responses (Fernández-Calvo et al., 2011). It was demonstrated that *MYC4* could form

complexes with glucosinolate-related MYBs to regulate glucosinolate biosynthesis (Schweizer et al., 2013). A recent study unraveled that a *MYC2*/*MYC3*/*MYC4* -controlled positive-feedback loop transcriptionally regulated spray-induced jasmonate accumulation (Van Moerkercke et al., 2019). These activities of *TaMYC4* in *A. thaliana* give additional credence to our speculation that *TaMYC4* is useful to pathogens in suppressing host defense in the early stages of wheat rust interaction. However, contrary to its implication in JA-pathways in *A. thaliana*, we found its negative function in the SA-dependent pathway as implicated in *PR5* suppression. This emphasizes our knowledge that a gene could function in different pathways in different species and under varying conditions. At this point, this study has not ascertained if this mechanism is similar for all the three rust types and has also not uncovered the detailed mode of action of *TaMYC4 -1BL* during compatible wheat-rust interaction. Hence further study is necessary to address these unknowns.

Cadenza P12S Mutant Showed Race Specific Resistance

Cadenza P12S mutant showed resistance to stem rust pathotypes TPMKC, leaf rust race PBJJG, and a field-collected uncharacterized stripe rust race, but not to stem rusts TTKSK (Ug99) and TKTTF GER. Also, the mutant was susceptible to all the stripe rust races tested except for two of them (Table 2.2). The race-specific resistance shown is partly supported by hypersensitive reaction and moderate resistance shown by the mutant (P12S) at the seedling stage during their interaction with the pathogens. Hypersensitivity is probably the most common race-specific resistance used in breeding and defined as “an active mechanism in which the rapid death of host cells around the point of infection

prevents colonization" (Robinson, 1976). In intermediate resistance, the pathogens penetrate the host and some rust development occurs before incompatible reaction resulting in the reduction of the spread of pathogens and the production of urediospores. The resulting infection types range from 1 to 3 in seedlings as confirmed by our results. We also noticed that the *Pgt* QFCSC and *Pst* PSTv-4 interactions with the mutant and wildtype cadenza enhanced resistance in both plant materials. Wheat cultivar Cadenza is known to possess some rust resistance genes including *Sr9g, Lr13* and *Yr7* (Singh, Park, McIntosh, & Bariana, 2008); and hence, these improved resistances could be attributed to these resistance genes.

An Approach to Create New Resistant Germplasm

Over the years, wheat rust resistance breeding has been focused on using Adult Plant Resistance (APR) and Resistance (R) genes from related species of wheat. While this is very resource consuming, these resistance genes are continually overcome by evolving virulent pathotypes of rust pathogens, particularly because most of these genes confer race-specific resistance. There have considerations for the building resistance gene cassettes to confer efficient resistances against different rust races. However, this requires multiple resistance genes cloning and effective means for inserting multiple genes at a single locus (Ellis, Lagudah, Spielmeyer, & Dodds, 2014). Undoubtedly, host genome editing is becoming very useful in hexaploid bread wheat as exemplified in (Wang et al., 2014). Still, the techniques, type of host gene to modify and the effect of the gene-editing on the host plant are major challenges.

Undeniably, mutation of some plant host S-genes may cause unintended effects due to pleiotropy. For example, the wheat *Lr34* gene causes leaf tip necrosis but also provides resistance to wheat leaf rust (Krattinger et al., 2009). A knockout of *Lr34* to reduce premature senescence of the flag leaf may result in susceptibility to leaf rust as an unintended effect. We observed that mutation in *TaMYC4 -1BL* did not cause any preliminary detrimental effect on the plant height, seed, and tiller morphology (Appendix A Figure 2.S4), but this would be evaluated quantitatively using near-isogenic lines. We suspect this could be because *TaMYC4 -1AL* and *1DL* homeologs' functions compensated for the loss of function in *TaMYC4 -1BL*.

This study has revealed an effective approach for identifying and manipulating a pathogen susceptible wheat gene to create new wheat rust resistance. Targeting and modification of *TaMYC4* using bioinformatics tools, VIGS and an available EMS mutagenized population of *TaMYC4 -1BL* sufficiently enhanced race-specific resistance among the three rust pathogens in bread wheat. With bioinformatics tools and the understanding of the interaction between effector-host genes, it was feasible to identify possible pathogen susceptible host genes. VIGS allows us to examine the importance of the candidate gene in rust disease pathogenesis. A loss-of-function mutation from the available EMS population and near-isogenic lines creation using single seed descent from a single cross between Alpowa and P12S enables us to evaluate the effect of the mutation on the host. This approach provides a means of navigating the challenges associated with germplasm creation and study of gene function in wheat, wheat transformation problems, and concerns with detrimental effects of a host gene mutation.

ACKNOWLEDGMENTS

We wish to thank the NSF BREAD program (Grant No. IOS-0965429) for financial support.

REFERENCES

Abdollahi, P., Kamiya, Y., Kawaura, K., & Ogihara, Y. (2012). Overexpression of Q/q-related homoeoalleles of hexaploid wheat reveals distinct recovery of flower transformation in the apetala2 mutant of Arabidopsis. *Plant Biotechnology*, 12.0224 a.

Anker, C. C., & Niks, R. E. (2001). Prehaustorial resistance to the wheat leaf rust fungus, Puccinia triticina, in Triticum monococcum (ss). *Euphytica, 117*(3), 209-215.

Appels, R., Eversole, K., Stein, N., Feuillet, C., Keller, B., Rogers, J., . . . Wang, L. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science, 361*(6403), eaar7191. doi:10.1126/science.aar7191

Atchley, W. R., & Fitch, W. M. (1997). A natural classification of the basic helix-loop-helix class of transcription factors. *Proceedings of the National Academy of Sciences of the United States of America, 94*(10), 5172-5176. doi:10.1073/pnas.94.10.5172

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol, 34*(5), 525-527. doi:10.1038/nbt.3519

Brooks, D. M., Bender, C. L., & Kunkel, B. N. (2005). The Pseudomonas syringae phytotoxin coronatine promotes virulence by overcoming salicylic acid-dependent defences in Arabidopsis thaliana. *Molecular Plant Pathology, 6*(6), 629-639. doi:10.1111/j.1364-3703.2005.00311.x

Browse, J. (2009). Jasmonate Passes Muster: A Receptor and Targets for the Defense Hormone. *Annual Review of Plant Biology, 60*(1), 183-205. doi:10.1146/annurev.arplant.043008.092007

Campbell, J., Zhang, H., Giroux, M. J., Feiz, L., Jin, Y., Wang, M., . . . Huang, L. (2012). A mutagenesis-derived broad-spectrum disease resistance locus in wheat. *Theoretical and Applied Genetics, 125*(2), 391-404. doi:10.1007/s00122-012-1841-7

Chandra, S., Singh, D., Pathak, J., Kumari, S., Kumar, M., Poddar, R., . . . Mukhopadhyay, K. (2016). De Novo Assembled Wheat Transcriptomes Delineate Differentially Expressed Host Genes in Response to Leaf Rust Infection. *PLoS One, 11*(2), e0148453. doi:10.1371/journal.pone.0148453

Chaudhary, B., Flagel, L., Stupar, R. M., Udall, J. A., Verma, N., Springer, N. M., & Wendel, J. F. (2009). Reciprocal silencing, transcriptional bias and functional divergence of homeologs in polyploid cotton (Gossypium). *Genetics, 182*(2), 503-517.

Chen, J., Clinton, M., Qi, G., Wang, D., Liu, F., & Fu, Z. Q. (2020). Reprogramming and remodeling: transcriptional and epigenetic regulation of salicylic acid-mediated plant defense. *Journal of Experimental Botany*. doi:10.1093/jxb/eraa072

Chen, X., Truksa, M., Snyder, C. L., El-Mezawy, A., Shah, S., & Weselake, R. J. (2011). Three homologous genes encoding sn-glycerol-3-phosphate acyltransferase 4 exhibit different expression patterns and functional divergence in Brassica napus. *Plant physiology, 155*(2), 851-865.

Dobon, A., Bunting, D. C. E., Cabrera-Quio, L. E., Uauy, C., & Saunders, D. G. O. (2016). The host-pathogen interaction between wheat and yellow rust induces temporally coordinated waves of gene expression. *Bmc Genomics, 17*, 380-380. doi:10.1186/s12864-016-2684-4

Dombrecht, B., Xue, G. P., Sprague, S. J., Kirkegaard, J. A., Ross, J. J., Reid, J. B., . . . Kazan, K. (2007). MYC2 Differentially Modulates Diverse Jasmonate-Dependent Functions in Arabidopsis. *The Plant cell, 19*(7), 2225. doi:10.1105/tpc.106.048017

Ellis, J. G., Lagudah, E. S., Spielmeyer, W., & Dodds, P. N. (2014). The past, present and future of breeding rust resistant wheat. *Frontiers in Plant Science, 5*(641). doi:10.3389/fpls.2014.00641

Felton, G. W., & Korth, K. L. (2000). Trade-offs between pathogen and herbivore resistance. *Current Opinion in Plant Biology, 3*(4), 309-314. doi:10.1016/s1369-5266(00)00086-8

Fernández-Calvo, P., Chini, A., Fernández-Barbero, G., Chico, J.-M., Gimenez-Ibanez, S., Geerinck, J., . . . Solano, R. (2011). The Arabidopsis bHLH Transcription Factors MYC3 and MYC4 Are Targets of JAZ Repressors and Act Additively with MYC2 in the Activation of Jasmonate Responses. *The Plant cell, 23*(2), 701. doi:10.1105/tpc.110.080788

Finver, S. N., Nishikura, K., Finger, L. R., Haluska, F. G., Finan, J., Nowell, P. C., & Croce, C. M. (1988). Sequence analysis of the MYC oncogene involved in the t(8;14)(q24;q11) chromosome translocation in a human leukemia T-cell line indicates that putative regulatory regions are not altered. *Proceedings of the National Academy of Sciences, 85*(9), 3052-3056. doi:10.1073/pnas.85.9.3052

Gemayel, R., Chavali, S., Pougach, K., Legendre, M., Zhu, B., Boeynaems, S., . . . Verstrepen, K. J. (2015). Variable Glutamine-Rich Repeats Modulate Transcription Factor Activity. *Molecular cell, 59*(4), 615-627. doi:10.1016/j.molcel.2015.07.003

Hahn, M., & Mendgen, K. (2001). Signal and nutrient exchange at biotrophic plant–fungus interfaces. *Current Opinion in Plant Biology, 4*(4), 322-327.

HEATH, M. C. (1997). Signalling between pathogenic rust fungi and resistant or susceptible host plants. *Annals of Botany, 80*(6), 713-720.

Holleley, C., & Sutcliffe, A. (2009). 96 Well DNA Extraction Protocol. In.

Hovav, R., Udall, J. A., Chaudhary, B., Rapp, R., Flagel, L., & Wendel, J. F. (2008). Partitioned expression of duplicated genes during development and evolution of a single cell in a polyploid plant. *Proceedings of the National Academy of Sciences, 105*(16), 6191-6195.

Hu, Z., Song, N., Xing, J., Chen, Y., Han, Z., Yao, Y., . . . Sun, Q. (2013). Overexpression of Three TaEXPA1 Homoeologous Genes with Distinct Expression Divergence in Hexaploid Wheat Exhibit Functional Retention in Arabidopsis. *PLoS One, 8*(5), e63667. doi:10.1371/journal.pone.0063667

Hussain, R. M. F., Sheikh, A. H., Haider, I., Quareshy, M., & Linthorst, H. J. M. (2018). Arabidopsis WRKY50 and TGA Transcription Factors Synergistically Activate Expression of PR1. *Frontiers in Plant Science, 9*(930). doi:10.3389/fpls.2018.00930

Jackie, C., & Huang, L. (2010). Silencing of multiple genes in wheat using Barley stripe mosaic virus. *Journal of Biotech Research, 2*.

Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B., & Weeks, D. P. (2013). Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic acids research, 41*(20), e188-e188.

Jin, Y., Singh, R., Ward, R., Wanyera, R., Kinyua, M., Njau, P., . . . Yahyaoui, A. (2007). Characterization of seedling infection types and adult plant infection responses of monogenic Sr gene lines to race TTKS of Puccinia graminis f. sp. tritici. *Plant Disease, 91*(9), 1096-1099.

Kazan, K., & Manners, J. M. (2013). MYC2: The Master in Action. *Molecular plant, 6*(3), 686-703. doi:https://doi.org/10.1093/mp/sss128

Krattinger, S. G., Lagudah, E. S., Spielmeyer, W., Singh, R. P., Huerta-Espino, J., McFadden, H., . . . Keller, B. (2009). A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science, 323*(5919), 1360-1363.

Li, T., Liu, B., Spalding, M. H., Weeks, D. P., & Yang, B. (2012). High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nature biotechnology, 30*(5), 390.

Line, R. F., & Qayoum, A. (1992). Virulence, aggressiveness, evolution, and distribution of races of Puccinia striiformis (the cause of stripe rust of wheat) in North

America, 1968-87. *Technical bulletin-United States Department of Agriculture*(1788).

McIntosh, R. A., Wellings, C. R., & Park, R. F. (1995). *Wheat rusts: an atlas of resistance genes*: Kluwer Academic Publishers.

Mittal, S., & Davis, K. R. (1995). Role of the phytotoxin coronatine in the infection of Arabidopsis thaliana by Pseudomonas syringae pv. tomato. *MPMI-Molecular Plant Microbe Interactions, 8*(1), 165-171.

Moore, J. W., Loake, G. J., & Spoel, S. H. (2011). Transcription dynamics in plant immunity. *The Plant cell, 23*(8), 2809-2820.

Muthukrishnan, S., Liang, G. H., Trick, H. N., & Gill, B. S. (2001). Pathogenesis-related proteins and their genes in cereals. *Plant Cell, Tissue and Organ Culture, 64*(2-3), 93.

Niks, R., & Dekens, R. (1991). Prehaustorial and posthaustorial resistance to wheat leaf rust in diploid wheat seedlings. *Phytopathology (USA)*.

Niu, Y., Figueroa, P., & Browse, J. (2011). Characterization of JAZ-interacting bHLH transcription factors that regulate jasmonate responses in Arabidopsis. *Journal of Experimental Botany, 62*(6), 2143-2154. doi:10.1093/jxb/erq408

Nomura, T., Ishihara, A., Yanagita, R. C., Endo, T. R., & Iwamura, H. (2005). Three genomes differentially contribute to the biosynthesis of benzoxazinones in hexaploid wheat. *Proceedings of the National Academy of Sciences, 102*(45), 16490-16495.

Pieterse, C. M., Ton, J., & Van Loon, L. (2001). Cross-talk between plant signalling pathways: boost or burden. *AgBiotechNet, 3*, 1-8.

Robinson, R. A. (1976). Plant Pathosystems. In *Plant Pathosystems* (pp. 15-31). Berlin, Heidelberg: Springer Berlin Heidelberg.

Schweizer, F., Fernández-Calvo, P., Zander, M., Diez-Diaz, M., Fonseca, S., Glauser, G., . . . Reymond, P. (2013). Arabidopsis Basic Helix-Loop-Helix Transcription Factors MYC2, MYC3, and MYC4 Regulate Glucosinolate Biosynthesis, Insect Performance, and Feeding Behavior. *The Plant cell, 25*(8), 3117. doi:10.1105/tpc.113.115139

Serfling, A., Templer, S. E., Winter, P., & Ordon, F. (2016). Microscopic and Molecular Characterization of the Prehaustorial Resistance against Wheat Leaf Rust (Puccinia triticina) in Einkorn (Triticum monococcum). *Frontiers in Plant Science, 7*(1668). doi:10.3389/fpls.2016.01668

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., . . . Qiu, J.-L. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. *Nature biotechnology, 31*(8), 686.

Shitsukawa, N., Tahira, C., Kassai, K.-i., Hirabayashi, C., Shimizu, T., Takumi, S., . . . Murai, K. (2007). Genetic and epigenetic alteration among three homoeologous genes of a class E MADS box gene in hexaploid wheat. *The Plant cell, 19*(6), 1723-1737.

Singh, D., Park, R. F., McIntosh, R. A., & Bariana, H. S. (2008). Characterisation of stem rust and stripe rust seedling resistance genes in selected wheat cultivars from the united kingdom. *Journal of Plant Pathology, 90*(3), 553-562.

Spoel, S. H., Koornneef, A., Claessens, S. M. C., Korzelius, J. P., Van Pelt, J. A., Mueller, M. J., . . . Pieterse, C. M. J. (2003). NPR1 Modulates Cross-Talk between Salicylate- and Jasmonate-Dependent Defense Pathways through a Novel Function in the Cytosol. *The Plant cell, 15*(3), 760. doi:10.1105/tpc.009159

Thomma, B. P. H. J., Penninckx, I. A. M. A., Cammue, B. P. A., & Broekaert, W. F. (2001). The complexity of disease signaling in Arabidopsis. *Current Opinion in Immunology, 13*(1), 63-68. doi:https://doi.org/10.1016/S0952-7915(00)00183-7

Tsuda, K., & Somssich, I. E. (2015). Transcriptional networks in plant immunity. *New Phytologist, 206*(3), 932-947. doi:10.1111/nph.13286

Van Loon, L. (1982). Regulation of changes in proteins and enzymes associated with active defence against virus infection. In *Active defense mechanisms in plants* (pp. 247-273): Springer.

Van Loon, L. C., & Van Strien, E. A. (1999). The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. *Physiological and Molecular Plant Pathology, 55*(2), 85-97. doi:https://doi.org/10.1006/pmpp.1999.0213

Van Moerkercke, A., Duncan, O., Zander, M., Šimura, J., Broda, M., Vanden Bossche, R., . . . Van Aken, O. (2019). A MYC2/MYC3/MYC4-dependent transcription factor network regulates water spray-responsive gene expression and jasmonate levels. *Proceedings of the National Academy of Sciences, 116*(46), 23345. doi:10.1073/pnas.1911758116

Voegele, R. T., & Mendgen, K. (2003). Rust haustoria: nutrient uptake and beyond. *New Phytologist, 159*(1), 93-100.

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C., & Qiu, J.-L. (2014). Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nature biotechnology, 32*(9), 947.

Williamson, M. P. (1994). The structure and function of proline-rich regions in proteins. *The Biochemical journal, 297 (Pt 2)*(Pt 2), 249-260. doi:10.1042/bj2970249

Xie, K., & Yang, Y. (2013). RNA-guided genome editing in plants using a CRISPR–Cas system. *Molecular plant, 6*(6), 1975-1983.

Zhang, H., Qiu, Y., Yuan, C., Chen, X., & Huang, L. (2018). Fine-Tuning of PR Genes in Wheat Responding to Different Puccinia Rust Species. *Journal of Plant Physiology & Pathology, 06*. doi:10.4172/2329-955X.1000178

Zhang, H., Wang, X., Giroux, M. J., & Huang, L. (2017). A wheat COP9 subunit 5-like gene is negatively involved in host response to leaf rust. *Molecular Plant Pathology, 18*(1), 125-133. doi:10.1111/mpp.12467

Zhang, Y., Bai, Y., Wu, G., Zou, S., Chen, Y., Gao, C., & Tang, D. (2017). Simultaneous modification of three homoeologs of Ta EDR 1 by genome editing enhances powdery mildew resistance in wheat. *The Plant Journal, 91*(4), 714-724.

Zheng, X.-y., Spivey, N. W., Zeng, W., Liu, P.-P., Fu, Z. Q., Klessig, D. F., . . . Dong, X. (2012). Coronatine promotes Pseudomonas syringae virulence in plants by activating a signaling cascade that inhibits salicylic acid accumulation. *Cell host & microbe, 11*(6), 587-596.

CHAPTER THREE

A NEW MODE OF NPR1 ACTION VIA AN NB-ARC-NPR1 FUSION PROTEIN REGULATES DEFENSE RESPONSE TO STEM RUST PATHOGEN IN WHEAT

Contribution of Author and Co-Authors

Manuscript in Chapter 3

Author: Xiaojing Wang*

Contributions: Planned and started the research, VIGS assays

Co-Author: Bernard Nyamesorto*

Contributions: conducted some of the knockdown assays, Kallisto transcript quantifications, and analysis, writing and editing of the manuscript

Co-Author: Hongtao Zhang*

Contributions: Characterization of knock-out mutants, Gene quantifications via Real-time PCRs

Co-Author: Yi Luo

Contributions: Identifying the *TaNPR1-7A* alternative-splicing isoform via RT

Co-Author: Xiaoqian Mu

Contributions: Identifying the *TaNPR1-7A* alternative-splicing isoform via cloning

Co-Author: Fangyan Wang

Contributions: SA, JA assays

Co-Author: Zhensheng Kang

Contributions: Designed the research jointly

Co-Author: Evans Lagudah

Contributions: Planned and designed the research jointly and editing the manuscript

Co-Author: Li Huang

Contributions: Planned and designed the research and wrote the manuscript

*Contributed equally.

Manuscript Information

Xiaojing Wang, Hongtao Zhang, Bernard Nyamesorto, Yi Luo, Xiaoqian Mu, Fangyan Wang, Zhensheng Kang, Evans Lagudah and Li Huang

New Phytologist

Status of the manuscript:

____ Prepared for submission to a peer-reviewed journal

_X___ Officially submitted to a peer-reviewed journal

____ Accepted by a peer-reviewed journal

____Published in a peer-reviewed journal

SUMMARY

NPR1 has been found to be a key transcriptional regulator in some plant defense responses. There are nine *NPR1* homologs (*TaNPR1*) in wheat, but little research has been done to understand the function of those *NPR1*-like genes in wheat defense response against stem rust (*Puccinia graminius* f. sp *tritici*) pathogens.

We used bioinformatics and reverse genetics approaches to study the expression and function of each TaNPR1. We found six members of TaNPR1 located on homeologous group 3 chromosomes (designated as *TaG3NPR1*) and three on homeologous group 7 chromosomes (designated as *TaG7NPR1*). Down-regulation of all nine of the *TaNPR1* homologs via virus induced gene co-silencing resulted in enhanced resistance to stem rust. More specifically down-regulating *TaG7NPR1* homeologs or *Ta7ANPR1* expression resulted in stem rust resistance phenotype. In contrast, knocking down *TaG3NPR1* alone did not show visible phenotypic changes to the rust pathogen.

An EMS-mutagenized population was used to screen the A genome mutation. Mutations derived in *Ta7ANPR1* enhanced resistance to stem rust. Transcriptome analysis showed that the *Ta7ANPR1* locus produced two alternative spliced variants under pathogen inoculated conditions.

Our research revealed a new mode of NPR1 action in wheat at the *Ta7ANPR1* locus regulating defense response to stem rust infection.

INTRODUCTION

Plants constantly battle with a variety of pathogens in the environment via their complex and effective innate immune systems (Spoel & Dong, 2012). Hypersensitive response (HR) is one of the strategies used to defend against biotrophic pathogens by which rapid programmed cell death occurs immediately surrounding the infection sites (Morel & Dangl, 1997) to restrict the pathogens from further spreading and replication. HR also activates a series of signals which are transduced to remote regions of the plants and generate systemic acquired resistance (SAR) with a broad-spectrum resistance to subsequent pathogen attacks (Pajerowska-Mukhtar, Emerine, & Mukhtar, 2013).

A central positive regulator of SAR signaling is NPR1 *(Non-expresser of Pathogenesis-Related genes 1*), also known as *Non-Immunity 1* [*NIM1*]). The gene is essential for transducing the SA signal to activate *Pathogenesis-Related* (*PR*) gene expression (Cao, Bowling, Gordon, & Dong, 1994; Dong, 2004). In addition, NPR1 is required by diverse immune signaling pathways, including basal defense, effector-triggered immunity (ETI) and induced systemic resistance (Pajerowska-Mukhtar et al., 2013; Rate & Greenberg, 2001; Shirano, Kachroo, Shah, & Klessig, 2002; Spoel et al., 2003). NPR1 also mediates crosstalk between SA- and JA-mediated signaling pathways (Spoel et al., 2003). The plant-specific transcription factor WRKY70 is identified as a common component downstream of NPR1 in both SA- and JA-mediated signal pathways (J. Li, Brader, & Palva, 2004). *WRKY70* expression is activated by SA and repressed by JA.

NPR1 contains an ankyrin repeat domain and a broad complex, tramtrack, and bric-a`-brac/poxvirus and zinc-finger (BTB/ POZ) domain (Aravind & Koonin, 1999; Cao, Glazebrook, Clarke, Volko, & Dong, 1997). Since the first cloning of *Arabidopsis thaliana NPR1* in 1997 (Cao et al., 1997), a significant amount of work has been done to understand the mode of NPR1 action. In the absence of infection, or at low concentration of SA, NPR1 predominantly exists as oligomers through intermolecular disulfide bonds and retained in the cytoplasm. After a pathogen challenge, with elevated SA level, NPR1 converts to a monomeric state by reduction of the redox-sensitive disulfide bonds. It is then translocated to the nucleus, where NPR1 physically interacts with TGA-bZIP transcriptional factors and activates the expression of defense response genes (Mou, Fan, & Dong, 2003). Nuclear accumulation of NPR1 is needed for basal defense gene expression and resistance, whereas its subsequent turnover is required for establishing SAR (Spoel et al., 2009). Two NPR1 paralogues, NPR3 and NPR4, are required to be the SA receptors (Fu et al., 2012). Both NPR3 and NPR4 contain the BTB domain and ankyrin repeats, which are typical adaptors for CUL3 substrate. Either NPR3 or NPR4 directly bind with SA thus modulating their interactions with NPR1 that result in NPR1 degradation through CUL3 mediated ubiquitination (Fu et al., 2012; Moreau, Tian, & Klessig, 2012).

Overexpression of *NPR1* orthologs in other plants, for example, in apple or grapevine, has been shown to provide broad-spectrum resistance (Le Henanff et al., 2011; Malnoy, Jin, Borejsza-Wysocka, He, & Aldwinckle, 2007). Also, in Gladiolus, GhNPR1 was shown to play a pivotal role in the SA-dependent systemic acquired resistance

(Zhong et al., 2015). In monocots, overexpression of a rice NPR1 homolog led to constitutive activation of defense response and hypersensitivity to light (Chern, Fitzgerald, Canlas, Navarre, & Ronald, 2005). However, the regulation of defense gene induction between rice and *Arabidopsis* is quite different (Silverman et al., 1995). In wheat, overexpression of *Arabidopsis NPR1* (*AtNPR1*) enhanced resistance to *Fusarium graminearum*, a necrotrophic pathogen causing wheat head blight (Makandar, Essig, Schapaugh, Trick, & Shah, 2006). Elevated *PR1* gene induction in the *AtNPR1* transgenic wheat lines suggested that a similar SA-dependent *NPR1*-mediated defense pathway exists in wheat. A study in wheat-stripe rust pathogen *Puccinia striinformis* f. sp. *tritici* (*Pst*) interaction indicated that *Pst* effectors interact with wheat NPR1 during infection, suggesting a role of wheat NPR1 during defense response to rust pathogens (X. Wang et al., 2016).

There are three fungal pathogens from the genus Puccinia that cause wheat rust diseases, namely leaf rust (*P. triticina)*, stem rust (*P. graminis* f. sp. *tritici)*, and stripe rust (*P. striiformis* f. sp. *Tritici).* These three diseases combined can cause estimated annual losses of $2~5 billion to wheat production worldwide (http://www.usda.gov/nass) depending on the varieties grown and developmental stage when infection occurred. Wheat rust pathogens are biotrophs that only survive in living cells and sequester nutrients from their host via haustoria (Peter N Dodds, Lawrence, Catanzariti, Ayliffe, & Ellis, 2004). HR is the most common phenotype observed among resistant wheat lines. In this study, we aimed to identify all the wheat homologs of *NPR1* (designated as *TaNPR1*) and explore their roles in the host defense response to rust pathogens.

MATERIALS AND METHODS

Plant Materials

Alpowa (PI 566596), a soft, white, spring wheat cultivar, was obtained from the USDA National Plant Germplasm System (NPGS). Chinese Spring (CS) and CS+Sr33 were obtained from Dr. Evans Lagudah at Commonwealth Scientific and Industrial Research Organization (CSIRO). Chinese Spring nulli-tetra lines and deletion lines were provided by Dr. Bikram S. Gill at Kansas State University and Dr. Adam Lukazewski at UC-Riverside. The EMS mutagenized Alpowa population was generated by the Giroux lab at Montana State University (Feiz, Beecher, Martin, & Giroux, 2009).

Mapping populations consist of 400 F_2 lines derived from the cross between the Cadenza mutant R805Q and Alpowa.

Sequence Analysis

All BLAST and all sequence downloads were conducted using the IWGSC (https://urgi.versailles.inra.fr/blast_iwgsc/blast.php) resources. The Softberry database (http://linux1.softberry.com/berry.phtml?topic=fgenesh&group=programs&subgroup=gfind) was used for the gene prediction. DNASTAR software (www.dnastar.com) was used to analyze the wheat-leaf rust RNA-sequence. InterProScan (http://www.ebi.ac.uk/InterProScan/), Pfam (http://pfam.xfam.org/), and PROSITE Scan (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_prosite.html) were used to predict the conserved domains and motifs. Multiple sequence alignments were created with ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/).

Plant Growth Conditions

For the race specificity screens, seeds were directly planted in 4-inch pots (5 seedlings/pot) containing SunGro Horticulture Sunshine mix (HeavyGardens Company, Denver, CO). Before inoculation, all the seedlings were grown in a growth room of the Plant Growth Center at Montana State University (PGC-MSU) under the following conditions: 22°C/14°C day/night temperatures and a 16-h photoperiod. Plants were watered every day and fertilized with Peters General Purpose Plant Food (Scotts-Miracle-Gro Company, Marysville, OH), at the concentration of 150 ppm every other day.

Pathogens

The *Pgt* race QFCSC was used for the stem rust assay and was provided by Dr. Yue Jin from Cereal Disease Laboratory, USDA-ARS, St. Paul, MN. Leaf rust race *Pt* PBJJG was provided by Dr. Robert Bowden, USDA-ARS Manhattan, KS and is maintained at Montana State University. Stripe rust isolates were from the Plant Pathology Research Center at Yangling, Shaanxi, in China.

Pathogen Inoculation

For leaf and stem rust evaluations, seedlings were inoculated with either *Pt* PBJJG or *Pgt* QFCSC at the 2-leaf stage. A video protocol detailing this process can be accessed at http://plantsciences.montana.edu/facultyorstaff/faculty/huang/lihuang.html. Urediospores were mixed in Soltrol 170 Isoparaffin (Chempoint, Bellevue, WA) at a concentration of 0.5mg/mL. The suspension was sprayed onto the leaves using a Badger 350 airbrush gun and Propel propellant (Badger Air-Brush Company, Franklin Park, IL). The inoculated seedlings were then placed in a Percival I-60D dew chamber (Percival

Scientific Inc., Perry, IA) with an ambient air temperature of 15-17°C for leaf rust and 19-22 °C for stem rust, respectively. After 24 hours of incubation, an additional 3 h high humidity and light intensity conditions were added for stem rust inoculated plants. All inoculated plants were then placed back in the growth chamber or greenhouse. Disease responses were assessed when rust symptoms were fully expressed on Alpowa 8-22 days post fungal pathogen inoculation (dpfi) using the seedling 0-4 IT scale (McIntosh, Wellings, & Park, 1995; Stakman, Stewart, & Loegering, 1962). In detail, the scale is as follows: IT0: no visible uredia; IT; hypersensitive flecks; IT1: small uredia with necrosis; IT2: small to medium-sized uredia with green islands surrounded by necrosis; IT3: medium-sized uredia without necrosis; IT4: large-sized uredia without necrosis. The variations within each class are indicated by the use of – (less than average for the class) and + (more than average for the class). When variable reactions were observed, IT ranges are listed from lowest to highest.

Stripe Rust Inoculation and Assessment

Stripe rust inoculations were conducted at the Institute of Plant Pathology of Northwest A & F University, China. Freshly collected urediospores were applied with a paintbrush to the surface of primary leaves of 7-day-old wheat seedlings. After inoculation, plants were incubated for 24 hours in dark in a 100% humidity dew chamber and were subsequently transferred to a growth chamber with a 16-hour photoperiod. Stripe rust infection types were assessed based on a 0 (immune)-9 scale (highest susceptible) (Line & Qayoum, 1992).

Virus-Induced Gene Silencing

The BSMV vectors utilized in these experiments were obtained from Dr. Andrew O. Jackson at UC Berkeley. The fragments used to silence *TaG7NPR1*, *TaG3NPR1* and *Ta7ANPR1* were generated by PCR amplification from two synthesized oligonucleotides primers containing 10 overlapping base pairs at the 3' terminus. Overlap Extension PCR amplification of dsDNA fragment using the program as follows： 8 min at 95°C, followed by 34 cycles of 30 s at 95°C, 30 s at 32°C and 40S at 72°C, 2 min at 72°C. The target fragments were inserted into the modified γ vector ready for direct PCR cloning described by Campbell and Huang (Campbell & Huang, 2010). Infectious RNA transcripts were synthesized in vitro using T7 RNA polymerase (New England Biolabs, Ipswich, MA) from linearized α, β, and γ plasmids. The BSMV inoculum was prepared with 1µl of each of the in vitro transcription reactions and 22.5 µl inoculation FES buffer. The inoculum was then used to rub-inoculated the first leaf of the two-leaf stage plants. For simplicity, the BSMV-derived construct with no insert was named as γ00, and each BSMV silencing construct was named as γtarget. For example, a BSMV silencing construct carried a 185-bp fragment of the wheat PDS gene was named as γPDS. The concurrent silencing BSMV inoculum was made by combining the α:β:(γtarget1:γtarget2) transcripts in a 2:2:(1+1) ratio with excess FES. For example, silencing multiple genes both PDS and G7ANPR1, BSMV inoculum was made by combining an equimolar ratio of α, β, and (γPDS: γG7ANPR1) at a 2:2:2(1+1) ratio with excess inoculation buffer (named as FES) containing a wounding agent.

Gene Expression Analysis by RT-qPCR and
RNA-seq

The expression of the genes targeted for silencing was quantified by comparative QRT-PCR. Sampled tissues for the time-course study were immediately frozen in liquid nitrogen and stored at -80°C prior to extraction of total RNA. Three independent biological replications were performed for each experiment. The time course of *G7ANPR1* gene expression was assessed by RT-qPCR. Wheat cultivar CS and CS+*Sr33*, was inoculated with the *Pgt* race QFCSC suspended in the inoculation buffer Soltrol 170 Isoparaffin (Chempoint, Bellevue, WA) and meanwhile, inoculation with Soltrol 170 Isoparaffin alone was used as mock control. Total RNA was isolated and treated with DNase I on column using the Qiagen RNeasy Plant Mini Kit (Qiagen, Valencia, CA) following the manufacturer's instruction. The quality and concentration of total RNA were assessed via agarose gels and $260/280_{ABS}$ measurements on a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific Inc., Wilmington, DE). Threshold values (Ct) generated from the CFX96 real-time PCR detection system (Bio-Rad, Hercules, CA) using iScript One-Step RT-PCR Kit with SYBR Green (Bio-Rad, Hercules, CA) following the manufacturers recommended protocol. We used gene-specific primers (Livak & Schmittgen, 2001) and relative gene expression $2^{-\Delta Ct}$ method for gene quantification. The amounts of RNA in each reaction were calculated using the average ΔCt normalized to three reference genes 18S rRNA, Actin and Glyceraldehyde 3-phosphate dehydrogenase (GAPDH). Primers specific for each gene are listed in Table S2. Each reaction was conducted with three technical replicates of the three biological replicates. So, relative expression of *G7ANPR1* gene is presented as expression level of

this gene in the leaf rust inoculated plants relative to that in the control (buffer Soltrol170 Isoparaffin inoculated plants). Standard deviations were calculated among different biological replicates. Mean relative expressions were calculated using the ΔCt method between biological replicates ± standard deviation.

Wheat cultivar CS and CS+*Sr33*, were inoculated with BSMV inoculum combining an equimolar ratio of α, β, and γ transcripts suspended in the inoculation buffer containing a wounding agent (FES), and inoculation with FES buffer alone was used as a mock control. All the leaf tissues snap-frozen in liquid nitrogen and stored at -80°C until the RNA isolation. The RT-qPCR was conducted similarly with the time course stem rust inoculation study.

Expression abundance of *TaG7NPR1* homeologs were based on RNA sequence data already available in the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI) via Kallisto software (Bray, Pimentel, Melsted, & Pachter, 2016). Sequence quality control checks were conducted via FastQC tool in interactive mode. Paired-end sequences were split using fastq-dump prior to transcript quantification. Wheat transcriptome data was downloaded from International Wheat Genome Sequencing Consortium (IWGSC)_v1.1, indexed and reads are pseudoaligned. Transcript abundance was then quantified and recorded in transcript per million (tpm).

Mutant Screen

The mutagenized population was generated by ethyl methane sulfonate (EMS) (Feiz et al., 2009). The population was allowed to self-pollinate and advanced to M_8 generation. The primers used in mutant identification were listed in Table S2.

For the A genome mutation screening, the A genome-specific primers G7A-MF + G7A-MR were designed to locate in the deletion variations among the A, B, and D genomes to ensure the A genome specificity (Appendix B Figure 3.S1). To detect the A genome mutation, G7A-MF + G7A-MR used to screen the EMS induced population first. The PCR products were then purified using the QIAGEN gel purification kit (Valencia, CA), sequenced and compared. First, the sequence from wild type Alpowa was compared with the *G7ANPR1* gene sequences of Chinese Spring from IWGSC, and then sequences from individual mutagenized lines were compared with the sequence of the wild type Alpowa for mutation identification.

SA/JA Level Analysis with the LC-MS

The extraction of SA/JA was according to Wang's method (L. Wang et al., 2017). Frozen samples were then ground in liquid N_2 with mortar and pestle. An amount of about 200 mg fresh leaves was extracted with 750 µL MeOH–H2O–HOAc (90:9:1, v/v/v) and centrifuged for 1 min at 10,000 rpm. The supernatant was collected, and the extraction was repeated twice. Pooled supernatants were dried in N_2, resuspended in 1000 µL of pure chromatographic grade MeOH, and finally filtered with a Millex- HV 0.22 µm filter from Millipore (Bedford, USA). Quantitation was done by the standard addition method by spiking control plant samples with SA and JA solutions (ranging from 50 to 1000 ng ml^{-1}), and extracting as described above. Analyses were carried out using an LC-30A+TripleTOF5600+ (AB SCIEX, Singapore) machine in Life Science Instrument Shared Platform of Northwest A & F University, China.

RESULTS

<u>Identification of Wheat *NPR1* Homologs</u>

To identify *NPR1*-like genes in bread wheat (*Triticum aestivum*, L.), a cDNA of *NPR1*-like sequence (named as W3SNPR1) was used as a probe to search its homologs via genomic blot ('Southern') hybridization. W3SNPR1 shares a 99% sequence identity to another mRNA sequence deposited in NCBI (accession XM_020328292.1) predicted as a BTB/POZ and ankyrin repeat-containing NPR1-like protein amplified from *Aegilops tauschii* (D genome donor of bread wheat). Nine hybridization fragments were detected, and the chromosome location of each fragment was determined using the wheat Chinese Spring (CS) nulli-tetrasomic (NT) lines (Figure 3.1). In each NT line, a pair of chromosomes was missing and compensated by a pair of its homeologous chromosomes. For example, in N3AT3B, the two 3A chromosomes are missing and compensated by four 3B chromosomes. When a fragment is missing in N3AT3B compared to the rest of the NT lines, it suggests the missing fragment is located on 3A chromosomes. By such an analysis, two fragments were assigned to each of the chromosomes 3B, 3D and 7A and one was assigned to each of the chromosomes 3A, 4A (7B translocated region) and 7D (Figure 3.1). The results revealed that all the wheat *NPR1* (referred as *TaNPR1* thereafter) homologs are located on six chromosomes of the two homeologous groups: 3A/3B/3D of group 3 (designated as *TaG3NPR1*) and 7A/4A/7D of group 7 (designated as *TaG7NPR1*). One *TaNPR1* homolog was found on chromosome 4A instead of 7B due to the ancient translocation between 4AL and 7BS in the tetraploid progenitor of

hexaploid wheat (Devos, Dubcovsky, Dvorak, Chinoy, & Gale, 1995; Liu, Atkinson, Chinoy, Devos, & Gale, 1992).

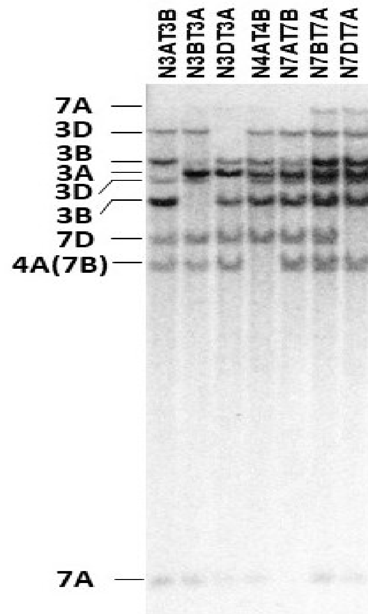


Figure 3.1: Genomic blot (Southern) hybridization result from genomic DNAs of seven wheat NT lines and W3SNPR1 as a probe. Arrows indicate the chromosome location of each fragment.

The sequence of W3SNPR1 was then used to BLAST search the International Wheat Genomic Sequence Consortium (IWGSC-CS RefSeq v1.0) database. Nine *NPR1*-like genes were identified. Chromosomes 3A, 3B and 3D each encode a gene that resembles a classical type of NPR1 protein with domains of BTB, ankyrin repeat and a highly conserved NPR1-like C terminal (Figure 3.2). In addition, a gene encoding for an integrated protein of a kinase fused with NPR1 is found about 17~113 kb proximal to the classical TaG3NPR1 in the same orientation on each of chromosomes 3B and 3D. At the

same homeologous locus of 3A there is only the C-terminal part of *NPR1*-like gene left, the sequence for protein kinase and most of the NPR1 domains are missing (Figure 3.2), explaining why only one fragment was detected on chromosome 3A (Figure 3.1). Additional search on the entire 3A sequences revealed no hits to the DNA sequence corresponding to the integrated kinase domain of the TaG3NPR1.

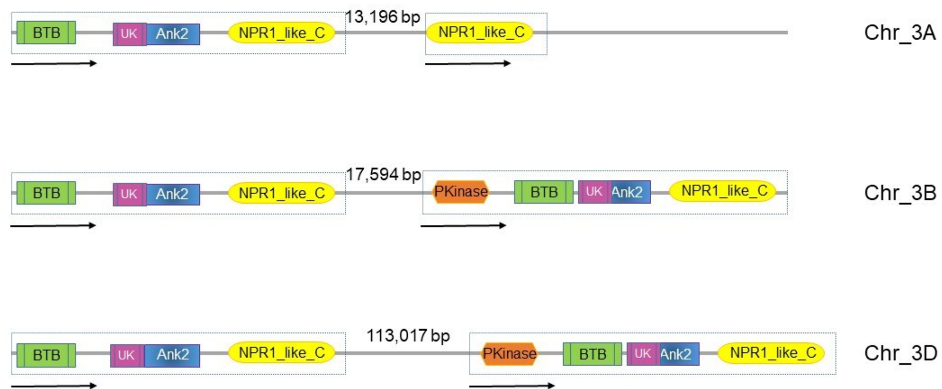


Figure 3.2 Arrangement and structures of six *NPR1* homologs in wheat group 3 chromosomes. One dashed box represents one predicted gene and the arrow under each box indicates the orientation of the gene. The distance between the two *NPR1* homologs in the same chromosome is indicated in base-pair but not according to scale in the figure. Each colored box represents a region of a known functional domain. BTB: Broad complex, Tramtrack, and Bric-a`-brac; UK: DUF3420 Unknown domain; Ank: Ankyrin repeats; NPR1_like_C: A region conserved at the NPR1 c-terminal; PKinase: Protein Kinase. The sizes of the domains are not to scale.

On 7A, 4A (7B) and 7D, each chromosome carries a different type of integrated NPR1. The N-terminal regions of the TaG7NPR1 proteins have domains of two consecutive DNA binding sites for MYC4 transcription factor and an NB-ARC (Figure

3.3). In close proximity (650~3000bp) to each of the integrated TaG7NPR1, a gene encoding for a CC-like + an NB-ARC domain is found in the opposite orientation (Figure 3.3). Similar *TaG7NPR1* genes in the same arrangement with a CC+NB-ARC were found in the A genome donor of *T. urartu* (http://plants.ensembl.org/Triticum_urartu/Info/Index) and *Ae. tauschii* (http://plants.ensembl.org/Aegilops_tauschii/Info/Index).

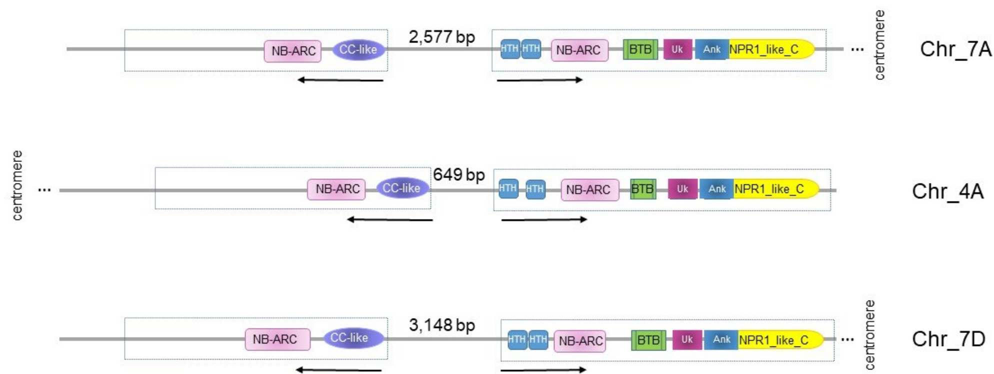


Figure 3.3 Arrangement and structures of three *NPR1* homologs and three *NB-CC*-like genes in wheat chromosomes 7A, 7D and 4A(7B). One dashed box represents one predicted gene and the arrow under each box indicates the orientation of the gene. The distance between the two genes in the same chromosome is indicated in base-pair but not according to scale. Because the orientation of *NPR1* homolog and the *NB-CC*-like genes on chromosome 4A is opposite to the genes on chromosome 7A and 7D, so the location of each chromosome's centromere is provided as a reference. BTB: Broad complex, Tramtrack, and Bric-a`-brac; UK: DUF3420 Unknown domain; Ank: Ankyrin repeats; NPR1_like_C: A region conserved at the NPR1 c-terminal. The sizes of the domains are not in scale.

Silencing of the *TaNPR1* Genes

After cloning of *Sr33* (Periyannan et al., 2013), we sought to test the involvement of the *TaNPR1* genes in the *Sr33*-mediated defense response in wheat; we knocked down the endogenous *TaNPR1* genes in two wheat lines, CS and CS+*Sr33*, using a barley stripe mosaic virus-induced gene silencing (BSMV-VIGS) assay. Several silencing constructs were made for the assay. A construct containing a 170-bp fragment conserved among the *TaG3NPR1* was used to silence all six *NPR1* homologs on group 3 chromosomes, and labeled as BSMV:G3. Another distinct construct containing a 170-bp fragment conserved among the *TaG7NPR1* was used to silence the *NPR1* homologs on chromosomes 7A, 4A(7B) and 7D, and labeled as BSMV:G7. The sequences (Appendix B Table 3.S2) and locations (Appendix B File 2.S2) of the primers are provided as supporting information. A construct carrying only the BSMV genome was used as a non-target control, and labeled as BSMV:00, and a construct carrying a 183-bp phytoene desaturase (*PDS*) gene was used as a non-target control for the assay and labeled as BSMV:PDS. Three rounds of silencing assays were conducted: Round 1, silence all the homologs simultaneously via co-inoculation of two silencing constructs of BSMV:G3 and BSMV:G7 (labeled as BSMV:G3+G7); Round 2, silence only one group of *TaNPR1* at a time with inoculation of either BSMV:G3 or BSMV:G7; Round 3, silence a specific gene when needed. In each assay, CS and CS+*Sr33* seedlings inoculated with BSMV:00, BSMV:PDS and only the inoculation buffer (mock) were included as controls. Six days post BSMV (dpb) inoculations, viral symptoms were visualized on the newly emerged leaves of plants inoculated with BSMV. At nine dpb, plants inoculated with BSMV:PDS started to show photo-bleaching phenotype, and plants inoculated with other BSMV constructs showed

viral-symptom-free leaf segments,. Three viral-symptom-free leaf segments were randomly sampled from plants inoculated with each targeting construct to check the expression level of the target genes through quantitative real-time PCR analysis with corresponding primers (Appendix B Table 3.S2). The results confirmed that about 30% reduction in relative expression of *TaG3NPR1* and *TaG7NPR1* (Table 3. 1) indicating BSMV induced gene silencing has been initiated. Stem rust pathogen *Puccinia graminis* f. sp. *tritici* (*Pgt*) race QFCSC was inoculated at 10 dpb. CS controls showed susceptible infection type scored as infection type (IT) 3 at 14 days post fungus inoculation (dpfi) (Figure 3.4). Resistant control CS+*Sr33* plants displayed a resistant phenotype score as IT1. A similar level of resistance was observed in CS+*Sr33* when both *TaG3NPR1* and *TaG7NPR1* were knocked down. Surprisingly, CS inoculated with either BSMV: G3+G7 or BSMV:G7 showed an enhanced level of resistance to the pathogen on the silenced leaf segments (Figure 3.4), suggesting knocking down *TaG7NPR1* genes enhanced CS resistance to the pathogen. When the six copies of *TaG3NPR1* were silenced in CS and CS+*Sr33*, infection types of both genotypes were the same as their corresponding none-silenced or mock controls, suggesting the *Sr33*-mediated resistance is *TaG3NPR1* independent.

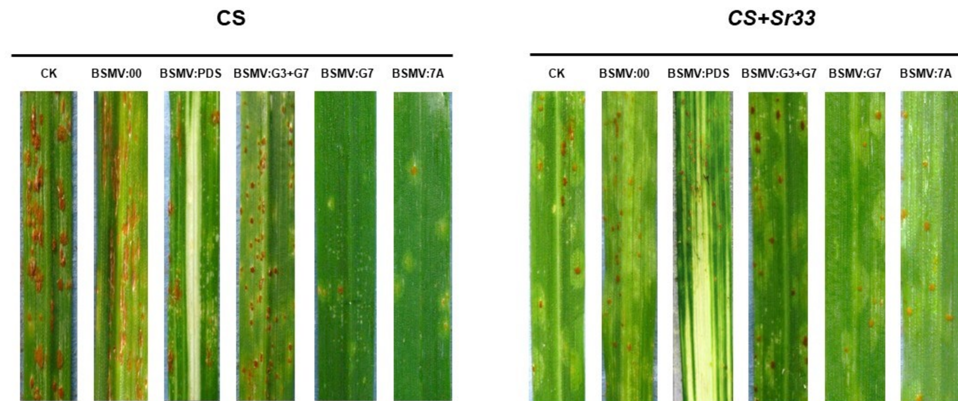


Figure 3.4 Infection types of *Sr33* near isogenic lines in Chinese Spring (CS) background when challenged with stem rust race QFCSC 14 days post inoculation. Leaves labeled with BSMV: G3+G7; BSMV: G7 and BSMV:7A were from the plants that *TaNPR1* of groups 3 & 7, only group 7 or only chromosome 7A were silenced, respectively. BSMV:00 plants were inoculated with BSMV without target gene. BSMV: PDS plants were inoculated with BSMV plus PDS gene. CK plants were not inoculated with BSMV, are the mock control.

TaG7NPR1 includes three homologs, *Ta7ANPR1*, *Ta4ANPR1* and *Ta7DNPR1*. To test which homeolog was vital for the enhanced resistance in CS, we aimed to silence each homeolog one at a time. To silence *Ta7ANPR1* specifically, a region-specific to this gene was synthesized using two overlapped oligos of G7AoligoF and G7AoligoR listed in Table S1 and was labeled as BSMV:7A. A similar approach was used for down-regulating *Ta4ANPR1* and *Ta7DNPR1* specifically. Only the *Ta7ANPR1* silenced plants showed enhanced resistance (Figure 3.4) whereas the *Ta4ANPR1* or *Ta7DNPR1* silenced plants had no changes of infection types in response to *Pgt* QFCSC (Data not shown). Real-time PCR assays revealed the transcript abundance of each target gene in silenced

CS plants was reduced about 20~41% (Table 3.1). The experiments suggested that the *Ta7ANPR1* gene is the one negatively involved in the defense response to *Pgt* QFCSC.

Table 3.1: qRT-PCR analysis of *TaNPR1*-like gene expression in silenced plants

Silencing constructs	NPR1 transcripts measured	Relative expression* in CS			Average	SD
		Exp1	Exp2	Exp3		
BSMV:(G3+G7)	Group3	0.76	0.70	0.76	0.74	0.03
	Group7	0.67	0.68	0.70	0.68	0.01
BSMV:G7	Group7	0.67	0.7	0.71	0.69	0.02
BSMV:G7A	NPR1-7A	0.69	0.59	0.8	0.69	0.08

*Relative expression was calculated by dividing the expression value of the target gene in silenced plants by the control plants infected with BSMV: 00. Each number is an average of triplicate.

Expression of *TaG7NPR1* Post Rust
Inoculations

Time course expression of the three *TaG7NPR1* genes were analyzed using previously produced RNA seq datasets. Notably, *Ta7ANPR1* was the only one among the three genes showing differential expression during stem rust infection in wheat plants (Appendix B Figure 3.S2). Transcript abundances of *Ta4ANPR1 a*nd *Ta7DNPR1* were low and unchanged during the time courses post three rust inoculation in both compatible and incompatible interactions (Appendix B Figure 3.S2 and S3).

To confirm the expression of *Ta7ANPR1* via real-time PCR, the transcript of the gene in wheat line CS was analyzed. In the absence of pathogen infection conditions, only a 5,785-bp cDNA (labeled as Ta7ANPR1 cDNA-1) was amplified (Figure 3.5). The protein encoded by this mRNA has only 881 amino acids with a stop codon appearing after exon 3, resulting in a short protein without any signature domains of NPR1 (Figure 3.5). To understand why the NPR1 domain was absent from the transcript, we

investigated possible alternative splicing at the locus using the primers flanking the stop codon and two RNA samples of CS grown under biotic stressed conditions such as rust/virus infection. An additional fragment was discovered from CS RNA extracted from leaf tissues 24 hours after inoculation with either stripe rust or barley stripe mosaic virus (Appendix B Figure 3.4S) or Cadenza RNA 24 hours after stem rust inoculation. This alternative spliced transcript of *Ta7ANPR1* has the intron (including the stop codon) that was retained between exons 3 and 4 in the cDNA-1 spliced out and translates to a protein of 1,437 amino acids with fused domains of NB-ARC and NPR1 (Figure 3.5). The same alternative spliced transcript was also identified from RNA-seq data generated from a pair of wheat *Lr47*-cultivar 'Scholar' near isogenic lines (NILs) post leaf rust pathogen inoculation (Appendix B Figure 3.S5). Interestingly, the transcript encoding for the protein without NPR1 domains was upregulated about eight-fold in the susceptible NIL (Scholar/-*Lr47*), whereas the levels of the two types of *Ta7ANPR1* transcripts were about the same in the resistant NIL (Scholar/+*Lr47*) with about one-fold increase at 2 days post-inoculation (Appendix B Figure 3.S5).

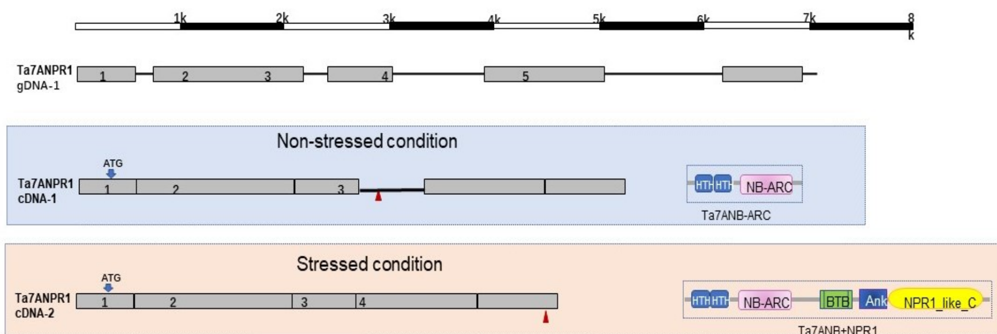


Figure 3.5: Genomic DNA and two transcripts of the Ta7ANPR1 gene. Ta7ANPR1 has five exons. Under non-stressed growth conditions, the transcript retains the intron between exons 3 and 4 including a stop codon. The encoded peptide has 881 amino

acids containing only the NB-ARC domain. Under stressed conditions, an additional transcript of Ta7ANPR1 was detected. The alternative spliced isoform has the retained intron and the stop codon removed so the encoded peptide is 1,431 amino acids containing NB-ARC and NPR1.

The transcript abundances of the mRNAs encoding for NB-ARC and NB-ARC-NPR1 were quantitatively measured via qRT-PCR in Cadenza during the time course of *Pgt* inoculation. Two sets of primers were designed as shown in Fig. S6. One pair is flanking the borders of the spliced-out region (925 bp), the other pair is inside the spliced fragment. During qRT-PCR, we set up the program with 20 second extension time to only allow ~200 bp fragment amplified. Using this strategy, we can measure each isoform of the *Ta7ANPR1* mRNAs. The transcript abundance of the mRNA encoding for the NB-ARC-NPR1 protein was not detectable at 0 dpi (Fig. S6). The two isoforms were at a similar expression level during *Pgt* TPMKC infection from 1~10 dpi (Appendix B Figure 3.S6).

Identification of Ta7ANPR1 Mutants

Six independent mutants of *Ta7ANPR1* were identified from two spring wheat backgrounds, three from ‘Alpowa’ and three from ‘Cadenza’. The Alpowa mutant was identified by a set of primers, G7A-MF1+G7A-MR1 (Table 3.S2), specifically designed for *Ta7ANPR1* after screening an ethyl methane sulfonate (EMS) mutagenized Alpowa population (Feiz et al., 2009). The specificity of the primers was verified on three group 7 NT lines. A 508-bp fragment was amplified from N7BT7D and N7DT7B but not from N7AT7D (Appendix B Figure 3.S1), confirming that the primers are specific to amplify *Ta7ANPR1*. After screening 576 individuals from the mutagenized population, a total of

three mutations were identified. The potential mutants were then tested with *Pgt* TPMKC, along with CS and Alpowa as susceptible controls, and CS+*Sr33* as resistance control, two mutants had similar infection type as CS and Alpowa (data not shown), and the locations of the mutations are all outside of the NB-ARC domain. However, one mutant with a missense mutation on *Ta7ANPR1* resulting in an R to Q at 805 (named as R805Q) was resistant to the pathogen (Figure 3.6a). This mutation was located within the NB-ARC domain (aa 525-814).

To confirm the new resistance of R805Q was due to the mutation, we made a cross between the mutant (as a male parent) and the wild type Alpowa. Ten F_1 individuals were tested with *Pgt* TPMKC and showed a susceptible infection phenotype at 14 days post-inoculation, whereas R805Q was highly resistant and wild type Alpowa was susceptible (Figure 3.6a). The ten F_1 individuals were bagged and self-pollinated to produce F_2 seeds. Four hundred F_2 were screened with *Pgt* TPMKC at the seedling stage, the segregation of resistant to susceptible fits 1:3 ratio. Seventy-five susceptible and 21 resistant individuals were sampled to amplify the region of mutation at *Ta7ANPR1* via PCR using the gene-specific primers. The sequence data confirmed the new resistance completely co-segregated with the SNP between R805Q and Alpowa.

The three mutants from wheat cultivar 'Cadenza' were identified from the wheat-Tilling database (https://www.seedstor.ac.uk/shopping-cart-tilling.php) using *Ta7ANPR1* sequence as a query. The database has more than 50 lines carrying a mutation on the *Ta7ANPR1* gene in a heterozygous state. Homozygous A529E and M357I mutants each

showed resistance to *Pgt* TPMKC (Figure 3.6b), whereas L1224R (mutation located in the DUF3420 unknown domain of the NPR1 portion) did not (Data not shown).

Homozygous mutant R805Q were also tested with *Pt* PBJJG and two *Pst* races. The mutant was as susceptible as the wild type to *Pt* PBJJG (Figure 3.6c) and *Pst* race CYR31 (Figure 3.6d) but was less susceptible to *Pst* CYR23 than the wild type (Figure 3.6e). These results suggested the resistance of the mutant was rust pathogen-species specific and race-specific.

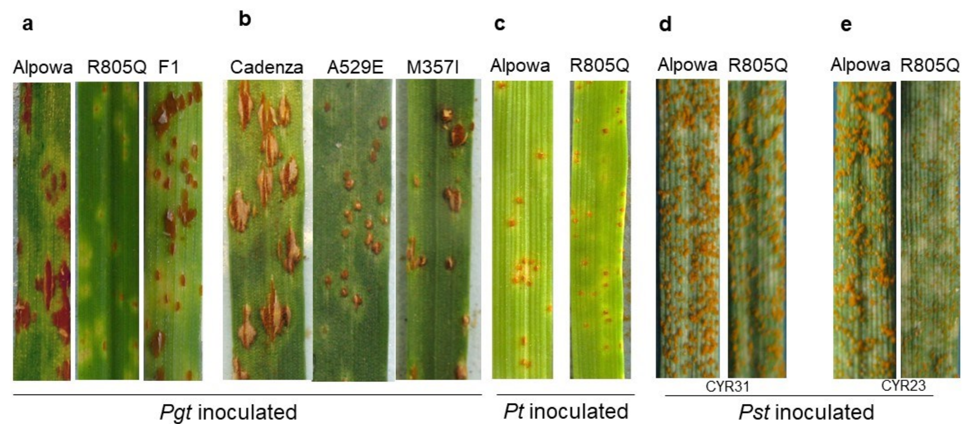


Figure 3.6: Infection types of *Ta7ANPR1* mutants. **a.** Infection types of wild type Alpowa, mutant R805Q and an F_1 of Alpowa/R805Q when challenged with stem rust race TPMKC 14 days post inoculation. **b.** Infection types of wild type Cadenza and mutants of A529E and M357I at 14 days post *Pgt* TPMKC inoculation. **c.** Infection types of wild type Alpowa and mutant R805Q at nine days post *Pt* PBJJG inoculation. **d.** Infection types of wild type Alpowa and mutant R805Q at 17 days post *Pst* CYR31 inoculation. **e.** Infection types of wild type Alpowa and mutant R805Q at 17 days post *Pst* CYR23 inoculation.

Levels of SA/JA and PR Genes when
TaG3NPR1 Genes were Down regulated

As explained earlier, NPR1 interacts with TGA transcription factors to activate SA-mediated signaling pathway genes. Hence, when no visible infection type changes were observed with *Pgt* inoculation after *TaG3NPR1* were down-regulated, we decided to measure the levels of SA/JA and five *PR* genes in CS leaf segments where the six copies of *TaG3NPR1* were silenced, and post-*Pst* inoculation. Leaf segments of the same stage as the silenced ones from CS seedlings without any treatments (mock) or treated with BSMV virus alone (BSMV:00) were used as controls. At ten days post BSMV inoculation, six viral-symptom-free leaf segments were sampled from each treatment. Each grounded leaf sample was divided into three portions; one of each was used for RNA extraction, SA and JA measurements. After confirming the *TaG3NPR1* levels were reduced at least 30% compared to the BSMV:00 control, the top three best *TaG3NPR1*-silenced samples (Appendix B Figure 3.S7a) were selected for further tests of SA, JA and the *PR* genes. We observed significant high levels of SA in the leaves with BSMV-inoculations (Appendix B Figure 3.S7b) compared with the mock control. However, the levels of three SA-mediated *PR* genes were similar among the high-SA and low-SA samples (Appendix B Figure 3.S7c). When we measured the SA/JA and *PR* genes at 24 hours post-*Pst* inoculation (hpi), the level of SA slightly reduced compared with the levels without *Pst* treatment (labeled as 0 hours post-*Pst* inoculation in the figures) but the three genes were significantly increased (Appendix B Figure 3.S7b). We learned from these observations that high-SA does not equal to high *PR1* level, but the expression of *PR1* did associate with a certain level of SA concentration, suggesting SA was essential

for PR1 expression. SA alone was not sufficient to achieve high levels of *PR1, PR2* and *PR5*. Interestingly, at 24 hpi, the SA levels detected in *TaG3NPR1* silenced leaves were significantly higher than that in the non-silenced leaves, but the levels of *PR1, PR2* and *PR5* genes were significantly lower than the levels in the non-silenced leaves (Appendix B Figure 3.S7c), suggesting some of the *TaG3NPR*1 genes were required for the SA-mediated *PR* gene expression.

In contrast, levels of JA were not significantly different among the different treatments without *Pst* inoculation, so were the levels of *PR3* and *PR10* genes (Appendix B Figure 3.S8a, b). Post-*Pst* inoculation, *PR3* expression level increased at 24 hpi, and the level was significantly higher in the *TaG3NPR1* silenced leaves than that in non-silenced control (Figure 3.S9b).

Transcript Abundance of *PR* Genes when *Ta7ANPR1* Gene was Knocked out

To explore why the knockout *Ta7ANPR1* enhanced resistance to stem rust, we analyzed five *PR* genes in the two mutants A529E and R805Q in the absence of pathogens. All five *PR* genes had low expression levels between the mutant R805Q and the wild-type Alpowa (Appendix B Figure 3.S10). The mutant A529E also had low expression levels in four of the five *PR* genes (Appendix B Figure 3.S10) except the *PR10* level in A529E was significantly higher than that in the wild-type Cadenza (Appendix B Figure 3.S10).

Because the *PR1* gene was undetectable in Cadenza and the mutant and *PR10* expression was different between the two lines without the pathogen, we measured the transcript abundances of the two *PR* genes post *Pgt* TPMKC at 5 time points (Appendix

B Figure 3.S11). *PR1* became detectable at 1 dpi but at very low levels until 5 dpi in both A529E and Cadenza (Appendix B Figure 3.S11a). A different pattern was observed with *PR10* with a relatively high base-level expression; mutant A529E had an up-regulated *PR10* expression during 1~2 dpi and then gradually returned to the base-level from 3 dpi (Appendix B Figure 3.S11b). In contrast, Cadenza had a low *PR10* level during the early time points (1~2 dpi) and slowly increased the level at later time points (3~5 dpi) (Appendix B Figure 3.S11b).

DISCUSSSION

A requirement of NPR1 in Wheat Defense Response to *Puccinia* is Selective

Bread wheat has nine homologs of the NPR1-like gene in the genome. Three of them, named as wNPR1 by Cantu et al. (Cantu et al., 2013), are a classical type of NPR1 with similar functional domains as the *Arabidopsis* AtNPR1. wNPR1 has a similar mode of action as the AtNPR1, interacting with a TGA transcription factor for the transduction of the SA signal (Cantu et al., 2013). During wheat-*Puccinia striiformis* interaction, wNPR1 was targeted by a stripe rust effector protein PNPi (for Puccinia NPR1 interactor) (X. Wang et al., 2016). The PNPi competes with wNPR1 for the interaction with TGA2.2 and reduced pathogenesis-related gene expression (X. Wang et al., 2016). Similarly, when the *wNPR1* gene was knocked down with a BSMV:G3NPR1 construct in our study, we found three SA-mediated *PR* genes (*PR1, PR2* and *PR5*) had significantly lower expression levels compared to the control (Appendix B Figure 3.S7). These results suggested that the *wNPR1* was involved in wheat defense response against stripe rust as

the classical *NPR1* gene for transducing the SA signal to activate *PR* gene expression (Cao et al., 1994; Dong, 2004). However, down-regulating *wNPR1* did not alter the infection type phenotype of the *Sr33*-mediated stem rust resistance, suggesting defense response to stem rust conferred by *Sr33* did not require wNPR1. Similarly, NPR1 is not always required in resistance in *Arabidopsis*. NPR1-independent resistance to a bacterial pathogen *Pseudomonas syringae* pv. *maculicola* and an oomycete *Peronospora parasitica* were identified in *Arabidopsis* (X. Li, Clarke, Zhang, & Dong, 2001). This pathway that bypasses NPR1 but requires SA is an EDS1-mediated pathway (X. Li et al., 2001). Our study suggested *Sr33*-mediated resistance is NPR1-independent, but it should not be generalized to imply that NPR1 is not required for immunity to stem rust based on only one gene. Clearly, more studies on other *Sr* genes are necessary to draw such a conclusion. Function of *AtNPR1* in other plant species may not be the same as observed in *Arabidopsis*. *AtNPR1* in rice has a similar function of its *NPR1* homolog as it could enhance resistance against bacterial pathogen *Xanthomonas oryzae* pv. *oryzae* (Xoo) (Chern et al., 2005), but was more susceptible to herbivore attack (Yuan et al., 2007). *AtNPR1*-expressing wheat was resistant to Fusarium head blight (Makandar et al., 2006), but was more susceptible to Fusarium seedling blight (Gao et al., 2013). Given the importance of NPR1 in host defense, the protein also becomes a target of pathogens. In rice, the NPR1 interacting protein NRR negatively regulates defense response to Xoo. In brief, NPR1 is an important component in plant defense, but its involvement in SA defense signaling is not a universal phenomenon.

NPR1 Proteins with Integrated Domains in the Wheat Genome

Besides the classical type of *NPR1*, five homologs of *NPR1*-like gene in the bread wheat genome occur as gene fusions with exogenous domains such as a protein kinase (Figure 3.2) or an NB-ARC domain of the most common plant immune receptors (Figure 3.3). The International Wheat Genome Sequence Consortium (IWGSC) and the EnsemblPlants databases have the sequences at all loci but the search tools fail to reveal them as fused NPR1 proteins. It was only in a recent study by Bailey et al. (2018) where three NB-ARC-NPR1 fused proteins in hexaploid wheat based on gene predictions were reported (Bailey et al., 2018). Our findings back up the prediction. A possible reason why the NPR1 fused proteins were overlooked may be attributed to the findings in our study where the mRNA of the *Ta7ANPR1* gene in the absence of pathogen infection has a stop codon before the sequences coding for the NPR1-like domains (Figure 3.6). Consequently, the gene was predicted only as an NB-ARC containing protein. The NPR1 proteins with integrated domains are produced only when the gene is alternatively spliced under certain stresses, for example biotic stresses (Appendix B Figure 3.S2). Because the sequences of *TaG7NPR1* in bread wheat can be traced to its diploid ancestors (alternatively spliced variants annotated for the diploid D genome as AETGV20038900.1, AETGV20038900.2, and AETGV20038900.3), it indicated the NB-ARC-NPR1 fused protein already existed in the diploid donor species of bread wheat. Evolutionary events to create this NB-ARC-NPR1 fusion and selection sweeps have maintained these NPR1 fused proteins, implying the benefits of these proteins for wheat. Several pieces of findings suggest that NPR1 might contribute to defense against

pathogens. The classic wNPR1 on chromosomes 3A, 3B and 3D have been shown to be involved in stripe rust resistance and a target of the pathogen (X. Wang et al., 2016). It raises the possibility that the NPR1 domain of the Ta7ANB-ARC-NPR1 is used as a decoy to monitor the wNPR1. The organization of the group 7 *NB-ARC-NPR1* genes in a head-to-head orientation with another *NB-ARC*-like gene (Figure 3.3) is indicative of some disease resistance gene pairs that require sensing and signaling partners to confer resistance function, for example, the *RRS1/RPS4* pair (Narusaka et al., 2009) and the *Pi5-1/Pi5-2* pair (Lee et al., 2009).

The Negative Regulation of Ta7ANBS-NPR1 is Specific

The function of the Ta7ANB-ARC-NPR1 fusion protein does not appear to enhance a general defense to all pathogens because the mutants showed resistance to stem rust but not to leaf rust pathogen (Appendix B Figure 3.S4). It appeared the mutations lifted the suppression on specific *Sr* genes. Among the six mutants identified, mutations found in the resistant mutants are all located in the NB-ARC region of the gene; the one mutation found in the NPR1 portion of the fusion protein did not affect infection type to the *Pgt* QFCSC (data not shown). This observation suggested the negative regulatory role of the fusion protein is more likely located in the NB-ARC region. The NPR1 domains of the fusion protein were less important for the acquired new resistance to stem rust or the specific condition to assay the function of the NPR1 in the fusion protein was not established in the study.

With the current studies, we can only speculate the possible modes of actions of the NB-ARC-NPR1 fusion protein as follows: (i) a “decoy” of the important NPR1 or (ii)

a region of ubiquitination. We hypothesize that some of the *R* genes monitor the integrity of the wNPR1 and defense activation is suppressed by the NB-ARC domain of the fusion protein. In the first 'decoy' scenario, the NPR1 portion is a decoy in the Ta7ANB-ARC-NPR1 protein. Alteration on the NPR1 domain of the fusion protein will release the suppression by the NB-ARC domain. There are two Helix-Turn-Helix (HTH) motifs in both NB-ARC and NB-ARC-NPR1 proteins (Figure 3.5). Based on the upregulated JA-responsive *PR* genes (*PR3* and *PR10*) expressions during the time course of Ta7ANPR1 mutant, we consider the function of these HTH motifs are for defense response. In the second scenario, under pathogen attack, the fusion protein is produced, the NB-ARC domain of the fusion protein may compete with the sole NB-ARC protein that suppresses the *R* genes, the NPR1 domain of the fusion protein could be ubiquitinated and subjected to degradation, and then release the suppression on the interacting R proteins. Further investigations are required to broaden our understanding of the benefits of this fusion protein and the function of each domain of the fusion protein.

Alternative Splicing and Innate Immunity

Alternative splicing has been reported to be associated with stresses including abiotic and biotic stresses (Filichkin, Priest, Megraw, & Mockler, 2015; Jordan, Schornack, & Lahaye, 2002; Zhang & Gassmann, 2007). Numerous TIR-NB-LRR and CC-NB-LRR plant R proteins have alternative isoforms (Ayliffe et al., 1999; Dinesh-Kumar & Baker, 2000; P. N. Dodds, Lawrence, & Ellis, 2001; Gassmann, Hinsch, & Staskawicz, 1999; Jordan et al., 2002). The location of the alternative splicing, so far, has only been seen between the sequence coding for NB and LRR domains (Jordan et al.,

2002). In some cases, not only the presence of the alternative isoforms of the gene but also their ratio is crucial for effective resistance against pathogen attack (Dinesh-Kumar & Baker, 2000). In other cases, for example, *L6*, although alternative form of *L6* was detected, no functional relevance could be assigned (P. N. Dodds et al., 2001). Our studies revealed that the alternative splicing from a locus of 7A is regulated by both the host and pathogen (Appendix B Figure 3.S2 and S3). The isoform of the Ta7ANB-NPR1 fusion protein is promoted by the host and suppressed by the pathogen during the interaction, suggesting alternative splicing is one of the strategies used by both host and pathogen during an interaction. Our findings reveal the complexity of host defense regulation and the exquisite nature of the pathogen surveillance systems in plants.

ACKNOWLEDGEMENTS

Xiaojing Wang wishes to thank the National Key Research and Development Program of China (No.2018YFD0200405) and the Shaanxi Provincial Postdoctoral fund (2017). Hongtao Zhang and Bernard Nyamesorto wish to thank the NSF BREAD program (Grant No. IOS-0965429) for financial support. The authors thank Dr. Hikmet Budak for help with the analysis of the *Ta7ANPR1* alternative isoform, and Dr. Michael Giroux for the use of his Alpowa mutagenesis population.

REFERENCES

Aravind, L., & Koonin, E. V. (1999). Fold prediction and evolutionary analysis of the POZ domain: Structural and evolutionary relationship with the potassium channel tetramerization domain. *Journal Of Molecular Biology, 285*(4), 1353-1361. doi:DOI 10.1006/jmbi.1998.2394

Ayliffe, M. A., Frost, D. V., Finnegan, E. J., Lawrence, G. J., Anderson, P. A., & Ellis, J. G. (1999). Analysis of alternative transcripts of the flax L6 rust resistance gene. *Plant Journal, 17*(3), 287-292. doi:DOI 10.1046/j.1365-313X.1999.00377.x

Bailey, P. C., Schudoma, C., Jackson, W., Baggs, E., Dagdas, G., Haerty, W., . . . Krasileva, K. V. (2018). Dominant integration locus drives continuous diversification of plant immune receptors with exogenous domain fusions. *Genome Biology, 19*. doi:ARTN 23 10.1186/s13059-018-1392-6

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol, 34*(5), 525-527. doi:10.1038/nbt.3519

Campbell, J., & Huang, L. (2010). Silencing of multiple genes in wheat using Barley stripe mosaic virus. *Journal of Biotech Research, 2*(1), 12-20.

Cantu, D., Yang, B. J., Ruan, R., Li, K., Menzo, V., Fu, D. L., . . . Dubcovsky, J. (2013). Comparative analysis of protein-protein interactions in the defense response of rice and wheat. *Bmc Genomics, 14*, 166. doi:Artn 166 10.1186/1471-2164-14-166

Cao, H., Bowling, S. A., Gordon, A. S., & Dong, X. N. (1994). Characterization Of an Arabidopsis Mutant That Is Nonresponsive To Inducers Of Systemic Acquired-Resistance. *Plant Cell, 6*(11), 1583-1592.

Cao, H., Glazebrook, J., Clarke, J. D., Volko, S., & Dong, X. N. (1997). The Arabidopsis NPR1 gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats. *Cell, 88*(1), 57-63. doi:Doi 10.1016/S0092-8674(00)81858-9

Chern, M., Fitzgerald, H. A., Canlas, P. E., Navarre, D. A., & Ronald, P. C. (2005). Overexpression of a rice NPR1 homolog leads to constitutive activation of defense response and hypersensitivity to light. *Molecular Plant-Microbe Interactions, 18*(6), 511-520. doi:10.1094/Mpmi-18.0511

Devos, K. M., Dubcovsky, J., Dvorak, J., Chinoy, C. N., & Gale, M. D. (1995). Structural Evolution Of Wheat Chromosomes 4a, 5a, And 7b And Its Impact on Recombination. *Theoretical And Applied Genetics, 91*(2), 282-288. doi:Doi 10.1007/Bf00220890

Dinesh-Kumar, S. P., & Baker, B. J. (2000). Alternatively spliced N resistance gene transcripts: Their possible role in tobacco mosaic virus resistance. *Proceedings Of the National Academy Of Sciences Of the United States Of America, 97*(4), 1908-1913. doi:DOI 10.1073/pnas.020367497

Dodds, P. N., Lawrence, G. J., Catanzariti, A.-M., Ayliffe, M. A., & Ellis, J. G. J. T. P. C. (2004). The Melampsora lini AvrL567 avirulence genes are expressed in haustoria and their products are recognized inside plant cells. *Plant Cell, 16*(3), 755-768.

Dodds, P. N., Lawrence, G. J., & Ellis, J. G. (2001). Six amino acid changes confined to the leucine-rich repeat beta-strand/beta-turn motif determine the difference between the P and P2 rust resistance specificities in flax. *Plant Cell, 13*(1), 163-178. doi:DOI 10.1105/tpc.13.1.163

Dong, X. N. (2004). NPR1, all things considered. *Current Opinion In Plant Biology, 7*(5), 547-552. doi:10.1016/j.pbi.2004.07.005

Feiz, L., Beecher, B. S., Martin, J. M., & Giroux, M. J. (2009). In Planta Mutagenesis Determines the Functional Regions of the Wheat Puroindoline Proteins. *Genetics, 183*(3), 853-860. doi:10.1534/genetics.109.106013

Filichkin, S., Priest, H. D., Megraw, M., & Mockler, T. C. (2015). Alternative splicing in plants: directing traffic at the crossroads of adaptation and environmental stress. *Current Opinion In Plant Biology, 24*, 125-135. doi:10.1016/j.pbi.2015.02.008

Fu, Z. Q., Yan, S. P., Saleh, A., Wang, W., Ruble, J., Oka, N., . . . Dong, X. N. (2012). NPR3 and NPR4 are receptors for the immune signal salicylic acid in plants. *Nature, 486*(7402), 228-+. doi:10.1038/nature11162

Gao, C. S., Kou, X. J., Li, H. P., Zhang, J. B., Saad, A. S. I., & Liao, Y. C. (2013). Inverse effects of Arabidopsis NPR1 gene on fusarium seedling blight and fusarium head blight in transgenic wheat. *Plant Pathology, 62*(2), 383-392. doi:10.1111/j.1365-3059.2012.02656.x

Gassmann, W., Hinsch, M. E., & Staskawicz, B. J. (1999). The Arabidopsis RPS4 bacterial-resistance gene is a member of the TIR-NBS-LRR family of disease-resistance genes. *Plant Journal, 20*(3), 265-277. doi:DOI 10.1046/j.1365-313X.1999.t01-1-00600.x

Jordan, T., Schornack, S., & Lahaye, T. (2002). Alternative splicing of transcripts encoding Toll-like plant resistance proteins - what's the functional relevance to innate immunity? *Trends Plant Sci, 7*(9), 392-398.

Le Henanff, G., Farine, S., Kieffer-Mazet, F., Miclot, A. S., Heitz, T., Mestre, P., . . . Chong, J. (2011). Vitis vinifera VvNPR1.1 is the functional ortholog of AtNPR1 and its overexpression in grapevine triggers constitutive activation of PR genes and enhanced resistance to powdery mildew. *Planta, 234*(2), 405-417. doi:10.1007/s00425-011-1412-1

Lee, S.-K., Song, M.-Y., Seo, Y.-S., Kim, H.-K., Ko, S., Cao, P.-J., . . . Lee, S. J. G. (2009). Rice Pi5-mediated resistance to Magnaporthe oryzae requires the presence of two coiled-coil–nucleotide-binding–leucine-rich repeat genes. *Genetics, 181*(4), 1627-1638.

Li, J., Brader, G., & Palva, E. T. (2004). The WRKY70 transcription factor: A node of convergence for jasmonate-mediated and salicylate-mediated signals in plant defense. *Plant Cell, 16*(2), 319-331. doi:10.1105/tpc.016980

Li, X., Clarke, J. D., Zhang, Y. L., & Dong, X. N. (2001). Activation of an EDS1-mediated R-gene pathway in the snc1 mutant leads to constitutive, NPR1-independent pathogen resistance. *Molecular Plant-Microbe Interactions, 14*(10), 1131-1139. doi:Doi 10.1094/Mpmi.2001.14.10.1131

Line, R. F., & Qayoum, A. J. T. b. (1992). Virulence, aggressiveness, evolution and distribution of races of Puccinia striiformis (the cause of stripe rust of wheat) in North America, 1968-87.

Liu, C., Atkinson, M., Chinoy, C., Devos, K., & Gale, M. (1992). Nonhomoeologous translocations between group 4, 5 and 7 chromosomes within wheat and rye. *Theor. Appl. Genet, 83*(3), 305-312.

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods., 25*(4), 402-408.

Makandar, R., Essig, J. S., Schapaugh, M. A., Trick, H. N., & Shah, J. J. M. P.-M. I. (2006). Genetically engineered resistance to Fusarium head blight in wheat by expression of Arabidopsis NPR1. *Mol. Plant Microbe Interact., 19*(2), 123-129.

Malnoy, M., Jin, Q., Borejsza-Wysocka, E., He, S., & Aldwinckle, H. (2007). Overexpression of the apple MpNPR1 gene confers increased disease resistance in Malus× domestica. *Mol. Plant Microbe Interact., 20*(12), 1568-1580.

McIntosh, R. A., Wellings, C. R., & Park, R. F. (1995). *Wheat rusts: an atlas of resistance genes*: Csiro Publishing.

Moreau, M., Tian, M., & Klessig, D. F. (2012). Salicylic acid binds NPR3 and NPR4 to regulate NPR1-dependent defense responses. *Cell Res., 22*(12), 1631.

Morel, J. B., & Dangl, J. L. (1997). The hypersensitive response and the induction of cell death in plants. *Cell Death And Differentiation, 4*(8), 671-683. doi:DOI 10.1038/sj.cdd.4400309

Mou, Z., Fan, W., & Dong, X. (2003). Inducers of plant systemic acquired resistance regulate NPR1 function through redox changes. *Cell, 113*(7), 935-944.

Narusaka, M., Shirasu, K., Noutoshi, Y., Kubo, Y., Shiraishi, T., Iwabuchi, M., & Narusaka, Y. (2009). RRS1 and RPS4 provide a dual Resistance-gene system against fungal and bacterial pathogens. *Plant J., 60*(2), 218-226.

Pajerowska-Mukhtar, K. M., Emerine, D. K., & Mukhtar, M. S. (2013). Tell me more: roles of NPRs in plant immunity. *Trends Plant Sci, 18*(7), 402-411.

Periyannan, S., Moore, J., Ayliffe, M., Bansal, U., Wang, X., Huang, L., . . . Bariana, H. (2013). The gene Sr33, an ortholog of barley Mla genes, encodes resistance to wheat stem rust race Ug99. *Science, 341*(6147), 786-788.

Rate, D. N., & Greenberg, J. T. (2001). The Arabidopsis aberrant growth and death2 mutant shows resistance to Pseudomonas syringae and reveals a role for NPR1 in suppressing hypersensitive cell death. *Plant J., 27*(3), 203-211.

Shirano, Y., Kachroo, P., Shah, J., & Klessig, D. F. (2002). A gain-of-function mutation in an Arabidopsis Toll Interleukin1 Receptor–Nucleotide Binding Site–Leucine-Rich Repeat type R gene triggers defense responses and results in enhanced disease resistance. *Plant Cell, 14*(12), 3149-3162.

Silverman, P., Seskar, M., Kanter, D., Schweizer, P., Metraux, J.-P., & Raskin, I. (1995). Salicylic acid in rice (biosynthesis, conjugation, and possible role). *J Plant Physiology, 108*(2), 633-639.

Spoel, S. H., & Dong, X. (2012). How do plants achieve immunity? Defence without specialized immune cells. *J Nature reviews immunology, 12*(2), 89.

Spoel, S. H., Koornneef, A., Claessens, S. M., Korzelius, J. P., Van Pelt, J. A., Mueller, M. J., . . . Kazan, K. (2003). NPR1 modulates cross-talk between salicylate-and jasmonate-dependent defense pathways through a novel function in the cytosol. *J The Plant Cell, 15*(3), 760-770.

Spoel, S. H., Mou, Z., Tada, Y., Spivey, N. W., Genschik, P., & Dong, X. (2009). Proteasome-mediated turnover of the transcription coactivator NPR1 plays dual roles in regulating plant immunity. *Cell, 137*(5), 860-872.

Stakman, E. C., Stewart, D. M., & Loegering, W. Q. (1962). Identification of physiologic races of *Puccinia graminis* var. *tritici*.

Wang, L., Li, H., Zhao, C., Li, S., Kong, L., Wu, W., . . . Zhang, H. (2017). The inhibition of protein translation mediated by AtGCN1 is essential for cold tolerance in Arabidopsis thaliana. *Plant, cell & environment, 40*(1), 56-68. doi:10.1111/pce.12826

Wang, X., Yang, B., Li, K., Kang, Z., Cantu, D., & Dubcovsky, J. (2016). A conserved Puccinia striiformis protein interacts with wheat NPR1 and reduces induction of pathogenesis-related genes in response to pathogens. *J Molecular Plant-Microbe Interactions, 29*(12), 977-989.

Yuan, Y., Zhong, S., Li, Q., Zhu, Z., Lou, Y., Wang, L., . . . Yang, D. (2007). Functional analysis of rice NPR1-like genes reveals that OsNPR1/NH1 is the rice orthologue conferring disease resistance with enhanced herbivore susceptibility. *J Plant biotechnology journal, 5*(2), 313-324.

Zhang, X.-C., & Gassmann, W. (2007). Alternative splicing and mRNA levels of the disease resistance gene RPS4 are induced during defense responses. *J Plant Physiology, 145*(4), 1577-1587.

Zhong, X., Xi, L., Lian, Q., Luo, X., Wu, Z., Seng, S., . . . Yi, M. (2015). The NPR1 homolog GhNPR1 plays an important role in the defense response of Gladiolus hybridus. *J Plant cell reports, 34*(6), 1063-1074.

CHAPTER FOUR

WHEAT *MYB21* TRANSCRIPTIONAL FACTOR GENE MODIFICATION ENHANCED HOST RESISTANCE AGAINST RUST PATHOGENS

This work is currently in progress

INTRODUCTION

Myeloblastosis (*Myb*) proteins are part of a large and functionally diverse protein family found in eukaryotes. Most Myb proteins function as transcription factors. The first *Myb* gene was identified as a viral "oncogene" (*v-Myb*) and named after a gene in *Avian myeloblastosis* virus. It causes myeloid leukemia in chicken (Klempnauer, Symonds, Evan, & Michael Bishop, 1984). In plants, the first cellular *Myb* (*c-Myb*) was characterized in *Zea mays* known as *C1* (Paz-Ares, Ghosal, Wienand, Peterson, & Saedler, 1987). Generally, *Myb* proteins are made up of a conserved N-terminal helix-turn-helix (HTH) DNA-binding domain of 52 amino acids, a central transcriptional activation domain and a diverse C-terminal regulatory region. The *Myb* family is divided into four classes based on the number of *Myb* domains namely, 1R-, R2R3-, 3R- and 4R-*Myb* proteins (Stracke, Werber, & Weisshaar, 2001). Even though plants have a large diversity of *Myb* family genes, R2R3 is the most abundant class of the *Myb* genes. This group of *Myb* family proteins has been studied extensively in plant species such as *Arabidopsis*, rice, maize, cotton, apple, petunia, snapdragon, grapevine and poplar (*Populus tremuloides*). They play important roles in primary metabolism, cell fate and

identity, developmental processes, secondary metabolism and responses to biotic and abiotic stresses (Cao, Li, Li, Zhao, & Wang, 2020).

In *Arabidopsis*, *AtMybL* modulates abiotic stress response through its function in the process of leaf senescence (Zhang et al., 2010). Also, *SiMyb75* is found to confer abiotic stress tolerance (Dossa et al., 2019). In biotic stress responses, *AtMyb44* is known to play a role in plant defense response against aphid (Liu et al., 2010). It has been suggested that *BjMyb1* is potentially involved in host defense against fungal attack through activating the expression of *BjCHI1* by binding to the *Wbl-4* element in the BjC-P promoter (Gao et al., 2016). In rice, *Myb1* was found to act as a regulator involved in both Pi starvation signaling and GA biosynthesis (Gu et al., 2017). Plant *R2R3-Myb* genes are involved in the signal transduction pathways of jasmonic acid (Lee, Qi, & Yang, 2001), salicylic acid (Raffaele, Rivas, & Roby, 2006), gibberellic acid (Murray, Kalla, Jacobsen, & Gubler, 2003) and abscisic acid (Devaiah, Madhuvanthi, Karthikeyan, & Raghothama, 2009). *Myb21* and *Myb24* were demonstrated to interact directly with JAZs to regulate male fertility (Song et al., 2011).These reemphasize the numerous roles of MYB family proteins in plants.

In wheat, *TaMyb* proteins are among the most abundant transcription factors. Studies of some *TaMyb* transcription factors in *Arabidopsis* revealed their function in drought and heat responses (Zhao et al., 2018; Zhao et al., 2017). While a lot of study of *TaMyb* transcription factors is carried out in model plants such as *Arabidopsis*, few studies have been completed in wheat. Moreover, no research has been done in relation to *TaMyb* genes and their role in wheat-rust pathogen interaction. In this study, we have

shown that *TaMyb21* knockdown and specific? knockout of *TaMyb21-3B* homeolog copy enhanced stem rust resistance in a susceptible wheat cultivar.

MATERIALS AND METHODS

For sources of plant materials, bioinformatics tools and analysis, plant growth conditions, gene silencing assay, PCR and RT-qPCR setups, leaf sampling and treatment, DNA and RNA extraction and assessment, gene knockout and genetic analysis procedure, refer to Chapter Two unless otherwise specified. All BLAST searches, sequence alignments, gene and domain predictions were conducted on the International Wheat Genome Sequence Consortium (IWGSC), ClustalW, Softberry, ExPASy bioinformatics portals and National Center for Biotechnology Information (NCBI) Conserved Domain Architecture Retrieval Tool. The expression profile of the target gene was analyzed via DNASTAR software (http://www.dnastar.com/). Gene silencing was done through barley stripe mosaic virus-induced gene silencing (BSMV-VIGS) assay. Rust pathogen inoculation was conducted as described in (Campbell et al., 2012). Stem rust infection types evaluation was carried out as described by (McIntosh, Wellings, & Park, 1995) 14 days after inoculation. Gene knock out was done via EMS mutagenesis, taking advantage of the already available wheat TILLING EMS population. Genetic analysis was conducted to test the genotype-phenotype association using F3 seeds from self-pollinated F1 progeny of a cross between Alpowa * *Myb21-3B* mutant.

RESULTS

Expression Profiles of *TaMyb21* in a Pair of NIL

The relative expression of the *TaMyb21* transcription factor was determined through RNA-seq generated from a near-isogenic line of MNR220 (Talajoor et al., 2016) at four time points post leaf rust inoculation. Upregulation of the gene in the susceptible NIL over the four time periods. However, in the resistant NIL, the expression of gene remained at very low level throughout the time period (Figure 4.1). The time points of the upregulation of *TaMyb21* in the susceptible NIL corresponds with critical time points of rust pathogen development in host plants. One day post fungal infections, haustoria begin forming and multiplies by three days post-infection. Eight days after leaf rust pathogen infection, spores become visible on the leaves. From this result, we hypothesized that *TaMyb21* was involved in the wheat-rust pathogen interaction.

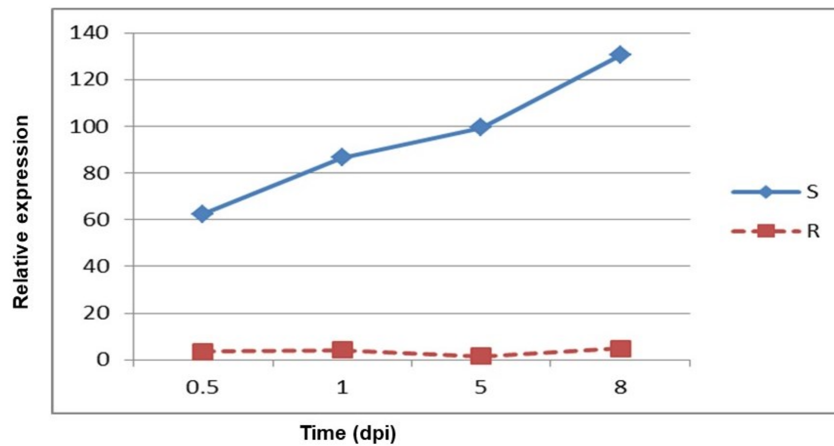


Figure 4.1: Relative expression of *TaMyb21*. RNA sequence data over four time points generated from leaf rust inoculated susceptible (S) and resistant (R) Alpowa cultivars were analyzed in DNASTAR software. No error bars since only one biological rep was used.

TaMyb21 Knockdown Enhanced Chinese
Spring Resistance to Stem Rust

There are three copies of *TaMyb21* in the bread wheat genome. We first identified copies of the gene using cDNA of the *TaMyb21*-like sequence named TRIUR3_09345 from *Aegilops tauschii* to BLAST search the IWGSC database. The BLAST result showed *TaMyb21*-like genes on 3A, 3B and 3D chromosomes with a 97-99% similarity and Expect value of 0.0. To test our hypothesis of possible upregulation of *TaMyb21* to benefit the pathogens, we silenced all three copies of *TaMyb21* (3A, 3B, 3D). The silencing of *TaMyb21* was completed using 287 bp sequences of the conserved region among the three gene copies of *TaMyb21* via BSMV-VIGS. We found out that the *TaMyb21* knockdown reduced the severity of stem rust race QFCSC infection in the

silenced leaves of susceptible Chinese spring cultivar (Figure 4.2) but did not enhance the resistance to leaf rust pathogen (Data not shown).

Figure 4.2*:* BSMV-VIGS of *TaMyb21* in Chinese spring (CS) cultivar. The first leaf of each plant was rub inoculated with the indicated BSMV constructs at the two-leaf stage and then spray inoculated with *Pgt* race QFCSC at 10 days after BSMV inoculation. The disease was assessed and photographed at 14 dpi. CS without any viral inoculation, used as a rust inoculation control; γ:00, CS inoculated with BSMV: Myb21; γ:Myb21.

<u>*TaMyb21-3B* Knockout and Confirmation</u>

We BLAST searched the wheat-tilling website using TaMyb21-3A, 3B and 3D cDNA sequences and identified several mutant lines. Three *TaMyb21-3A* mutants and two *TaMyb21-3B* mutant lines contain missense mutations with a low SIFT score, and high confidence of sequence at the mutation sites for confirmation. We could not find complete information on the *TaMyb21-3D* mutant lines for a downstream analysis; hence we did not examine any of *TaMyb21-3D* mutant lines. After genotyping the mutants with homeolog-specific primers, only one of the *TaMyb21-3B* mutant lines were confirmed to have the claimed mutation, hereafter referred to as S91L-MYB21-3B or S91L (Table 4.1)

with primer BN3B (**F:** GGTGGCTGAACTACCTGAAG **R:** GCGGCCATGGACATCTT).

Table 4.1: Summary of *TaMyb21-3B* Cadenza mutant identified from wheat TILLING website

Chromo-some	**Nucleotide change**	**Mutated amino acid position**	**Amino acid changes**	**Type of mutation**	**SIFT score***	**Mutated ID****
3B	C to T	91	S to L	missense	0	S91L-MYB21-3B

* SIFT score predicts whether an amino acid substitution affects protein function, and ranges from 0 to 1. The amino acid substitution is predicted to be damaging if the score is ≤0.05 and tolerated if the score is >0.05.

**Gene IDs were given based on amino acid change and position of change, gene name and chromosome

Also, *TaMyb21-3B* protein was predicted to have a *Myb-type Helix-turn-helix DNA-binding* domain. The change of amino acid from serine to leucine in the mutant occurred in the conserved DNA binding domain (Figure 4.3).

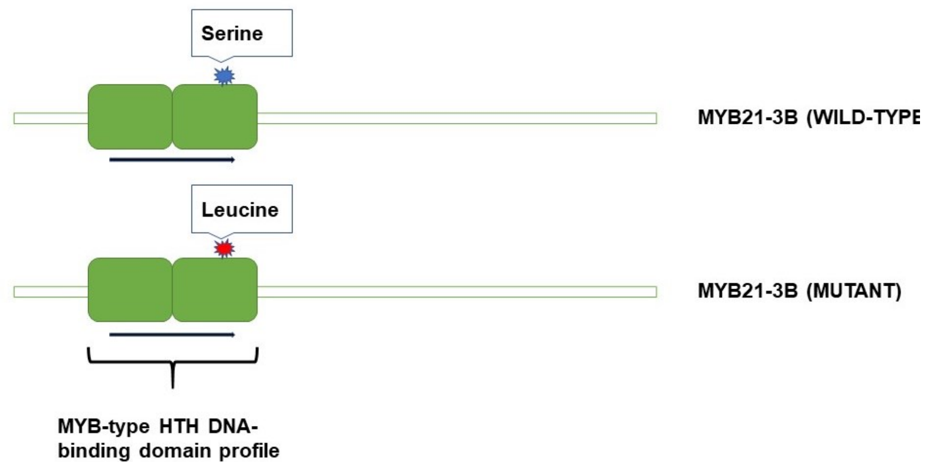


Figure 4.3: Predicted conserved domain of *TaMyb21-3B* homeolog in wild type and mutant. Gene and conserved domains were predicted using scaffold sequences of wheat *TaMyb21-3B* gene copy from IWGSC via Softberry.com and NCBI's CDART and ExPASy bioinformatics portal, respectively. Mutation in the conserved domain caused an amino acid change from Serine to Leucine. Arrows show gene orientation. Diagram not drawn to scale.

Infection Type of Cadenza S91L Mutant

The S91L mutant is in the Cadenza background and was phenotyped using stem rust race TPMKC. As Cadenza is known to have *Sr9g*, *Lr13* and *Yr7* genes (Singh, Park, McIntosh, & Bariana, 2008), we used the *Pgt* TPMKC because the wild type Cadenza showed resistance to *Pgt* QFCSC at the seedling stage. The mutant showed enhanced resistance against the pathogen (Figure 4.4).

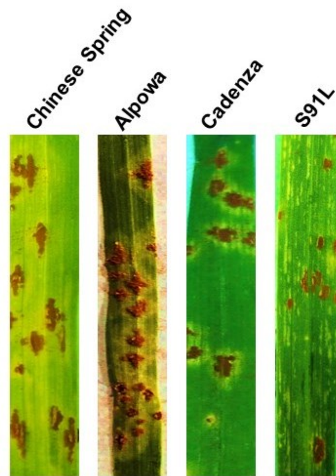


Figure 4.4: Infection type of the S91L mutant. Infection types of wildtypes Chinese Spring, Alpowa and Cadenza and S91L (Cadenza mutant) at 14 days post stem rust race TPMKC inoculation.

A comparison of the mutant and the wild type Cadenza indicated no virtual variations in plant height, seed and tiller morphology (Figure 4.5). At present, we have not confirmed an association between the resistant phenotype and the mutation genotype by genetic analysis. We could conclude that the resistance to *Pgt* TPMKC was possibly due to the mutation in *TaMyb21-3B*. Future rust phenotyping of more *TaMyb21-3B* missense mutants will be required.

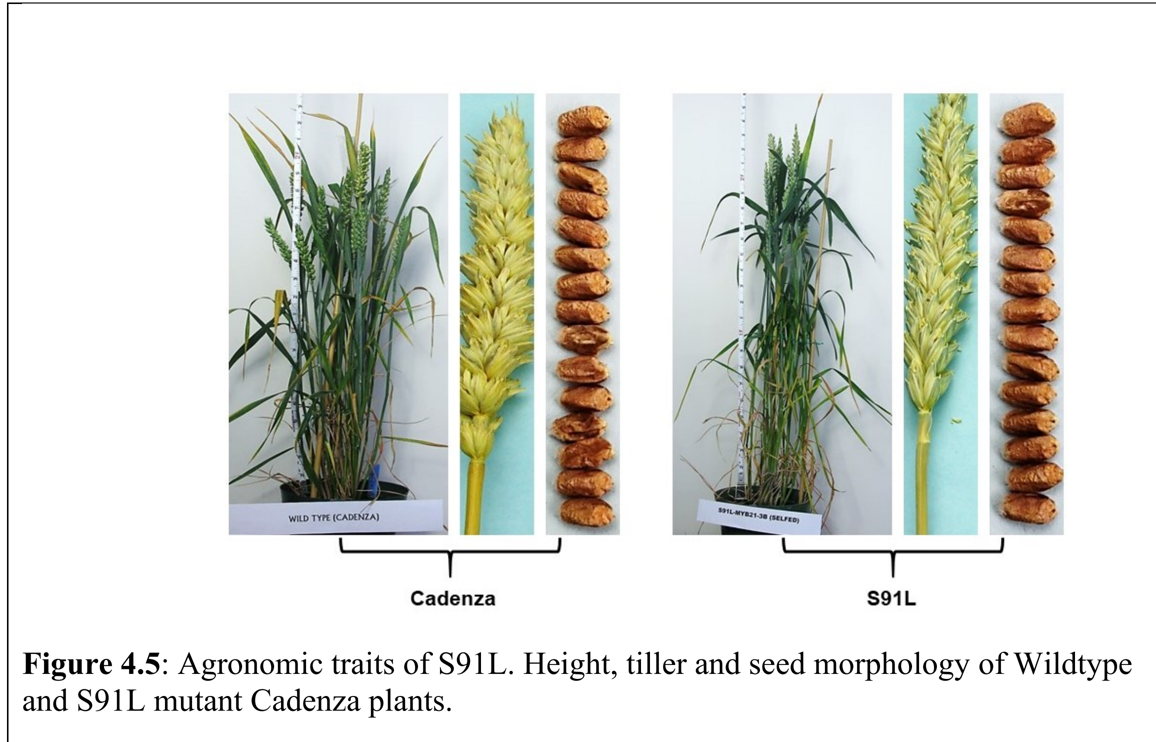


Figure 4.5: Agronomic traits of S91L. Height, tiller and seed morphology of Wildtype and S91L mutant Cadenza plants.

DISCUSSION

TaMyb21 acts as a potential negative regulator of wheat defense during wheat-stem rust interaction based on the infection type of *TaMyb21* knock-down plants. The increase in expression of *TaMyb21* during leaf rust interaction in the susceptible Alpowa line supported this assertion. In rice (*O. sativa spp. japonica*), R2R3-MYB gene *OsJaMyb* was shown to be responsive to infection by the blast fungus *Magnaporthe oryzae* (Lee et al., 2001). Even though the research showed that *OsJaMyb* was upregulated in both susceptible and resistant cultivars, the upregulation was more pronounced in the susceptible cultivar. The study could not, however, ascertain whether the gene's upregulation positively or negatively impact the host-microbe interaction. Our study demonstrated that the reduction in the expression of the *TaMyb21* gene enhanced

Chinese spring's resistance against stem rust QFCSC. However, preliminary genetic analysis on the *TaMyb21-3B* knock-out mutant did not confirm phenotype-genotype relation. Future investigation is required.

ACKNOWLEDGMENTS

We wish to thank the NSF BREAD program (Grant No. IOS-0965429) for financial support.

REFERENCES

Campbell, J., Zhang, H., Giroux, M. J., Feiz, L., Jin, Y., Wang, M., . . . Huang, L. (2012). A mutagenesis-derived broad-spectrum disease resistance locus in wheat. *Theoretical and Applied Genetics, 125*(2), 391-404. doi:10.1007/s00122-012-1841-7

Cao, Y., Li, K., Li, Y., Zhao, X., & Wang, L. (2020). MYB Transcription Factors as Regulators of Secondary Metabolism in Plants. *Biology, 9*(3). doi:10.3390/biology9030061

Devaiah, B. N., Madhuvanthi, R., Karthikeyan, A. S., & Raghothama, K. G. (2009). Phosphate Starvation Responses and Gibberellic Acid Biosynthesis are Regulated by the MYB62 Transcription Factor in Arabidopsis. *Molecular plant, 2*(1), 43-58. doi:https://doi.org/10.1093/mp/ssn081

Dossa, K., Mmadi, M. A., Zhou, R., Liu, A., Yang, Y., Diouf, D., . . . Zhang, X. (2019). Ectopic expression of the sesame MYB transcription factor SiMYB305 promotes root growth and modulates ABA-mediated tolerance to drought and salt stresses in Arabidopsis. *AoB PLANTS, 12*(1). doi:10.1093/aobpla/plz081

Gao, Y., Jia, S., Wang, C., Wang, F., Wang, F., & Zhao, K. (2016). BjMYB1, a transcription factor implicated in plant defence through activating BjCHI1 chitinase expression by binding to a W-box-like element. *Journal of Experimental Botany, 67*(15), 4647-4658. doi:10.1093/jxb/erw240

Gu, M., Zhang, J., Li, H., Meng, D., Li, R., Dai, X., . . . xu, G. (2017). Maintenance of phosphate homeostasis and root development are coordinately regulated by MYB1, an R2R3-type MYB transcription factor in rice. *Journal of Experimental Botany, 68*(13), 3603-3615. doi:10.1093/jxb/erx174

Klempnauer, K.-H., Symonds, G., Evan, G. I., & Michael Bishop, J. (1984). Subcellular localization of proteins encoded by oncogenes of avian myeloblastosis virus and avian leukemia virus E26 and by the chicken c-myb gene. *Cell, 37*(2), 537-547. doi:https://doi.org/10.1016/0092-8674(84)90384-2

Lee, M.-W., Qi, M., & Yang, Y. (2001). A Novel Jasmonic Acid-Inducible Rice myb Gene Associates with Fungal Infection and Host Cell Death. *Molecular Plant-Microbe Interactions®, 14*(4), 527-535. doi:10.1094/MPMI.2001.14.4.527

Liu, R., Lü, B., Wang, X., Zhang, C., Zhang, S., Qian, J., . . . Dong, H. (2010). Thirty-seven transcription factor genes differentially respond to a harpin protein and affect resistance to the green peach aphid in Arabidopsis. *Journal of Biosciences, 35*(3), 435-450.

McIntosh, R. A., Wellings, C. R., & Park, R. F. (1995). *Wheat rusts: an atlas of resistance genes*: Kluwer Academic Publishers.

Murray, F., Kalla, R., Jacobsen, J., & Gubler, F. (2003). A role for HvGAMYB in anther development. *The Plant Journal, 33*(3), 481-491. doi:10.1046/j.1365-313X.2003.01641.x

Paz-Ares, J., Ghosal, D., Wienand, U., Peterson, P. A., & Saedler, H. (1987). The regulatory c1 locus of Zea mays encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. *The EMBO journal, 6*(12), 3553-3558.

Raffaele, S., Rivas, S., & Roby, D. (2006). An essential role for salicylic acid in AtMYB30-mediated control of the hypersensitive cell death program in Arabidopsis. *FEBS Letters, 580*(14), 3498-3504. doi:10.1016/j.febslet.2006.05.027

Singh, D., Park, R. F., McIntosh, R. A., & Bariana, H. S. (2008). Characterisation of stem rust and stripe rust seedling resistance genes in selected wheat cultivars from the united kingdom. *Journal of Plant Pathology, 90*(3), 553-562.

Song, S., Qi, T., Huang, H., Ren, Q., Wu, D., Chang, C., . . . Xie, D. (2011). The Jasmonate-ZIM Domain Proteins Interact with the R2R3-MYB Transcription Factors MYB21 and MYB24 to Affect Jasmonate-Regulated Stamen Development in Arabidopsis. *The Plant cell, 23*(3), 1000. doi:10.1105/tpc.111.083089

Stracke, R., Werber, M., & Weisshaar, B. (2001). The R2R3-MYB gene family in Arabidopsis thaliana. *Current Opinion in Plant Biology, 4*(5), 447-456. doi:https://doi.org/10.1016/S1369-5266(00)00199-0

Talajoor, M., Wang, X., Zhang, H., Zhang, P., Bhavani, S., Tabe, L., . . . Huang, L. (2016). Wheat mutant MNR220 delays haustoria formation during leaf rust pathogenesis at the seedling stage. *Canadian Journal of Plant Pathology, 38*(3), 338-347. doi:10.1080/07060661.2016.1205663

Zhang, X., Ju, H.-W., Chung, M.-S., Huang, P., Ahn, S.-J., & Kim, C. S. (2010). The R-R-Type MYB-Like Transcription Factor, AtMYBL, is Involved in Promoting Leaf Senescence and Modulates an Abiotic Stress Response in Arabidopsis. *Plant and Cell Physiology, 52*(1), 138-148. doi:10.1093/pcp/pcq180

Zhao, Y., Cheng, X., Liu, X., Wu, H., Bi, H., & Xu, H. (2018). The Wheat MYB Transcription Factor TaMYB31 Is Involved in Drought Stress Responses in Arabidopsis. *Frontiers in Plant Science, 9*(1426). doi:10.3389/fpls.2018.01426

Zhao, Y., Tian, X., Wang, F., Zhang, L., Xin, M., Hu, Z., . . . Peng, H. (2017). Characterization of wheat MYB genes responsive to high temperatures. *BMC plant biology, 17*(1), 208-208. doi:10.1186/s12870-017-1158-4

CHAPTER FIVE

CONCLUSION

Generally, these studies from wheat *MYC4, NPR1* to *MYB21* transcriptional genes clearly demonstrated that editing host genes targeted by a pathogen in wheat could be a rapid and effective approach for creating wheat rust-resistant materials. We learned from these studies that some of the wheat genes are contributing to the susceptibility of the host. These host genes permit pathogens to colonize their hosts, most likely through the manipulation of different signaling proteins that usually act to suppress the host defense response. This uncovers some of the unknown functions of the wheat genes and paves the way for their modification for improved plant resistance.

All these studies began with transcriptomic studies via *in silico* bioinformatics investigation and identification of host candidate genes which has proven successful. This confirms the fact that bioinformatics tools could conveniently be useful in our quest to explore the vast repertoire of wheat genes using the already available vast data. Likewise, the successful use of BSMV-VIGS has once again demonstrated the importance of this tool in wheat reverse genetics studies. Also, the availability of EMS mutagenized TILLING populations proves very vital in creating gene mutations by fast means instead of resorting to other methods that come with challenges such as wheat transformation problems. Useful mutations discovered could easily be transferred into more elite lines using markers and phenotypic screens.

Loss-of-function mutations in *MYC4-1BL*, *Ta7ANBS-NPR1* and *MYB21-3B* all conferred enhanced resistance of wheat to rust pathogens. However, these resistances are

race-specific which confirms a characteristic of edited pathogen susceptible host genes. This is a limitation that is also associated with some R genes. One way to resolve this is to begin to create double, triple mutants in one material since we have fully discovered some of these pathogen targeted host genes. It is also worth noting that mutations in these genes have not shown any initial virtually recorded detrimental effect on the plants (pending quantitative evaluations using near-isogenic lines). This could be attributed to the functional compensation of genes in the polyploid wheat genome. It further proves the effectiveness and harmlessness of this work in creating resistance within bread wheat's genome.

In summary, these studies have unequivocally demonstrated a fast approach to targeting and editing wheat genes for creating rust resistance within wheat's own genome as well as creating rust-resistant germplasms for usage in wheat resistance breeding. For the future study of these genes, I suggest further in-depth studies to elucidate the mode of action of these genes in wheat rust interactions.

CUMULATIVE REFERENCES

Abdollahi, P., Kamiya, Y., Kawaura, K., & Ogihara, Y. (2012). Overexpression of Q/q-related homoeoalleles of hexaploid wheat reveals distinct recovery of flower transformation in the apetala2 mutant of Arabidopsis. *Plant Biotechnology*, 12.0224 a.

Aktar-Uz-Zaman, M., Tuhina-Khatun, M., Hanafi, M. M., & Sahebi, M. (2017). Genetic analysis of rust resistance genes in global wheat cultivars: an overview. *Biotechnology & Biotechnological Equipment, 31*(3), 431-445. doi:10.1080/13102818.2017.1304180

Anikster, Y. (1986). Teliospore Germination in Some Rust Fungi. *Phytopathology, 76*(10), 1026-1030. doi:DOI 10.1094/Phyto-76-1026

Anker, C. C., & Niks, R. E. (2001). Prehaustorial resistance to the wheat leaf rust fungus, Puccinia triticina, in Triticum monococcum (ss). *Euphytica, 117*(3), 209-215.

Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., . . . Poland, J. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science, 361*(6403), eaar7191.

Appels, R., Eversole, K., Stein, N., Feuillet, C., Keller, B., Rogers, J., . . . Wang, L. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science, 361*(6403), eaar7191. doi:10.1126/science.aar7191

Aravind, L., & Koonin, E. V. (1999). Fold prediction and evolutionary analysis of the POZ domain: Structural and evolutionary relationship with the potassium channel tetramerization domain. *Journal Of Molecular Biology, 285*(4), 1353-1361. doi:DOI 10.1006/jmbi.1998.2394

Atchley, W. R., & Fitch, W. M. (1997). A natural classification of the basic helix-loop-helix class of transcription factors. *Proceedings of the National Academy of Sciences of the United States of America, 94*(10), 5172-5176. doi:10.1073/pnas.94.10.5172

Ayliffe, M. A., Frost, D. V., Finnegan, E. J., Lawrence, G. J., Anderson, P. A., & Ellis, J. G. (1999). Analysis of alternative transcripts of the flax L6 rust resistance gene. *Plant Journal, 17*(3), 287-292. doi:DOI 10.1046/j.1365-313X.1999.00377.x

Bailey, P. C., Schudoma, C., Jackson, W., Baggs, E., Dagdas, G., Haerty, W., . . . Krasileva, K. V. (2018). Dominant integration locus drives continuous diversification of plant immune receptors with exogenous domain fusions. *Genome Biology, 19*. doi:ARTN 2310.1186/s13059-018-1392-6

Boller, T., & He, S. Y. (2009). Innate immunity in plants: an arms race between pattern recognition receptors in plants and effectors in microbial pathogens. *Science (New York, N.Y.), 324*(5928), 742-744. doi:10.1126/science.1171647

Bowman, S. M., & Free, S. J. (2006). The structure and synthesis of the fungal cell wall. *Bioessays, 28*(8), 799-808.

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol, 34*(5), 525-527. doi:10.1038/nbt.3519

Brooks, D. M., Bender, C. L., & Kunkel, B. N. (2005). The Pseudomonas syringae phytotoxin coronatine promotes virulence by overcoming salicylic acid-dependent defences in Arabidopsis thaliana. *Molecular Plant Pathology, 6*(6), 629-639. doi:10.1111/j.1364-3703.2005.00311.x

Browse, J. (2009). Jasmonate Passes Muster: A Receptor and Targets for the Defense Hormone. *Annual Review of Plant Biology, 60*(1), 183-205. doi:10.1146/annurev.arplant.043008.092007

Campbell, J., & Huang, L. (2010). Silencing of multiple genes in wheat using Barley stripe mosaic virus. *Journal of Biotech Research, 2*(1), 12-20.

Campbell, J., Zhang, H., Giroux, M. J., Feiz, L., Jin, Y., Wang, M., . . . Huang, L. (2012). A mutagenesis-derived broad-spectrum disease resistance locus in wheat. *Theoretical and Applied Genetics, 125*(2), 391-404. doi:10.1007/s00122-012-1841-7

Cantu, D., Yang, B. J., Ruan, R., Li, K., Menzo, V., Fu, D. L., . . . Dubcovsky, J. (2013). Comparative analysis of protein-protein interactions in the defense response of rice and wheat. *Bmc Genomics, 14*, 166. doi:Artn 16610.1186/1471-2164-14-166

Cao, H., Bowling, S. A., Gordon, A. S., & Dong, X. N. (1994). Characterization Of an Arabidopsis Mutant That Is Nonresponsive To Inducers Of Systemic Acquired-Resistance. *Plant Cell, 6*(11), 1583-1592.

Cao, H., Glazebrook, J., Clarke, J. D., Volko, S., & Dong, X. N. (1997). The Arabidopsis NPR1 gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats. *Cell, 88*(1), 57-63. doi:Doi 10.1016/S0092-8674(00)81858-9

Cao, Y., Li, K., Li, Y., Zhao, X., & Wang, L. (2020). MYB Transcription Factors as Regulators of Secondary Metabolism in Plants. *Biology, 9*(3). doi:10.3390/biology9030061

Chandra, S., Singh, D., Pathak, J., Kumari, S., Kumar, M., Poddar, R., . . . Mukhopadhyay, K. (2016). De Novo Assembled Wheat Transcriptomes Delineate Differentially Expressed Host Genes in Response to Leaf Rust Infection. *PLoS One, 11*(2), e0148453. doi:10.1371/journal.pone.0148453

Chaudhary, B., Flagel, L., Stupar, R. M., Udall, J. A., Verma, N., Springer, N. M., & Wendel, J. F. (2009). Reciprocal silencing, transcriptional bias and functional divergence of homeologs in polyploid cotton (Gossypium). *Genetics, 182*(2), 503-517.

Chen, J., Clinton, M., Qi, G., Wang, D., Liu, F., & Fu, Z. Q. (2020). Reprogramming and remodeling: transcriptional and epigenetic regulation of salicylic acid-mediated plant defense. *Journal of Experimental Botany*. doi:10.1093/jxb/eraa072

Chen, W. J., & Zhu, T. (2004). Networks of transcription factors with roles in environmental stress response. *Trends in Plant Science, 9*(12), 591-596. doi:https://doi.org/10.1016/j.tplants.2004.10.007

Chen, X., Truksa, M., Snyder, C. L., El-Mezawy, A., Shah, S., & Weselake, R. J. (2011). Three homologous genes encoding sn-glycerol-3-phosphate acyltransferase 4 exhibit different expression patterns and functional divergence in Brassica napus. *Plant physiology, 155*(2), 851-865.

Chen, X. M. (2014). Integration of cultivar resistance and fungicide application for control of wheat stripe rust. *Canadian Journal of Plant Pathology, 36*(3), 311-326. doi:10.1080/07060661.2014.924560

Chern, M., Fitzgerald, H. A., Canlas, P. E., Navarre, D. A., & Ronald, P. C. (2005). Overexpression of a rice NPR1 homolog leads to constitutive activation of defense response and hypersensitivity to light. *Molecular Plant-Microbe Interactions, 18*(6), 511-520. doi:10.1094/Mpmi-18.0511

de Toledo Thomazella, D. P., Brail, Q., Dahlbeck, D., & Staskawicz, B. (2016). CRISPR-Cas9 mediated mutagenesis of a DMR6 ortholog in tomato confers broad-spectrum disease resistance. *bioRxiv*, 064824.

Devaiah, B. N., Madhuvanthi, R., Karthikeyan, A. S., & Raghothama, K. G. (2009). Phosphate Starvation Responses and Gibberellic Acid Biosynthesis are Regulated by the MYB62 Transcription Factor in Arabidopsis. *Molecular plant, 2*(1), 43-58. doi:https://doi.org/10.1093/mp/ssn081

Devos, K. M., Dubcovsky, J., Dvorak, J., Chinoy, C. N., & Gale, M. D. (1995). Structural Evolution Of Wheat Chromosomes 4a, 5a, And 7b And Its Impact on Recombination. *Theoretical And Applied Genetics, 91*(2), 282-288. doi:Doi 10.1007/Bf00220890

Dinesh-Kumar, S. P., & Baker, B. J. (2000). Alternatively spliced N resistance gene transcripts: Their possible role in tobacco mosaic virus resistance. *Proceedings Of the National Academy Of Sciences Of the United States Of America, 97*(4), 1908-1913. doi:DOI 10.1073/pnas.020367497

Dobon, A., Bunting, D. C. E., Cabrera-Quio, L. E., Uauy, C., & Saunders, D. G. O. (2016). The host-pathogen interaction between wheat and yellow rust induces temporally coordinated waves of gene expression. *Bmc Genomics, 17*, 380-380. doi:10.1186/s12864-016-2684-4

Dodds, P. N., Lawrence, G. J., Catanzariti, A.-M., Ayliffe, M. A., & Ellis, J. G. J. T. P. C. (2004). The Melampsora lini AvrL567 avirulence genes are expressed in haustoria and their products are recognized inside plant cells. *Plant Cell, 16*(3), 755-768.

Dodds, P. N., Lawrence, G. J., & Ellis, J. G. (2001). Six amino acid changes confined to the leucine-rich repeat beta-strand/beta-turn motif determine the difference between the P and P2 rust resistance specificities in flax. *Plant Cell, 13*(1), 163-178. doi:DOI 10.1105/tpc.13.1.163

Dombrecht, B., Xue, G. P., Sprague, S. J., Kirkegaard, J. A., Ross, J. J., Reid, J. B., . . . Kazan, K. (2007). MYC2 Differentially Modulates Diverse Jasmonate-Dependent Functions in Arabidopsis. *The Plant cell, 19*(7), 2225. doi:10.1105/tpc.106.048017

Dong, X. N. (2004). NPR1, all things considered. *Current Opinion In Plant Biology, 7*(5), 547-552. doi:10.1016/j.pbi.2004.07.005

Dossa, K., Mmadi, M. A., Zhou, R., Liu, A., Yang, Y., Diouf, D., . . . Zhang, X. (2019). Ectopic expression of the sesame MYB transcription factor SiMYB305 promotes root growth and modulates ABA-mediated tolerance to drought and salt stresses in Arabidopsis. *AoB PLANTS, 12*(1). doi:10.1093/aobpla/plz081

Dvorak, J. (1972). Genetic Variability in Aegilops-Speltoides Affecting Homoeologous Pairing in Wheat. *Canadian Journal of Genetics and Cytology, 14*(2), 371-&.

Dvorak, J., Diterlizzi, P., Zhang, H. B., & Resta, P. (1993). The Evolution of Polyploid Wheats - Identification of the a-Genome Donor Species. *Genome, 36*(1), 21-31. doi:DOI 10.1139/g93-004

Ellis, J. G., Lagudah, E. S., Spielmeyer, W., & Dodds, P. N. (2014). The past, present and future of breeding rust resistant wheat. *Frontiers in Plant Science, 5*(641). doi:10.3389/fpls.2014.00641

Feiz, L., Beecher, B. S., Martin, J. M., & Giroux, M. J. (2009). In Planta Mutagenesis Determines the Functional Regions of the Wheat Puroindoline Proteins. *Genetics, 183*(3), 853-860. doi:10.1534/genetics.109.106013

Felton, G. W., & Korth, K. L. (2000). Trade-offs between pathogen and herbivore resistance. *Current Opinion in Plant Biology, 3*(4), 309-314. doi:10.1016/s1369-5266(00)00086-8

Fernández-Calvo, P., Chini, A., Fernández-Barbero, G., Chico, J.-M., Gimenez-Ibanez, S., Geerinck, J., . . . Solano, R. (2011). The Arabidopsis bHLH Transcription Factors MYC3 and MYC4 Are Targets of JAZ Repressors and Act Additively with MYC2 in the Activation of Jasmonate Responses. *The Plant cell, 23*(2), 701. doi:10.1105/tpc.110.080788

Filichkin, S., Priest, H. D., Megraw, M., & Mockler, T. C. (2015). Alternative splicing in plants: directing traffic at the crossroads of adaptation and environmental stress. *Current Opinion In Plant Biology, 24*, 125-135. doi:10.1016/j.pbi.2015.02.008

Finver, S. N., Nishikura, K., Finger, L. R., Haluska, F. G., Finan, J., Nowell, P. C., & Croce, C. M. (1988). Sequence analysis of the MYC oncogene involved in the t(8;14)(q24;q11) chromosome translocation in a human leukemia T-cell line indicates that putative regulatory regions are not altered. *Proceedings of the National Academy of Sciences, 85*(9), 3052-3056. doi:10.1073/pnas.85.9.3052

Flor, H. H. (1971). Current status of the gene-for-gene concept. *Annual Review of Phytopathology, 9*(1), 275-296.

Fu, Z. Q., Yan, S. P., Saleh, A., Wang, W., Ruble, J., Oka, N., . . . Dong, X. N. (2012). NPR3 and NPR4 are receptors for the immune signal salicylic acid in plants. *Nature, 486*(7402), 228-+. doi:10.1038/nature11162

Gao, C. S., Kou, X. J., Li, H. P., Zhang, J. B., Saad, A. S. I., & Liao, Y. C. (2013). Inverse effects of Arabidopsis NPR1 gene on fusarium seedling blight and fusarium head blight in transgenic wheat. *Plant Pathology, 62*(2), 383-392. doi:10.1111/j.1365-3059.2012.02656.x

Gao, Y., Jia, S., Wang, C., Wang, F., Wang, F., & Zhao, K. (2016). BjMYB1, a transcription factor implicated in plant defence through activating BjCHI1 chitinase expression by binding to a W-box-like element. *Journal of Experimental Botany, 67*(15), 4647-4658. doi:10.1093/jxb/erw240

Garnica, D. P., Nemri, A., Upadhyaya, N. M., Rathjen, J. P., & Dodds, P. N. (2014). The Ins and Outs of Rust Haustoria. *PLOS Pathogens, 10*(9), e1004329. doi:10.1371/journal.ppat.1004329

Gassmann, W., Hinsch, M. E., & Staskawicz, B. J. (1999). The Arabidopsis RPS4 bacterial-resistance gene is a member of the TIR-NBS-LRR family of disease-resistance genes. *Plant Journal, 20*(3), 265-277. doi:DOI 10.1046/j.1365-313X.1999.t01-1-00600.x

Gemayel, R., Chavali, S., Pougach, K., Legendre, M., Zhu, B., Boeynaems, S., . . . Verstrepen, K. J. (2015). Variable Glutamine-Rich Repeats Modulate Transcription Factor Activity. *Molecular cell, 59*(4), 615-627. doi:10.1016/j.molcel.2015.07.003

Gu, M., Zhang, J., Li, H., Meng, D., Li, R., Dai, X., . . . xu, G. (2017). Maintenance of phosphate homeostasis and root development are coordinately regulated by MYB1, an R2R3-type MYB transcription factor in rice. *Journal of Experimental Botany, 68*(13), 3603-3615. doi:10.1093/jxb/erx174

Hahn, M., & Mendgen, K. (2001). Signal and nutrient exchange at biotrophic plant–fungus interfaces. *Current Opinion in Plant Biology, 4*(4), 322-327.

HEATH, M. C. (1997). Signalling between pathogenic rust fungi and resistant or susceptible host plants. *Annals of Botany, 80*(6), 713-720.

Holleley, C., & Sutcliffe, A. (2009). 96 Well DNA Extraction Protocol. In.

Hovav, R., Udall, J. A., Chaudhary, B., Rapp, R., Flagel, L., & Wendel, J. F. (2008). Partitioned expression of duplicated genes during development and evolution of a single cell in a polyploid plant. *Proceedings of the National Academy of Sciences, 105*(16), 6191-6195.

Hu, Z., Song, N., Xing, J., Chen, Y., Han, Z., Yao, Y., . . . Sun, Q. (2013). Overexpression of Three TaEXPA1 Homoeologous Genes with Distinct Expression Divergence in Hexaploid Wheat Exhibit Functional Retention in Arabidopsis. *PLoS One, 8*(5), e63667. doi:10.1371/journal.pone.0063667

Huang, L., Brooks, S. A., Li, W., Fellers, J. P., Trick, H. N., & Gill, B. S. (2003). Map-based cloning of leaf rust resistance gene Lr21 from the large and polyploid genome of bread wheat. *Genetics, 164*(2), 655-664.

Huang, S., Sirikhachornkit, A., Su, X., Faris, J., Gill, B., Haselkorn, R., & Gornicki, P. (2002). Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the Triticum/Aegilops complex and the evolutionary history of polyploid wheat. *Proc Natl Acad Sci U S A, 99*(12), 8133-8138. doi:10.1073/pnas.072223799

Hussain, R. M. F., Sheikh, A. H., Haider, I., Quareshy, M., & Linthorst, H. J. M. (2018). Arabidopsis WRKY50 and TGA Transcription Factors Synergistically Activate Expression of PR1. *Frontiers in Plant Science, 9*(930). doi:10.3389/fpls.2018.00930

International Wheat Genome Sequencing, C. (2014). A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. *Science, 345*(6194), 1251788. doi:10.1126/science.1251788

Jackie, C., & Huang, L. (2010). Silencing of multiple genes in wheat using Barley stripe mosaic virus. *Journal of Biotech Research, 2*.

Jenkins, J. A. (1929). Chromosome homologies in wheat and Aegilops. *American Journal of Botany, 16*(4), 238-245. doi:Doi 10.2307/2435839

Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B., & Weeks, D. P. (2013). Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic acids research, 41*(20), e188-e188.

Jin, Y., Singh, R., Ward, R., Wanyera, R., Kinyua, M., Njau, P., . . . Yahyaoui, A. (2007). Characterization of seedling infection types and adult plant infection responses of monogenic Sr gene lines to race TTKS of Puccinia graminis f. sp. tritici. *Plant Disease, 91*(9), 1096-1099.

Jordan, T., Schornack, S., & Lahaye, T. (2002). Alternative splicing of transcripts encoding Toll-like plant resistance proteins - what's the functional relevance to innate immunity? *Trends Plant Sci, 7*(9), 392-398.

Kazan, K., & Manners, J. M. (2013). MYC2: The Master in Action. *Molecular plant, 6*(3), 686-703. doi:https://doi.org/10.1093/mp/sss128

Kihara, H. (1954). Considerations on the Evolution and Distribution of Aegilops Species Based on the Analyser-Method. *Cytologia, 19*(4), 336-357. doi:DOI 10.1508/cytologia.19.336

Kim, D., Alptekin, B., & Budak, H. (2018). CRISPR/Cas9 genome editing in wheat. *Functional & Integrative Genomics, 18*(1), 31-41. doi:10.1007/s10142-017-0572-x

Klarzynski, O., Plesse, B., Joubert, J. M., Yvin, J. C., Kopp, M., Kloareg, B., & Fritig, B. (2000). Linear beta-1,3 glucans are elicitors of defense responses in tobacco. *Plant physiology, 124*(3), 1027-1038. doi:10.1104/pp.124.3.1027

Klempnauer, K.-H., Symonds, G., Evan, G. I., & Michael Bishop, J. (1984). Subcellular localization of proteins encoded by oncogenes of avian myeloblastosis virus and avian leukemia virus E26 and by the chicken c-myb gene. *Cell, 37*(2), 537-547. doi:https://doi.org/10.1016/0092-8674(84)90384-2

Krasileva, K. V., Vasquez-Gross, H. A., Howell, T., Bailey, P., Paraiso, F., Clissold, L., . . . Dubcovsky, J. (2017). Uncovering hidden variation in polyploid wheat. *Proceedings of the National Academy of Sciences of the United States of America, 114*(6), E913-E921. doi:10.1073/pnas.1619268114

Krattinger, S. G., Lagudah, E. S., Spielmeyer, W., Singh, R. P., Huerta-Espino, J., McFadden, H., . . . Keller, B. (2009). A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science, 323*(5919), 1360-1363.

Kumar, R., Kaur, A., Pandey, A., Mamrutha, H. M., & Singh, G. P. (2019). CRISPR-based genome editing in wheat: a comprehensive review and future prospects. *Molecular Biology Reports, 46*(3), 3557-3569. doi:10.1007/s11033-019-04761-3

Lawrence, G. J., Finnegan, E. J., Ayliffe, M. A., & Ellis, J. G. (1995). The L6 gene for flax rust resistance is related to the Arabidopsis bacterial resistance gene RPS2 and the tobacco viral resistance gene N. *The Plant cell, 7*(8), 1195-1206.

Le Henanff, G., Farine, S., Kieffer-Mazet, F., Miclot, A. S., Heitz, T., Mestre, P., . . . Chong, J. (2011). Vitis vinifera VvNPR1.1 is the functional ortholog of AtNPR1 and its overexpression in grapevine triggers constitutive activation of PR genes and enhanced resistance to powdery mildew. *Planta, 234*(2), 405-417. doi:10.1007/s00425-011-1412-1

Lee, M.-W., Qi, M., & Yang, Y. (2001). A Novel Jasmonic Acid-Inducible Rice myb Gene Associates with Fungal Infection and Host Cell Death. *Molecular Plant-Microbe Interactions®, 14*(4), 527-535. doi:10.1094/MPMI.2001.14.4.527

Lee, S.-K., Song, M.-Y., Seo, Y.-S., Kim, H.-K., Ko, S., Cao, P.-J., . . . Lee, S. J. G. (2009). Rice Pi5-mediated resistance to Magnaporthe oryzae requires the presence of two coiled-coil–nucleotide-binding–leucine-rich repeat genes. *Genetics, 181*(4), 1627-1638.

Li, J., Brader, G., & Palva, E. T. (2004). The WRKY70 transcription factor: A node of convergence for jasmonate-mediated and salicylate-mediated signals in plant defense. *Plant Cell, 16*(2), 319-331. doi:10.1105/tpc.016980

Li, T., Liu, B., Spalding, M. H., Weeks, D. P., & Yang, B. (2012). High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nature biotechnology, 30*(5), 390.

Li, X., Clarke, J. D., Zhang, Y. L., & Dong, X. N. (2001). Activation of an EDS1-mediated R-gene pathway in the snc1 mutant leads to constitutive, NPR1-independent pathogen resistance. *Molecular Plant-Microbe Interactions, 14*(10), 1131-1139. doi:Doi 10.1094/Mpmi.2001.14.10.1131

Line, R. F., & Qayoum, A. (1992). Virulence, aggressiveness, evolution, and distribution of races of Puccinia striiformis (the cause of stripe rust of wheat) in North America, 1968-87. *Technical bulletin-United States Department of Agriculture*(1788).

Line, R. F., & Qayoum, A. J. T. b. (1992). Virulence, aggressiveness, evolution and distribution of races of Puccinia striiformis (the cause of stripe rust of wheat) in North America, 1968-87.

Ling, H. Q., Zhao, S., Liu, D., Wang, J., Sun, H., Zhang, C., . . . Wang, J. (2013). Draft genome of the wheat A-genome progenitor Triticum urartu. *Nature, 496*(7443), 87-90. doi:10.1038/nature11997

Liu, C., Atkinson, M., Chinoy, C., Devos, K., & Gale, M. (1992). Nonhomoeologous translocations between group 4, 5 and 7 chromosomes within wheat and rye. *Theor. Appl. Genet, 83*(3), 305-312.

Liu, R., Lü, B., Wang, X., Zhang, C., Zhang, S., Qian, J., . . . Dong, H. (2010). Thirty-seven transcription factor genes differentially respond to a harpin protein and affect resistance to the green peach aphid in Arabidopsis. *Journal of Biosciences, 35*(3), 435-450.

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *Methods., 25*(4), 402-408.

Loegering, W. Q., & Ellingboe, A. (1987). HH Flor: pioneer in phytopathology. *Annual Review of Phytopathology, 25*(1), 59-66.

Luo, M. C., Gu, Y. Q., Puiu, D., Wang, H., Twardziok, S. O., Deal, K. R., . . . Dvorak, J. (2017). Genome sequence of the progenitor of the wheat D genome Aegilops tauschii. *Nature, 551*(7681), 498-+. doi:10.1038/nature24486

Makandar, R., Essig, J. S., Schapaugh, M. A., Trick, H. N., & Shah, J. J. M. P.-M. I. (2006). Genetically engineered resistance to Fusarium head blight in wheat by expression of Arabidopsis NPR1. *Mol. Plant Microbe Interact., 19*(2), 123-129.

Malnoy, M., Jin, Q., Borejsza-Wysocka, E., He, S., & Aldwinckle, H. (2007). Overexpression of the apple MpNPR1 gene confers increased disease resistance in Malus× domestica. *Mol. Plant Microbe Interact., 20*(12), 1568-1580.

Malnoy, M., Viola, R., Jung, M.-H., Koo, O.-J., Kim, S., Kim, J.-S., . . . Nagamangala Kanchiswamy, C. (2016). DNA-free genetically edited grapevine and apple protoplast using CRISPR/Cas9 ribonucleoproteins. *Frontiers in Plant Science, 7*, 1904.

McCallum, C. M., Comai, L., Greene, E. A., & Henikoff, S. (2000). TILLING for Plant Functional Genomics. *Plant physiology, 123*(2), 439. doi:10.1104/pp.123.2.439

Mcfadden, E. S., & Sears, E. R. (1946a). The Origin of Triticum Spelta and Its Free-Threshing Hexaploid Relatives. *Journal of Heredity, 37*(3), 81-89.

McFadden, E. S., & Sears, E. R. (1946b). The origin of triticum spelta and its free-threshing hexaploid relatives*. *Journal of Heredity, 37*(3), 81-89. doi:10.1093/oxfordjournals.jhered.a105590

McIntosh, R. A., Wellings, C. R., & Park, R. F. (1995). *Wheat rusts: an atlas of resistance genes*: Csiro Publishing.

Mendgen, K., Hahn, M., & Deising, H. (1996). Morphogenesis and mechanisms of penetration by plant pathogenic fungi. *Annual Review of Phytopathology, 34*(1), 367-386. doi:10.1146/annurev.phyto.34.1.367

Miki, Y., Yoshida, K., Mizuno, N., Nasuda, S., Sato, K., & Takumi, S. (2019). Origin of wheat B-genome chromosomes inferred from RNA sequencing analysis of leaf transcripts from section Sitopsis species of Aegilops. *DNA research : an international journal for rapid publication of reports on genes and genomes, 26*(2), 171-182. doi:10.1093/dnares/dsy047

Mittal, S., & Davis, K. R. (1995). Role of the phytotoxin coronatine in the infection of Arabidopsis thaliana by Pseudomonas syringae pv. tomato. *MPMI-Molecular Plant Microbe Interactions, 8*(1), 165-171.

Moore, J. W., Loake, G. J., & Spoel, S. H. (2011). Transcription dynamics in plant immunity. *The Plant cell, 23*(8), 2809-2820.

Moreau, M., Tian, M., & Klessig, D. F. (2012). Salicylic acid binds NPR3 and NPR4 to regulate NPR1-dependent defense responses. *Cell Res., 22*(12), 1631.

Morel, J. B., & Dangl, J. L. (1997). The hypersensitive response and the induction of cell death in plants. *Cell Death And Differentiation, 4*(8), 671-683. doi:DOI 10.1038/sj.cdd.4400309

Mou, Z., Fan, W., & Dong, X. (2003). Inducers of plant systemic acquired resistance regulate NPR1 function through redox changes. *Cell, 113*(7), 935-944.

Murray, F., Kalla, R., Jacobsen, J., & Gubler, F. (2003). A role for HvGAMYB in anther development. *The Plant Journal, 33*(3), 481-491. doi:10.1046/j.1365-313X.2003.01641.x

Muthukrishnan, S., Liang, G. H., Trick, H. N., & Gill, B. S. (2001). Pathogenesis-related proteins and their genes in cereals. *Plant Cell, Tissue and Organ Culture, 64*(2-3), 93.

Narusaka, M., Shirasu, K., Noutoshi, Y., Kubo, Y., Shiraishi, T., Iwabuchi, M., & Narusaka, Y. (2009). RRS1 and RPS4 provide a dual Resistance-gene system against fungal and bacterial pathogens. *Plant J., 60*(2), 218-226.

Nekrasov, V., Wang, C., Win, J., Lanz, C., Weigel, D., & Kamoun, S. (2017). Rapid generation of a transgene-free powdery mildew resistant tomato by genome deletion. *Scientific reports, 7*(1), 482.

Niks, R., & Dekens, R. (1991). Prehaustorial and posthaustorial resistance to wheat leaf rust in diploid wheat seedlings. *Phytopathology (USA)*.

Niu, Y., Figueroa, P., & Browse, J. (2011). Characterization of JAZ-interacting bHLH transcription factors that regulate jasmonate responses in Arabidopsis. *Journal of Experimental Botany, 62*(6), 2143-2154. doi:10.1093/jxb/erq408

Nomura, T., Ishihara, A., Yanagita, R. C., Endo, T. R., & Iwamura, H. (2005). Three genomes differentially contribute to the biosynthesis of benzoxazinones in hexaploid wheat. *Proceedings of the National Academy of Sciences, 102*(45), 16490-16495.

Olivera, P. D., Rouse, M. N., & Jin, Y. (2018). Identification of New Sources of Resistance to Wheat Stem Rust in Aegilops spp. in the Tertiary Genepool of Wheat. *Frontiers in Plant Science, 9*(1719). doi:10.3389/fpls.2018.01719

Pajerowska-Mukhtar, K. M., Emerine, D. K., & Mukhtar, M. S. (2013). Tell me more: roles of NPRs in plant immunity. *Trends Plant Sci, 18*(7), 402-411.

Pathak, G. N. (1940). Studies in the cytology of cereals. *Journal of Genetics, 39*(3), 437-467. doi:Doi 10.1007/Bf02982884

Paz-Ares, J., Ghosal, D., Wienand, U., Peterson, P. A., & Saedler, H. (1987). The regulatory c1 locus of Zea mays encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. *The EMBO journal, 6*(12), 3553-3558.

Peng, A., Chen, S., Lei, T., Xu, L., He, Y., Wu, L., . . . Zou, X. (2017). Engineering canker-resistant plants through CRISPR/Cas9-targeted editing of the susceptibility gene Cs LOB 1 promoter in citrus. *Plant biotechnology journal, 15*(12), 1509-1519.

Periyannan, S., Moore, J., Ayliffe, M., Bansal, U., Wang, X., Huang, L., . . . Bariana, H. (2013). The gene Sr33, an ortholog of barley Mla genes, encodes resistance to wheat stem rust race Ug99. *Science, 341*(6147), 786-788.

Petersen, G., Seberg, O., Yde, M., & Berthelsen, K. (2006). Phylogenetic relationships of Triticum and Aegilops and evidence for the origin of the A, B, and D genomes of common wheat (Triticum aestivum). *Mol Phylogenet Evol, 39*(1), 70-82. doi:10.1016/j.ympev.2006.01.023

Pieterse, C. M., Ton, J., & Van Loon, L. (2001). Cross-talk between plant signalling pathways: boost or burden. *AgBiotechNet, 3*, 1-8.

Pretorius, Z., Singh, R., Wagoire, W., & Payne, T. (2000). Detection of virulence to wheat stem rust resistance gene Sr31 in Puccinia graminis. f. sp. tritici in Uganda. *Plant Disease, 84*(2), 203-203.

Raffaele, S., Rivas, S., & Roby, D. (2006). An essential role for salicylic acid in AtMYB30-mediated control of the hypersensitive cell death program in Arabidopsis. *FEBS Letters, 580*(14), 3498-3504. doi:10.1016/j.febslet.2006.05.027

Rate, D. N., & Greenberg, J. T. (2001). The Arabidopsis aberrant growth and death2 mutant shows resistance to Pseudomonas syringae and reveals a role for NPR1 in suppressing hypersensitive cell death. *Plant J., 27*(3), 203-211.

Riley, R., Chapman, V., & Kimber, G. (1961). Origin of Genetic Control of Diploid-Like Behavior of Polyploid Wheat. *Journal of Heredity, 52*(1), 22-&. doi:DOI 10.1093/oxfordjournals.jhered.a107015

Robinson, R. A. (1976). Plant Pathosystems. In *Plant Pathosystems* (pp. 15-31). Berlin, Heidelberg: Springer Berlin Heidelberg.

Rouse, M. N., Olson, E. L., Gill, B. S., Pumphrey, M. O., & Jin, Y. (2011). Stem Rust Resistance in Aegilops tauschii Germplasm. *Crop Science, 51*(5), 2074-2078. doi:10.2135/cropsci2010.12.0719

Salamini, F., Ozkan, H., Brandolini, A., Schafer-Pregl, R., & Martin, W. (2002). Genetics and geography of wild cereal domestication in the near east. *Nat Rev Genet, 3*(6), 429-441. doi:10.1038/nrg817

Sánchez-León, S., Gil-Humanes, J., Ozuna, C. V., Giménez, M. J., Sousa, C., Voytas, D. F., & Barro, F. (2018). Low-gluten, nontransgenic wheat engineered with CRISPR/Cas9. *Plant biotechnology journal, 16*(4), 902-910. doi:10.1111/pbi.12837

Sanjana, N. E., Cong, L., Zhou, Y., Cunniff, M. M., Feng, G., & Zhang, F. (2012). A transcription activator-like effector toolbox for genome engineering. *Nature protocols, 7*(1), 171-192. doi:10.1038/nprot.2011.431

Sarkar, P., & Stebbins, G. L. (1956). Morphological Evidence Concerning the Origin of the B-Genome in Wheat. *American Journal of Botany, 43*(4), 297-304. doi:Doi 10.2307/2438947

Sax, K. (1922). Sterility in wheat hybrids. II. Chromosome behavior in partially sterile hybrids. *Genetics, 7*(6), 513-552.

Schweizer, F., Fernández-Calvo, P., Zander, M., Diez-Diaz, M., Fonseca, S., Glauser, G., . . . Reymond, P. (2013). Arabidopsis Basic Helix-Loop-Helix Transcription Factors MYC2, MYC3, and MYC4 Regulate Glucosinolate Biosynthesis, Insect Performance, and Feeding Behavior. *The Plant cell, 25*(8), 3117. doi:10.1105/tpc.113.115139

Scofield, S. R., Huang, L., Brandt, A. S., & Gill, B. S. (2005). Development of a Virus-Induced Gene-Silencing System for Hexaploid Wheat and Its Use in Functional Analysis of the Lr21-Mediated Leaf Rust Resistance Pathway. *Plant physiology, 138*(4), 2165. doi:10.1104/pp.105.061861

Serfling, A., Templer, S. E., Winter, P., & Ordon, F. (2016). Microscopic and Molecular Characterization of the Prehaustorial Resistance against Wheat Leaf Rust (Puccinia triticina) in Einkorn (Triticum monococcum). *Frontiers in Plant Science, 7*(1668). doi:10.3389/fpls.2016.01668

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., . . . Qiu, J.-L. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. *Nature biotechnology, 31*(8), 686.

Shirano, Y., Kachroo, P., Shah, J., & Klessig, D. F. (2002). A gain-of-function mutation in an Arabidopsis Toll Interleukin1 Receptor–Nucleotide Binding Site–Leucine-Rich Repeat type R gene triggers defense responses and results in enhanced disease resistance. *Plant Cell, 14*(12), 3149-3162.

Shitsukawa, N., Tahira, C., Kassai, K.-i., Hirabayashi, C., Shimizu, T., Takumi, S., . . . Murai, K. (2007). Genetic and epigenetic alteration among three homoeologous genes of a class E MADS box gene in hexaploid wheat. *The Plant cell, 19*(6), 1723-1737.

Shrawat, A. K., & Armstrong, C. L. (2018). Development and application of genetic engineering for wheat improvement. *Critical Reviews in Plant Sciences, 37*(5), 335-421.

Silverman, P., Seskar, M., Kanter, D., Schweizer, P., Metraux, J.-P., & Raskin, I. (1995). Salicylic acid in rice (biosynthesis, conjugation, and possible role). *J Plant Physiology, 108*(2), 633-639.

Singh, D., Park, R. F., McIntosh, R. A., & Bariana, H. S. (2008). Characterisation of stem rust and stripe rust seedling resistance genes in selected wheat cultivars from the united kingdom. *Journal of Plant Pathology, 90*(3), 553-562.

Song, S., Qi, T., Huang, H., Ren, Q., Wu, D., Chang, C., . . . Xie, D. (2011). The Jasmonate-ZIM Domain Proteins Interact with the R2R3-MYB Transcription Factors MYB21 and MYB24 to Affect Jasmonate-Regulated Stamen Development in Arabidopsis. *The Plant cell, 23*(3), 1000. doi:10.1105/tpc.111.083089

Spoel, S. H., & Dong, X. (2012). How do plants achieve immunity? Defence without specialized immune cells. *J Nature reviews immunology, 12*(2), 89.

Spoel, S. H., Koornneef, A., Claessens, S. M., Korzelius, J. P., Van Pelt, J. A., Mueller, M. J., . . . Kazan, K. (2003). NPR1 modulates cross-talk between salicylate-and jasmonate-dependent defense pathways through a novel function in the cytosol. *J The Plant Cell, 15*(3), 760-770.

Spoel, S. H., Koornneef, A., Claessens, S. M. C., Korzelius, J. P., Van Pelt, J. A., Mueller, M. J., . . . Pieterse, C. M. J. (2003). NPR1 Modulates Cross-Talk between Salicylate- and Jasmonate-Dependent Defense Pathways through a Novel Function in the Cytosol. *The Plant cell, 15*(3), 760. doi:10.1105/tpc.009159

Spoel, S. H., Mou, Z., Tada, Y., Spivey, N. W., Genschik, P., & Dong, X. (2009). Proteasome-mediated turnover of the transcription coactivator NPR1 plays dual roles in regulating plant immunity. *Cell, 137*(5), 860-872.

Stakman, E. C., Stewart, D. M., & Loegering, W. Q. (1962). Identification of physiologic races of *Puccinia graminis* var. *tritici*.

Stracke, R., Werber, M., & Weisshaar, B. (2001). The R2R3-MYB gene family in Arabidopsis thaliana. *Current Opinion in Plant Biology, 4*(5), 447-456. doi:https://doi.org/10.1016/S1369-5266(00)00199-0

Szurman-Zubrzycka, M. E., Zbieszczyk, J., Marzec, M., Jelonek, J., Chmielewska, B., Kurowska, M. M., . . . Szarejko, I. (2018). HorTILLUS-A Rich and Renewable Source of Induced Mutations for Forward/Reverse Genetics and Pre-breeding Programs in Barley (Hordeum vulgare L.). *Frontiers in Plant Science, 9*, 216-216. doi:10.3389/fpls.2018.00216

Talajoor, M., Wang, X., Zhang, H., Zhang, P., Bhavani, S., Tabe, L., . . . Huang, L. (2016). Wheat mutant MNR220 delays haustoria formation during leaf rust pathogenesis at the seedling stage. *Canadian Journal of Plant Pathology, 38*(3), 338-347. doi:10.1080/07060661.2016.1205663

Thomma, B. P. H. J., Penninckx, I. A. M. A., Cammue, B. P. A., & Broekaert, W. F. (2001). The complexity of disease signaling in Arabidopsis. *Current Opinion in Immunology, 13*(1), 63-68. doi:https://doi.org/10.1016/S0952-7915(00)00183-7

Tsuda, K., & Somssich, I. E. (2015). Transcriptional networks in plant immunity. *New Phytologist, 206*(3), 932-947. doi:10.1111/nph.13286

Van Der Biezen, E. A., & Jones, J. D. (1998). Plant disease-resistance proteins and the gene-for-gene concept. *Trends in biochemical sciences, 23*(12), 454-456.

van der Hoorn, R. A. L., & Kamoun, S. (2008). From Guard to Decoy: A New Model for Perception of Plant Pathogen Effectors. *The Plant cell, 20*(8), 2009. doi:10.1105/tpc.108.060194

Van Loon, L. (1982). Regulation of changes in proteins and enzymes associated with active defence against virus infection. In *Active defense mechanisms in plants* (pp. 247-273): Springer.

Van Loon, L. C., & Van Strien, E. A. (1999). The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. *Physiological and Molecular Plant Pathology, 55*(2), 85-97. doi:https://doi.org/10.1006/pmpp.1999.0213

Van Moerkercke, A., Duncan, O., Zander, M., Šimura, J., Broda, M., Vanden Bossche, R., . . . Van Aken, O. (2019). A MYC2/MYC3/MYC4-dependent transcription factor network regulates water spray-responsive gene expression and jasmonate levels. *Proceedings of the National Academy of Sciences, 116*(46), 23345. doi:10.1073/pnas.1911758116

van Schie, C. C., & Takken, F. L. (2014). Susceptibility genes 101: how to be a good host. *Annual Review of Phytopathology, 52*, 551-581.

Voegele, R. T., & Mendgen, K. (2003). Rust haustoria: nutrient uptake and beyond. *New Phytologist, 159*(1), 93-100.

Wang, J. R., Luo, M. C., Chen, Z. X., You, F. M., Wei, Y. M., Zheng, Y. L., & Dvorak, J. (2013). Aegilops tauschii single nucleotide polymorphisms shed light on the origins of wheat D-genome genetic diversity and pinpoint the geographic origin of hexaploid wheat. *New Phytologist, 198*(3), 925-937. doi:10.1111/nph.12164

Wang, L., Li, H., Zhao, C., Li, S., Kong, L., Wu, W., . . . Zhang, H. (2017). The inhibition of protein translation mediated by AtGCN1 is essential for cold tolerance in Arabidopsis thaliana. *Plant, cell & environment, 40*(1), 56-68. doi:10.1111/pce.12826

Wang, M. N., & Chen, X. M. (2015). Barberry Does Not Function as an Alternate Host for Puccinia striiformis f. sp tritici in the U. S. Pacific Northwest Due to Teliospore Degradation and Barberry Phenology. *Plant Disease, 99*(11), 1500-1506. doi:10.1094/Pdis-12-14-1280-Re

Wang, M. N., Wan, A. M., & Chen, X. M. (2015). Barberry as Alternate Host Is Important for Puccinia graminis f. sp tritici But Not for Puccinia striiformis f. sp tritici in the US Pacific Northwest. *Plant Disease, 99*(11), 1507-1516. doi:10.1094/Pdis-12-14-1279-Re

Wang, X., Yang, B., Li, K., Kang, Z., Cantu, D., & Dubcovsky, J. (2016). A conserved Puccinia striiformis protein interacts with wheat NPR1 and reduces induction of pathogenesis-related genes in response to pathogens. *J Molecular Plant-Microbe Interactions, 29*(12), 977-989.

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C., & Qiu, J.-L. (2014). Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nature biotechnology, 32*(9), 947.

Williamson, M. P. (1994). The structure and function of proline-rich regions in proteins. *The Biochemical journal, 297 (Pt 2)*(Pt 2), 249-260. doi:10.1042/bj2970249

Xie, K., & Yang, Y. (2013). RNA-guided genome editing in plants using a CRISPR–Cas system. *Molecular plant, 6*(6), 1975-1983.

Yuan, Y., Zhong, S., Li, Q., Zhu, Z., Lou, Y., Wang, L., . . . Yang, D. (2007). Functional analysis of rice NPR1-like genes reveals that OsNPR1/NH1 is the rice orthologue conferring disease resistance with enhanced herbivore susceptibility. *J Plant biotechnology journal, 5*(2), 313-324.

Zhang, H., Qiu, Y., Yuan, C., Chen, X., & Huang, L. (2018). Fine-Tuning of PR Genes in Wheat Responding to Different Puccinia Rust Species. *Journal of Plant Physiology & Pathology, 06*. doi:10.4172/2329-955X.1000178

Zhang, H., Wang, X., Giroux, M. J., & Huang, L. (2017). A wheat COP9 subunit 5-like gene is negatively involved in host response to leaf rust. *Molecular Plant Pathology, 18*(1), 125-133. doi:10.1111/mpp.12467

Zhang, W., Zhang, M., Zhu, X., Cao, Y., Sun, Q., Ma, G., . . . Cai, X. (2018). Molecular cytogenetic and genomic analyses reveal new insights into the origin of the wheat B genome. *Theor Appl Genet, 131*(2), 365-375. doi:10.1007/s00122-017-3007-0

Zhang, X.-C., & Gassmann, W. (2007). Alternative splicing and mRNA levels of the disease resistance gene RPS4 are induced during defense responses. *J Plant Physiology, 145*(4), 1577-1587.

Zhang, X., Ju, H.-W., Chung, M.-S., Huang, P., Ahn, S.-J., & Kim, C. S. (2010). The R-R-Type MYB-Like Transcription Factor, AtMYBL, is Involved in Promoting Leaf Senescence and Modulates an Abiotic Stress Response in Arabidopsis. *Plant and Cell Physiology, 52*(1), 138-148. doi:10.1093/pcp/pcq180

Zhang, Y., Bai, Y., Wu, G., Zou, S., Chen, Y., Gao, C., & Tang, D. (2017). Simultaneous modification of three homoeologs of Ta EDR 1 by genome editing enhances powdery mildew resistance in wheat. *The Plant Journal, 91*(4), 714-724.

Zhao, J., Wang, M. N., Chen, X. M., & Kang, Z. S. (2016). Role of Alternate Hosts in Epidemiology and Pathogen Variation of Cereal Rusts. *Annual Review of Phytopathology, Vol 54, 54*, 207-228. doi:10.1146/annurev-phyto-080615-095851

Zhao, Y., Cheng, X., Liu, X., Wu, H., Bi, H., & Xu, H. (2018). The Wheat MYB Transcription Factor TaMYB31 Is Involved in Drought Stress Responses in Arabidopsis. *Frontiers in Plant Science, 9*(1426). doi:10.3389/fpls.2018.01426

Zhao, Y., Tian, X., Wang, F., Zhang, L., Xin, M., Hu, Z., . . . Peng, H. (2017). Characterization of wheat MYB genes responsive to high temperatures. *BMC plant biology, 17*(1), 208-208. doi:10.1186/s12870-017-1158-4

Zheng, X.-y., Spivey, N. W., Zeng, W., Liu, P.-P., Fu, Z. Q., Klessig, D. F., . . . Dong, X. (2012). Coronatine promotes Pseudomonas syringae virulence in plants by activating a signaling cascade that inhibits salicylic acid accumulation. *Cell host & microbe, 11*(6), 587-596.

Zhong, X., Xi, L., Lian, Q., Luo, X., Wu, Z., Seng, S., . . . Yi, M. (2015). The NPR1 homolog GhNPR1 plays an important role in the defense response of Gladiolus hybridus. *J Plant cell reports, 34*(6), 1063-1074.

APPENDICES

APPENDIX A

SUPPORTING INFORMATION FOR CHAPTER TWO

File 2.S1: Alignment of *TaMYC4* homeologs in ClustalW. VIGS insert highlighted green.

```
cT3 ATGGCCATGGTGGTTGATGGTGCAGGCCCCGCGGGGCTCGCCAGCGCGCCCTGCGAGCGG 60
1AL -------------------------GCCCCGCGGGGCTCGCCAGCGCGCCCTGCGAGCGG 35
1DL -------------------------------CGGGGCTCGCCAGCGCGCCCTGCGAGCGG 29
1BL ------------------------------------------------------------ 0


cT3 GCCAGGCAGGCCTACACCTTCGGCCTCCGCACCATGGTCTGCATCCCCCTCGGCACCGGC 120
1AL GCCAGGCAGGCCTACACCTTCGGCCTCCGCACCATGGTCTGCATCCCCCTCGGCACCGGC 95
1DL GCCAGGCAGGCCTACACCTTCGGCCTCCGCACCATGGTCTGCATCCCCCTCGGCACCGGC 89
1BL ------------------------------------------------------------ 0


cT3 GTGCTCGAGCTCGGCGCCACCGAGGTCATCTTCCAGACCAACGATAGCTTGGGGAGGATC 180
1AL GTGCTCGAGCTCGGCGCCACCGAGGTCATCTTCCAGACCAACGATAGCTTGGGGAGGATC 155
1DL GTGCTCGAGCTCGGCGCCACCGAGGTCATCTTCCAGACCAACGATAGCTTGGGGAGGATC 149
1BL ------------------------------------------------------------ 0


cT3 CGCTCGCTCTTCAACCTCAACGGCGGAGGAGGGGGCTCTGGATCCTGGCCGCCCATCGCC 240
1AL CGCTCGCTCTTCAACCTCAACGGCGGAGGAGGGGGCTCTGGATCCTGGCCGCCCATCGCC 215
1DL CGCTCGCTCTTCAACCTCAACGGCGGAGGAGGGGGCTCTGGATCCTGGCCGCCCATCGCG 209
1BL ------------------------------------------------------------ 0


cT3 CCGCCGCCCCAGGAGGCGGAGACGGATCCGTCCGTGCTCTGGCTCGCCGACGCGCCGGCC 300
1AL CCGCCGCCCCAGGAGGCGGAGACGGATCCGTCCGTGCTCTGGCTCGCCGACGCGCCGGCC 275
1DL CCGCCGCCGCAGGAGGCGGAGACGGATCCGTCCGTGCTCTGGCTCGCCGACGCGCCGGCC 269
1BL ------------------------------------------------------------ 0


cT3 GGGGACATGAAGGAGTCGCCGCCGTCCGTCGAGATCTCCGTCTCCAAGCCGCCGCAGCCA 360
1AL GGGGACATGAAGGAGTCGCCGCCGTCCGTCGAGATCTCCGTCTCCAAGCCGCCGCAGCCA 335
1DL GGGGACATGAAGGAGTCGCCGCCGTCCGTCGAGATCTCCGTCTCCAAACCGCCGCCGCCA 329
1BL ------------------------------------------------------------ 0


cT3 CAGCCGCCGCAGATCCATCAGTTCGAGAACGGGAGCACCAGCACGCTCACGGAGAACCCC 420
1AL CAGCCGCCGCAGATCCATCAGTTCGAGAACGGGAGCACCAGCACGCTCACGGAGAACCCC 395
1DL CAACCGCCGCAGATCCATCACTTCGAGAACGGGAGCACCAGCACGCTCACGGAGAATCCC 389
1BL ------------------------------------------------------------ 0


cT3 GGCCTCTCCGTGCACGCGCAGCAGCCTCCGCCGCAGCAGGCGGCCGCGGCGGCGCAGAGG 480
1AL GGCCTCTCCGTGCACGCGCAGCAGCCTCCGCCGCAGCAGGCGGCCGCGGCGGCGCAGAGG 455
1DL AGTCTCTCCGTGCACGCGCAGCAGCCTCCGCCCCAGCAGGCGGCTGCGGCGGCGCAGAGG 449
1BL ------------------------------------------------------------ 0


cT3 CAGAACCAGC------ACCAGCTCCAGCATCAGCACCAGCTCCAGCTCCAGCACCAGCAC 534
1AL CAGAACCAGC------ACCAGCTCCAGCATCAGCACCAGCTCCAGCTCCAGCACCAGCAC 509
1DL CAGAACCAGCACCAGCAGCAGCTCCAGCATCAGCACCAGCTCCAGCTCCAGCACCAGCAC 509
1BL ------------------------------------------------------------ 0
```

```
cT3 AATCAGGGTCCTTTCCGCCGGGAGCTCAATTTCTCAGATTTCGCGTCCAACGCATCCGTC 594
1AL AATCAGGGTCCTTTCCGCCGGGAGCTCAATTTCTCAGATTTCGCGTCCAACGCATCCGTC 569
1DL AACCAGGGTCCTTTCCGCCGGGAGCTCAATTTCTCAGATTTCGCGTCCAACGCATCCGTC 569
1BL ------------------------------------------------------------ 0

cT3 ACGGCGGAGCGGCAGCGGCGAGAGAAGCTGAACCAGCGGTTCTACACGCTCCGCGCCGTG 654
1AL ACGGCGGAGCGGCAGCGGCGAGAGAAGCTGAACCAGCGGTTCTACACGCTCCGCGCCGTG 629
1DL ACGCGC-------------------------------------------------ACGAG 580
1BL ------------------------------------------------------------ 0

cT3 GTGCCCAACGTGTCCAAGATGGACAAGGCCTCGCTGCTCGGCGACGCCATCTCCTACATC 714
1AL GTGCCCAACGTGTCCAAGATGGACAAGGCCTCGCTGCTCGGCGACGCCATCTCCTACATC 689
1DL GTGCCCAACGTGTCCAAGATGGACAAGGCCTCGCTGCTGGGCGACGCCATCTCCTACATC 640
1BL ------------------------------------------------------------ 0

cT3 AACGAGCTTCGCGGCAAGATGACGGCGCTGGAGTCGGACAAGGAGACGCTCCATTCCCAA 774
1AL AACGAGCTTCGCGGCAAGATGACGGCGCTGGAGTCGGACAAGGAGACGCTCCATTCCCAA 749
1DL AACGAGCTCCGCGGCAAGATGACGGTGCTGGAGTCTGACAAGGAGACGCTCCACTCCCAA 700
1BL ------------------------------------------------------------ 0

cT3 ATCGAGGCGCTCAAGAAGGAGCGCGACGCCCGGCCGGCCGCGCCGTCGTCGGGGATGCAC 834
1AL ATCGAGGCGCTCAAGAAGGAGCGCGACGCCCGGCCGGCCGCGCCGTCGTCGGGGATGCAC 809
1DL ATCGAGGCGCTCAAGAAGGAGCGCGACGCCCGGCCGGCCGCGCCGTCGTCGGGGATGCAC 760
1BL ----------------------------------------------CGTCGGGGATGCAC 14
                                                 *************

cT3 GACAACGGGGCGCGGTGCCACGCGGTGGAGATCGAGGCCAAGATCCTGGGGCTGGAGGCG 894
1AL GACAACGGGGCGCGGTGCCACGCGGTGGAGATCGAGGCCAAGATCCTGGGGCTGGAGGCG 869
1DL GACAACGGGGCGCGGTGCCACGCGGTCGAGATCGAGGCCAAGATCCTGGGGCTGGAGGCG 820
1BL GACAACGGGGCGCGGTGCCACGCGGTGGAGATCGAGGCCAAGATCCTGGGGCTGGAGGCG 74
    ************************** *********************************

cT3 ATGATCCGCGTGCAGTGCCACAAGCGCAACCACCCGGCGGCGAAGCTGATGACGGCGCTG 954
1AL ATGATCCGCGTGCAGTGCCACAAGCGCAACCACCCGGCGGCGAAGCTGATGACGGCGCTG 929
1DL ATGATCCGCGTGCAGTGCCACAAGCGCAACCACCCGGCGGCGAAGCTGATGACGGCGCTG 880
1BL ATGATCCGCGTGCAGTGCCACAAGCGCAACCACCCGGCGGCGAAGCTGATGACGGCGCTG 134
    ************************************************************

cT3 CGGGAGCTGGACCTGGACGTGTACCACGCCAGCGTCTCGGTGGTGAAGGACATCATGATC 1014
1AL CGGGAGCTGGACCTGGACGTGTACCACGCCAGCGTCTCGGTGGTGAAGGACATCATGATC 989
1DL CGGGAGCTGGACCTGGACGTGTACCACGCCAGCGTGTCCGTGGTGAAGGACATCATGATC 940
1BL CGGGAGCTGGACCTGGACGTGTACCACGCCAGCGTCTCCGTGGTGAAGGACATCATGATC 194
    *********************************** ** *********************

cT3 CAGCAGGTGGCGGTGAAGATGGCCACCCGGGTCTACTCCCAGGACCAGCTCAACGCGGCG 1074
1AL CAGCAGGTGGCGGTGAAGATGGCCACCCGGGTCTACTCCCAGGACCAGCTCAACGCGGCG 1049
1DL CAGCAGGTGGCGGTGAAGATGGCCACCCGGGTCTACTCCCAGGACCAGCTCAACGCGGCG 1000
1BL CAGCAGGTGGCGGTGAAGATGGCCACCCGGGTGTACTCGCAGGACCAGCTCAACGCGGCG 254
```

```
******************************* ***** ********************

cT3 CTCTACGGCCGCCTCGCCGAGCCGGGCACCGCGATGCAAATCCGCTAA  1122
1AL CTCTACGGCCGCCTCGCCGAGCCGGGCACCGCGATGCAAATCCGGTAA  1097
1DL CTCTACGGCCGCCTCGCCGAGCCGGGCACCGCGATGCAAATCCGGTAA  1048
1BL CTCTACGGCCGCCTCGCCGAGCCGGGCACCGCGATGCAAATCCGGTAA  302
    ******************************************* ***
```

File 2.S2: *TaMYC4* homeologs protein sequence for domain prediction.

>1AL-Protein (946 aa); 704-753: bHLH domain, 753-774: Leucine zipper domain

MHKAPQPKVYLFYLEPATQSWNAGDMSFLHACERSLFNPAATDLLLACMLFPM

ILASLLDPMHDSSFFSMDKLANKTHNRISHWWLTGGPIHTGYPKAGEAKSDGW

VKSVTDNPGPRSRQEATSTAVTPKFGQPSKEQEREKKRQLLQEQQRSTISATS

TATKPAQFTTLGLAPRLSSPPIHSHLSHLPLSTLPPRLLSLPPIHSPSSPSPD

PGPGPGPMNLWTDDNASMMEAFMASADMPAFPWGAAATPPPPAAVPQQPAFNQ

DTLQQRLQAIIEGSRETWTYAIFWQSSTDAGASLLGWGDGYYKGCDDADKRRQ

QPTPASAAEQEHRKRVLRELNSLIAGGGAAAPDEAVEEEVTDTEWFFLVSMTH

APCERARQAYTFGLRTMVCIPLGTGVLELGATEVIFQTNDSLGRIRSLFNLNG

GGGGSGSWPPIAPPPQEAETDPSVLWLADAPAGDMKESPPSVEISVSKPPQPQ

PPQIHQFENGSTSTLTENPGLSVHAQQPPPQQAAAAAQRQNQHQLQHQHQLQL

QHQHNQGPFRRELNFSDFASNASVKVTPPFFKPESGEILNFGADSTSRRNPSP

APPAATASLTTAPGSLFSQHTATVTAPSNDAKNNPKRSMEATSRASNTNHHQT

ATANEGMLSFSSAPTTRPSTGTGAPAKSESDHSDLEASVREVESSRVVPPPEE

KRPRKRGRKPANGREEPLNHVEAERQRREKLNQRFYALRAVVPNVSKMDKASL

LGDAISYINELRGKMTALESDKETLHSQIEALKKERDARPAAPSSGMHDNGAR

CHAVEIEAKILGLEAMIRVQCHKRNHPAAKLMTALRELDLDVYHASVSVVKDI

MIQQVAVKMATRVYSQDQLNAALYGRLAEPGTAMQIRRMDGWMHVYIIIIIII

IIIILLCFCKTKMDCPSPFSSSVVVAVVCMHASLLVLAERASMDD

>1BL-Protein (1282 aa); 1060- 1110: bHLH domain, 1110-1131: Leucine zipper domain, 1179: mutation site

MLGNGVDAGFNSSANMAYRGRGCGYSRGHAARLQQRPWPRQAVLPKSIWRQQRQRRRQS

YDEDEHEDKGANMVTGSYGVDTNWYADTGATHHITRELDKLTIRDKYHRHDQVHTASGF

ASPLHPYNLYSQMYGDQLPLQLVDTHIIPLHKGVKCLDVTSGRVYISRDVVFDETVFPF

ESLHPNAGARLKQEILLLPPNLWSFDQGGNNCTDHYDCSTSSGVAFDHVQVHGENAEEN

GVENDQNPVQNEGILHVEEEDEADAELEDDSLQPATPVRQDPSGSSRGSAPDHPRCGPA

ANCHIARQRASPARSHTQAGDSDRASSVGRLGIEPVRAHAAACRISCGERLCGNQLQQY

KCSHTAGHRFWDRISCAKQRIWCILRILCDIDSYVFSCAATSSTSTSDDVTCYNSPTKR

QEATNTAVTPKFGQPNEEEREKKRQLLQEQQSSTISAVIPLFYPGQLNKQRAGERRIGG

RRRRRRPNPHNLLLWASPPIFPHHQSIPTFPISLSPPSPALAFSPIHSPSSPSPDPGPG

PMNLWTDDNASMMEAFMASADMPAFPWGAAATPPPPAAVPQQPAFNQDTLQQRLQAIIE

GSRETWTYAIFWQSSTDAGASLLGWGDGYYKGCDDADKRRQQPTPASAAEQEHRKRVLR

ELNSLIAGGGAAAPDEAVEEEVTDTEWFFLVSMTQSFPNGMGLPGQALFAGQPTWIATG

LASAPCERARQAYTFGLRTMVCIPLGTGAYTFGLRTMVCIPLGTGVLELGATEVIFQTN

DSLGRIRSLFNLNGGGGGSGSWPPVAPPPQEAETDPSVLWLADAPAGDMKESPPSVEIS

VSKPPPPQPPQIHHFENGSTSTLTENPSLSVHAQQPPPQQAAAAAQRQNQHQLQHQHQL

QLQHQHNQGPFRRELNFSDFASNASVTVTPPFFKPESGEILNFGADSTSRRNPSPAPPA

ATASLTTAPGSLFSQHTATVTAPSNDAKNNPKRSMEATSRASNTNHHQTATANEGMLSF

SSAPTTRPSTGTGAPAKSESDHSDLEASVREVESSRVVPPPEEKRPRKRGRKPANGREE

PLNHVEAERQRREKLNQRFYALRAVVPNVSKMDKASLLGDAISYINELRGKMTALESDK

ETLHSQIEALKKERDARPAAPSSGMHDNGARCHAVEIEAKILGLEAMIRVQCHKRNHPA

AKLMTLRELDLDVYHASVSVVKDIMIQQVAVKMATRVYSQDQLNAALYGRLAEPGTAMQ

IRPIDIEMEKEGDDHVVVKSLFRRMDGWLHVYIIIIIILLCFCK

>1DL-Protein (1197 aa); 835- 884: bHLH domain, 884-905: Leucine zipper domain

MSIRRLGASDFRGVRERRSGAFSSKIWFREKRLILGTFDTAEEAARAHDATAW

RLLRPRRDMNFPDVSSQRAQDLAPLQRLFTDEDRRVHRRRQRRLAIAEMDVKT

LVLWRERFPQDIVDERQFYKQRRLERDARRRERAAYREDRRSRKQATQLKLKL

RETSGWDFEDEQHADAYIQTSEQEATNTAVTPKFGQPTRRRKERRNGSCCKNS

RAAQISAVIPLFYPGQLNKQRGSERRIGGRRRRRRPNPHNLLLWASPPIFPPH

QSIPTFPISSSLSTHPPPALAFPPIHSPSFPSPDPGPGPMNLWTDDNASMMEA

FMASADMPAFPWGAAATPPPPAAVPQQPAFNQDTLQQRLQAIIEGSRETWTYA

IFWQSSTDAGASLLGWGDGYYKGCDDADKRRQQPTPASAAEQEHRKRVLRELN

SLIAGGGAAAPDEAVEEEVTDTEWFFLVSMTQSFPNGMGLPGQALFAGQATWI

ATGLASAPCERARQAYTFGLRTMVCIPLGTGAYTFGLRTMVCIPLGTGVLELG

ATEVIFQTNDSLGRIRSLFNLNGGGGGSGSWPPIAPPPQEAETDPSVLWLADA

PAGDMKESPPSVEISVSKPPPPQPPQIHHFENGSTSTLTENPSLSVHAQQPPP

QQAAAAAQRQNQHQQQLQHQHQLQLQHQHNQGPFRRELNFSDFASNASVTVTP

PFFKPESGEILNFGADSTSRRNPSPAPPAATASLTTAPGSLFSQHTATVTAPS

NDAKNNPKRSMEATSRASNTNHHQNATANEGMLSFSSAPTTRPSTGTGAPAKS

ESDHSDLEASVREVESSRVVPPPEEKRPRKRGRKPANGREEPLNHVEAERQRR

EKLNQRFYALRAVVPNVSKMDKASLLGDAISYINELRGKMTALESDKETLHSQ

IEALKKERDARPAAPSSGMHDNGARCHAVEIEAKILGLEAMIRVQCHKRNHPA

AKLMTALRELDLDVYHASVSVVKDIMIQQVAVKMATRVYSQDQLNAALYGRLA

EPGTAMQIRTKMDCPSPFSSSVVAVVCMHASLLVLAERAWMIDMWQHWWKKRR

ISEEGEERQSKTLNEPGTQTLLATRSQAYCSALPLHMPMPPPDPCRDSPNFVV

PCPQIGASFFSTAQCRVVLARERFDNRQRVDILQSMVVIILVNAWSPISRAHG

GRCPCWQAVLGHVVGSLLTVLANERFQNDS

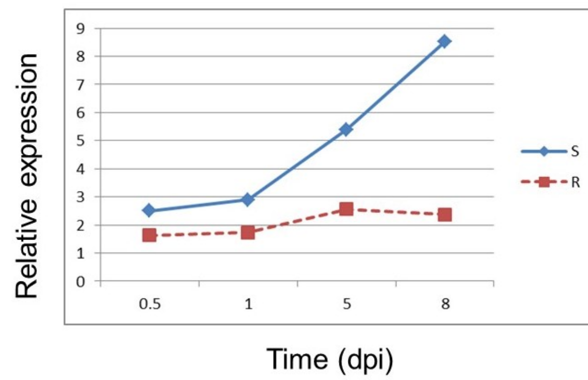


Figure 2.S1: Relative expression of *TaMYC4*. RNA sequence data over four time points generated from leaf rust inoculated susceptible (S) and resistant (R) Alpowa cultivars was analyzed in DNASTAR software.

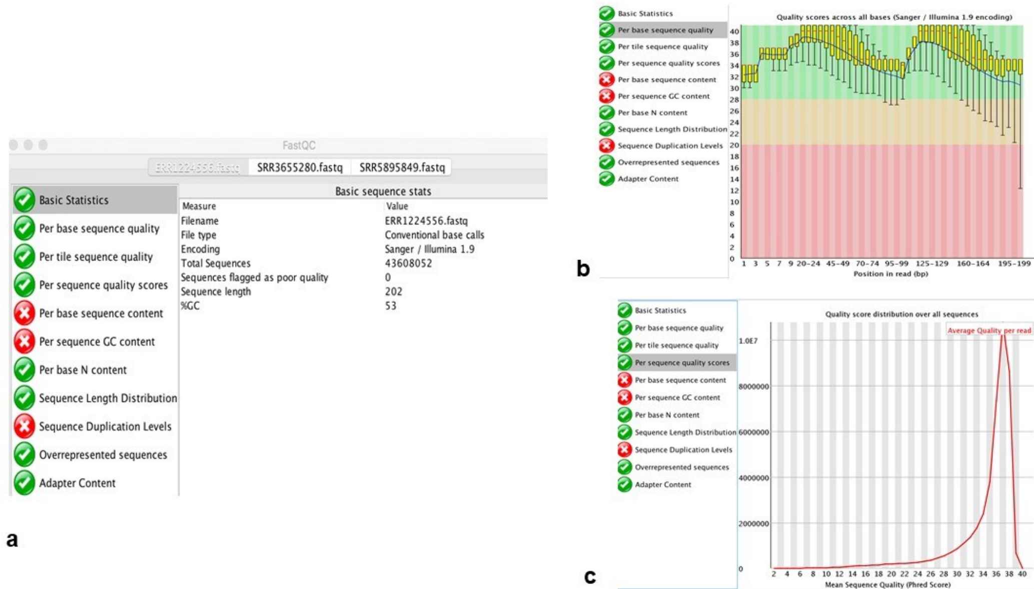


Figure 2.S2: FastQC RNA-seq quality check output. The quality of RNA-seq data from NCBI-SRA was checked via FastQC. **a)** Summary of quality result, **b)** good Per base sequence quality and **c)** good per sequence quality score.

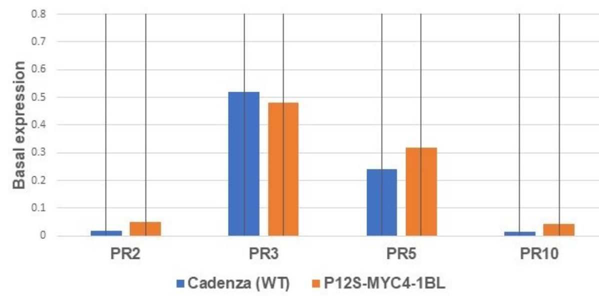


Figure 2.S3: Basal expressions of PR genes responding to stem rust (race TPMKC) in Wildtype and mutant Cadenza. Basal expressions of the PR genes were determined at 0 dpi in Cadenza wildtype and mutant. Error bars represent standard deviation among three biological reps.

Figure 2.S4: Agronomic traits of P12S. Height, tiller and seed morphology of Wildtype and P12S mutant Cadenza plants.

Table 2.S1: Summary of primers

Primer name	Purpose	Tm (°C)	Sequence 5' – 3'
VIGS-F/R	Gene knockdown		F: GACGCTCCATTCCCAAATCG R: TCATGATGTCCTTCACCACG
BN4RT	qRT-PCR	56	F: TCCGCACCATGGTCTGCAT R: TCCGCCGTTGAGGTTGAAGAG
BN1BL	PCR	55	F: ATGATCCGCGTGCAGTG R: GAGTAATAGGATCGCGGTTTCC
ACTIN	qRT-PCR	57	F: CCAGCAATGTATGTCGCAATCC R: CCAGCAAGGTCCAAACGAAGG
PR2	qRT-PCR	57	F: CTCGACATCGGTAACGACCAG R: GCGGCGATGTACTTGATGTTC
PR3	qRT-PCR	57	F: AGAGATAAGCAAGGCCACGTC R: GGTTGCTCACCAGGTCCTTC
PR5	qRT-PCR	57	F: ACAGCTACGCCAAGGACGAC R: CGCGTCCTAATCTAAGGGCAG
PR10	qRT-PCR	57	F: TTAAACCAGCACGAGAAACAT R: ATCCTCCCTCGATTATTCTCACG
Allele1	KASP assay	70	GAAGTTGACCAAGTTCATGCTTGCAGTG CCACAAGCGCAACCAC**C**
Allele2	KASP assay	70	GAAGGTCGGAGTCAACGGATTTGCAGTG CCACAAGCGCAACCAC**T**
CP*	KASP assay	70	CCCGGGTGGCCATCTTCACCGCCACC

*CP= common primer

APPENDIX B

SUPPORTING INFORMATION FOR CHAPTER THREE

File 3.S1. Amino acid sequences of proteins described in Figures 3.2 and 3.3. The location of each predicted functional domain is provided.

```
NPR1-like proteins on group 3 chromosomes
>G105400_Chr.3AS (400 aa)
MQRHLLDFLDNVEVDNLPLILSVANLCNKSCVKLFERCLEIVVRSNLDMITLEKALPED
VIKQIIDSRITLGLASPEDNGFPNKHVRRILKALDSDDVELVRMLLTEGQTNLDDAFAL
HYAVEHCDSKITTELLDIALADVNLRNPRGYTVLHIAARRRDPKIVVSLLTKGARPSDI
TFDGRKAVQIAKRLTKHGDYFGNTEEGKPSPNDKLCIEILEQAERRDPQLGEASVSLAL
AGDCLRGKLLYLENRVALARIMFPIEARVAMDIAQVDGTLEFTLGPSTNPPLEITTVDL
NDTSFKMKEEHLARMRALSKTVEVGKRFFPRCSNKLDKIMDDEPELASLGRDASSERRR
RFHDLQDALLKAFSEDKEEFNKTTTLSSSSSSTSTVARNLTGRPRR

 46-90:DUF3420 unknown domain
 89-192: Ankyrin repeats (3 copies)
 189-387: NPR1-like-C terminal
 -----------------------------------------------------------
>G123800_Chr.3B (633 aa)
MEAPSSHVTASFSDCDDSVSMEDAAPDADVEALRRLSDNLAAAFRSPDDFAFLADALVA
VPGAPDLRVHRCVLSARSPFLRALFKRRAAAAGSSGGVEGNRLELRELLGDEVEVRYEA
LELVLDYLYSGRVRDLPKSACACVDEGGCAHVGCHPAVSFMAQVLFAASTFQVGELANL
FQESMYSTFPPDIIGKHVHHLSPDNSQHRESATFSIVLLFHVFEKRVLTHVMPFFSMQR
HLLDFLDKVEVDNLPLILSVANLCNKSCVKLFERCMEMVVRSNLDMITLEKALPQDVIK
QITDLRITLGLASPEDNGFPNKHVRRILRALDSDDVELVRMLLTEGQTNLDDAFALHYA
VEHCDSKITTELLDIALADVNLRNPRGYTVLHIAAKRRDPKIVVSLLTKGARPSDFTFD
GRKAVQISKRLTKHGDYFGNTEEGKPSPNDKLCIEILEQAERRDPQLGEASLSLALAGD
CLRGKLLYLENRVALARIMFPIEARVAMDIAQVDGTLEFTLGSSTNPPLEITTVDLNDT
SFKMKEEHLARMRALSKTVELGKRFFPRCSNVLDKIMDDEPELASLGRDASSERKRRFH
DLQDTLLKAFSEDKEEFNRTTTLSSSSSSTSTVARNLAGRTRR

45-185: BTB domain
279-323: DUF3420 unknown domain
322-411: Ankyrin repeat
422-620: NPR1-like-C
-----------------------------------------------------------
>G107500_Chr.3D-1(704 aa)
MEAPSSHVTASFSDCDDSVSMGDAAPDADVEALRRLSDNLAAAFRSPDDFAFLADALVA
VPGAPDLRVHRCVLSARSPFLRALFKRRAAAAAATGSSGGAEGNRVELRELLGEEVEVG
YEALELVLDYLYSGRVRDLPKSACACVDVDGCAHVGCHPAVSFMAQVLFAASTFQVGEL
ASLFQVRLPPPLLHHDAFLVSDIYGACWEYIFARTFLCSDSKESMYSTFPPDTIGKHVH
HLSSDNSQHRESATILYRYLNLCISDAEAGDELRVYHLMITPVGLLLVTVFCRCIDKPV
LTHGMPFFSMQRHLLDFLDKVEVDNLPLVLSVANLCNKSCVKLFERCLEMVVQSDLDMI
TLEKALPQDVIKQITDSRITLGLASPEDNGFPNKHVRRILRALDSDDVELVRMLLTEGQ
TNLDDAFALHYAVEHCDSKITTELLDIALADVNLRNPRGYTVLHIAAKRRDPKIVVSLL
TKGARPSDLTFDGRKAVQISKRLTKHGDYFGNTEEGKPSPNDKLCIEILEQAERRDPQF
```

GEASVSLALAGDCLRGMLLYLENRVALARIMFPIEARVAMDIAQVDGTLEFTLGYGTNP
PLEITTVDLNDTSFKMKEEHLARMRALSKTVELGKRFFPRCSNVLDKIMDDEPELASLG
TDACSERKRRFHDLQDTLLKAFSEDKEEFNRTTTLSSSSSSTSTVARNLAGRTRR

45-185:BTB domain
350-394:DUF3420 unknown domain
395-481:Ankyrin repeat
691-727:NPR1-like C

>G117002LC_Chr.3D-2(902aa)
MVATRPLAERDAMFVAIDAAAEELDNETEDVISSTSAIPPPPFAPLRHDDEDTASMYT
AVYNEEKYNIGGTPPLEFQEGHHQRLAGPHSRYAGMIWMAGSRIIVCGDYHGSSQLDW
HTCYKIIEGICNGLHYLHEQIDKSVLHLDLKPANVLLDNNLAPKLTDFGLSKLLDQYQ
TMFTPNRFGTLGYMAPEYIDEGTITPKTDIFSFGVIIMELITGHRDYPSVAGTSSDGF
IELALKKWRNTLEKAQGYTKIEVDCQQIKRCIQIGLICVNPECSKRPTTAKIIKMLKG
LEISDCRISNEATSSAHQTSNGVSEQGKPHPKNTLCIEISEEADRQDLALVSLAIAGD
CLCGNLKLLYLKNGEHISPQASIVRIVVPGSPAGTELHVDPYMLSMRSPYLRAIFARH
NNHYGASGGSAEGNRMDLQEFLGKEVKVEYDALLLVLLYLYSGRVGDLPESAYICADE
NGCAHLGCHPTVSFKVQVLFVAFTFQVPELTSLFQVCISLSLLPMRDLFDVLHKVEVD
NLPLILSVANLCKKSCRKLLERCLEMVVQSNLDMITLEKVLPPDVVKQITDLRLSFGL
ASSEDMGFPNKHVRRILRALDSNDVELVRILIKKGQTNLDDAFALHYAVEHCDSKITT
ELLDIALADVSLRNPRGYTVLHIAARRKDPKIIVSLLTKGARPSDFTLDGRKAVQISK
RLTKHGDYFGSTEAGKPSPKDKLCIKILEQAERRDPQLGEASVSLALAGDCLLGKLLY
LENRVALARIMFPIEARVAMDIAQVDGTLEFTLGSSTNPPLEITTVDLNDTSFKMKEE
HLARMRALSKTDKLNSANVSSRAVQLYWTSSWMMNLSWLHSEEMHLPRFHDLQDTLLK
VFSEDKEEFARTTALSSSSSSTSLSKEFDRSA
1-290: Protein Kinase domain
370-451: BTB domain
620-695: Ankyrin repeat
654-679: Ankyrin repeat
696-891: NPR1-like_C

>G158300LC_Chr.3B 96831605-96839117 (932 aa)
MPAMASSSKALWGSLSTEQKLACRVALQDADELAPPPLPVDITLGQARAHYTDMRGKS
RRRHSGRRHRTKGTTYGSSTSTDSRRKKSQSLRDPPRVMVATRPLAERDAMFVAIDAA
AEELDNETEDVISSTSAIPPPPFAPLRHDDEDTASMYTAVYNEEKYNIGGTPPLEFQE
GHHQRLAGPHSRYAGMIWMAGSRIIVCGDYHGSSQLDWHTCYKIIEGICNGLHYLHEQ
IDKSVLHLDLKPANVLLDNNLAPKLTDFGLSKLLDQYQTMFTPNRFGTLGYMAPEYID
EGTITPKTDIFSFGVIIMELITGHRDYPSVAGTSSDGFIELALKKWRNTLEKAQGYTK
IEVDCQQIKRCIQIGLICVNPECSKRPTTAKIIKMLKGLEISDCRISNEATSSAHQTS
NGVSEQGKPHPKNTLCIEISEEADRQDLALVSLAIAGDCLCGNLKLLYLKNGEHISPQ
ASIVRIVVPGSPAGTELHVDPYMLSMRSPYLRAIFARHNNHYGASGGSAEGNRMDLQE
FLGKEVKVEYDALLLVLLYLYSGRVGDLPESAYICADENGCAHLGCHPTVSFKVQVLF
VAFTFQVPELTSLFQVCISLSLLPMRDLFDVLHKVEVDNLPLILSVANLCKKSCRKLL
ERCLEMVVQSNLDMITLEKVLPPDVVKQITDLRLSFGLASSEDMGFPNKHVRRILRAL

DSNDVELVRILIKKGQTNLDDAFALHYAVEHCDSKITTELLDIALADVSLRNPRGYTV
LHIAARRKDPKIIVSLLTKGARPSDFTLDGRKAVQISKRLTKHGDYFGSTEAGKPSPK
DKLCIKILEQAERRDPQLGEASVSLALAGDCLLGKLLYLENRVALARIMFPIEARVAM
DIAQVDGTLEFTLGSSTNPPLEITTVDLNDTSFKMKEEHLARMRALSKTDKVEILDFI
PFSC

197-384: Protein Kinase
649-692:Unknown function domain DUF3420
693-781:Ankyrin repeat
792-924: NPR1-like C termina

>Chr.3A 69307893-69310005 (the gene has not been predicted by IWGSC) (125 aa)(The DNA sequence corresponding to the Kinase region of the 3B and 3D Kinase fused NPR1 gene was used to as a query to search the IWGSC, and no homologs found on 3A)
MRGLRLIFGWAGVSEFYVADVRTPAKPPRLLLFCGTSDIRAHLRTNGYRIGWHNASGQ
LV
ASARILFPIEARIAMNIAQVDGALEFTLGSSTNPPPAITTVDLNYTPFKLKDVHLARM
RA
LSQNI

1-132: NPR1-like C termina

NPR1-like proteins on group 7 and 4A chromosomes
>G023600LC+G023700LC+G021800_Chr.7A (1,437 aa)
MGRSPSSTGVAGVKKGPWTVEEDILLINYFQRHGNGGNWRTLPQCAGLNRSGKSCRLR
WTNYLRPNIKRGTFTDDEEKTIIHLHSIHGNKWSAIATHLPGRTGNFIKNYWNTNLRK
KLLQMGINPVTHRHTDLSMLKGLPGLHAAAPGNSLSSGVSMGTRAQPHTNAASFAGSS
SDMNTLGIKDKAASFAGSSSSDMTTLGVKDEAASFTGSHSDMKPPRIKDKAASFTGNY
SKMKPFRIKDEALSFTGSYSDMNSLRIKDEATSFTGSYSDIKSVSLQDETTSFTGSYS
ATGTGWDMNALSGLQADPDKFQLLQDHCWVISASLGAMRALVEKLDMLLLAYSSPQEC
SSKRVKDGMHLLKDDLEEISSYLDELLEVEDPPPMAMCWMNEARELSYDMEDYIDSLL
LCVPPDHFNKNKKRKKNKKKIIKKRLKWHRQITYIAQVSEHGVRTSKRIHVTVVPPLP
KKSKIAETISEFRIYVQEAIERHDRYLIVIDDLWDASAWDIIKCAFPKGSHGSRIIIT
TQIEDVALTCCCDHWEHVFEMKPLDDDHSRKLFFNRLFGSESDCPEEFKQVSNEIVDI
CGGLPLATINIASHLANQQTGLALDLLAYMCDSLRSHSWSSSTLERTRKVLNLSYNNL
PQHLKTCLLYFHMYPEGSIIWKDDLVKQWVAEGFVATGKGKEQDKETTEKTAGIYFDA
LVDRRFIQPLYINYNNKVRRLSLIFGDAKYAKTPENITRSQVRSLRFFGLFKCMPCVR
DFKVLRVLNLQLSGHCGAHDPIDLTGISELFQLRYLKITSDICIKLPNSIRGLQCLKT
LDLMDATRVTAVPWDIIHLPHLLHLTLPVDTNLLDWIGSMTDSVISLWSLGKLNYLQD
LHLNMSCIPSLLGVEALAHLIGGHGNLKTIVMSHGSSVIPGASKAIISWDDLEPLPLL
QRFECSPHSCITFSRIPKWIKELSHLCILKIAVVELQISWVDILRGLPALTALSLYVR
RAPIERIIFNKADGFSVLKYFKLTCTSGIACLKFEAQAMPNLWKLKLGFNATPRMDQH
QLIRIEHMPNLKGISVKFGGVAAHIEYARSVVTSHPRNPTINMQLVNFCSNCDGSTKQ

NLLTNWAVKSRAAASAPMEDPPSHFTTTFSSSDGIFMADMNLEALCRLSDNLAAVFRS
PDVFAFLVDMHIVVPDAPDLPVLFAASTFQIAKLTNLFQEKGKWRLLDVLDDVEVDNL
PLILSLVANLCNKSCMKLLERCLEMVVRSNLDMITLEKALPPDIIKQIADSRLSLGLV
SAEDKGFPNKHVRRILQALDSNDAEVATILHQEGQTNLDDAFALHYAVEHCDSKIITE
ILDITLADVNHRNPRGSLMKILFPIEATVAMEIPQVNVQLGKRFFPCCSNVLKKIMDD
ESELDSIGRDTSTERKRRFHELQDLLQKAFSEDKESSSLTSLRFI

15-63:HTH myb-type DNA binding motif
69-114:HTH myb-type DNA binding motif
525-814:NB-ARC domain
1085-1154:BTB domain
1248-1292: DUF3420 unknown domain
1315-1390:Ankyrin repeats

>G703400LC+G470500_Chr.4A (1,314 aa)
MGRSPSCGSEAPVKKGPWTEEEDRLLVDYIKLHGTGGNWRTIPKRAGLNRCGKSCRLR
WTNYLHPDIKRGPFTDDEEKTTIHLHSMLGNKYVYCSLLLLLYTIYLSTIFFSNSGAS
YSVAADDLSLKQVVGDCNPSAGKDRQFYQELLEHEPAQEATPHGHRPHHAQPARRPQP
AYRAPQSPRHRRRPQMGRSPSCGSEAPVKKGPWTEEEDRLLVDYIKLHGTGGNWRTIP
KRAGLNRCGKSCRLRWTNYLHPDIKRGPFTDDEEKTTIHLHSMLGNKYVYCSLLLLLY
TIYLSTIFFSNSGASYSVAADDLSLKQVVGDCNPSAGKDRQFYQELLEHEPAQEATPH
GHRPHHAQPARRPQPAYRAPQSPRHRRRPQMKPLDEDHSRKLFFNRLFGSGSDCPEEF
KQVSNEIVDVCGGLPLATINIATHLANQQKAVSLDFLTYIRDSLRSQSWSSSASERTR
QVLNLSYNDLPHHLKTCLLYLHMYPEGSIVWKDDLVKQWVAEAFINTRKGKGKDQNWM
EKAAGIYFDELVDRRFIQPLNINYNNKVLSCTVHEVVRDLLAHKSVVVDYNRKNISLS
HKARRLSLLFANASYAKTPVNITKAQVRSLNFVGLFECMPCIGEFKVLRVLNLQLSGH
CGDHDPIDLTGISELFQLRYLKITSDVCIKLPNQMRKLQCLETLDIIDAPRVTAIPWD
IIYLPHLLHLTLPVDTNLLDWIGSMTGSIISLWSLGNLNYLQDLYLTISSTHPSGHPE
KNMEALGSLLGGHGNLKIIVVSHGPLVMRAYRSMLSVQSPFLDAVFARRAAEGEDDPL
DLRELLGEEVEVGYEALQLVLEYLYTGRIRDLPKSACVCADVDGCAHVGCLPAISFMA
QVIFAASIFEVAVLTNHFQVRLFLLLSVAMISSYHLFIMPAGMHIGKRKGNCSKHGND
VLPVLFNKNLFEAVWRGADMRLLLDVLDDVEVDNLPLILSVANLCNKSCMHLLERCLE
MVVRSNLDMITLEKALPPDVIKQITDSRISLGLISPKDNGFPNKHVRRIFGALDSYDV
ELVRLLLYEGQTNLDDAFALHYAVEHCDPKITTELLDLELADVNRRNQRGYTVLHIAA
RRRNPKILVSVLTKGARPSDLTFDGRKAIQISKRLTKHGDYFGITEEGKPSPEYSLCI
EILEQAERRGPQLGEASVSLAIAGDCQRGTLLHLENRVTLMRIMFPTEARVAMDIAQV
DCTLELTLDSEAKPPSEKEVATIDLNETRFQMNDEHLARMSALFKTDKIMNDEPELAS
LQRDTSTERNRRFGDMHDTLQKAFSEDKGSSSFDKSGS

15-63:HTH myb-type DNA binding motif
204-252: HTH myb-type DNA binding motif
369-514:NB-ARC
768-841:BTB domain
992-1036:DUF3420 unknown function domain
1035-1123:Ankyrin repeats

1135-1314:NPR1-like-C

>G023000LC+G019000_Chr.7D (1,351 aa)
MGRSPSCSSEPPLKKGPWTEEEDRLLVDYIKLHGTGGNWRTIPKRAGLNRCGKSCRLR
WTNYLHPDIKRGPFTDEEEKTTIHLHSMLGNKWSAIATHLPGRTGNFIKNYWNTNLRK
KLLHMGIDPITHNPRTDLSQLTGLPSLPATAAAPMREFLACRYLIIIDDLWDASAWDI
IKYAFPNGNRGSRIIITTQIEDVALTCCCDHSEHVFMMKPLDDDHSRKLFFNRLFGSG
SDCPEEFKQVSNEIVDVCGGLPLAMINIASHLANQQTAVSLDFLTYIRDSLRSQSWSS
STSERTRQVLNLSYNNLPHHLKTCLLYLHMYPEGSIVWKDDLVKQWVAEAFINTRKGK
EKDQNWMEKAAGIYFDELIDRRFIQPLNINYNNKVLSCTVHEVVRDLLAHKSAEENFI
VVVDYNRKNISLSHKARRLSLLFATARYAKTPVNITKPQVRSLNFVGLFECMPCIGEF
KVLRVLNLQLSGHCGDHDPIDLTGISELFQLRYLKITSDVCIKLPNQMRKLQCLETLD
IIDAPRVTAIPWDIIYLPHLLHLTLPVDTNLLDWIGSMTDSIISLWSLGNLNYLQDLY
LTISSTHPSGHPEKNMEALGSLLGGHGNLKIIVVSHGPSVKDIVVPGASKVIISWDEL
EPLPLLQRFECSPHSCVIFSRIPKWVKKLGNLCILKIAVVELQMSCVDILRGLPALTA
LSLYVRCAPAQRILFDKMAGFSVLKYFKLRFTSGIPWLKFEADAMPNLWKLKLGFNAI
PRMDQHQLIRIEHMPDLKEISVKFGGIAALIEYAVKTVISNHLRNPRVHVCLVTSTSY
GDESTKEKPPTNSAVEMRAYRSVLSVRSPFLDAIFARRAAEGEGNPLDLRELLGEEVE
VGYEALQLVLEYLYTGCIRDLPKSACVCADVDGCAHVGCLPAISFMAQVIFAASIFQV
AVLTNHFQVRLFLLFSVAMISSYHLFIMPAGMHIGILVLTFHDDLCVYHLRIILVESL
LRLLLDVLDDVEVDNLPLILSVANLCNKSCMHLLERCLEMVVRSNLDMITLEKALPSD
IIEQITDSRLSLGLVSPEDKGFPNKHVRRILRALDSDDVCLVRMLLKEGRTNLDGAFA
LHHAVEHCDSKVTMELLDIGLADVNHRNPRGYTVLHVAARRRDPKILVSVLTKGARHS
DLTFDGRKAVQISKRLTKHGDYFGITEEGKPSREDRLCIKILEQAERRDPQLGEASVS
LAIAGDRQRGKLLYLENRVALMRIMFPTEARIAMDIAQVDCTLKLTLDSGAKPPPEKE
LATIDLNETPFHMNEEHLARMSTLSKTDKIMDDEPELAPLQRDASTERTRRFGYLQDA
VQKALSEDKASFDKSGS

15-63: HTH Myb-like DNA binding domain
69-114: HTH Myb-like DNA binding domain
146-347:NB-ARC domain
828-886:BTB domain
1031-1075:DUF3420 unknown domain
1098-1173:Ankyrin repeats
1174-1351:NPR1-like domain

>G021700_Chr.7A (1,107 aa)
MEAAAATAFVGRIAPKLLEFLAANHKLRQNLEHDITYIQREFALISAAIQQDDDCRWR
SGDHVKRAWIQIIRDLAHAIEDCIDRFMHRVTISGASTWIRQAVHRVQTVTVRKEFAK
AIRELKKISQESSKLRETYYSANIGAGTSSSSVASSVMACETATQMVIDDTLSAGQPV
GMDAPWEELLELIQQQQQQLKVISIVGFDGIGKTLLARCVYDTIENQYEARAWVSAAE
QGVPTNVIKEILQQFAIPTNGGGNLSKLCAILRLYLGTKRFFIVIDDMRTEFWHDIKD
AFVGLSGRVLVTTAIHSVANACSSSAAHDHVYAMKTLADEHSRLLFFKEAFQDDNPPI
DKEDQLGSEALKKCDGLPLALVTTARYLQSTGNPTHGNWATLCHNLGAHLETKEMLAR
MKRVLVHSYTSLVKHDVKTCLLYLGIYRSGRTVRRGSLIRKWCAEGFIQGDYMCNALD

AAKANFKELLNRSIIKHTDASSKNNKDQVKTYHTHGMMLEFILHMSKCDNFITLLYDQ
MAPPPPPSKIRWLSLHDASARVVNDLSLVRSLTVFGKAHDSVLDFSKYELLRVLDLEE
CSNHLEDKHLREICSNLLLLRYLSLGAAHKVAVLPKEIKKLQLLETLDVRKTKIEVLP
TQVMELPCLIHLFGKFKLQQGVGGRKMHKLQIWFSENSKLETVAGFVVDSNKSQGFAQ
FMEHMKHLTKVKIWCEQSSNNSRDPTASGSSSNTNNYTHVSKAIKGFIKRSTDVKKAH
SLSLHCNDKWFQDLLVNLSLEKEEASSCYLSSLKLQGGNICSLPPFVTMLGGLTKLCL
SSPHHQLGGDILVALSRVRCLAYLKMIASQLDNLVIVEGALGNLRQLCIVVEVMTELE
VQEGALPLLESLQLLCKDLNGFCSKMIQSLRRIKEVTLHDGVNGETKQKWKEAAKKHP
RRPNLLFVKTAEDADMGSEPADNSESPVAQTTATTVSVITTQDAISTGQSVQVDGADL
QQGDEKEHAYKIDILEDFASKTCLDPPMNTESFEQGMEGMVGLEDERMEDVTHSTDQA
DQNVVLSVVGENRRKRARLDIGEDNSMDKVVDRVKRKKPEDVQEPRLKARPRGRQVAV
GAEDL
6-119:Rx-CC-like domain
176-458:NB-ARC domain

>G018900_Chr.7D (1,103 aa)
MEAAAATAFACRIAPKLLAFLATNHKLRQNLEHDITYIRNEFALISAAIQQDDDHRWR
SRPSRDHMQRAWVKIIRDLAHAVEDCIDRFMHRVTISGTSTWLRQAVHRVQTVTVRKE
FAKAIREIKKISQESSKLRETYYSASIAAGTSSCSAASSSVASETTTQMAIDDTLSAG
RSVGMDAPQDELLELIQQQQQLKVISIVGFHGIGKTLLARRVYHRMENQYEARAWVS
AAKQGVPTNVLKQILQQLGIPTNGGGNFNKLCAVLRLYLGSKRFFIVIDDMQTEFWHD
IKDIFVGLSGRVLVTTAIQSVANACSSSAAHDHVYAMKTLADEHSRLLFFKEAFQDDN
PPVNKEDQLGSEALKKCDGLPLALVTTARYLQSTGNPTRENWATLCHNLGVHLETKEM
LARMKCALVHSYTSLVKHDVKTCLLYLGIYHTGRTVRRGNLIRKWCAEGFIQGDYMCN
ALDAAVANFRELVNRSIIQRTDASSKNIKDQVKTYRTHGMMLEFILHMSKCDNFITLL
YDQLAPPPPPSKVRWLSLHDASARVANDLSLVRSLTVFGKAHESVLYFSKYELLRVLD
LEECGNHLEDKHLREICNNLLLLRYLSLGAALTVTVLPKEIKKLQLLETLDVRRTRIE
ILPTQVMELPCLIHLFGKFKLQQGVGGRKMLKLQTWCSENSKLETVAGFVVDNNKSQG
FAQLMEHMKHLTKVKIWCQQSTNNSMDPISSGSSSSSKYTHLSEAIKGFIKRSTDKKA
HSLSLSFNDRWCQDLLVNLSLEKEQASSCYLSSLKVQGGNICSLPPFVTMLGGLTSLC
LSSPHHQLSGDILAALSRVRCLAYLRLIASQLGKLVIVRGALGSLRRLCIVVEVMTEL
EVQEGALPLLESLQLLCKDLNGFCSIMIQSLGRIKEVTLHDGVNDETRQKWKEAAKNH
PRRPKLLFVKTAEDVDMGSEPADNSESPTVPTNDTTLPVTAPHDAISTGQSVQVDGDD
LQQDDDEKEDTDKIDMLVDFVSKTCLGTMMNKESFEQGMEGMVVGENRRKRARTVGEG
NSMDKVFDRVKRKNPEDIPVKLGSRLEVIFWSWALKFTGDVRNLVAVGDVALIWTIWR
C

6-119:Rx-CC-like domain
179-461: NB-ARC domain

>G470600_Chr.4A (1,057 aa)
MEAAAATAFACRIAPKLLAFLAANHKLRQNLEHDITYIRNEFALISAAIQQDDDHRWR
SGPSRDYMQRAWVKIIRDLAHAVEDCIDRFMHRVTISGTSTWLRQAVHRVQTVTVRKE
FAKTIRELKKISQESSKLRETYYSASIAAGTSSYSAASSSVASETMTQMAKDDTLSAG
RSVGMDAPQDELLELIQQQQQLKVISIVGFHGIGKTLLARRVYHTIENQYQARAWVS

AAEQGFPTNVLKQILQQLGIPSNGGGNFNKLCTVLRLYLRSKRFFIVIDDMQTEFWHD
IKDIFVGLSGRILVTTAIQSVANACSSSAAHDHVYAMKTLADEHSRLLFFKEAFQDDN
PPVNKEDQLGSEALKKCDGLPLALVTTARYLQSTGIYPTGRTVRRGNLIRKWCAEGFI
QGDYMCNALDTAVANFKELVNRSIIQRTDASSKNIKDQVKTYHTHGMMLEFILHMSKC
DNFITLLYDQLAPPPPPSKIRWLSLHDASARVANDLSLVRSLTVFGKAHESVLYFSKY
ELLRVLDLEECGNHLEDKHLREICNNLLLLRYLSLGATLTVTVLPKEIKKLQLLETLD
VRRTRIEILPTQVMELPCLIHLFGKFKLQQGVGGRKMHKLQTWSSKNSKLETVAGFVV
DNNKSQGFAQLMEHMKHLTKVKIWCQQSTNNSMDPISSGSSSSSKYTHLSEAIKGFIK
RSTDKKAHSLSLNFNDRWCQDLLVNLSLEKEEASSCYLSSLKVQGGNICSLPPFVTML
GGLTSLCLSSPHHQLSGDILAALSRVRCLAYLKLIASQLGKLVIVKGALGSLRRLCIV
VEVMTELEVQEGALPLLESLQLLCKDLNGFCSMMIQSLGRIKEVTLHDGVNDETKQKW
KEAAKKHPRYPKLLFVKTTEDVDMGSEPADNSESPTVPTNGRTLPMTAPHDAISTGQS
VQVDGDDLQQDDDEKEDTDKIDMLADFASKTCLGTMMNKESFEQGMEGMVGLDDQQME
DVTPSTDQADQSLVLLVVGENRRKRARTVGEGNSMDKVFDRVKRKNLEDVPETRPTKQ
TAPGLAVAISDGN

6-119:Rx-CC-like domain
179-383:NB-ARC domain

File 3.S2. The alignment of G3NPR1-like homologs with clustalw2

```
G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       ------------------------------------------------------------
W3NPR           ------------------------------------------------------------
G3DNPR1         ------------------------------------------------------------
G3BNPR1-1       ATGGTGGCTACGCGTCCTTTGGCGGAGAGAGACGCCATGTTCGTAGCCATCGATGCAGCG 60

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       ------------------------------------------------------------
W3NPR           ------------------------------------------------------------
G3DNPR1         ------------------------------------------------------------
G3BNPR1-1       GCTGAGGAATTGGATAACGAGACCGAGGACGTCATCAGCTCCACGTCGGCTATTCCGCCA 120

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       ------------------------------------------------CACAC------- 5
W3NPR           ------------------------------------------------------------
G3DNPR1         ----------------------------------------------------C------- 1
G3BNPR1-1       CCACCTTTTGCGCCCCTCCGCCATGACGACGAGGACACCGCGTCCATGTACACCGCCGTC 180

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       -----------AAGGCATCCCTCTGGG---------CAGTTTTCCTTCAATGAATGGTAT 45
W3NPR           -------------------------------------------------ATGGA------ 5
G3DNPR1         -----------AGG--------------------------------TCGGCGAG------ 12
G3BNPR1-1       TACAACGAGGAAAAGTACAACATTGGGGGGACGCCGCCGCTTGAGTTCCAGGAAGGCCAC 240

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       CCTAGTCAACTTCTCGGATGCCTCA-TGGGCCTGAAAACATGTCAATCAAGTTCAGTACT 104
W3NPR           ---------------------------GGCCCCGAG-----------------CAG---- 17
G3DNPR1         ------------------------------CTCG------------------CCAGCCTC 24
G3BNPR1-1       CACCAACGTCTTGCCGGGCCTCACAGCCGGTATGCGGGCATGATTTGGATGGCCGGCTCC 300

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       TGTTTTGGCAAATTTTG-----TCCATG-------------------------------- 127
W3NPR           -----------------------CCACG-------------------------------- 22
G3DNPR1         T---------------------TCCAGG-------------------------------- 31
G3BNPR1-1       CGTATCATTGTCTGTGGAGACTACCATGGATCTTCCCAACTTGATTGGCACACGTGCTAC 360

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       ----------------TTTCGAAGGGTTACAACTGCCATCGCTGG--------------T 157
W3NPR           ----------------TC-------------ACTGCC----------------------- 30
G3DNPR1         ----------------T---------------CCGCCTCC-------------------- 40
G3BNPR1-1       AAAATAATTGAGGGGATTTGCAATGGTTTGCACTACCTTCATGAGCAAATTGACAAGTCT 420

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       TCCCCTCGCCT--------------------------TTAGGTGCCTTCCGACACCA--- 188
W3NPR           -TCCTTC----------------------------------------TCCGAC------- 42
G3DNPR1         -CCCCTCCCCT--------------------------TC--------TCCACC------- 58
G3BNPR1-1       GTTCTTCACTTGGACCTTAAGCCTGCGAACGTATTGCTTGATAACAATTTGGCACCAAAA 480

G3ANPR1         ------------------------------------------------------------
```

```
G3BNPR1-2       ----CCGAGGTGGGGGAATCCCAGTT-TGTGAT-AGCGTC--------TTCTTTCAAGTT 234
W3NPR           ----------TGCGA--------------CGAC-AGCGTC--------TCC--------- 60
G3DNPR1         -------------------------------AT-GATGCT--------TTC--------- 69
G3BNPR1-1       CTTACAGATTTTGGGCTGTCAAAACTGCTTGATCAATATCAAACTATGTTCACTCCAAAT 540

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       TGT-----------GAACAACGAAATACCGG---------------------TGGCTCCT 262
W3NPR           -AT-----------GGGGGACGCGGCGCCGG---------------------ACGC---- 83
G3DNPR1         ---------------------------TTGG---------------------TTTC---- 77
G3BNPR1-1       CGTTTTGGTACATTAGGCTACATGGCACCAGAATACATCGATGAAGGTACAATCACACCT 600

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       -----GGCCGTAC----------------------------------------------- 270
W3NPR           -----GGACGT------------------------------------------------- 89
G3DNPR1         -----AGACAT------------------------------------------------- 83
G3BNPR1-1       AAGACAGACATATTTAGTTTCGGTGTGATAATCATGGAGCTAATAACGGGACACAGGGAC 660

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       -------------------------CCGACGGC--------------------------G 279
W3NPR           ---------------------------GGAGGC--------------------------G 96
G3DNPR1         --------------------------TTATGGT--------------------------G 91
G3BNPR1-1       TACCCAAGTGTTGCTGGAACATCTTCCGATGGCTTTATTGAGCTTGCACTTAAAAAATGG 720

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       ACAAATACGCCGGTCAACGGCCTGA------------------TCGATTTGCGTCGGGCC 321
W3NPR           C----TCCGCCG--------CCTCT------------------CCGAC------------ 114
G3DNPR1         ------------------------------------------------------------
G3BNPR1-1       AGAAACACACTGGAGAAAGCACAGGGGTACACAAAGATAGAAGTAGATTGCCAACAAATA 780

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       AAATTGCAAGATTGCCATGAATCCCACCCTCGTTTAATTTATTTGTTTAAAA---TCTAC 378
W3NPR           -------------------AACCTCGCCGCCG---------------------------- 127
G3DNPR1         ------------------------------------------------------------
G3BNPR1-1       AAACGATGC-ATTCAAATAGGTCTAATCTGCGTTAATCCTGAGTGCTCCAAAAGGCCTAC 839

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       AAAAATGGTGTGGCTCTTGGCTTTGATATTAAACCTCCCTTTGGCGAGCTGCG------- 431
W3NPR           -------------------------------------CCTTCCGCTCGCCGGA------- 143
G3DNPR1         --------------------------------------CTT------GCTGGG------- 100
G3BNPR1-1       AACAGCAAAAATTATCAAGATGCTTAAAGGACTAGAAATTTCAGATTGTAGGATTAGCAA 899

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       --------CGTCTCGCTCACTTC-GACTTCC------------------------TCGCG 458
W3NPR           --------CG--------ACTTC-GCCTTCC------------------------TCGC- 161
G3DNPR1         --------AAT-------ACAT---ATTTGC------------------------GCGC- 117
G3BNPR1-1       TGAGGCAACGTCATCAGCACATCAGACTTCAAATGGGGTTAGTGAACAAGGAAAACCACA 959

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       CC-------TTCGCT-----------------------CCCACCAACTGGCC---GCTCT 485
W3NPR           -C-------GACGCG-----------------------CTCGTCGCCGTGCC---GGGC- 186
G3DNPR1         ----------ACGTT-----------------------CTTGT----------------- 127
```

```
G3BNPR1-1       TCCAAAAAATACACTATGTATTGAGATTTCGGAGGAAGCTGATAGACAGGATCTAGCATT 1019

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       CCCGACTCCCG-AGTCGCACGCGGCCGCC-------------GCAACAGCAGCGCC---- 527
W3NPR           ----GCGCCCG-ACCTGCGCGTG--CACC-------------GCTGC----GTGCT---- 218
G3DNPR1         ------GCTCA-GACAGCAAG--------------------------------------- 141
G3BNPR1-1       GGTCTCTCTTGCAATTGCTGGTGACTGTCTATGTGGAAACTTGAAGTTGTTGTACCTGAA 1079

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       -AATGGAGG----------CCCCGAGCAGCCA-CGTCACCGCCTCCTTCTCCGACTGCGA 575
W3NPR           -------GT----------CCGCGCGGAGCC-------------CCTTC-----CTGC-- 241
G3DNPR1         -----GAAA----------GCATGTACAGCA-------------CCTTT----------- 162
G3BNPR1-1       AAACGGAGAACACATTTCTCCACAGGCATCCATCGTCCGCATCGTTGTCCCAGGCTCTCC 1139

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       CGACAGCGT----CTCCATGGAGG----------------ACGCGGCGCCC-------GA 608
W3NPR           -----GCGC----CCTCTTCAAG-----------------CGCCGCGCC--------GC 266
G3DNPR1         --------C----CCCCAGACACA----------------ATAGGAAAGC---------A 185
G3BNPR1-1       CGCTGGTACTGAACTTCATGTGGACCCCTACATGCTCTCCATGCGGAGTCCCTACCTTCG 1199

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       CGCG------GACCGCCGC-GC--------CGCCGCCGCCGGGT-CGGTG--GAGGGGAA 650
W3NPR           CGCC------G-CCGCCGCTAC--------CGGTTCGTCCGG---CGGCGCCGAGGGCAA 308
G3DNPR1         TGTG------CACCACC---------------TTTCCTC---------------AGACAA 209
G3BNPR1-1       TGCGATCTTTGCCCGCCACAACAATCATTATGGTGCATCTGGGGGCAGTGCGGAGGGCAA 1259

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       CCGGCTGGAGCTCCGGGAGCTTCTCGGCGACGAGGTCGAGGTCAGGTACGAGGCGCTGGA 710
W3NPR           CCGGGTGGAGCTCCGGGAGCTTCTCGGCGAGGAGGTGGAGGTCGGGTACGAGGCGCTGGA 368
G3DNPR1         TTCAC---AACACCGGGAG----TCTGCCACAATATT-ATATC--------GCTACTTAA 253
G3BNPR1-1       CAGGATGGACCTCCAAGAATTTCTAGGCAAGGAAGTGAAGGTCGAGTATGATGCGCTGTT 1319

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       GCTGGTGCTCGACTACCTGTACAGCGGCCGCGTCCGCGACCTCCCCAAGTCGGCGTGCGC 770
W3NPR           GCTGGTGCTCGACTACCTGTACAGCGGCCGCGTTCGCGACCTCCCCAAGTCGGCGTGCGC 428
G3DNPR1         ATTTATGC---ATCTCCGATGCAGAGGCTG-----GGGAT------GAGTTG-CGTGTAT 298
G3BNPR1-1       ACTGGTACTCTTGTACCTGTACAGTGGCCGTGTCGGAGACCTTCCCGAGTCGGCGTACAT 1379

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       CTGCGTCGACGAGGGCGGCTGCGCGCACGTCGGCT-GCCACCCCGCCGTCTCCTTCATGG 829
W3NPR           CTGCGTCGACGTCGACGGATGCGCCCACGTCGGCT-GCCACCCCGCCGTCTCCTTCATGG 487
G3DNPR1         ATCATTTGATGATAACGCCT--------GTTGGCTTGCTAC-------------TTGTAA 337
G3BNPR1-1       ATGTGCCGACGAGAACGGTTGTGCGCACCTCGGCT-GCCACCCCACCGTCTCCTTCAAGG 1438

G3ANPR1         ------------------------------------------------------------
G3BNPR1-2       CGCAGGTCCTCT-TCGCCGCATCCACCTTCCAGGTCGGCGAGCTCGCCAACCTCTTCCAG 888
W3NPR           CGCAGGTCCTCT-TCGCCGCATCCACCTTCCAGGTCGGCGAGCTCGCCAGCCTCT----- 541
G3DNPR1         CA---GTTTTCTGTCGCTGTATTGAC-----AAGCCAGT--ACTTACC------------ 375
G3BNPR1-1       TGCAGGTTCTCT-TTGTTGCCTTCACCTTCCAGGTCCCCGAGCTGACCAGCCTCTTCCAG 1497

G3ANPR1         ------------------------------------------------------------
```

```
G3BNPR1-2     GAAAGCATGTACAGCACCTTTCCCCCGGATATAATAGGAAAGCATGTGCACCACCTTTCC 948
W3NPR         ------------------------------------------------------------
G3DNPR1       ------------------------------------------------------------
G3BNPR1-1     GT---------------------------------------------------------- 1499

G3ANPR1       ------------------------------------------------------------
G3BNPR1-2     CCAGACAATTCACAACACCGGGAGTCTGCGACATTCTCTATTGTTCTATTGTTTCATGTT 1008
W3NPR         ------------------------------------------------------------
G3DNPR1       ------------------------------------------------------------
G3BNPR1-1     ------------------------------------------------------------

G3ANPR1       ------AAGCCAGCACTAACCCATGGTATGTCCTTCTTTTCCATGCA---GCGGCATCTC 51
G3BNPR1-2     TTTGAGAAGCGAGTACTAACCCATGTTATGCCCTTCTTTTCCATGCA---GCGGCATCTC 1065
W3NPR         -------------------------------------------TGCA---GCGGCATCTC 555
G3DNPR1       ---------------------CATGGTATGCCATTCTTTTCCATGCA---GCGGCATCTC 411
G3BNPR1-1     -----------------------CTGCATCTCCCTTTCCCTTTTGCCAATGCGGGATCTC 1536
                                                         ***    **** *****
G3ANPR1       CTTGATTTCCTTGATAATGTTGAAGTGGATAACCTTCCGTTGATCTTATCTGTTGCAAAC 111
G3BNPR1-2     CTTGATTTCCTTGATAAAGTTGAAGTGGATAACCTTCCGTTGATCTTATCTGTTGCAAAC 1125
W3NPR         CTTGATTTCCTTGATAAAGTTGAAGTGGATAACCTTCCATTGGTCTTATCTGTTGCAAAC 615
G3DNPR1       CTTGATTTCCTTGATAAAGTTGAAGTGGATAACCTTCCATTGGTCTTATCTGTTGCAAAC 471
G3BNPR1-1     TTTGATGTCCTTCATAAGGTTGAAGTGGATAACCTTCCACTGATCTTGTCTGTTGCAAAC 1596
               ***** ***** **** ********************  ** **** ************
G3ANPR1       TTATGCAACAAATCTTGCGTGAAACTGTTTGAGAGATGCCTGGAGATAGTAGTCCGGTCA 171
G3BNPR1-2     TTATGCAACAAATCTTGCGTGAAACTGTTCGAGAGATGCATGGAGATGGTAGTCCGGTCA 1185
W3NPR         TTATGCAACAAATCTTGCGTCAAACTGTTCGAGAGATGCCTGGAGATGGTAGTCCAGTCA 675
G3DNPR1       TTATGCAACAAATCTTGCGTCAAACTGTTCGAGAGATGCCTGGAGATGGTAGTCCAGTCA 531
G3BNPR1-1     TTATGCAAAAAATCTTGCCGGAAACTGCTCGAGAGATGCCTTGAGATGGTAGTCCAGTCA 1656
              ******** *********   ****** * ********* * ***** ******* ****
G3ANPR1       AATCTTGACATGATTACTCTTGAGAAAGCATTGCCTGAAGATGTTATCAAGCAAATTATT 231
G3BNPR1-2     AATCTTGACATGATTACTCTAGAGAAAGCATTGCCTCAAGATGTCATCAAGCAAATTACT 1245
W3NPR         GATCTTGACATGATTACTCTAGAGAAAGCATTGCCTCAAGATGTTATCAAGCAAATTACT 735
G3DNPR1       GATCTTGACATGATTACTCTAGAGAAAGCATTGCCTCAAGATGTTATCAAGCAAATTACT 591
G3BNPR1-1     AATCTTGACATGATTACTCTTGAGAAAGTATTGCCTCCAGATGTTGTTAAGCAAATTACT 1716
               ******************* ******* *******  ******  * ********** *
G3ANPR1       GATTCACGGATAACTCTTGGATTAGCTTCACCCGAAGACAATGGCTTTCCTAACAAACAC 291
G3BNPR1-2     GATTTACGGATAACTCTTGGATTAGCTTCACCCGAAGACAATGGCTTTCCTAACAAACAC 1305
W3NPR         GATTCACGGATAACTCTTGGATTAGCTTCACCCGAAGACAATGGCTTTCCTAACAAACAC 795
G3DNPR1       GATTCACGGATAACTCTTGGATTAGCTTCACCCGAAGACAATGGCTTTCCTAACAAACAC 651
G3BNPR1-1     GACTTACGGCTAAGTTTTGGATTAGCTTCATCCGAGGACATGGGCTTTCCTAACAAACAT 1776
              ** * **** *** * ************** **** ****  *****************
G3ANPR1       GTAAGAAGGATACTCAAGGCACTTGATTCTGATGATGTGGAGCTTGTCAGGATGCTGCTC 351
G3BNPR1-2     GTAAGAAGGATACTCAGAGCACTTGATTCTGATGATGTGGAGCTTGTCAGGATGCTGCTC 1365
W3NPR         GTAAGAAGGATACTCAGAGCACTTGATTCTGATGATGTGGAGCTTGTCAGGATGCTGCTC 855
G3DNPR1       GTAAGAAGGATACTCAGAGCACTTGATTCTGATGATGTGGAGCTTGTCAGGATGCTGCTC 711
G3BNPR1-1     GTAAGAAGGATACTCAGAGCTCTTGATTCCAATGATGTGGAGCTAGTAAGGATTCTGATC 1836
              ****************  ** ********  ************* ** ***** *** **
G3ANPR1       ACAGAAGGGCAGACTAACCTTGACGATGCATTTGCATTGCACTATGCTGTAGAACACTGT 411
G3BNPR1-2     ACAGAAGGGCAGACTAACCTTGATGATGCATTTGCATTGCACTATGCTGTAGAACACTGT 1425
W3NPR         ACAGAAGGGCAGACTAACCTTGATGATGCATTTGCATTGCACTATGCTGTAGAACACTGT 915
G3DNPR1       ACAGAAGGGCAGACAAACCTTGATGATGCATTTGCATTGCACTATGCTGTAGAACACTGT 771
```

```
G3BNPR1-1      AAGAAAGGGCAGACCAATCTTGATGATGCATTTGCGCTCCATTATGCTGTAGAACACTGT 1896
               *   ********** ** ***** ***********   * ** *****************
G3ANPR1        GACTCAAAAATTACAACAGAACTTCTGGACATCGCACTTGCGGATGTTAATCTCAGAAAC 471
G3BNPR1-2      GACTCAAAAATTACAACAGAACTTCTGGACATCGCACTTGCGGATGTTAATCTCAGAAAC 1485
W3NPR          GACTCAAAAATTACAACAGAACTTCTGGACATCGCACTTGCGGATGTTAATCTCAGAAAC 975
G3DNPR1        GACTCAAAAATTACAACAGAACTTCTGGACATCGCACTTGCGGATGTTAATCTCAGAAAC 831
G3BNPR1-1      GACTCAAAAATTACAACAGAACTTCTGGACATCGCACTTGCAGATGTTAGTCTCAGAAAC 1956
               ***************************************** ******* **********
G3ANPR1        CCAAGAGGTTATACTGTTCTTCACATCGCTGCTAGGCGGAGAGATCCTAAAATTGTTGTC 531
G3BNPR1-2      CCAAGAGGTTATACTGTTCTTCACATCGCCGCTAAGCGGAGAGATCCTAAAATCGTTGTC 1545
W3NPR          CCAAGAGGTTATACTGTTCTTCACATCGCCGCTAAGCGGAGAGATCCTAAAATCGTTGTC 1035
G3DNPR1        CCAAGAG----------------------------------------------------- 838
G3BNPR1-1      CCAAGAGGTTACACTGTTCTTCACATTGCCGCCAGACGGAAGGATCCTAAAATCATTGTC 2016
               *******
G3ANPR1        TCCCTTTTAACCAAAGGTGCTCGGCCTTCTGATATTACATTTGATGGAAGAAAAGCAGTT 591
G3BNPR1-2      TCCCTTTTAACCAAAGGTGCCCGGCCTTCAGATTTTACATTTGATGGAAGAAAAGCAGTT 1605
W3NPR          TCCCTTTTAACCAAAGGTGCCCGGCCTTCGGATTTAACATTTGATGGAAGAAAAGCAGTT 1095
G3DNPR1        ------------------------------------------------------------
G3BNPR1-1      TCCCTTTTAACCAAAGGTGCTCGGCCTTCCGATTTTACACTTGATGGAAGAAAAGCAGTT 2076

G3ANPR1        CAAATCGCAAAGAGACTCACAAAACATGGGGATTATTTGGGAATACTGAAGAAGGGAAG 651
G3BNPR1-2      CAAATCTCAAAGAGACTCACAAAACATGGTGATTATTTGGGAATACTGAAGAAGGAAAG 1665
W3NPR          CAAATCTCAAAGAGACTCACAAAACATGGGGATTATTTGGGAATACTGAAGAAGGAAAG 1155
G3DNPR1        ----------------------------------------------------AAGGAAAA 846
G3BNPR1-1      CAAATCTCAAAGAGACTCACAAAACATGGGGATTATTTGGGAGTACTGAAGCAGGAAAG 2136
                                                                    *** **
G3ANPR1        CCGTCTCCCAATGATAAATTATGCATTGAGATACTGGAGCAAGCTGAAAGAAGGGATCCA 711
G3BNPR1-2      CCGTCTCCCAATGATAAATTATGCATTGAGATATTGGAGCAAGCTGAAAGAAGGGATCCA 1725
W3NPR          CCGTCTCCCAATGATAAATTATGCATTGAGATATTGGAGCAAGCTGAAAGAAGGGATCCA 1215
                                           G3oligor
G3DNPR1        CCGTCTCCCAATGATAAATTATGCATTGAGATATTGGAGCAAGCTGAAAGAAGGGATCCA 906
G3BNPR1-1      CCGTCTCCCAAGGATAAATTATGCATTAAGATACTGGAGCAAGCTGAAAGAAGGGATCCA 2196
               *********** *************** ***** **************************
G3ANPR1        CAACTTGGAGAAGCATCGGTTTCTCTTGCATTGGCTGGTGACTGTCTTCGTGGGAAGTTA 771
G3BNPR1-2      CAACTTGGAGAAGCATCACTTTCTCTTGCATTGGCTGGTGACTGTCTTCGTGGAAAGTTA 1785
W3NPR          CAATTTGGAGAAGCATCGGTTTCTCTTGCATTGGCTGGTGACTGTCTTCGTGGGATGTTA 1275
G3DNPR1        CAATTTGGAGAAGCATCGGTTTCTCTTGCATTGGCTGGTGACTGTCTTCGTGGGATGTTA 966
G3BNPR1-1      CAACTTGGAGAAGCGTCGGTTTCTCTTGCATTGGCTGGTGACTGTCTTCTTGGGAAGTTA 2256
               *** ********** **  ****************************** *** * ****
G3ANPR1        CTGTACCTTGAAAACCGAGTTGCTTTGGCAAGGATAATGTTTCCAATTGAGGCAAGAGTA 831
G3BNPR1-2      CTGTACCTTGAAAACCGAGTTGCTTTGGCAAGGATAATGTTTCCAATTGAGGCAAGAGTA 1845
W3NPR          CTGTACCTTGAAAACCGAGTTGCTTTGGCAAGGATAATGTTTCCAATTGAGGCAAGAGTA 1335
G3DNPR1        CTGTACCTTGAAAACCGAGTTGCTTTGGCAAGGATAATGTTTCCAATTGAGGCAAGAGTA 1026
G3BNPR1-1      TTGTACCTTGAAAACCGAGTTGCTTTGGCAAGGATAATGTTTCCAATTGAGGCAAGAGTA 2316
                ***********************************************************
G3ANPR1        GCAATGGACATTGCTCAAGTGGATGGTACTTTGGAATTTACCCTTGGTCCTAGTACAAAT 891
G3BNPR1-2      GCAATGGACATTGCTCAAGTGGATGGTACTTTGGAATTTACCCTTGGTTCTAGTACAAAT 1905
W3NPR          GCAATGGACATTGCTCAAGTGGATGGTACTTTGGAATTTACCCTTGGTTATGGTACAAAT 1395
                                                                G3oligo
G3DNPR1        GCAATGGACATTGCTCAAGTGGATGGTACTTTGGAATTTACCCTTGGTTATGGTACAAAT 1086
G3BNPR1-1      GCAATGGACATTGCTCAAGTGGATGGTACTTTGGAATTTACCCTTGGTTCTAGTACAAAT 2376
               ************************************************  * ********
G3ANPR1        CCACCTCTGGAGATAACAACCGTTGATCTAAATGATACTTCTTTCAAAATGAAGGAGGAA 951
```

```
G3BNPR1-2       CCACCTCTGGAGATAACAACCGTTGATCTGAATGATACTTCTTTCAAAATGAAGGAGGAA 1965
W3NPR           CCACCTCTGGAGATAACAACCGTTGATCTAAATGATACTTCTTTCAAAATGAAGGAGGAA 1455
G3DNPR1         CCACCTCTGGAGATAACAACCGTTGATCTAAATGATACTTCTTTCAAAATGAAGGAGGAA 1146
G3BNPR1-1       CCACCTCTGGAGATAACAACCGTTGATCTGAATGATACTTCTTTCAAAATGAAGGAGGAA 2436
                ***************************** ******************************
G3ANPR1         CACTTAGCTCGGATGAGAGCCCTCTCCAAAACAG-----TCGAAGTCGGCAAACGTTTTT 1006
G3BNPR1-2       CACTTAGCTCGGATGAGAGCCCTCTCCAAAACAG-----TTGAACTCGGCAAACGTTTCT 2020
W3NPR           CACTTAGCTCGGATGAGAGCCCTCTCCAAAACAG-----TGGAACTCGGCAAACGTTTCT 1510
G3DNPR1         CACTTAGCTCGGATGAGAGCCCTCTCCAAAACAG-----TGGAACTCGGCAAACGTTTCT 1201
G3BNPR1-1       CACTTAGCTCGGATGAGAGCCCTCTCCAAAACAGATAAGTTGAACTCGGCAAACGTTTCT 2496
                **********************************     * *** ************* *
G3ANPR1         TCCCGCGCTGTTCAAATAAATTGGACAAGATCATGGATGATGAACCTGAGCTGGCTTCGC 1066
G3BNPR1-2       TCCCACGCTGTTCAAATGTGCTGGACAAGATCATGGACGATGAACCTGAGCTGGCTTCCC 2080
W3NPR           TCCCGCGCTGTTCAAATGTGCTGGACAAGATCATGGACGATGAACCTGAGCTGGCTTCGC 1570
G3DNPR1         TCCCGCGCTGTTCAAATGTGCTGGACAAGATCATGGACGATGAACCTGAGCTGGCTTCGC 1261
G3BNPR1-1       TCCCGAGCTGTTCAATTGTATTGGACAAGTTCCTGGATGATGAACCTGAGCTGGCTTCAC 2556
                ****  ********* *    ******** ** **** ******************** *
G3ANPR1         TCGGAAGAGATGCATCCTCAGAGAGGAGGAGGAGGTTTCATGACCTGCAAGATGCGCTCC 1126
G3BNPR1-2       TCGGAAGAGATGCATCCTCCGAGAGGAAGAGGAGGTTTCACGACCTGCAAGATACGCTTC 2140
W3NPR           TCGGAACAGATGCATGCTCCGAGAGGAAGAGGAGGTTTCACGACCTGCAAGATACGCTTC 1630
G3DNPR1         TCGGAACAGATGCATGCTCCGAGAGGAAGAGGAGGTTTCACGACCTGCAAGATACGCTTC 1321
G3BNPR1-1       TCGGAAGAGATGCATCTTCC-----------TCGGTTTCATGACCTGCAAGATACGCTCC 2605
                ****** ********  **              ******* ************ **** *
G3ANPR1         TGAAGGCGTTCAGCGAGGACAAGGAGGAGTTTAACAAAACGACAACCCTTTCATCTTCCT 1186
G3BNPR1-2       TGAAGGCGTTCAGCGAGGACAAGGAGGAGTTTAACAGAACGACAACCCTTTCATCTTCGT 2200
W3NPR           TGAAGGCGTTCAGCGAGGACAAGGAGGAGTTTAACAGAACGACAACCCTTTCATCTTCGT 1690
G3DNPR1         TGAAGGCGTTCAGCGAGGACAAGGAGGAGTTTAACAGAACGACAACCCTTTCATCTTCGT 1381
G3BNPR1-1       TGAAGGTTTTCAGCGAGGACAAGGAGGAGTTTGCCAGAACGACAGCCCTTTCATCTTCCT 2665
                ******  ************************  ** ******* ************* *
G3ANPR1         CATCGTCGACGTCCACTGTAGCAAGGAACTTGACAGGCCGGCCTAGGAGATGA 1239
G3BNPR1-2       CATCGTCGACGTCCACTGTAGCAAGGAACTTGGCAGGTCGAACTAGGAGATGA 2253
W3NPR           CATCGTCGACGTCTACTGTAGCAAGGAACTTGGCAGGTCGAACTAGGAGATGA 1743
G3DNPR1         CATCGTCGACGTCCACTGTAGCAAGGAACTTGGCAGGTCGAACTAGGAGATGA 1434
G3BNPR1-1       CATCGTCGACGTCC-CT-CAGCAAGGAATTTGACAGGTCGGCCTAG------- 2709
                *************  **  ********* *** **** **  ****
```

File S3   The alignment of G7NPR1-like homologs with clustalw2

```
G4ANPR1         ATGGGGAGGTCACCGTCCTGCGGTAGTGAGGCCCCCGTCAAGAAGGGCCCGTGGACGGAG 60
G7DNPR1         ATGGGGAGGTCGCCGTCCTGCAGCAGTGAGCCCCCCTTGAAGAAGGGCCCATGGACGGAG 60
G7ANPR1         ATGGGTCGGTCGCCATCCTCCACCGGTGTCGCCGGCGTGAAGAAGGGCCCGTGGACGGTA 60
                *****  **** ** **** *    ***   **  * * *********** *******
G4ANPR1         GAGGAGGACAGGCTGCTCGTGGACTACATCAAGTTGCATGGCACCGGCGGCAACTGGCGA 120
G7DNPR1         GAGGAGGACAGGCTGCTCGTGGACTACATCAAGTTGCATGGCACCGGCGGCAACTGGCGA 120
G7ANPR1         GAGGAGGATATACTGCTCATAAACTACTTCCAGAGGCATGGCAACGGCGGCAACTGGCGA 120
                ******** *  ****** *  ***** ** **  ******** ****************
G4ANPR1         ACCATCCCCAAGCGCGCTGGCCTCAACCGCTGCGGCAAGAGCTGCCGCCTCCGTTGGACC 180
G7DNPR1         ACCATCCCCAAGCGCGCCGGCCTCAACCGCTGCGGCAAGAGCTGCCGCCTCCGTTGGACC 180
G7ANPR1         ACCTTGCCCCAGTGCGCCGGCCTGAACCGCTCCGGCAAGAGCTGCCGCCTCCGCTGGACC 180
                *** * *** ** **** ***** ******* ********************* ******
G4ANPR1         AACTACCTCCACCCCGACATCAAGCGCGGACCCTTCACCGACGATGAGGAGAAAACCACC 240
```

```
G7DNPR1         AACTACCTCCACCCCGACATCAAGCGCGGACCCTTCACCGATGAGGAGGAGAAAACCACC 240
G7ANPR1         AACTACCTCCGCCCCAACATCAAGCGCGGAACCTTCACCGACGATGAGGAGAAAACCATC 240
                ********** **** ************** ********** ** ************* *
G4ANPR1         ATCCACCTCCACTCCATGCTCGGCAACAAGTGGTCGGCGATTGCAACCCATCTGCCGGGA 300
G7DNPR1         ATCCACCTCCACTCCATGCTCGGCAACAAGTGGTCGGCGATTGCAACCCATCTGCCAGGA 300
G7ANPR1         ATACACCTCCACTCCATTCACGGGAACAAGTGGTCGGCGATCGCGACTCATTTGCCGGGG 300
                ** ************** * *** ***************** ** ** *** **** **
G4ANPR1         AGGACAGGCAATTTTATCAAGAATTACTGGAACATGAACCTGCGCAAGAAGCTACTCCAC 360
G7DNPR1         AGGACAGGCAATTTTATCAAGAATTACTGGAACACAAACCTGCGCAAGAAGCTGCTCCAC 360
G7ANPR1         AGAACAGGCAATTTTATCAAGAATTACTGGAACACGAACCTGCGCAAGAAGCTGCTCCAG 360
                ** *******************************  ***************** *****
G4ANPR1         ATGGGCATCGACCCCATCACGCACAGCCCGCGCGCCGACCTCAGCCAGCTTACAGGGCTC 420
G7DNPR1         ATGGGCATCGACCCCATCACGCACAACCCGCGCACCGACCTCAGCCAGCTTACAGGGCTC 420
G7ANPR1         ATGGGCATCAACCCCGTCACGCAC---CGACACACTGACCTCAGCATGCTCAAGGGGCTC 417
                ********* ***** ********   *  * * * *********  *** *  ******
G4ANPR1         CCCAGTCTCCCCGCCACCGCCGCCGCCCCCAGTAATAGCCTCTCC---GGCGTCAGCATG 477
G7DNPR1         CCCAGTCTCCCCGCCACCGCCGCCGCCCCCAGTAATAGCCTCTCC---AGCGTCAGCATG 477
G7ANPR1         CCCGGCCTGCACGCC------GCTGCCCCCGGTAATAGCCTCTCCTCTGGTGTCAGCATG 471
                *** * ** * ****      ** ****** **************    * *********
G4ANPR1         GGCCCCAGCGCCTGGGACATGAATGCTGTCCAGGCCCACGCCGACGTGGCCAGCTTCGCC 537
G7DNPR1         CCCCCCAGCACCTGGGACATGAATGCTGCCCAGGCCCACGCTGACGTGGCCAGCTTCGCC 537
G7ANPR1         GGCACCAG-------GGCA-----------CAGCCCCACACCAACGCGGCCAGCTTCGCC 513
                  * ****       * **           *** ***** *  *** *************
G4ANPR1         GGGGGCGGCTCGGACATGAACACTCTCAGGATCAAGGACGAAGCTACCAGCTTCACCGG- 596
G7DNPR1         GGGGGCGG---------------TCTCAGGATCAAGGACGAAGCTACCAGCTTCACCGG- 581
G7ANPR1         GGGAGCAGCTCGGACATGAACACTCTCGGGATCAAGGACAAAGCAGCCAGCTTCGCCGGG 573
                *** ** *               **** *********** ****  ******** ****
G4ANPR1         --CAGCTACTCAGACATGAAACCTCCCAGGATCAAGGATGAAGCCAGCAGCTTCACCGGT 654
G7DNPR1         --CAGTTACTCAGACATGAAACCTCCTATGATCAAGGATGAAGCCAGCAGCTTTACCGGC 639
G7ANPR1         AGCAGCAGCTCGGACATGACTACTCTCGGGGTCAAGGACGAAGCCGCCAGCTTCACTGGC 633
                  ***   *** *******   ***    * ******* ******  ****** ** **
G4ANPR1         AGCCACTCGGAAATGAAACCTCCCAGAATCAAGTACGAAGCCGCCAGCTTAATCAGTAGC 714
G7DNPR1         AGTCACTCGGAACTGAAACCTCCCAGGATCAAGTACGAAGCCACTAGTTTCACCGGTAGC 699
G7ANPR1         AGCCACTCAGACATGAAACCTCCCAGGATCAAGGACAAAGCCGCCAGCTTTACCGGCAAC 693
                ** ***** **  ************* ****** ** ***** * ** ** * * * * *
G4ANPR1         CACTCGGACATGAAACCTCCCAGGATCAAGGACGAAGCCGCCAGCTTCACTGGAAGACAC 774
G7DNPR1         CACTCGGACATGAAACCTCACAGGATCAAGGACGAAGCCGCCAGCTTCACTGGCAGACAC 759
G7ANPR1         TACTCGAAGATGAAACCTTTCAGGATCAAGGACGAAGCACTCAGCTTCACTGGCAGCTAC 753
                 ***** * *********  ******************   ************ **  **
G4ANPR1         TCGGACATGAAACCTCTCAGGGTCAAGGACGAAGCCGCCTGCTTCACCGGCAGCTACTCG 834
G7DNPR1         TCGGACATGAAACCTCTCAGGATCAAGGACGAAGCCGCCAGCTTCACCGGCAGTTGCTCG 819
G7ANPR1         TCGGACATGAATTCTCTCAGGATCAAGGACGAAGCCACCAGCTTCACAGGCAGCTACTCG 813
                ***********  ******** ************** ** ******* ***** * ****
G4ANPR1         GACAGGAAACCTCCCAGGATCCAGGACGAA---------------GGCAGCTACTCAGGC 879
G7DNPR1         GACAGGAAACCTCCCAGGATCCAGAACGAA---------------GGCAGCTACTCGGGC 864
G7ANPR1         GACATTAAATCTGTGAGTCTCCAGGACGAAACCACCAGCTTCACTGGCAGCTACTCGGCC 873
                ****  *** **   **  ***** *****               *********** * *
```

```
G4ANPR1         GCTGTTGCTGGCTGGGACATGAACACTCTCTCCATGCTCCAGCCCAACACCGCCAAGTTC 939
G7DNPR1         GCTGGTGCTGGTTGGGACATGAACACTCTCTCCATGCGCCAGTCCAACGCCGCCAAGTTC 924
G7ANPR1         ACCGGCACTGGCTGGGACATGAATGCTCTCTCCGGGCTCCAGGCCGACCCCGACAAGTTC 933
                 * *   **** ***********  ********  ** **** ** ** *** *******
G4ANPR1         CATCTCCTCCAGGGCCATTGTTGGGTGATCAGTGCTTCCTTGGGCGCCATGAGAGCTCTT 999
G7DNPR1         CATCTCCTCCAGGGCCATTGTTGGGTGATCAGTGCTTCCTTGGGCGCCATGAGAGCTCTT 984
G7ANPR1         CAGCTCCTCCAGGACCATTGTTGGGTGATCAGTGCTTCGTTGGGCGCCATGAGAGCCCTT 993
                ** ********** ************************ ***************** ***
G4ANPR1         GTTGGGAAGCTGGACGTGCTGCTGCTAGATGATTCATCTTCTCAGAAATTCTTCTCCAAG 1059
G7DNPR1         GTTGGGAAGCTGGACATGCTGCTGCTAGATGATTCATCTTCTCAGAAATTCTTCTCCAAG 1044
G7ANPR1         GTTGAGAAGCTGGACATGCTGCTGCTAGCTTATTCATCTCCTCAGGAATGCTCCTCCAAG 1053
                **** ********** ************ * ******** ***** *** ** *******
G4ANPR1         AGGGTGATGGATGGGTTGCGCCTCCTCAAAGATGATGTGGAAGAGCTAAGCTCCTACCTT 1119
G7DNPR1         AGGGTGATGGATGGGATGCGTCTCCTCAAAGATGATGTTGAAGAGATAAGCTCCTACCTT 1104
G7ANPR1         AGGGTGAAGGATGGGATGCACCTTCTCAAAGATGACCTTGAAGAGATAAGCTCCTACCTT 1113
                ******* ******* ***  ** ***********  * ****** **************
G4ANPR1         GATGAACTGTCAGAGGTGGATGACCATCCCCCAATAGCCATGTGCTGGATGAATGAGGCG 1179
G7DNPR1         GATGAACTGTCAGAGGTGGATGACCATCCCCCAATAGCCATGTGCTGGATGAATGAGGTG 1164
G7ANPR1         GATGAACTTTTAGAGGTTGAGGACCCTCCCCCGATGGCCATGTGCTGGATGAATGAGGCA 1173
                ******** * ****** ** **** ****** ** **********************
G4ANPR1         CGTGATCTGTCTTACGACATGGAGGATTACATTGACAGCCTTTTATC---TGTGCCACCT 1236
G7DNPR1         CGTGACCTGTCTTACGACATGGAGGATTACATTGACAGCCTTTTATC---TGTGCCACCT 1221
G7ANPR1         CGCGAACTGTCTTATGACATGGAGGATTACATTGACAGCTTATTATTATGTGTGCCACCT 1233
                ** ** ******** ************************ * ****    **********
G4ANPR1         GACCATTTCAACAAAAACAAGAAGAAGAAAAACAAGAAGATGATGACCATAAAGAAGAGG 1296
G7DNPR1         GACCATTTCAACAAAAACAAGAAGAAGAAAAACAAGAAGATGATGACCATAAAGAAGAGG 1281
G7ANPR1         GACCATTTCAACAAAAATAAGAAGAGGAAGAAGAATAAAAAGAAGATAATAAAGAAGAGG 1293
                ***************** ******* *** ** ** ** * ** **   ***********
G4ANPR1         CTCAAATGGTACAAACAGATCACATACGTCACGCAAGTGTCAGAACATGGTGTCAAGACC 1356
G7DNPR1         CTCAAATGGTACAAACAGATCACATACGTCATGCAAGTGTCAGAACATGGTGTCAAGACC 1341
G7ANPR1         CTCAAGTGGCACAGACAGATCACATACATTGCGCAAGTGTCAGAACATGGTGTCAGGACC 1353
                ***** *** *** ************* *   *********************** ****
G4ANPR1         AGCAAACGCGTTCATGTTAGTGTTGTCCCTCGTCTTCCCAAGAAGCCCAAGATTGCGGAA 1416
G7DNPR1         AGCAAACGCGTTCATGTTAGTGTTGTCCCTCGTCTTCCCAAGAAGCCCAAGATTGCGGAA 1401
G7ANPR1         AGCAAACGCATTCATGTTACTGTTGTCCCTCCTCTTCCCAAAAAGTCCAAGATTGCGGAA 1413
                ********* ********* *********** ********* *** **************
G4ANPR1         ACAATTTCAGATTTCAGGATCTATGTCCAGGAGGCCATTGAACGCAACTTGAGGAGACGT 1476
G7DNPR1         ACAATTTCAGATTTCAGGATCTATGTCCAGGAGGCTATTGAACGCAACTTGAGGACACGT 1461
                                                                 G7AoligoF
G7ANPR1         ACAATATCAGAATTCAGGATCTATGTCCAGGAGGCAATTGAACGGCAT-----GACAGGT 1468
                ***** ***** *********************** ********  *      ** * **
G4ANPR1         ACATTCTTGTCCACTGGCCATATGCTTCCAATGCACTATGAAGAAACCAATGATCATGTT 1536
G7DNPR1         ACATTCTTGTCCACTGGCCATATGCTTCCAATGCACTATGAAGAAACCAATGATCACGTT 1521
G7ANPR1         ATCTAATAGTC-ATTGATGACTTG--TGGGATGCATCAGCATGGGAT--ATTATTAAAT- 1522
                *  *  * *** * **   *  **  *   *****  *  * *  *   ** ** *  *
G4ANPR1         GTAATTGATGGCCGGATGAATGAGTTTATCGACTCACTGGTGGCTGATATTGTGGCAGAT 1596
G7DNPR1         GTAATTGATGGCCGGATGAATGAGTTTATCGACTCACTGGTGGC------------AGAT 1569
G7ANPR1         GTGCTT---TTCCGAAGGGAAG-----------CCA-TGGCAGT------------AGAA 1555
```

```
                     **  **     *** * * * *          ** ***  *           ***
G4ANPR1     CAGCAGCAGCT---CAAGGTGGTATCTGTCCTTGGATCTGGATGTCTGGGTAAAACTACA 1653
G7DNPR1     CAGCAGCAGCT---CAAGGTG--------CCT---------------------------- 1590
G7ANPR1     TAATAATAACCACACAGATTGA-------------------------------------- 1577
             *  *  * *    **   **
G4ANPR1     CTTGCCAAAGTGTTGTATCTAATTGTTATTGATGATTTGTGGGACGCATCAGCATGGGAT 1713
G7DNPR1     ------------------------------------------GACACA----------AT 1598
G7ANPR1     -----------------------------------------AGATGTT----------GC 1586
                                                      **
G4ANPR1     ATTATCAAGTATGC-CTTTCCAAATGGAAACCGTGGCAGCAGAATAATAATAACTACACA 1772
G7DNPR1     AT-ATCTGATATGC-GT---------GAATTCCTGGCA---------------------- 1625
G7ANPR1     ATTAACATGT-TGCTGT----------GATCACTGG------------------------ 1611
            ** * *   * ***  *           *    ***
G4ANPR1     GATTGAAGATGTTGCATTAACATGTTGCTCTGATCACTCAGAACATGTTTTCGAGATGAA 1832
G7DNPR1     ------------TGCA------------------------GAACATGTTTTCATGATGAA 1649
G7ANPR1     ----------------------------------------GAGCATGTTTTCGAGATGAA 1631
                                                    ** *********  ******
G4ANPR1     ACCTCTTGACGAAGATCACTCAAGAAAGCTATTCTTTAACAGGCTTTTTGGTTCTGGAAG 1892
G7DNPR1     ACCTCTTGATGACGATCACTCAAGAAAGCTGTTCTTTAACAGACTTTTTGGTTCTGGAAG 1709
                    G7A
G7ANPR1     ACCTCTTGACGACGATCACTCAAGAAAGCTATTCTTTAACAGACTTTTTGGTTCTGAAAG 1691
                    RTF
            ********* ** ***************** *********** ************ ***
G4ANPR1     TGACTGTCCTGAAGAATTTAAACAAGTTTCAAATGAAATTGTTGATG------------- 1939
G7DNPR1     TGACTGCCCTGAAGAATTTAAACAAGTTTCAAATGAAATTGTTGATGTCACAATCGTGGT 1769
G7ANPR1     TGACTGTCCTGAAGAATTCAAACAAGTTTCAAATGAAATTGTTGATATAT--GTGGTGGT 1749
            ****** *********** **************************
G4ANPR1     ------------------------------------GC---TCCATAGTCTGGA--AGGA 1958
G7DNPR1     CAAGTTCTA----CTTCGGAAAGAACAAGACAAAGGGC---TCCATAGTCTGGA--AGGA 1820
G7ANPR1     TTGCCGCTAGCAACAATCAATATAGCTAGTCATTTGGCAAACCAACAGACAGGATTAGCA 1809
                                                 **    * * ** * ***  ** *
G4ANPR1     --TGATCTGGTGAAGCA-ATGGGTGG----CTGAAGCC------TTCATCAA--TACAC- 2002
G7DNPR1     --TGATCTGGTGAAGCA-ATGGGTGG----CTGAAGCC------TTCATCAA--TACAC- 1864
G7ANPR1     CTTGATTTGCTGGCATACATGTGTGATTCACTGAGGTCCCATTCTTGGTCAAGTTCTACC 1869
              **** ** **    * *** ***     **** * *     **  ****  *  **
G4ANPR1     --GAAAAGGGAAAGGAAAAGATCAAAACT-GGATGGAGAAAGCTGCA--GGAATCT---- 2053
G7DNPR1     --GAAAAGGGAAAGAAAAAGATCAAAACT-GGATGGAGAAAGCTGCA--GGAATCT---- 1915
                    G7Aoligo
G7ANPR1     TTGGAAAGAACGAGAAAAGTACTGAACCTTAGCTACAATAATCTTCCTCAGCATCTGAAG 1929
              * ****    ** ***  *   ** **  * *  *  ** ** *    * ****
G4ANPR1     --AT-TTCGATGAACTT----GTTGAT--AGAAGATTCATCCAACCCCTAAATATCAACT 2104
G7DNPR1     --AT-TTCGATGAACTT----ATTGAT--AGAAGATTCATCCAACCTCTAAATATCAACT 1966
G7ANPR1     ACATGTTTGCTGTATTTTCATATGTATCCAGAGGGCTC-CATAATCTGGAAGGATGATCT 1988
              ** ** * ** * **     *  **  *** *  **     ** *   **  ** * **
G4ANPR1     AC---AACAATAAGGTGTTGTCATGT--ACTGTT-CATGAAGTGGTA-----CGTGATCT 2153
G7DNPR1     AC---AACAATAAGGTGTTGTCCTGT--ACTGTT-CATGAAGTGGTA-----CGTGATCT 2015
G7ANPR1     GGTGAAGCAATGGGTGGCTGAAGGGTTTGTTGCTACAGGAAAAGGGAAAGAGCAAGATAA 2048
                 * ****  *  * **    **    ** * ** ***  ** *    *  ***
G4ANPR1     TC---TTGCACACAAGTCTG-----------------------TGGTAGTAGATTATA 2185
G7DNPR1     TC---TTGCACACAAGTCTGCAGAAGAGAATTTCATTG-------TGGTAGTAGATTATA 2065
```

```
G7ANPR1     AGAAACGACAGAGAAAACTGCAGGAATCTATTTTGATGCGCTTGTTGATAGAAGATT-CA 2107
                    ** * **  ***                         ** *** *****  *
G4ANPR1     ATCGAAAGA-ATATATCACTTTC--TCATAAGGCCCGTCGGCTATCTCTCCTCTTTGCCA 2242
G7DNPR1     ATCGAAAGA-ATATATCACTTTC--TCATAAGGCCCGTCGGCTATCTCTCCTCTTTGCCA 2122
                                             G7A-MF
G7ANPR1     TTCAACCTTTATATATCAACTACAACAATAAGGTTCGTCGTCTATCTCTCATCTTTGGTG 2167
             ** *     ********  * *     ******  ***** ********* ******
G4ANPR1     ATGCAAGTTATGCCAAGACACCTGTAAACATCACAAAGGCGCAGGTTCGGTCACTTAATT 2302
G7DNPR1     CTGCAAGGTATGCCAAGACACCAGTAAACATCACAAAGCCGCAAGTTCGGTCACTTAATT 2182
                                                           G7A-RTR
G7ANPR1     ATGCAAAATATGCCAAGACACCAGAAAACATCACAAGGTCACAAGTTCGGTCACTCAGAT 2227
             *****  ************** * *********** * * ** *********** *  *
G4ANPR1     TTGTTGGATTGTTCGAGTGCATGCCTTGTATTGGAGAGTTCAAGGTTCTCCGTGTTCTGA 2362
G7DNPR1     TTGTTGGATTGTTCGAGTGCATGCCTTGTATTGGAGAGTTCAAGGTTCTCCGTGTTCTGA 2242
G7ANPR1     TTTTTGGATTATTCAAGTGTATGCCTTGTGTTAGAGACTTCAAGGTTCTCCGTGTCCTAA 2287
            ** ******* *** **** ********* ** **** **************** ** *
G4ANPR1     ACCTTCAACTATCCGGTCATTGTGGCGACCATGACCCCATAGACCTCACTGGGATTTCAG 2422
G7DNPR1     ACCTTCAACTATCCGGTCATTGTGGCGACCATGACCCCATAGACCTCACTGGGATTTCAG 2302
G7ANPR1     ACCTTCAACTATCTGGTCATTGTGGCGCCCATGACCCCATAGACCTCACTGGCATTTCAG 2347
            ************* ************* *********************** *******
G4ANPR1     AACTGTTTCAGTTGAGATATTTGAAGATTACAAGTGATGTGTGCATAAAACTACCAAACC 2482
G7DNPR1     AACTGTTTCAGTTGAGATATTTGAAGATTACAAGTGATGTGTGCATAAAACTACCAAACC 2362
G7ANPR1     AGCTGTTTCAGTTAAGATATTTGAAGATCACAAGTGATATCTGCATAAAACTACCAAACA 2407
            * *********** ************** ********* * *****************
G4ANPR1     AAATGCGAAAGCTGCAATGTTTGGAAACGCTGGATATTATAGATGCACCAAGAGTCACTG 2542
G7DNPR1     AAATGCGAAAGCTGCAATGTTTGGAAACGCTGGATATTATAGATGCACCAAGAGTCACTG 2422
G7ANPR1     GTATACGAGGGCTGCAATGTTTGAAAACGCTCGATCTTATGGATGCAACAAGAGTCACTG 2467
              ** ***  ************* ******* *** **** ****** ************
G4ANPR1     CTATCCCGTGGGATATTATATATCTCCCACACTTGTTGCACTTGACGCTTCCTGTTGACA 2602
G7DNPR1     CTATCCCATGGGATATTATATATCTCCCACACTTGTTGCACCTAACCCTTCCTGTTGACA 2482
G7ANPR1     CTGTTCCATGGGATATTATACATCTCCCACACTTGTTGCACCTGACTCTTCCGGTTGACA 2527
            ** * ** ************ ******************** * ** ***** *******
G4ANPR1     CAAATCTGCTGGATTGGATTGGCAGCATGACTGGCTCCATCATCAGTCTGTGGAGCCTTG 2662
G7DNPR1     CAAATCTGCTGGATTGGATTGGCAGCATGACTGACTCCATCATCAGTCTGTGGAGCCTTG 2542
G7ANPR1     CAAATCTGCTGGATTGGATAGGCAGCATGACTGACTCAGTCATCAGTCTGTGGAGCCTTG 2587
            ******************* ************* ***  *********************
G4ANPR1     GCAACCTGAACTACCTGCAAGATCTTTATCTTACCATTTCTTCTACACATCCTTCTGGCC 2722
G7DNPR1     GCAACCTGAACTACCTGCAAGATCTTTATCTTACCATTTGTTCTACACATCCTTCTGGCC 2602
                                                    G7A-MR
G7ANPR1     GCAAGCTGAACTACCTGCAGGATCTTCATCTTAACATGTCGTGTAT---TCCTTCA---C 2641
            **** ************** ****** ****** *** ** * **     ******    *
G4ANPR1     ATCCAGAGAAAAACATGGAAGCTCTGGGTTCTTTACTAGGAGGACATGGCAACCTGAAAA 2782
G7DNPR1     ATCCAGAGAAAAACATGGAAGCTCTGGGTTCTTTACTAGGAGGACATGGCAACCTGAAAA 2662
G7ANPR1     TTCTGGG------CGTGGAAGCTCTGGCTCATTTGATCGGAGGGCATGGTAACCTGAAAA 2695
             **  *      * ************ *  ***  * ***** ***** **********
G4ANPR1     TTATAGTAGTGTCTCATGGCCCATTGGTTAAAGATATTGTGGTTCCTGGTGCCTCAAAAG 2842
G7DNPR1     TTATAGTAGTGTCTCATGGCCCATCGGTTAAAGATATTGTGGTTCCTGGTGCCTCAAAAG 2722
G7ANPR1     CTATAGTAATGTCTCATGGCTCATCGGT------------GATTCCTGGCGCCTCAAAAG 2743
             ******* *********** *** ***            * ******* **********
G4ANPR1     TAATCATTTCCTGGGATGAGTTGGAACCTCTCCCACTTCTCCAGAGATTTGAATGCTCGC 2902
```

```
G7DNPR1         TAATCATTTCCTGGGATGAGTTGGAACCTCTCCCACTTCTCCAGAGATTTGAATGCTCAC 2782
G7ANPR1         CAATAATTTCCTGGGATGACTTGGAACCTCTCCCCCTTCTCCAGAGATTTGAATGCTCGC 2803
                 *** ************** ************** *********************** *
G4ANPR1         CACACAGCTGCGTCATATTTTCCCGAATTCCTAGGTGGATTAAAAAACATGGTAACTTGT 2962
G7DNPR1         CGCACAGCTGCGTCATATTTTCCCGAATTCCTAAGTGGGTTAAAAAACTTGGTAACTTGT 2842
G7ANPR1         CGCACAGCTGCATCACATTTTCCCGAATTCCTAAGTGGATTAAAGAACTTAGCCACCTGT 2863
                * ********* *** ***************** **** ***** *** * *  ** ***
G4ANPR1         GCATTCTGAAGATTGCAGTGGTGGAACTGCAGATGAGTTGTGTTGATATTCTCAGAGGAT 3022
G7DNPR1         GCATTCTGAAGATTGCAGTGGTGGAACTGCAGATGAGTTGTGTTGATATTCTCAGAGGAT 2902
G7ANPR1         GCATTCTGAAGATTGCAGTGGTGGAACTGCAGATTAGTTGGGTTGATATTCTGAGAGGAT 2923
                ********************************** ***** *********** *******
G4ANPR1         TGCCTGCCCTCACTGCTCTGTCGCTGTATGTGCGGTGCGCGCCTGCTCAAAGGATCCTCT 3082
                                                                    G7oligoF
G7DNPR1         TGCCTGCCCTTACTGCTCTGTCGCTGTATGTGCGGTGCGCGCCTGCTCAAAGGATCCTCT 2962
G7ANPR1         TGCCTGCCCTCACTGCTCTGTCTCTGTATGTGCGGCGTGCGCCCATTGAAAGGATCATCT 2983
                ********** *********** ************ * *****   * ******** ***
G4ANPR1         TTGACAAGATGGCCGGGTTTTCAGTTCTCAAGTACTTCAAGTTGAGGTGCACGAGTGGTA 3142
G7DNPR1         TTGACAAGATGGCCGGGTTCTCAGTTCTCAAGTACTTCAAGTTGAGGTTCACGAGTGGTA 3022
G7ANPR1         TCAACAAGGCAGACGGGTTCTCAGTTCTCAAGTACTTCAAGTTGACTTGCACGAGTGGTA 3043
                *  *****   * ****** *************************  * ***********
G4ANPR1         TCCCATGGCTGAAATTTGAGGCGGATGCAATGCCTAACCTCTGGAAGCTGAAGCTAGGTT 3202
                                           G7oligoR
G7DNPR1         TCCCATGGCTGAAATTTGAGGCGGATGCAATGCCTAATCTCTGGAAGCTGAAGCTGGGTT 3082
G7ANPR1         TAGCGTGTCTAAAATTTGAGGCGCAAGCAATGCCTAATCTCTGGAAGCTCAAGCTAGGTT 3103
                *  * ** ** ************ * *********** *********** ***** ****
G4ANPR1         TCAATGCAATCCCCCGAATGGACCAACATCAACTAATCTGCATCGAGCATATTCCAGACC 3262
G7DNPR1         TCAATGCAATCCCCCGAATGGACCAACATCAACTAATCCGCATCGAGCATATGCCAGACC 3142
G7ANPR1         TCAATGCCACCCCCCGAATGGACCAACATCAACTTATCCGCATCGAGCATATGCCAAACC 3163
                ******* * ************************ *** ************* *** ***
G4ANPR1         TTAAAGAGATCTCTGTAAAATTTGGGGGTGTAGCTGCTCTTATAGAGTATGCCGTGAAGA 3322
G7DNPR1         TTAAAGAGATCTCTGTAAAATTTGGGGGTATAGCTGCTCTTATAGAGTATGCCGTGAAGA 3202
G7ANPR1         TTAAAGGGATATCCGTAAAATTTGGGGGTGTAGCTGCTCATATAGAGTATGC---GAGGT 3220
                ****** *** ** *************** ********* ************   ** *
G4ANPR1         CTGTCATTAGTAATCATCTGAGAAATCCTAGAGTCCATGTGTGTTTGGTGACTTCTACTT 3382
G7DNPR1         CGGTCATTAGTAATCATCTGAGAAATCCTAGAGTCCATGTGTGTTTGGTGACTTCTACTT 3262
G7ANPR1         CTGTTGTAACTAGTCATCCGAGAAATCCTACAATCAATATGCAATTGGTGAATTTTTGTT 3280
                * **  * * ** ***** *********** * ** ** **   ******* ** *  **
G4ANPR1         TCTATGGTGCCGAGAGGACAAAAGAGAAACCTCCCACCAACTCGGCGATCAAGTCGCGCA 3442
G7DNPR1         CCTATGGTGATGAAAGCACAAAAGAGAAACCTCCCACCAACTCGGCGGTCGAGTCGCACA 3322
G7ANPR1         CAAATTGTGATGGAAGCACAAAACAGAACCTTCTCACCAACTGGGCTGTCAAGTCACGTG 3340
                   ** ***  *  ** ****** **** * ** ******** ***  ** **** *
G4ANPR1         TGGCCGCCACAACGCCAATGGAGACTCCGACCAGCCACGTCACCACCGCCTTCGCCCACT 3502
G7DNPR1         TGGCCGCCGCAACGTCAACGGAGACTCCGACCAGCCAGGTCACCACCGCCTTCGCCCACT 3382
G7ANPR1         CCGCAGCCTCAGCACCAATGGAGGATCCTCCCAGTCACTTCACCACCACCTTCTCCAGCT 3400
                  ** *** ** *  *** ****  ***  **** **  ******** ***** **  **
G4ANPR1         GCAACAGTGTCTCGACAGAGGGTGCTGCTGCTTCCGGGGACGCAGATGTGGAGACGCTCT 3562
G7DNPR1         GCGACAGTGTCTCGACAGAGGGTGCTGCTGCTTCCGGGGACGCAGATGTGGAGACGCTCT 3442
G7ANPR1         CCGATGGCATCTTCATGGCAG------------------ACATGAACTTGGAGGCGCTCT 3442
                 * *  *  ***  *  *  *                  **    *  ***** ******
```

```
G4ANPR1         GCCACCTCTCCGACAATCTCGCCACCGCCTTTCGCTCTCCAGAAGACTTCACCTTCCTTG 3622
G7DNPR1         GCCACCTCTCCAACAATCTCGCCACCGCCTTTCGCCCTCCAGAAGACTTCACCTTCCTCG 3502
G7ANPR1         GCCGCCTATCTGACAACCTCGCCGCCGTCTTCCGCTCGCCGGATGTCTTCGCCTTCCTTG 3502
                *** *** **  **** ****** *** *** *** * ** ** * **** ******* *
G4ANPR1         TCGACATGCGTGCGTACCGCTCCATGCTCTCCGTCCAGAGTCCCTTCCTGGATGCCGTCT 3682
G7DNPR1         TCGACATGCGTGCGTACCGCTCCGTGCTCTCCGTCCGGAGTCCCTTCCTTGATGCCATCT 3562
G7ANPR1         TTGACATGCA---------------------------------------------CATC- 3516
                * *******                                              * **
G4ANPR1         TCGCGCGTCGTGCCGCTGAAGGGGAGGACGACCCGCTGGACCTCCGGGAGCTTCTGGGTG 3742
G7DNPR1         TCGCGCGTCGTGCCGCTGAAGGGGAGGGCAACCCGCTGGACCTCCGGGAGCTTCTGGGTG 3622
G7ANPR1         ------GTCGTGCC--------------CGACGCGCCCGACCT----------------- 3539
                      ********              * ** ***  *****
G4ANPR1         AGGAGGTGGAGGTCGGGTATGAGGCACTGCAGTTGGTGCTCGAGTACCTGTACACCGGCC 3802
G7DNPR1         AGGAGGTGGAGGTCGGGTATGAGGCACTGCAGTTGGTGCTCGAGTACCTGTACACCGGCT 3682
G7ANPR1         ------------------------------------------------------------

G4ANPR1         GCATCCGCGACCTTCCCAAGTCAGCATGCGTTTGCGCCGACGTGGACGGCTGCGCCCACG 3862
G7DNPR1         GCATCCGCGACCTTCCCAAGTCGGCATGCGTTTGCGCAGACGTGGACGGCTGCGCCCATG 3742
G7ANPR1         ------------------------------------------------------------

G4ANPR1         TCGGTTGCCTTCCCGCCATCTCCTTCATGGCGCAGGTCATCTTTGCCGCATCCATCTTCG 3922
G7DNPR1         TCGGCTGCCTCCCTGCCATCTCGTTCATGGCGCAGGTCATCTTTGCCGCATCCATCTTCC 3802
G7ANPR1         ------GCCT-------------------------GTCCTCTTTGCCGCATCCACCTTCC 3568
                      ****                         *** *************** ****
G4ANPR1         AGGTTGCGGTGCTCACCAACCACTTCCAGTTTCAAAACTGTCCGCTTAAATGTGTTCAAT 3982
G7DNPR1         AGGTTGCCGTGCTCACCAACCACTTCCAG------------------------------- 3831
G7ANPR1         AGATTGCCAAGCTCACCAACCTCTTCCAG------------------------------- 3597
                ** ****   *********** *******
G4ANPR1         ATGGAAAGGTGCTGTGTGATTTTAATGGGAAAAGGAAAGGTAACTGTTCAAAGCATGGAA 4042
G7DNPR1         ------------------------------------------------------------
G7ANPR1         ---------------------------GAAAAAGGAAAG--------------------- 3609

G4ANPR1         ATGATGTGCTTCCAGTTTTGTTCAATAAAAATCTGTTTGAAGCAGTATGGAGAGGTGCTG 4102
G7DNPR1         ------------------------------------------------------------
G7ANPR1         ------------------------------------------------------------

G4ANPR1         ATATGCGGCTTCTCCTTGATGTCCTTGATGATGTTGAAGTGGATAACCTTCCATTGATCT 4162
G7DNPR1         -----CGGCTTCTCCTTGATGTCCTTGATGATGTTGAAGTGGATAACCTTCCATTGATCT 3886
G7ANPR1         -----TGGCGTCTCCTTGATGTCCTTGATGATGTTGAAGTTGATAACCTTCCATTGATCT 3664
                      *** ****************************** *******************
G4ANPR1         TATCTGTTGCAAACTTATGCAATAAATCTTGCATGCATCTGCTTGAGAGATGCCTTGAGA 4222
G7DNPR1         TATCTGTTGCGAACTTATGCAATAAATCTTGCATGCATCTGCTCGAGAGATGCCTTGAGA 3946
G7ANPR1         TATCTGTTGCAAACTTATGCAATAAATCATGCATGAAACTTCTCGAGAGATGCCTTGAGA 3724
                ********** ***************** ****** * ** *  ****************
G4ANPR1         TGGTAGTCCGGTCAAATCTTGACATGATTACTCTTGAGAAAGCATTGCCTCCAGATGTTA 4282
G7DNPR1         TGGTAGTCCGGTCAAATCTTGACATGATTACTCTTGAGAAAGCATTGCCTTCAGATATCA 4006
G7ANPR1         TGGTAGTTCGGTCAAATCTTGACATGATTACTCTTGAGAAAGCATTGCCTCCAGATATTA 3784
```

```
                    ******* ******************************************* ***** * *
G4ANPR1             TCAAGCAAATTACTGACTCGCGGATAAGTCTTGGGCTAATTTCACCCAAGGACAACGGAT 4342
G7DNPR1             TCGAGCAGATTACTGATTCACGGCTAAGTCTTGGACTAGTTTCACCCGAAGATAAGGGAT 4066
G7ANPR1             TCAAGCAAATTGCTGATTCACGCCTAAGTCTTGGGTTAGTTTCAGCCGAAGACAAGGGAT 3844
                    ** **** *** **** ** **  **********  ** ***** ** * ** ** ****
G4ANPR1             TTCCTAACAAACATGTAAGAAGGATATTCGGAGCTCTCGATTCTTATGATGTGGAGCTAG 4402
G7DNPR1             TTCCTAACAAACATGTAAGAAGGATACTCAGAGCTCTTGATTCTGATGATGTGTGTTTAG 4126
G7ANPR1             TTCCTAACAAACATGTAAGAAGGATACTCCAAGCTCTTGACAGTAATGATGCGGAGGTAG 3904
                    ************************** **  ****** **   * ****** *    ***
G4ANPR1             TGAGGTTGCTACTCTACGAAGGGCAGACTAATCTTGATGATGCATTTGCATTGCACTATG 4462
G7DNPR1             TGAGGATGCTGCTCAAAGAAGGGCGGACTAATCTTGATGGTGCATTTGCATTGCACCATG 4186
G7ANPR1             CGACGATCCTGCACCAGGAAGGGCAGACTAATCTTGATGATGCGTTTGCGTTGCACTATG 3964
                     ** * * ** * * * ******* ************** *** ***** ****** ***
G4ANPR1             CTGTAGAACATTGTGACCCAAAGATCACAACAGAACTTCTGGACCTCGAACTTGCAGATG 4522
G7DNPR1             CTGTAGAACATTGTGACTCGAAGGTCACAATGGAACTTCTGGATATCGGACTTGCAGATG 4246
G7ANPR1             CTGTTGAACATTGTGACTCAAAGATCATAACAGAAATTTTGGACATCACACTTGCAGATG 4024
                    **** ************ * *** *** **  *** ** ****  **  ***********
G4ANPR1             TTAATCGTAGAAATCAAAGAGGTTATACTGTTCTTCACATTGCTGCCAGGCGAAGAAATC 4582
G7DNPR1             TTAATCATAGAAATCCAAGAGGTTATACTGTTCTTCACGTTGCTGCCAGGCGAAGAGATC 4306
G7ANPR1             TTAATCATAGAAATCCTAGAGGTT------------------------------------ 4048
                    ****** ********  *******
G4ANPR1             CTAAAATTCTTGTCTCTGTTTTAACCAAAGGAGCTCGGCCTTCCGATCTTACATTTGATG 4642
G7DNPR1             CTAAAATTCTTGTCTCTGTTTTAACCAAAGGTGCTCGGCATTCTGATCTTACATTTGATG 4366
G7ANPR1             ------------------------------------------------------------

G4ANPR1             GAAGAAAAGCAATACAAATCTCAAAGAGACTCACAAAACATGGGGATTACTTTGGGATTA 4702
G7DNPR1             GAAGAAAAGCAGTTCAAATCTCAAAGAGACTCACAAAACATGGGGATTACTTTGGGATTA 4426
G7ANPR1             ------------------------------------------------------------

G4ANPR1             CTGAAGAAGGAAAACCGTCTCCTGAATATAGTTTATGCATTGAGATACTGGAGCAAGCTG 4762
G7DNPR1             CAGAAGAAGGAAAACCGTCTCGTGAAGATAGGTTATGCATTAAGATACTGGAGCAAGCTG 4486
G7ANPR1             ------------------------------------------------------------

G4ANPR1             AAAGAAGGGGTCCACAACTTGGAGAAGCATCGGTTTCTCTTGCAATAGCTGGTGACTGTC 4822
G7DNPR1             AAAGAAGGGATCCACAACTTGGAGAAGCATCGGTTTCTCTTGCAATAGCTGGTGACCGTC 4546
G7ANPR1             ------------------------------------------------------------

G4ANPR1             AACGTGGGACGCTGTTGCACCTCGAAAATCGAGTTACTTTGATGAGGATAATGTTTCCAA 4882
                                                  ExG7A-RTF
G7DNPR1             AACGTGGAAAGCTGTTGTACCTCGAAAATCGAGTTGCTTTGATGAGGATAATGTTTCCAA 4606
G7ANPR1             ------------------------------------CTTTGATGAAAATACTCTTTCCGA 4072
                                                        *********  *** * ***** *
G4ANPR1             CTGAGGCAAGAGTAGCAATGGACATCGCTCAAGTGGATTGCACTTTGGAGTTAACCCTTG 4942
G7DNPR1             CTGAGGCAAGAATAGCAATGGACATCGCTCAAGTGGATTGTACTTTGAAGTTAACCCTTG 4666
G7ANPR1             TTGAGGCAACAGTAGCAATGGAAATTCCTCAAGTGAAT---------------------- 4110
                     ******** * ********** **  ******** **
G4ANPR1             ATTCTGAGGCAAAACCTCCTTCGGAGAAAGAAGTGGCAACCATTGATCTAAATGAAACTC 5002
                                         ExG7A-RTR
G7DNPR1             ATTCTGGGGCAAAACCTCCTCCGGAGAAAGAACTGGCAACCATTGATCTAAATGAAACTC 4726
```

```
G7ANPR1         ------------------------------------------------------------

G4ANPR1         GTTTCCAAATGAATGACGAACACTTAGCTCGGATGAGTGCCCTCTTCAAAACAGTTGAAC 5062
G7DNPR1         CTTTCCATATGAATGAAGAACACTTAGCTCGGATGAGTACCCTCTCCAAAACAGTTGAAC 4786
G7ANPR1         -----------------------------------------------------GTCCAAC 4117
                                                                     **  ***
G4ANPR1         TCTGTAAACGCTTCTTTCCGCGTTGTTCAAATGTCATAGACAAGATTATGAATGATGAAC 5122
G7DNPR1         TCTGTAAACGCTTCTTTCCGCGTTGTTCAAATGTGATAGACAAGATTATGGATGATGAAC 4846
G7ANPR1         TCGGCAAACGCTTCTTCCCATGTTGTTCGAATGTGCTGAAGAAGATCATGGATGATGAAT 4177
                ** * *********** **  ******* *****  *  * ***** *** ********
G4ANPR1         CTGAGCTGGCTTCCCTTCAAAGAGACACATCCACCGAGAGGAATCGGAGGTTCGGTGACA 5182
G7DNPR1         CTGAGCTGGCTCCTCTTCAAAGAGACGCATCCACCGAGAGGACTCGGAGGTTCGGTTACC 4906
G7ANPR1         CTGAGCTGGATTCCATCGGAAGAGACACATCCACCGAGAGGAAGCGGAGGTTCCATGAGC 4237
                ********* * *  *   ******* ***************  *********  * *
G4ANPR1         TGCACGACACGCTTCAGAAGGCATTCAGCGAGGACAAGGGTTCGTCTAGTTTTGACAAAT 5242
G7DNPR1         TGCAGGACGCGGTTCAGAAGGCATTAAGCGAGGACAAGG------CTAGTTTTGACAAAT 4960
G7ANPR1         TGCAGGACTTGCTTCAGAAGGCATTCAGCGAGGATAAGGAGTCGTCTAGTTT-GACAAGT 4296
                **** ***  * ************* ******** ****      ******* ***** *
G4ANPR1         CT--GGGTCATAG-- 5253
G7DNPR1         CT--GGGTCATAG-- 4971
G7ANPR1         CTCAGGTTTATTTGA 4311
                **   ** * **
```

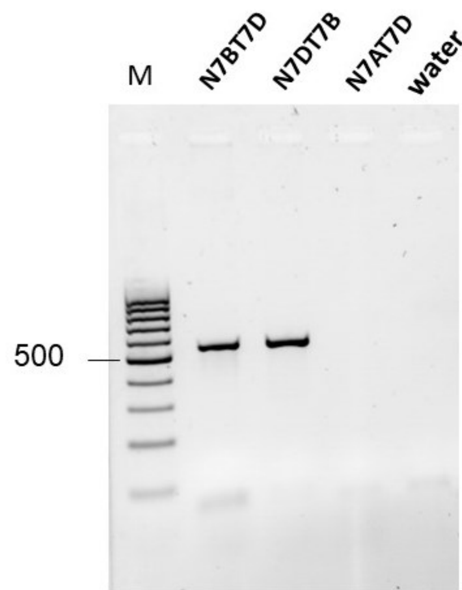


Figure 3.S1: Specificity of primers for screening mutations on *Ta7ANPR1*. The primers were designed based on the specific sequences of *Ta7ANPR1* and only amplified a fragment when 7A chromosome is present. Nothing was amplified when 7A chromosome is absent in N7AT7D or in negative water control.

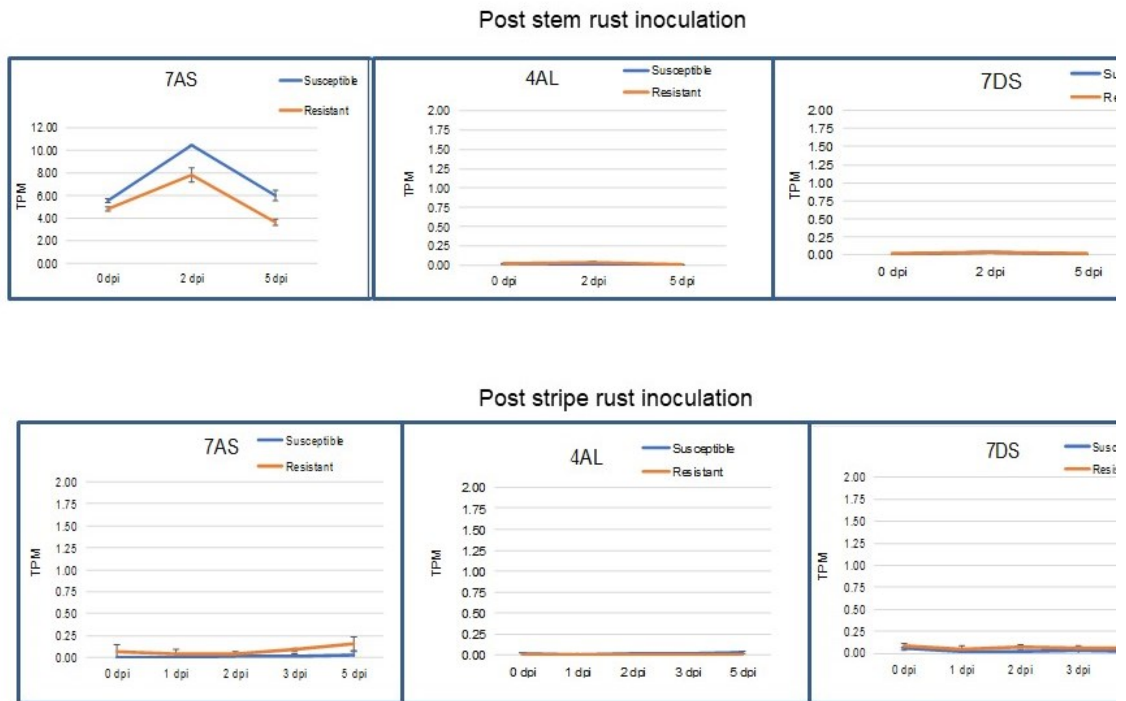


Figure 3.S2: Transcripts of *Ta7ANPR1, Ta4ANPR1* and *Ta7DNPR1* in the RNAseq data generated from *Pgt* and *Pst* inoculated wheat. TPM: Transcripts Per Million.

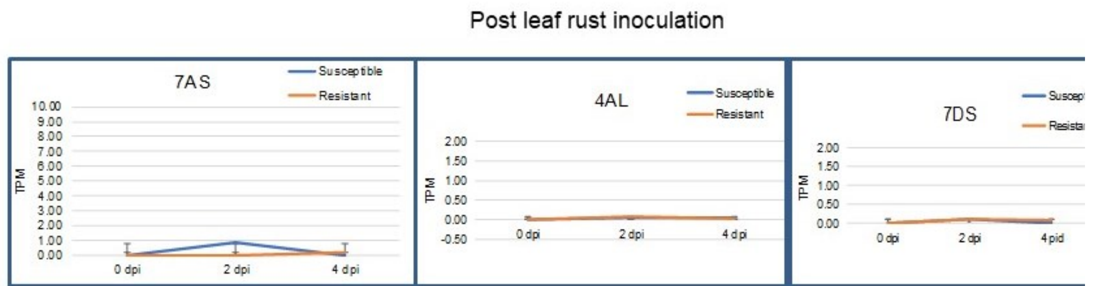


Figure 3.S3: Transcripts of *Ta7ANPR1, Ta4ANPR1* and *Ta7DNPR1* in the RNAseq data generated from *Pst* inoculated wheat. TPM: Transcripts Per Million.

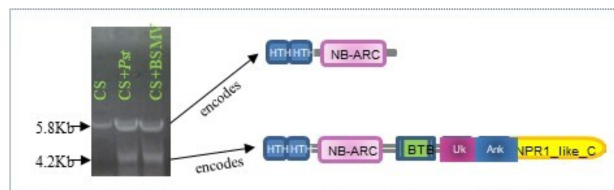


Figure 3.S4: Reverse transcription amplified cDNAs from total RNAs extracted from CS under no-stressed condition, stripe rust inoculated or BSMV inoculated condition. An additional fragment was amplified from CS under stressed conditions.

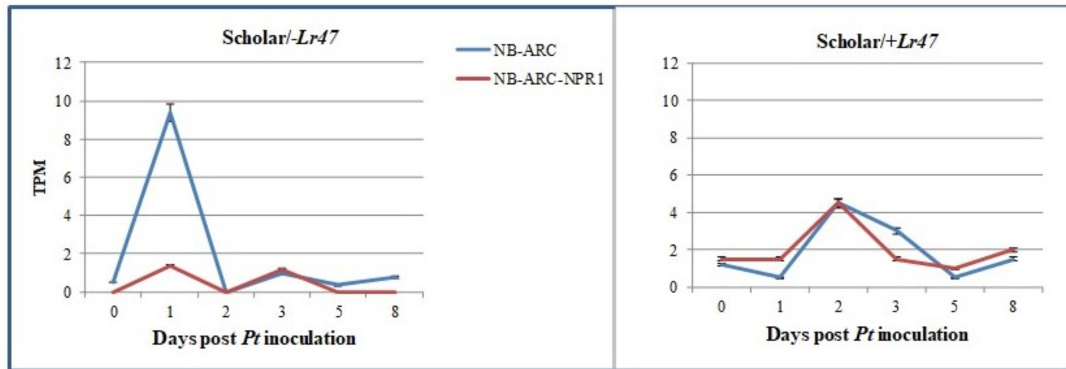


Figure 3.S5: Detection of two transcripts of *Ta7ANPR1* in the RNAseq data generated from *Lr47* near-isogenic lines post leaf rust inoculation at 6 time points. TPM: Transcripts Per Million.

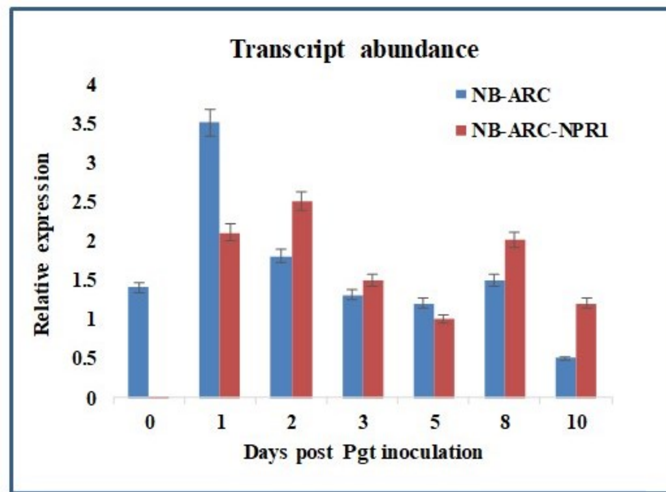


Figure 3.S6: Transcript of two isoforms of Ta7ANPR1 in Cadenza during the time course of Pgt *TPMKC* infection.

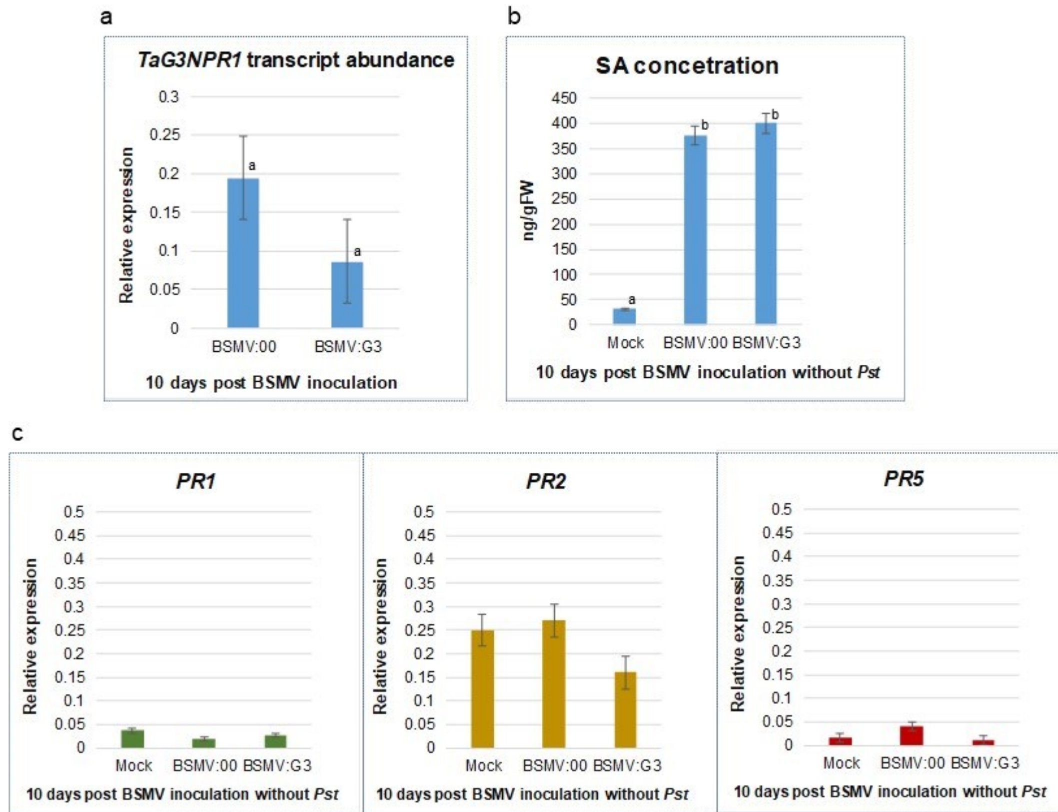


Figure 3.S7: Transcript abundances. **a.** Transcript abundances of six copies of NPR1 on group 3 homeologous chromosomes at 10 days post BSMV inoculations with or without a silencing target. **b.** SA concentrations in the silenced leaf segments and the corresponding segments of controls. **c.** Transcript abundances of three SA-mediated *PR* genes.

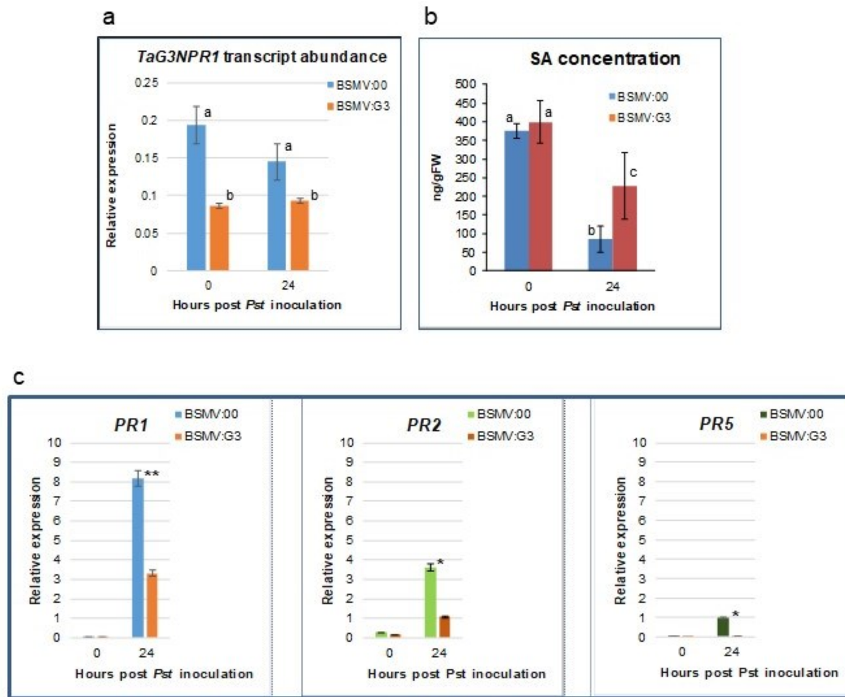


Figure 3.S8: a. Transcript abundances of six copies of NPR1 on group 3 homeologous chromosomes during a time-course study. **b.** SA concentrations in the silenced leaf segments and the corresponding segments of a none-silencing control during the time-course study of post-*Pst* inoculation. **c.** Transcript abundances of three SA-mediated *PR* genes. The 0 dpi is corresponding to the 10 days post BSMV inoculations. The same letters indicate the differences are not significant, and different letters indicate the differences are significant. * indicates the ρ value < 0.5 and ** indicates the ρ value < 0.01.

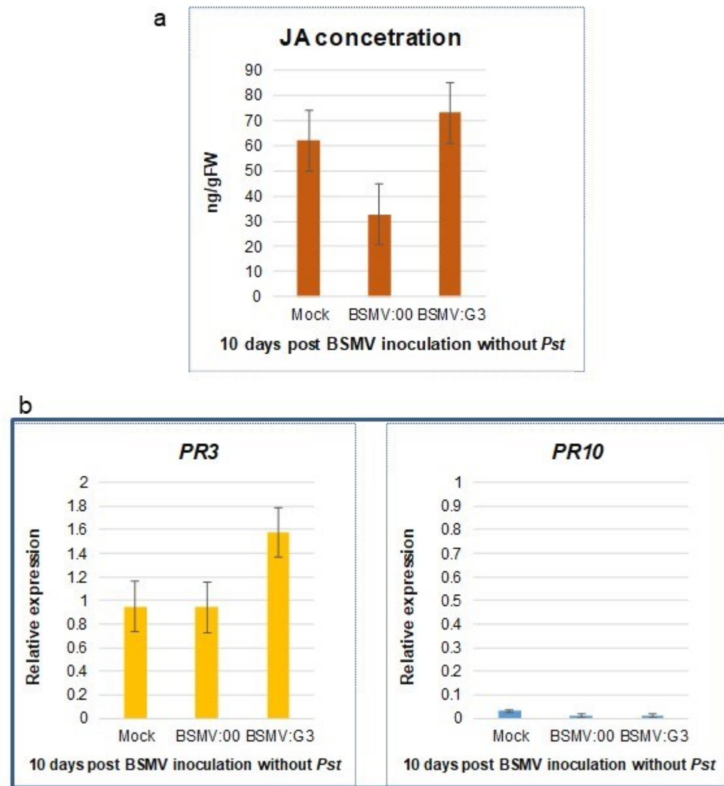


Figure 3.S9: JA concentrations and PR genes expression. a**.** JA concentrations in the leaf segments of six copies of NPR1 on group 3 homeologous chromosomes were silenced and the corresponding segments of controls at 10 days post BSMV inoculations with or without a silencing target. b**.** Transcript abundances of two JA-mediated *PR* genes.

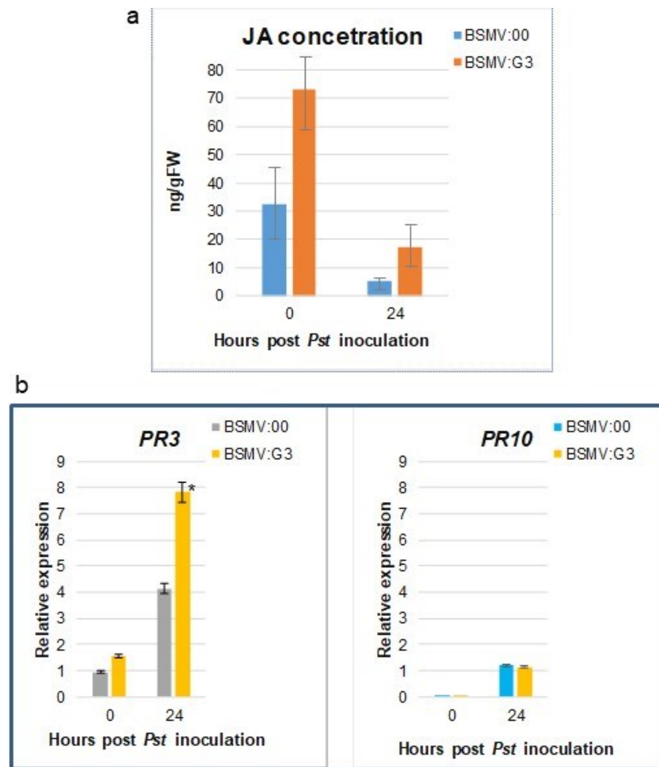
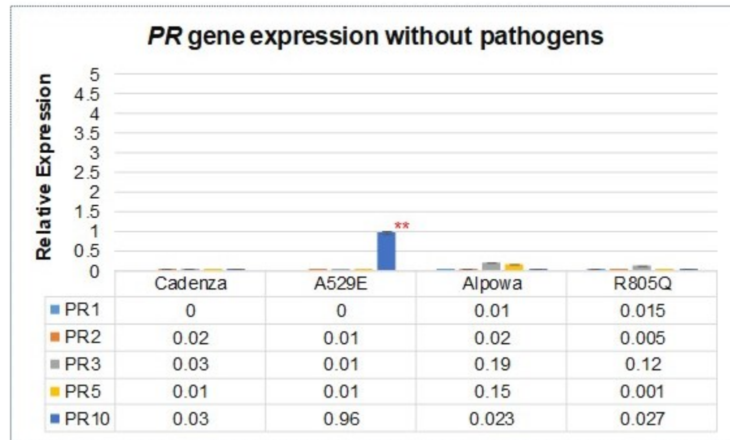


Figure 3.S10: JA concentrations and PR genes expression a. JA concentrations in the silenced leaf segments and the corresponding segments of a none-silencing control during the time-course study of post-*Pst* inoculation. b. Transcript abundances of two JA-mediated *PR* genes. The 0 dpi is corresponding to the 10 days post BSMV inoculations. The same letter indicates the differences are not significant, and different letters indicate the differences are significant. * indicates the ρ value < 0.5.



	Cadenza	A529E	Alpowa	R805Q
PR1	0	0	0.01	0.015
PR2	0.02	0.01	0.02	0.005
PR3	0.03	0.01	0.19	0.12
PR5	0.01	0.01	0.15	0.001
PR10	0.03	0.96	0.023	0.027

Figure 3.S11: Transcript abundances of five *PR* genes in the wild types and mutants without pathogen inoculation. ** indicates the ρ value < 0.01.

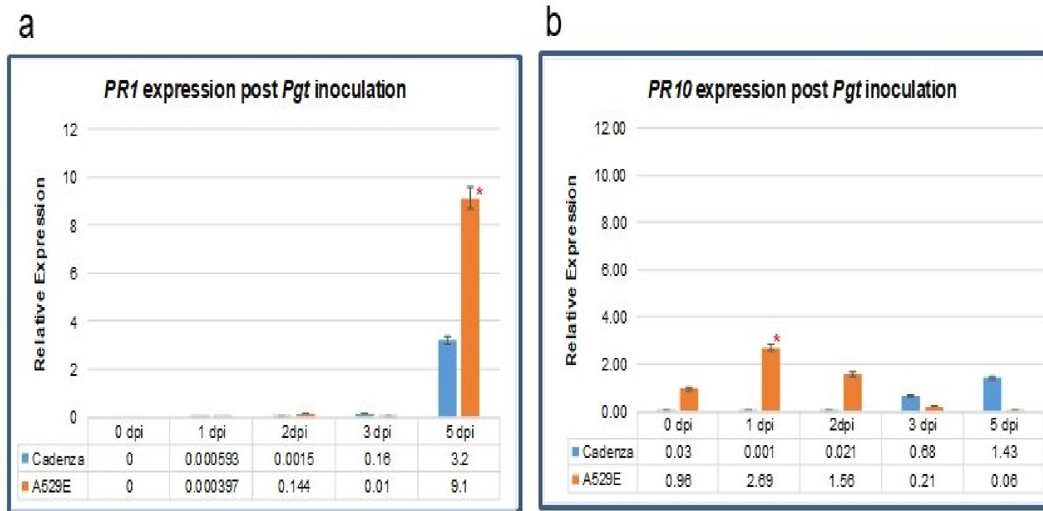


Figure 3.S12: Transcript abundances of two PR1 (a) and PR10 (b) genes during a time-course study of post-Pgt inoculation. * indicates the ρ value < 0.5.

Table 3.S1: Sequences of the Oligos used to construct silencing constructors

Primers	Sequences (5'-3')
G3-oligoF	AGATATTGGAGCAAGCTGAAAGAAGGGATTCACAATTTGTAGAAGCATCGGTTTCTCTTGCATTGGCTGGTGACTGTCTTCGTGGGATGT
G3-oligoR	GAGCTAAGTGTTCCTCCTTCATTTTGAAAGAAGTATCATTTAGATCAACGGTTGTTATCTCCAGAGGTGGATTTGTACCAACATCCCACG
G7-oligoF	TCTTTGACAAGATGGCCGGGTTCTCAGTTCTCAAGTACTTCAAGTTGAGGTTCACGAGTGGTATCCCATGGCTGAAATTTGAGGCGGATG
G7-oligoR	CTCGATGCGGATTAGTTGATGTTGGTCCATTCGGGAAATTGCATTGAAACCCAGCTTCAGCTTCCAGAGATTAGGCATTGCATCCGCCTC
G7A-oligoF	GACAGGTATCTAATAGTCATTGATGACTTGTGGGATGCATCAGCATGGGATATTATTAAATGTGCTTTTCCGAAGGGAAGCCATGGCA
G7A-oligoR	GATACATATGAAAATACAGCAAACATGTCTTCAGATGCTGAGGAAGATTATTGTAGCTAAGGTTCAGTACTTTTCTCGTTTGCCATGGCT

Table 3.S2: Sequences of the primers used for quantitative real-time PCR

Primers	Sequences (5'-3')	Genes measured
G3-RTF	GGCTTTCCTAACAAACAC	NPR1-like on 3A,3B,3D
G3-RTR	AAGTGCGATGTCCAGAAGTT	
G7-RTF	TGACCCCATAGACCTCACTGG	NPR1-like on 7A,4A,7D
G7-RTR	CAAGGCTCCACAGACTGATGA	
G7A-MF	GTTCGTCGTCTATCTCTCATCTTTGG	NPR1-like on 7A,4A,7D
G7A-MR	CCAGAAGTGAAGGAATACACGACAT	
18S-F	GTGACGGGTGACGGAGAATT	18S
18S-R	GACACTAATGCGCCCGGTAT	

Table 3.S3: Genes silenced by each of the constructors

Constructor	Silenced Gene IDs	Chromosome	Functional domain
BSMV:G3	G105400	3A	NPR1-like
	G123800	3B	NPR1-like
	G158300LC	3B	NPR1-like
	G107500	3D	NPR1-like
	G117002LC	3D	NPR1-like
BSMV:G7	G023700LC	7A	NB-ARC
	G703500LC	4A	NB-ARC
	G023000LC	7D	NB-ARC
	G093700LC	6A	NB-ARC
	G099900	6B	NB-ARC
BSMV:7A	G023700LC	7A	Between NB and NPR1