Towards map-based cloning of a Hessian fly resistance gene *H34* in wheat

by

Nida Ghori

B.S., University of Agriculture Faisalabad, 2012
M.S., University of Agriculture Faisalabad, 2014

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Agronomy
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2022

Abstract

Wheat is a staple food crop worldwide and insect damage is a major constraint for its production. Among the insects, Hessian fly (HF, *Mayetiola destructor*) is a destructive pest that significantly reduces wheat grain yield. To date, 37 HF resistant genes have been named, but diagnostic markers for these genes are lacking, which hampers their deployment in wheat breeding. HF resistance gene *H34* on the short arm of chromosome 6B was one of the genes from a U.S. winter wheat Clark. To finely map *H34*, a cross was made between two F_{12} recombinant inbred lines (RIL115-S and RIL118-R) derived from Ning7840 x Clark. RIL118-R carries the resistance allele and RIL118-S carries the susceptibility allele at *H34*. Screening 286 (RIL115-S x RIL118-R) F_3 lines using flanking Kompetitive Allele Specific PCR (KASP) markers identified five heterogenous inbred families (HIFs) segregating at *H34*. The first round of screening of 159 homozygous recombinant plants from five different HIFs using the KASP markers delimited *H34* to a 5.0 Mb interval. Genotype-by-sequencing (GBS) analysis of the four pairs of near-isogenic lines (NILs) from the selected HIFs identified additional SNPs in the *H34* region that further narrowed the *H34* region to 1.3 Mb after screening 75 additional homozygous recombinant NILs. RNA-sequencing (RNA-seq) of the four pairs of NILs identified three differentially expressed genes (DEGs) in the *H34* interval and they were considered as the putative *H34* candidate genes for further study. Using the sequences of the DEGs and GBS-SNPs identified in the *H34* interval, seven KASP markers were designed and validated to be diagnostic in a US winter wheat panel of 203 lines. These markers can be used in gene pyramiding of *H34* with other HF resistance genes using marker-assisted selection (MAS) in the U.S. wheat-breeding programs. Furthermore, studying mechanism of HF resistance in wheat using RNA-seq data revealed that genes encoding defense proteins, stress-regulating transcription factors, and secondary metabolites were strongly up regulated

within the first 48 hours of larval feeding, revealing an early defense in resistant wheat plants in response to larval attack. Also, HF feeding on resistant plants triggered the secretions of R-gene receptors by HF to initiate a hypersensitive response (HR) in the plants. This HR response resulted in production of reactive oxygen species (ROS) to up regulate the downstream genes involved in cell wall fortification and activation of different transcription factors (TFs), which prevents HF to access the nutrients in the resistant plants and eventually results in the death of HF larvae. The new knowledge generated in this study will aid in better understanding of HF-resistant mechanisms and developing new crop improvement strategies to increase HF resistance in wheat.

Towards map-based cloning of a Hessian fly resistance gene *H34* in wheat

by

Nida Ghori

B.S., University of Agriculture Faisalabad, 2012
M.S., University of Agriculture Faisalabad, 2014

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Agronomy
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2022

Approved by:

Co-Major Professor
Dr. Allan K. Fritz

Approved by:

Co-Major Professor (primary)
Dr. Guihua Bai

Copyright

Abstract

Wheat is a staple food crop worldwide and insect damage is a major constraint for its production. Among the insects, Hessian fly (HF, *Mayetiola destructor*) is a destructive pest that significantly reduces wheat grain yield. To date, 37 HF resistant genes have been named, but diagnostic markers for these genes are lacking, which hampers their deployment in wheat breeding. HF resistance gene *H34* on the short arm of chromosome 6B was one of the genes from a U.S. winter wheat Clark. To finely map *H34*, a cross was made between two F_{12} recombinant inbred lines (RIL115-S and RIL118-R) derived from Ning7840 x Clark. RIL118-R carries the resistance allele and RIL118-S carries the susceptibility allele at *H34*. Screening 286 (RIL115-S x RIL118-R) F_3 lines using flanking Kompetitive Allele Specific PCR (KASP) markers identified five heterogenous inbred families (HIFs) segregating at *H34*. The first round of screening of 159 homozygous recombinant plants from five different HIFs using the KASP markers delimited *H34* to a 5.0 Mb interval. Genotype-by-sequencing (GBS) analysis of the four pairs of near-isogenic lines (NILs) from the selected HIFs identified additional SNPs in the *H34* region that further narrowed the *H34* region to 1.3 Mb after screening 75 additional homozygous recombinant NILs. RNA-sequencing (RNA-seq) of the four pairs of NILs identified three differentially expressed genes (DEGs) in the *H34* interval and they were considered as the putative *H34* candidate genes for further study. Using the sequences of the DEGs and GBS-SNPs identified in the *H34* interval, seven KASP markers were designed and validated to be diagnostic in a US winter wheat panel of 203 lines. These markers can be used in gene pyramiding of *H34* with other HF resistance genes using marker-assisted selection (MAS) in the U.S. wheat-breeding programs. Furthermore, studying mechanism of HF resistance in wheat using RNA-seq data revealed that genes encoding defense proteins, stress-regulating transcription factors, and secondary metabolites were strongly up regulated

within the first 48 hours of larval feeding, revealing an early defense in resistant wheat plants in response to larval attack. Also, HF feeding on resistant plants triggered the secretions of R-gene receptors by HF to initiate a hypersensitive response (HR) in the plants. This HR response resulted in production of reactive oxygen species (ROS) to up regulate the downstream genes involved in cell wall fortification and activation of different transcription factors (TFs), which prevents HF to access the nutrients in the resistant plants and eventually results in the death of HF larvae. The new knowledge generated in this study will aid in better understanding of HF-resistant mechanisms and developing new crop improvement strategies to increase HF resistance in wheat.

Table of Contents

List of Figures xi
List of Tables xv
Acknowledgements xvi
Dedication xviii
Chapter 1 - Review of literature 1
1.1 Introduction of wheat 1
1.1.1 Wheat production and market classes in the United States 2
1.1.2 Limiting factors for wheat yield 3
1.2 Hessian fly in wheat 4
1.2.1 Phenotypic symptoms of HF infestation on wheat 5
1.2.2 HF biotypes and host range 6
1.2.3 Highly diversified Avr-effector genes as a source of resistance breakage 7
1.3 Control measures 9
1.3.1 Delayed wheat plantation 9
1.3.2 Destruction of volunteer wheat and crop rotations 9
1.3.3 Application of insecticides 10
1.3.4 Biological control 10
1.3.5 Genetic control 11
1.3.6 Integrated pest management 12
1.4 Mechanisms of HF resistance in wheat 12
1.4.1 Wheat-HF system fits the gene-for-gene model 12
1.4.2 Cellular mechanisms of Hessian fly resistance 14
1.4.3 Reactive oxygen species (ROS) accumulation in compatible interaction 15
1.4.4 Molecular mechanisms of plants defense against Hessian fly 16
1.5 Mapping of HF resistance genes in wheat 17
1.5.1 Physical and genetic mapping of resistance genes in wheat 17
1.5.2 Reported R genes and their sources 19
1.5.3 Fine mapping and gene cloning in wheat 20

1.5.4 RNA-seq: an indispensable and robust tool for profiling differentially expressed genes 21
1.5.5 Development of diagnostic markers 22
1.6 References 24
Chapter 2 - Fine mapping of the Hessian fly resistance gene *H34* 49
2.1 Introduction 49
2.2 Materials and Methods 51
2.2.1 Materials 51
2.2.2 Evaluation of Hessian fly resistance 51
2.2.3 DNA extraction and KASP markers development 52
2.2.4 Recombinant isolation and high-resolution map construction 53
2.2.5 Candidate gene prediction 54
2.2.6 RNA sample preparation and sequencing 54
2.2.7 RNA-seq library preparation and data analysis 55
2.2.8 Cloning of full-length genomic and coding sequences of DEGs 55
2.3 Results 56
2.3.1 Construction of a linkage map and development of KASP markers in *H34* region 56
2.3.2 Fine mapping to delimit candidate interval of *H34* 56
2.3.3 Putative annotated genes underlying *H34* 58
2.3.4 Identification of differentially expressed genes in *H34* interval 59
2.3.5 KASP markers validation in diversity panel 59
2.4 Discussion 61
2.4.1 Fine mapping of *H34* 61
2.4.2 Prediction of *H34* candidate genes 62
2.4.3 Development of tightly linked diagnostic markers for *H34* 64
2.5 References 66
Chapter 3 - Transcriptomics to unravel mechanisms of Hessian fly resistance in wheat 85
3.1 Introduction 85
3.2 Materials and methods 87
3.2.1 Evaluation of HF resistance 87
3.2.2 RNA sample preparation and sequencing 87

3.2.3 RNA-seq library preparation and data analysis 87
3.2.4 Function and pathway analyses of DEGs 87
3.2.5 Validation of RNA-seq data using qRT-PCR 88
3.2.6 Neutral red staining to assess cell wall permeability 88
3.3 Results 89
3.3.1 Differentially expressed genes in response to Hessian fly attack identified by RNA-seq analysis 89
3.3.2 Gene ontology (GO) classification 90
3.3.3 Distribution of DEGs across the wheat genome 90
3.3.4 Biotic stress and defense adaptation related DEGs induced after Hessian fly attack 91
3.3.5 DEGs involved in hormonal homeostasis, photosynthesis, and secondary metabolite production 91
3.3.6 DEGs involved in redox reaction and cell wall integrity 92
3.3.7 Differential expression of transcription factors involved in HF resistance 93
3.3.8 Verification of differentially expressed genes via qRT-PCR 93
3.4 Discussion 94
3.4.1 Fitness cost of defense regulated by phytohormones and their interaction with photosynthesis, and secondary metabolites production 95
3.4.2 Resistant wheat plants strengthen cell wall and use HR-mediated key defense strategy 96
3.4.3 Wheat transcription factors mediated plant defense against HF 97
3.4.5 Conclusion 99
3.5 References 100
Appendix A - Supplementary material chapter 3 119

List of Figures

Figure 1.1. Diagrammatic illustration of effector triggered immunity (ETI) in the wheat-Hessian fly (HF) system. HF injects effector proteins through saliva just below the cell wall of a wheat plant to interfere plant immune signaling; in return, the effector recognition R-proteins produced in a host cell trigger plant resistance response to kill the insect larvae. However, if the effector recognition R-proteins are absent in wheat plants, effector proteins will interfere with host immune signaling to induce plant susceptibility, a phenomenon called effector triggered susceptibility (ETS). 45

Figure 1.2. Chromosomal distribution of different Hessian fly genes (*H1-H37, Hdic*) in wheat. 46

Figure 2.1. Hessian fly (HF)-infested wheat seedlings. (a) A susceptible plant is stunted without newly grown leaves (left) but resistant plant shows normal growth three weeks after infestation; b) HF resistant plant does not carry living larvae; c) HF susceptible plant carries living larvae between leaf sheaths. 72

Figure 2.2. The overall procedure of fine mapping and marker development for *H34*. RILs = recombinant inbred lines, NILs = near isogenic lines, GBS = genotyping by sequencing, QTL= quantitative trait locus. 73

Figure 2.3. A portion of chromosome 6BS linkage map showing the position of *H34* interval and physical position based on IWGSC RefSeq. v2.1. The markers on the right were named by combination of HF (Hessian fly) with the physical position of respective SNPs in IWGSC RefSeq v2.1. 74

Figure 2.4. KASP assay profiles of selected markers on a) selected homozygous and heterozygous recombinants and b) near-isogenic lines (NILs) selected from the HIFs. Blue dots represent the Ning (susceptibility) allele, green dots represent the Clark (resistance) allele, and red dots represent the heterozygotes with both alleles. Light blue dots in the green cluster are for positive control (RIL118-R) and yellow dots in blue cluster are for negative control (RIL115-S) parent. Pink dots represent the genotypes with poor PCR amplification and black dots represent the water control. 75

Figure 2.5. Chromosome walking of *H34* using recombinants identified from heterogeneous inbred families (HIFs). a) Recombinants identified in the *H34* region from the first round of screening using nine new markers and NILs selected from different HIFs; b) Recombinants

in the *H34* region identified from the second round of screening progeny from HIF-58 using nine new markers identified by genotype-by-sequencing (GBS). Black and grey bars represent chromosome segments from Clark and Ning 7840 respectively. R and S represent resistant and susceptible phenotypes of each line, respectively. The four markers highlighted in red color are the flanking markers. 76

Figure 2.6. High-confidence (HC) genes and differentially expressed genes (DEGs) in the 1.3 Mb *H34* interval on chromosome arm 6BS based on IWGSC RefSeq v2.1 (IWGSC, 2018). (a) The genetic (left) and physical (right, indicated by the number after HF) map locations of the markers in the recombinant lines (RIL) selected from different heterogeneous inbred families (HIFs). Flanking markers are highlighted in yellow color. (b) Gene identity (IDs) of HC genes and DEGs (highlighted in green) in the *H34* interval as annotated in IWGSC RefSeq v2.1. 77

Figure 2.7. Sequence alignment of DEGs in *H34* interval of RIL115-S and RIL118-R against CS reference genome IWGSC RefSeq v2.1 to identify the SNPs for development of diagnostic markers. (a-c) KASP primers sequences and locations for gene TraesCS6B03G0019400, TraesCS6B03G0019500 and TraesCS6B03G0022700, respectively. In each gene structure black boxes represent 3' and 5' UTR regions, green boxes represent exons and black lines represent introns. Red boxes showed the SNP locations in the DEGs. Red arrows labeled the primer sequences of KASP marker in the DEGs. 78

Figure 2.8. Allele separation of KASP markers in the diversity panel of AM203. (a) Allele separation of KASP marker HF6028581. The green dots represent the accessions with the Clark allele, the blue dots represent the accessions with the Ning7840 allele, the yellow dot is Clark (control), the light blue dot is Ning7840 (control) and the black dots are NTC (Non-template control). (b) HF resistance score (%) of the accessions from the 203 U.S. diversity panel. Grey bar (dots) represents the accessions carrying Clark (R) marker allele and white boxes represent the accessions carrying Ning7840 (S) allele. 79

Figure 3.1. Responses of resistant and susceptible NILs to HF infestation two weeks after attack by HF larvae. A) Hessian fly-infested wheat seedlings showing normal growth in resistant NIL-R (right), but stunted plant in the susceptible NIL-S (left) (b) Living larvae lived on second leaf, close to the root crown of the susceptible NIL-S 12-h after initial attack by HF

larvae. (c) Resistant line with no living larvae on NIL-R 12-h after initial attack by HF larvae. 107
Figure 3.2. (a) Number of significantly (≥1 or ≤−1-fold change, $p < 0.05$) differentially expressed genes (DEGs) in the four pairs of resistant NIL-R compared to their susceptible NIL-S at 12-h (blue circle) and 48-h (orange circle) after HF larval attack. B) Bar chart comparison of total numbers of DEGs in NIL-R at 12-h and 48-h after infestation. The numbers of DEGs are marked on the top of each bar. 108
Figure 3.3. Gene ontology (GO) distribution of differentially expressed genes (DEGs) in the four pair of resistant NIL-R identified at 12-h and 48-h after larval attack on wheat. Blue bars represent DEGs at the 12-h and orange bars represent DEGs after the 48-h treatment..... 109
Figure 3.4. Chromosome distribution of the differentially expressed genes (DEGs) in the four pair of resistant NIL-R at 12-h and 48-h after larval attack on wheat. Blue bars represent DEGs after 12 hours and orange bars represent DEGs after 48 hours of Hessian fly larval attack. 110
Figure 3.5. Pathway analysis highlighted significant differentially expressed genes (DEGs) ($p < 0.05$) involved in resistance against Hessian fly in the four pairs of resistant NILs at (a) 12-h and (b) 48-h after larval attack. Blue and red colors represent down and up-regulated DEGs, respectively. 111
Figure 3.6. Differential expressed genes (DEGs) involving in hormonal homeostasis, secondary metabolite production and photosynthesis after 12-h and 48-h of HF larval attack in the four pairs of NIL-R. The DEGs are grouped within a heatmap based on their associated functions. Scales indicate log2-fold changes (≤−1 or ≥1). Gene identity (IDs) were named using their number only after removal of the prefix “TraesCS”. 112
Figure 3.7. Expression profile of wheat differentially expressed R-genes and differentially expressed genes (DEGs) involved in redox reaction, defense signaling and cell wall strengthening 12-h and 48-h after HF infestation in the four pairs of resistant NIL-R. The DEGs are grouped within a heatmap based on their associated function. Red and blue represent up-regulated and down-regulated DEGs, respectively. Scales indicating log2 fold change (≤−1 or ≥1) shown on right side of the figure. Gene identification number (IDs) was written with their number only and prefix “TraesCS” is removed to make them short. ROS; Reactive oxygen species. 113

Figure 3.8. Neutral red stained wheat plant represents change in plant cell permeability in the wheat seedling at the bottom of second leaf right above the root crown harboring the HF larvae at 12-h and 48-h of larval attack. Left panel represent a susceptible (NIL-S) NIL-1; 115/118_F_3_43 and right panel is a resistant NIL-1; 115/118_F_3_43 (NIL-R). (a) NIL-S before staining; living larvae can be seen clearly (b) NIL-R before staining with no visible larvae. (c-d) Uninfested NIL-S and NIL-R punctured with a minute pin (diameter 200 µm) immediately before staining and taken as negative control (e-f) NIL-S and NIL-R after 12-h of HF infestation. (g-h) NIL-S and NIL-R after 48-h of HF infestation..... 114
Figure 3.9. Differential expression of transcription factors in the four pairs of resistant NILs (NIL-R) attached by HF. Scales indicate log2-fold change (≤−1 or ≥1). shown on right side of the figure. Gene identification number (IDs) was written with their number only and prefix “TraesCS” is removed to make them short. 115
Figure 3.10. Verification of differentially expressed genes (DEGs) by qRT-PCR using NIL-1; 115/118_F_3_43 (a) TraesCS4B02G086600 (b) TraesCS1B02G245800 and (c) TraesCS4B02G070300 were verified. The x-axis in each chart representing the three time points 0-h: before attack of larvae; 12-h: 12 hours after attack of HF larvae; 48-h: 48 hours after attack of HF larvae and the y-axis representing log fold change value of expression determined by qRT-PCR. Purple bars represent susceptible NIL-1; 115/118_F_3_43 and green bars represent resistant NIL-1; 115/118_F_3_43. Asterisks representing significant differences between each comparison using a two-tailed Student's t-test. ns; non-significant difference between each comparison. 116
Figure 3.11. General figure showing mechanism of Hessian fly resistance in wheat. HSPs; Heat Shock proteins..... 117

List of Tables

Table 1.1. Hessian fly genes sources, their location and marker types in wheat genome 47

Table 2.1. Marker haplotypes of the 14 recombinants selected from F_3 for heterogenous inbred families (HIFs) development 80

Table 2.2. A list of KASP primers developed from 90K SNP arrays, RNA-sequencing and genotyping-by-sequencing of near-isogenic lines contrasting in *H34* alleles..... 81

Table 2.3. Annotated genes present within the finely-mapped *H34* interval based on IWGSC RefSeq v2.1..... 83

Table 2.4. A list of gene-specific primers used to amplify full-length sequence of deferentially expressed genes..... 84

Table 3.1. Primers used in this study for verification of the deferentially expressed genes by qRT-PCR..... 118

Acknowledgements

A small-town resilient girl, who came to U.S.A to chase her dreams.

I have no words to describe my gratitude to my supervisor Dr. Guihua Bai who instilled in me with skills for being a better scientist, taught me how to do quality research, and conducting collaborative work in research. I am thankful to him for giving me an opportunity to work in USDA Central Small Grain Genotyping Lab and equipping me with the skill set that I am going to carry with me in future career. I would like to extend my thanks to my co-major advisor Dr. Allan Fritz and committee members Drs. Ming-Shun Chen, Guorong Zhang, Harold Trick and Jesse Poland for their valuable suggestions and guidance in my research, and time to review my dissertation.

For the fine mapping work, I cannot thank enough to Dr. Ming-Shun Chen for helping me to prepare soil and training me for Hessian fly phenotyping in greenhouse. I would also like to extend my thanks to my lab mate Dr. Yunfeng Xu to help me in genotyping and data analysis, Dr. Hui Chen and Lanfei Zhao for teaching me lab techniques, Drs. Amy Bernardo and Paul St. Amand for assisting me with GBS library preparation and SNP data analysis, and Dr. Lanfei Zhao and my friend Sumera Asghar and Rehan Saeed from China Agriculture University for helping me in RNA-seq data analyses.

I really appreciate the support from all former and present lab members in Dr. Bai's lab, especially Drs. Umara Sahar Rana, Abdulrehman Hashimi, Yaoguang Li, Lirong Zhang, and graduate students and friends Mr. Yuzhou Xu, Ms Ruolin Bian, Ms. Xiaoting Xu, Mr. Volodymyr Kavetskyi for their valuable suggestions in my research and making me feel like home in Manhattan. Without them, I would never be able to balance my life and work in Manhattan during these years.

I am thankful to my sponsors and funding agencies for their generous financial support to complete my Ph.D. degree, in particular the Foreign Student Scholarship from the Fulbright to cover my tuition and part of living expense, and the funding from the National Research Initiative Competitive Grants 2017-67007-25939 and 2022-68013-36439 from the USDA National Institute of Food and Agriculture to support part of my research work. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture. The USDA is an equal opportunity provider and employer.

This work was not possible without the encouragement and motivation of my dear dad Mr. Mukhtar Ahmad Ghouri and mom Mrs. Razia Mukhtar who always believe in me and give me wings to chase my dreams. I have no words to describe their love and affection they have given to me throughout my life. A big fat thanks to my brothers Awais Ghouri, Hammad Ghouri and sister Farah Ghouri who were miles away, but they were here with me all the time when I needed their emotional support and for cheering me up during this hard journey. So many names outside the lab in Manhattan who remained available and provided me a homely environment.

Lastly, I am highly indebted to say thanks to ALMIGHTY Allah for always showering his countless blessings on me and helping me with ease throughout my life.

Dedication

To mom and dad

who always picked me up on time,

and encouraged me to purse my dreams

against all the odds of our society

Chapter 1 - Review of literature

1.1 Introduction of wheat

Wheat (*Triticum aestivum* L.) is one of the staple food crops worldwide. It provides major energy requirement to the world's population. Wheat ranks the third most important cereal crop next to only maize (Zea mays L.) and rice in annual production (Graybosch & Peterson, 2010). One-fifth of the total calories of the world's population comes from wheat (FAO, 2010), making wheat an important component of food security at the global level.

Among different types of wheat planted in the worldwide, bread wheat (*Triticum aestivum*, 2n = 6x = 42, genome formula AABBDD) occupies about 95% of the wheat area and is used for bread, noodles, pastry, cookies, and other related products; the remaining 5% is planted with durum wheat (*Triticum turgidum* ssp. *durum*, 2n = 4x = 28, genome formula AABB) mainly used for pasta (Dubcovsky & Dvorak, 2007).

Based on ploidy level, wheat can also be divided into diploid (*Triticum monococcum*), tetraploid (*Triticum dicoccum*) and hexaploid bread wheat (*Triticum aestivum*). The hexaploid wheat was developed from two independent events of hybridization. The allopolyploidization merged the A and B genomes to form tetraploid wheat *Triticum turgidum* spp. *dicoccoides* (2n =28, AABB). Domestication of *Triticum dicocoides* and its rapid expansion in Europe and Asia lead to a second independent hybridization event with *Aegilops tauschii* to form the hexaploid wheat *Triticum aestivum* (2n = 6x = 42, AABBDD) (Sorrells & Yu, 2009). Hexaploid (bread) wheat has a large genome size of about 16 Gb and was predicted to have at least 107,891 high-confidence genes (Hussain and Rivandi, 2007, IWGSC 2018).

1.1.1 Wheat production and market classes in the United States

The United States (U.S.) produced over 49.7 million metric tons of wheat in 2021, putting the country in fifth place in terms of global wheat production. Within the same period, the United States was the third-largest exporter of wheat grain, flour, and wheat products after Russia and the European Union (Statista 2021, https://www.statista.com/topics/6034/wheat-in-the-us/). Bread wheat can also be classified into four major classes: hard red winter (HRW), hard red spring (HRS), soft red winter (SRW), and white wheat. Each class has different requirements of end-use quality and the region-specific growing conditions. Hard red winter wheat is mainly grown in the Great Plains area from Montana to Texas and is primarily used to make bread and related products such as hard rolls, flat-bread, and croissants. Hard red spring wheat is mostly cultivated in the Northern Great Plains area and is mainly used for protein blending purposes to be blended with other wheat classes to improve the flour strength due to its high protein and gluten contents. Soft red winter wheat is the third largest class of wheat in the U.S., and it is mostly grown in the eastern part of the U.S. Commonly it is used for specialty products such as crackers, cookies and other confectionary products. While the white wheat class is primarily grown in the Pacific Northwest. This class delivers excellent milling results and is often mixed with HRS wheat for improving bread color and quality. In addition to bread wheat, the tetraploid durum wheat grown in North Dakota and Montana is known for their excellent qualities for producing pasta (U.S. Wheat Associates, https://www.uswheat.org/working-with-buyers/wheat-classes/).

Most states in the U.S. produce wheat with Kansas as the largest wheat producing state with total production of 364 million bushels followed by North Dakota and Oklahoma with 196 and 115 million bushels produced, respectively, in the year 2021 (FAO, 2022, Statista, 2022).

1.1.2 Limiting factors for wheat yield

It has been estimated that world wheat production would require to be increased from its current production level of 642 million tons to around 840 million tons by 2050 because of the quick population growth rate along with the increase in per capita wheat consumption. To meet this target, it is dire need to increase the wheat yield by 77% through genetic, agronomic, and physiological interventions (FAO, 2010). The yield potential refers to maximum yield that can be produced by a cultivar in the presence of optimal environmental conditions with the best management practices. This potential starts decreasing by the sub-field scale variability in growth conditions and by the onset of biotic and abiotic stresses during the growing seasons. In the southern Great Plains with rainfed agriculture system, maximum yield potential cannot be achieved due to various stresses such as drought, pathogens, and insects (Connor et al., 2011; Evans & Fischer, 1999). Improving wheat adaptation to biotic and abiotic stresses along with understanding of plant stress responses is important for wheat genetic improvement (Zhang et al., 2011).

Biotic stresses including diseases and insect pests drastically affect wheat yield. Viruses and fungi are mainly responsible for wheat diseases. Hollandbeck et al. (2018) estimated the average yield losses of 10.7% in the past five years due to diseases including stripe rust, leaf rust, wheat streak mosaic virus, barley yellow dwarf, Fusarium head blight (FHB), powdery mildew, tan spot, and bunt. Moreover, insect pests also pose a serious threat in attaining the desired yield. However, the effects of insects on yield losses have been neglected in recently reported studies. The yield loss due to insect pests was up to 5.1% in pre-green revolution era, but it has been increased up to 9.3% after green revolution (Dhaliwal et al. 2010). Hessian fly (*Mayetiola destructor* Say) is among one of the major destructive pests and it alone causes 5 to 6% annual

yield losses to U.S. farmers (David Buntin, 1999; UNO, 2019). In special cases, yield losses up to 35% were ascribed in semiarid Morocco and the south of the Iberian Peninsula (Amri et al., 1990). To ensure food security, there is an urgent need to breed wheat varieties with not only high yield potential, but also resistance to different pests.

1.2 Hessian fly in wheat

Hessian fly is one of the most invasive pests of wheat in Southwest Asia, New Zealand, Europe, North Africa, and North America (Harris et al., 2003). It belongs to the family Cecidomyiidae, order: Diptera and is known to be a member of a large group of insects called gall midges because of its gall forming characteristics. Unlike other gall midges which produce outgrowth galls in plants, HF converts the whole plant into gall and inhibits the plant growth by formation of nutritive cells at the feeding sites which is just above the base of wheat seedling. Inhibition of plant growth results in death of susceptible plants. Also, it is an important pest for genetic studies due to its evolutionary position, relationship with host plant and an unusual chromosome number including polytene chromosomes. Moreover, its short life cycle of almost 30 days and small genome of 158 Mb (Harris et al., 2015; Johnston et al., 2004) are its additional features which make it a primary pest for genetic studies. HF is an insect sexually reproduced and female fly can lay approximately 100 to 400 eggs on the adaxial part of the wheat leaves which are hatched at 20 ^{o}C within 3-4 days.

Like other flies, HF has a life cycle of four stages: egg, larvae, pupae, and adult stage. After hatching, newly hatched larvae crawl down towards the base of the young leaves to establish their feeding sites by formation of gall (Anderson & Harris, 2006). First instar larvae are red in color and then turn white as second instar larvae within few days. Mature larvae are

hardened and change into dark brown color puparium called flex seed because of its resemblance to seed of flex plants. Moreover, there is no need for continuous breeding of HFs for maintaining the insect because its larvae can be stored at 4 ^{o}C for up to one year (Sardesai et al., 2005).

1.2.1 Phenotypic symptoms of HF infestation on wheat

HF infestation not only reduces grain yield but also grain quality. Its larvae cause irreversible and massive morphological changes when they infest wheat plants at seedling and stem elongation stages. Wheat may only reduce seed count and weight if infestation occurs at a late growth stage, but plants will die if the infestation occurs at one-leaf stage (Buntin & Chapin, 1990; Byers & Gallun, 1972).

Two types of interactions exist between wheat and HF: compatible interaction and incompatible interaction. In the compatible interaction at the seedling stage, HF larvae rupture the epidermal and mesophyll cells near the feeding sites (gal) by converting them into nutritive cells with large nuclei and thin cell walls which facilitate insect mandibles to get accesses to plant cell sap, which results in stunted plant growth and development of dark green leaves and shortened internodes in infested primary tillers that may lead to the death of young seedlings. If plants are attacked at stem elongation stage, similar symptoms can be observed along with lodging, and fewer seeds with lower seed weight produced (Harris et al., 2003). In the incompatible HF-wheat interaction, HF larvae usually die after 5 days of infestation with no larval growth and white instar can be seen on the basal part of the young seedlings. It has been reported previously that rapid mobilization of resources such as lipids and other molecules strengthens the cell wall of the plants, which prevents the HF mandibles to access the stem tissues, so larvae are unable to induce the susceptibile symptoms such as formation of sink

(nutritive tissue) and stunted growth. Moreover, the resistant plants remain normal with light green leaves and tiller development and normal internodal distance between leaves (Harris et al., 2010).

1.2.2 HF biotypes and host range

To date, 16 HF biotypes have been reported based on their reactions to a differential set of wheat cultivars carrying different HF resistance genes. Those biotypes are named from A to O, Great Plains (GP) and biotype *vH9* and *vH13* (virulent to *H9* and *H13*). In 1977, responses of female HF biotypes were evaluated by infesting four wheat varieties containing *H3*, *H5*, *H6*, and *H7/H8* resistant genes, respectively (Gallun & Patterson, 1977; Ratcliffe et al., 2000). Among them, GP biotype is avirulent to all the four genes while biotype L is virulent on all the four cultivars. Since then, 37 genes have been reported and it is impossible to test the responses of all HF resistance genes to each of the biotypes because there could be 237 biotypes of HFs in response to 37 known HF genes (Cambron et al., 2010; Chen, 2008). Also, the new biotypes are unceasingly emerging in response to new *H* genes that are deployed in field (Anderson & Harris, 2019).

Although wheat is preferred host, HF also infests other small grain crops including barley, rye, triticale, brachypodium, oat and 17 other grass species in tribe Triticeae such as *Aegilops* and *Agropyron* as alternate hosts. Rye and barley are the preferred hosts of HF following wheat. Several grass species from Bromeae tribe have been reported as potential hosts of HF in New Zealand (Chen, 2008; Stewart, 1992). After mating, female HFs used both chemical and physical cues from the host plants to find their preferred host with perfect oviposition sites (Harris et al., 1993; Painter, 1951). They prefer to lay eggs on adaxial, rather

than the abaxial, side of wheat leaves because of presence of parallel grooves and ridges on adaxial leaf surfaces (Kanno & Harris, 2000).

1.2.3 Highly diversified Avr-effector genes as a source of resistance breakage

New HF biotypes are rapidly emerging in responses to new H genes evolving in host plants. These H genes may recognize specific effectors in pests to induce the effector triggered immunity in plants (Jones & Dangl, 2006). Among hundreds of effector genes identified from the insect pests, only a few are Avr-encoded effectors (Bos et al., 2010; Zhao et al., 2015b). The effector genes have N-terminal signal peptides (SPs) and rapidly evolving sequences in conserved coding regions. These characteristics were previously detected in salivary gland transcripts encoding small secreted salivary gland proteins (SSSGPs) (Kämper et al., 2006; Spanu et al., 2010). In most of the protein coding genes, the coding regions are conserved while the non-coding regions (introns) are highly variable because of selective constraints. However, the reverse pattern has been noticed in HF where coding regions have higher sequence variation compared to non-coding regions including introns and untranslated regions (Chen et al., 2010). 2010). Chen et. al (2010) surveyed the SSSGP-1 gene family to identify the unique structure of this effector gene family, and found that among seven genes in the family, six are tandemly repeated genes, one (SSSGP-1A2) is inverted repeat of SSSGP-1A1 gene. All these genes reside within a 15 kb genomic region and contain a conserved promoter region, a 5'-untranslated region (UTR), a single peptide coding region (SPCR), an intron, a mature protein coding region (MPCR) and 3'-UTR. Most of SSSGP gene families carry an intron except one group with no intron and a few with multiple introns. These introns are present within SPCR or between SPCR and MPCR. Like SSSGP-1 group, members in all these groups are clustered within small regions in HF genome (Chen et al., 2010). Moreover, protein alignment combined with structural

analyses of SSSGP-71, the largest effector gene family, showed that their mature proteins contain a cyclin like F-Box domain in the N-terminus and a series of 13 leucine rich repeats (LRRs) and N-terminal signal peptides (SPs) (Zhao et al., 2015b).

Hessian fly effector genes are continuously evolving to give the pest an advantage over its host. However, the effectors, on the other hand, also expose the insect to plant defense mechanisms. The continuous emergence of virulent biotypes in response to selection pressure from resistance genes deployed in wheat varieties is a great threat to the wheat production in the United States. Hessian fly effectors are exceptionally diversified in their coding regions (MPCRs), and rapid diversification in mature SSSGPs suggests that the genes are under selection pressure for functional adaptation and allows the insect to counter changes in host plants for virulence (Chen et al., 2010; Lehmann, 2006).

Although strong positive selection is the main reason for rapid emergence of new biotypes, the high homology between other conserved regions of the genes leads to unequal recombination between the gene family members, particularly those arranged in tandem arrays. Moreover, homogenization is not specific to a specific domain in the genes but occurs throughout the gene and in the inter-genic regions as well. Furthermore, unequal crossovers led to homogenization of the whole array while smaller gene conversions possibly homogenize small regions (Hilliker et al., 1994). For gene conversion, sequence heterogeneity within MPCRs could change the recombination hotspots and in turn can affect the length of conversion tracks or crossover events (Jeffreys & May 2004). Therefore, in plant-insect interaction, successful insect pest biotypes require many genes that encode effector proteins along with ability to change themselves for functional adaptation in response to shift in host plant population.

1.3 Control measures

1.3.1 Delayed wheat plantation

Delayed wheat planting can reduce HF damage. Planting on fly-free dates when the HF larvae died before the emergence of winter wheat is one of the most effective approaches to avoid its infestation. However, this strategy is not effective in Southern United States where HF larvae emerge in the entire winter season (Horton et al., 1945). Moreover, owing to climate change and annual fluctuation in temperatures and rainfall patterns, planting date may vary every year. Also, delayed planting may result in forage losses for dual purpose wheat, increase in winter cold injury and reduction in grain yield (Buntin et al., 1992).

1.3.2 Destruction of volunteer wheat and crop rotations

Volunteer wheat is a major source of HF for initial infestation in wheat because it emerges early as early season HF populations (David G. Buntin & Hudson, 1991; Parks, 1917). Destruction of volunteer wheat has been considered as an effective control measure since early 1900s. HF can move from volunteer wheat to regular wheat after regular wheat plants emergence, which reduces the benefit of planting at or after fly-free date.

The previous year wheat stubbles can also be a source HF for infestation in wheat because it can act as a bridge between two seasons. Therefore, continuous wheat plantation should be avoided to break the chain of HF infestation, and crop rotation can help in killing HF third instar or larvae that are dormant in stubbles. Those third instar or larvae are potential source of infestation in next wheat season. In addition, major hosts of HF are only from tribe Triticeae,

therefore, crops from non-Triticeae tribe can be rotated with wheat to break the chain infestation from one season to another (Whitworth et al., 2009).

1.3.3 Application of insecticides

Systematic seed treatment before germination and foliar application of insecticides at seedling stage are the two main methods to control HF infestation. Seed treatment is one type of impermanent control over a period of approximately 30 days and remains effective only if there is one fall generation of HF in the fields, but not effective against subsequent generations. This method could be beneficial when other management practices such as delayed planting and resistant cultivars are not available because of its high cost/benefit ratio (Wilde et al., 2001).

Foliar insecticides are applied at the seedling stage of 2 -3 leaves to control adults and new larvae before they reach the leaf sheaths and became established in the stems where insecticides may not be able to reach. Moreover, multiple applications are required as multiple HF broods come out throughout the wheat growing season. Although, foliar application on crop may provide some HF control, limited window of effectiveness, narrow range of application timing, economic threshold and associated cost of several applications make the control method one of the least used management options (Buntin, 2007; VanDuyn et al., 2003).

1.3.4 Biological control

Several species of HF enemies have been reported to be effective in fields. Most of the parasitoids belong to superfamily Chalcidocidae and phylum Hymenopteran. These parasitoids mostly affect puparia of spring generation of HF and few of them attack at the egg stage of fall generation (Gahan, 1933). Although 41 species of HF parasitoids are described in North America

and Europe, only three, *P. hiemalis*, *H. destructor* and *E. allynii*, are widespread and considered lethal to HF in laboratory experiments. Unfortunately, none of them were effective in fields (Hill, 1933; Schuster & Lidell, 1990).

1.3.5 Genetic control

Although many pest management strategies can be deployed to reduce HF population, use of host plant resistance can provide economic and eco-friendly approaches for HF control (Harris et al., 2015; Ratcliffe et al., 2000). Many HF genes have been identified on different chromosomes and cultivars carrying those genes have been widely planted over large areas for many years (Zhao et al., 2020). But these genes may lose their effectiveness after 6 to 8 years of deployment due to high genetic variations in HF virulence genes that leads to high heterogeneity in different HF populations (Chen et al., 2010, 2014; Ratcliffe et al., 2000; Zhao et al., 2015). As a result, a wheat cultivar resistant to a HF biotype may quickly loss their resistance after planting for several years. Moreover, temperature changes can significantly reduce the effectiveness of some resistance genes. Chen et al. (2014) documented that when cultivars carrying *H13* are exposed to 40 ^{o}C for few hours, they become susceptible to an avirulent HF biotype. Also, resistance genes *H3*, *H5*, *H10*, *H11*, *H12* and *H18* lost their resistance above certain temperatures (Buntin et al., 1992; Cambron et al., 2010; Sosa & Foster, 1976; Tyler & Hatchett, 1983). Plant resistance is also affected by the order of wheat infestation by avirulent and virulent HF genotypes because infestation by a virulent genotype may develop the systematic susceptibility which provides refuge for later infested avirulent larvae, resulting in the survival of larvae of both HF biotypes.

To mitigate the loss of cultivar resistance, cultivar rotation with different sources of resistance genes should be practiced (Gould 1986; Tooker 2012). Area-wide crop losses can also be avoided by monitoring the virulent HF biotypes. When a virulent HF biotype starts increasing, it is time to deploy new R genes that are resistant to the emerging HF biotypes (Chen, Liu, et al., 2009; Stuart et al., 2012). Furthermore, the levels and durability of HF resistance can be improved by stacking multiple R-genes in a single cultivar to protect wheat against HF (Bassi et al., 2019a; Stuart et al., 2012).

1.3.6 Integrated pest management

Although many HF control measures can be used individually to reduce HF infestation, it can be more effective when they are used in a combination. For instance, where highly resistant cultivars are not available, delayed planting coupling with systematic seed treatment can be an option to minimize HF damage. Moreover, with moderately resistant cultivars, parasitoids can reduce HF populations in a field (Buntin, 2007; Chen et al., 1991). Hence, a wheat producer can adopt a combination of several appropriate control measures to manage HF population with limited inputs based on the field conditions and geographical regions.

1.4 Mechanisms of HF resistance in wheat

1.4.1 Wheat-HF system fits the gene-for-gene model

Wheat and HF interaction fits gene-for-gene model as proposed by Dr. Flor (Flor, 1971). In this model, when a resistance (*R*) gene in plant recognizes the intrusion of a specific pathogen or insect pest, a complementary gene from the pathogen or insect interacts with the resistance gene in plants to trigger the disease resistance in the plants (Flor, 1971). During the past several

years, several studies have intended to validate the gene-for-gene interaction between plants and insects' effectors. However, lack of reference genome in insects makes it difficult to establish an effective system to study such interaction between plant and insects (Bent & Mackey, 2007; Jones & Dangl, 2006).

So far, only few insect resistance genes have been cloned in plants. Among them, the potato Mi resistance (R) gene against aphid and whitefly showed a gene-for-gene interaction. Also, Asian rice gall midge interaction is also followed the gene-for-gene model in rice. HF shares the common feeding mechanisms with Asian rice gall midges, and both share unconventional conservation pattern in effector-coding genes (Behura et al., 2000; Bentur et al., 2003; Gould, 1986; Rossi et al., 1998), suggesting wheat and HF system most likely fits a gene-for-gene model (Byers and Gallun 1970) (Figure 1.1). To further investigate the hypothesis, genetic analyses have been performed to determine if avirulence in HF is attributed to the effector-coding avirulence (*Avr*) genes and found that the virulence to R genes in the coordinated gene pair (*H7H8*) and *H3* is simply inherited and controlled by independent recessive genetic factors (effectors) (Byers & Gallun, 1972). Gallun (1977) extended the gene-for-gene association to two new genes *H5* and *H6* after obtaining evidence of the first X-linked HF *Avr* gene. High resolution mapping and finger printing contig (FPC) based physical mapping of six *Avr* genes (*vH5, vH6, vH9, vH13, vH24,* and *vHdic*) on HF polytene chromosomes provided further evidence in support of the hypothesis.

Later, substantial studies on specific small secreted salivary gland protein (SSSGP-1) gene family have demonstrated that gall midges injected SSSGPs into wheat and SSSGPs act as effectors to suppress plant defense and reprogram physiological pathways in the attacked plants (Byers & Gallun, 1972; Grant et al., 2006; Tjallingii, 2006; Weech et al., 2008). Furthermore,

Zhao et al., (2015) discovered the SSSGP-71 gene family, the largest gene family identified to date, which resembles ubiquitin E3 ligase effectors in plant pathogenic bacteria. These effectors interact with wheat R-genes, and mutations in the SSGP-71 genes render the effector triggered immunity in wheat genotypes that carry resistance genes *H6* or *H9*. Thus, the way HF feeds on wheat, presence of putative effector-coding genes in its salivary glands, its ability to modulate gene expressions and wheat plant's responses to HF attack through cultivar specific NB-LRR R genes all support that wheat-HF interaction fits the gene-for-gene model as proposed by Hatcher and Gallun (1977).

1.4.2 Cellular mechanisms of Hessian fly resistance

Besides R-gene-mediated defense, other resistance mechanisms include synthesis of toxic chemicals and nutrient inhibitory enzymes, and some other mechanisms that change plant cellular structure to enhance mechanical barrier to insect probing and feeding. Cellular defense often started with the recognition of pathogen associated molecular patterns (PAMPs) present on the surface of plant cells. PAMPs triggered immunity induces reactive oxygen species (ROS), Ca+ production and callose deposition in plasmodesmata of the cells (Tuteja, 2007). Also, with the recognition of effectors, different transcripts encoding proteins and enzymes such as cellulose synthase, pectic esterase, glycosylases and hydrolases are promptly up-regulated in the cells of the resistant plants to prevent the HF larval attack. These genes are involved in strengthening and fortification of plant cell wall and cuticle, thus may play an important role in abstaining the insects to access the nutrients, resulting in the death of insects. The fortified cell walls can also decrease the digestibility of plant tissues by insects with their mouth parts known as mandibles. Moreover, production of various phenolic compounds increases in response to escalated gene expression in resistant plants and these compounds are deposited into the cell

walls to strengthen and thicken the cell walls following HF attacks. The highly cross-linked lipids fortified cell wall could inhibit the further spread of toxins and effectors to other plant cells (Khajuria et al., 2013).

1.4.3 Reactive oxygen species (ROS) accumulation in compatible interaction

Hypersensitive response (HR) is the main form of resistance in plants' direct responses against pathogens, but it is also a type of defense mechanism in gall-producing insects which have fixed feeding site (Goodman & Novacky, 1994). Three major forms of ROS molecules, hydrogen peroxide (H_2O_2), superoxide anion (O^{2-}) and hydroxyl radicals, are primarily involved in cell oxidative reduction at insect feeding sites, which lead to the HR and initiation of plant cell death around the infectious site. This phenomenon eventually stops the further spread of HF by starving them to death in 3 to 5 days as seen in interactions between some plants and pathogens (Gechev et al., 2006; Heath, 2000).

ROS molecules are involved in defense responses because they serve as signals to activate several other defense mechanisms such as NADPH-dependent oxidase system, Germin-like oxidase system, amine oxidase, and glycolate oxidase systems. Moreover, in addition to these oxidases, class III peroxidases such as pH dependent cell wall peroxidases also play an important role in ROS secretion in response to insect attack under certain conditions. All these systems are associated with H_2O_2 and O^{2-} production in responses to HF attack during incompatible interactions (Apel & Hirt, 2004; Mika et al., 2004; Mittler et al., 2004; Razem & Bernards, 2003). Liu et al. (2009) provided more insight into pathways contributing to accumulation of ROS molecules in resistant plants during HF larval attack. A 2-fold increase in peroxidases activity was observed in resistant plants carrying resistant genes (*H9* and *H13*)

during incompatible reactions, but no change has been detected in the controls and infested plants during compatible interactions (Liu et al., 2010).

Like other galling insects, ROS defense is among one of the prominent defense systems against HF in plants. However, in contrast to strong evidence of the major roles of ROS in plant defense against pathogens, evidence of roles of ROS defense against insect attack in plants is quite preliminary, erratic, and sporadic (Chen, 2008).

1.4.4 Molecular mechanisms of plants defense against Hessian fly

Previous studies have confirmed that genes involved in defense mechanisms encode lectin like protein, protein inhibitors, transmembrane proteins, proteinase inhibitors and other enzymes and proteins that are involved in primary metabolisms and ROS production (Liu et al., 2010; Zhu et al., 2008). Moreover, recent studies on HF resistance genes shown that expressed gene sets in resistant plants were different from these in the susceptible plants when they are attacked by HF. This difference indicates the rapid mobilization and reuse of plant resources, suppression of primary metabolisms and reduction in plant growth to trigger the direct defenses against HF. The reserved resources are mainly used to up regulate the enzymes that are involved in catabolic pathways to release substance and energy required for plant defense (Zhu et al., 2011 and 2012). Moreover, these mobilized resources are then converted into toxic compounds which do not bring immediate lethality of HF larvae but slow down the larval attack to extend the time required for plants to trigger a defense mechanism. In response to these toxic defense compounds, various genes encoding enzymes and proteins involved in remodeling and fortification of cell walls become active to starve the HF larvae to death (Lamb & Dixon, 1997; McMullen et al., 2009; Shukle et al., 1990; Zhu et al., 2011).

To sum up, upon insect attack, various insect resistance genes (R-genes) in plants differentially expressed between resistant and susceptible plants in responses to insect effectors (Avr-genes). The up- and down-regulation of the genes trigger different molecular pathways for plant defense. The decrease in metabolic rates following HF infestation saves the resources for plants to initiate the defense. These metabolites cascade a signaling pathway to prevent the delivery of effectors into the host cells and inhibit the formation of HF nutritive tissues in the plants. The absence of nutritive tissue at feeding sites deprives HF larvae of nutrient uptake, resulting in the death of larvae due to malnutrition (Khajuria et al., 2013).

1.5 Mapping of HF resistance genes in wheat

1.5.1 Physical and genetic mapping of resistance genes in wheat

Genetic map has been used to identify location, linear order of genes or quantitative trait loci (QTLs) for different traits (Yim et al., 2002). One of the main objectives of QTL mapping is to identify tightly linked markers to resistance genes or QTLs to monitor their movement in progeny when they are transferred into well-adapted cultivars through breeding. Usually, a QTL can be mapped using a biparental population to identify the region associated with a target trait and then the tightly linked markers can be converted into breeder-friendly markers for breeding selection. These QTL maps also facilitate calculation of the physical and genetics distances between genes and DNA markers on the chromosomes (Buerstmayr et al., 2020).

Genetic linkage maps are usually developed based on segregation and recombination frequency of marker/gene loci during meiosis in various populations such as F2, recombinant inbred lines (RILs), backcross and double haploid (Erayman, 2004). To obtain high quality maps with high marker coverage on genomes, high marker polymorphisms and recombination rates

between parents of mapping populations are essential (Vuylsteke et al., 1999). The genetic maps can be constructed using different types of markers such as restriction fragment length polymorphism (RFLP), simple sequence repeat (SSR), diversity arrays technology (DArT), and single nucleotide polymorphism (SNPs). However, single nucleotide polymorphism (SNP) markers are more suitable for high-resolution gene mapping and functional marker development due to its abundance in genomes. Moreover, rapid development of next-generation sequencing (NGS) technologies makes it possible to generate thousands of SNPs per sample at an affordable cost at present. The recombination frequencies among markers and genes in the segregating mapping population are used to calculate genetic distances (cM) among markers or genes. The positions of markers and QTL intervals can be located in the maps. Integration of genetic information from multiple genetic maps from the same species allows the construction of a unique consensus map to generate more accurate estimation of marker positions for QTL mapping. These maps can be further used to aid marker assisted selection for gene pyramiding, fine-mapping and gene cloning by tracking the markers tightly linked to target genes or QTLs (Collard et al., 2005; Paterson et al., 2000).

However, genetic distance does not necessarily reflect the corresponding physical distance between markers or genes. It can vary not only among populations but also among chromosomes, or even among different regions within a chromosome. For example, a 1 centimorgan (cM) region can harbor only several genes in one population, but maybe several hundred genes in another population. Therefore, physical maps can be used to identify putative candidate genes in a targeted region for fine mapping and gene cloning as physical distance does not vary with chromosome regions or populations where the genotype is used as one of the parents (Ashikari et al., 2005; Schmalenbach et al., 2011; Yang et al., 2012). A physical map

provides the distinct position of a locus on a particular chromosome, chromosome arm or deletion bin and the distance between loci in the number of nucleotide base pairs (bp). Physical mapping of QTLs is the best route towards detailed molecular characterization and functional studies of the relationship between genetic polymorphisms and trait variation (Buerstmayr et al., 2020). It is also of direct relevance to breeding since it makes QTLs more easily integrated into new cultivars by marker-assisted breeding and genomic selection.

1.5.2 Reported R genes and their sources

To date, 37 R-genes (*H1- H37* and *Hdic*) have been documented underlying the incompatible interactions between wheat and HF (Zhang et al., 2021). Majority of these genes are from wheat including *H1, H2, H3, h4, H5, H7, H8, H12, H32*, and *H34* from common wheat and *H6, H9, H10, H11, H14, H15, H16, H17, H18, H19, H20, H28, H29*, and *H31* from durum wheat (Liu, et al., 2005; Liu, et al., 2005b). Moreover, many are from wheat relative species including *H13, H22, H23, H24, and H26* from *Aegilops tauschii*, *H21* and *H25* from Secale cereal (rye), *H27* from *Aegilops ventricose*, *H30* from *Aegilops triuncialis*, and *Hdic* from emmer wheat (Li et al., 2013; Liu et al., 2005; Sardesai et al., 2005). These genes are mapped on different chromosomes of all three genomes. Among 18 genes from A genome, 14 (*H3, H5, H6, H9, H10, H11, H12, H14, H15, H16, H17, H19, H28, H29* and *Hdic*) formed a cluster in a ~1 cM gene rich region on the distal end of chromosome 1A (Liu et al., 2005). Only seven genes were mapped on D genome, including *h4, H7, H8, H13, H22, H23, H24, H26, H30* and *H32* (Wang et al., 2006), and eight genes (*H8, H18, H20, H21, H25, H31* and *H34*) were mapped in the B genome (Li et al., 2013; Wang et al., 2006). Recently, two new major QTLs for HF resistance (*QH.icd-2A* and *QHara.icd-6B*) were identified by association mapping, and they were derived from *T. dicoccum* and *T. araraticum*, respectively (Bassi et al., 2019). In wheat, HF

resistance is mostly controlled by dominant genes except for *h4* as a recessive gene (Gallun & Patterson, 1977; Zhao et al., 2015). These genes may differ in resistance to different HF biotypes in different geographical regions (Jones & Dangl, 2006).

1.5.3 Fine mapping and gene cloning in wheat

After QTL mapping that located a QTL to an interval on a chromosome using flanking markers, fine map can narrow down the QTL region to the smallest possible interval for candidate gene cloning. Fine mapping requires that marker-trait associations are tested in the populations with large numbers of recombination. This can be achieved by increasing the size of mapping populations or by increasing the number of generations of selfing to create the populations (Buerstmayr et al., 2020). Developing heterogeneous inbred families (HIFs) to select recombinants and near isogenic lines for fine mapping of the targeted QTL is proved to be a time-saving approach in cereal crops (Liu et al., 2020; Merry et al., 2019; Zhang et al., 2013). Fine mapping and cloning of a gene can facilitate understanding of the gene actions and their interaction with other genes so they can be used in crop improvement programs (Thind et al., 2017). After precise mapping of a QTL, putative candidate gene sequences can be identified by blasting the flanking markers for the QTL in the wheat reference genome to identify possible high confidence candidate genes underlying the QTL. The sequence of these genes can be used to develop gene-specific diagnostic markers for breeding application after validating them in diversified association mapping panels (Schweiger et al., 2016; Wang et al., 2006).

In wheat, positional cloning has lagged many other major crops because of its multiple ploidy level, big genome, less recombination rate, large number of transposable elements, homeologous genes on different genomes. Moreover, delayed availability of annotated genome

sequence and difficulty in designing of gene specific markers to amplify targeted genes also make it difficult to clone a gene in wheat (Appels et al., 2018; Dvorak et al., 2018; Quinn et al., 2013; Sandhu & Gill, 2002; Schweiger et al., 2016; Taagen et al., 2021). Recently, genes for several important agronomic traits have been cloned and characterized, including plant height and grain size and for disease resistance such as *Fhb7* (Wang et al., 2020) and *Fhb1* (Su et al., 2019) for fusarium head blight, *Sr60* (Chen et al., 2020), *Stb6* (Saintenac et al., 2018) against *Septoria tritici* blotch, *Pm5e* (Xie et al., 2020), *Pm60* (Zou et al., 2018), *Pm3b* (Yahiaoui et al., 2004) against powdery mildew, *Lr14a* (Kolodziej et al., 2021) against leaf rust, *Yr15* (Klymiuk et al., 2018), *Yr36* (Fu et al., 2009) and *Yr10* (Liu et al., 2014) against strip rust have been cloned in wheat. The genes for resistance to abiotic stresses including *TaEXPB23* (Han et al., 2012), *TaSnRK2.7* (Zhang et al., 2011), *TaWRKY44* (Wang et al., 2015), *TaMYBsm3* (Li et al., 2019) against drought stress, and for other plant development traits such as *VRN1, VRN2* and *VRN3* (Yan et al., 2003, 2004, 2006) conditioning wheat vernalization. These cloned genes have been used in different studies to develop diagnostic markers for selecting these genes in breeding and to investigate interactions with other genes to identify related pathways controlling these traits (Krasileva et al., 2017; Lord et al., 2019; Taagen et al., 2021).

1.5.4 RNA-seq: an indispensable and robust tool for profiling differentially expressed genes

Accurate and precise identification of differentially expressed genes (DEGs) in the QTL region is a key step for understanding relationship between genetic and phenotypic variation to select final candidate gene for cloning. RNA sequencing (RNA-seq) has become a popular choice to measure gene expression levels in various organisms (Costa-Silva et al., 2017; Mortazavi et al., 2008) because it can not only differentiate the levels of cDNA transcripts

between two or more contrasting genotypes/conditions at different time points, but also identify chromosome locations of the DEGs using a reference genome (Bai et al., 2016; Bao et al., 2014). RNA-seq does not require prior knowledge of genes of interest and can be performed with or without availability of reference genome (de-novo). RNA-seq data sets can be used for sequence assemblies for subsequent mapping of RNA-seq reads to a physical map, and for detecting exon/intron boundaries, alternative splicing, and novel transcribed regions in a single sequencing run (Schliesky et al., 2012). It can also provide significant information about sequence structural variation and methylation pattern in the genes (Cokus et al., 2008). Moreover, gene ontology (GO) enrichment analysis can evaluate if the candidate genes are critical components of gene co-expression functional networks (Tohge & Fernie, 2012).

Furthermore, by comparing RNA-seq data between contrasting genotypes, SNPs can be discovered and the expression levels of the functional genes with sequence variations can be observed simultaneously at a reasonable cost. Also, the locations of causal variations in coding regions of the genes of interests may be identified and used to predict the phenotypes (Yu et al., 2014). In addition, RNA-seq is also a useful platform for gene characterization, gene expression quantification as well as post translational process analysis (Quinn et al., 2013; Zhao et al., 2019).

1.5.5 Development of diagnostic markers

One of the most often cited benefits of genetic markers for plant breeding has been their use in marker-assisted selection (MAS), in which the markers are used as indirect selection tools to select target traits in breeding programs. This allows the breeder to achieve early selection of a trait, a combination of many traits, or traits that are usually not visually detectable in field (Koebner & Summers, 2002). Development of functional markers can improve effectiveness of

MAS (Zhong et al., 2018). Development of such markers requires identification of contrasting allele sequences of functional genes affecting the plant phenotypes or the sequences that are in strong linkage disequilibrium with the QTLs. Moreover, the identification and validation of diagnostic markers requires association analysis on the populations of elite cultivars to directly correlate marker haplotypes with the target trait phenotypes (Prodhomme et al., 2020). Markers are also essential for candidate gene mapping, gene pyramiding, and the map-based cloning of genes underlying target traits. Gene specific markers based on known functional polymorphisms within the target gene(s) are the first choice for MAS breeding programs (Bini et al., 2022; Hasan et al., 2021). However, such diagnostic markers may be difficult and expensive to be identified, markers at a sub-centimorgan distance from the target QTL are often the most efficient substitutes for indirect selection for the target traits (Rafalski, 2002).

1.6 References

Amri, A., Cox, T. S., Gill, B. S., & Hatchett, J. H. (1990). Chromosomal location of the Hessian fly resistance gene *H20* in 'Jori' Durum wheat. *Journal of Heredity*, *81*(1), 71–72. https://doi.org/10.1093/oxfordjournals.jhered.a110929

Anderson, K. G., & Harris, M. O. (2006). Does R gene resistance allow wheat to prevent plant growth effects associated with Hessian fly (Diptera: Cecidomyiidae) Attack? *Journal of Economic Entomology*, *99*(5), 1842–1853. https://doi.org/10.1093/jee/99.5.1842

Anderson, K. M., & Harris, M. O. (2019). Susceptibility of North Dakota Hessian fly (Diptera: Cecidomyiidae) to 31 *H* genes mediating wheat resistance. *Journal of Economic Entomology*. https://doi.org/10.1093/jee/toz121

Apel, K., & Hirt, H. (2004). Reative oxygen species: metabolism, oxidative stress, and signal transduction. *Annual Review of Plant Biology*, *55*(1), 373–399. https://doi.org/10.1146/annurev.arplant.55.031903.141701

Appels, R., Eversole, K., Stein, N., Feuillet, C., Keller, B., Rogers, J., Pozniak, C. J., Choulet, F., Distelfeld, A., Poland, J., Ronen, G., Sharpe, A. G., Barad, O., Baruch, K., Keeble-Gagnère, G., Mascher, M., Ben-Zvi, G., Josselin, A.-A., Himmelbach, A., … Wang, L. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, *361*(6403). https://doi.org/10.1126/science.aar7191

Bai, H., Harmancı, A. S., Erson-Omay, E. Z., Li, J., Coşkun, S., Simon, M., Krischek, B., Özduman, K., Omay, S. B., Sorensen, E. A., Turcan, Ş., Bakırcığlu, M., Carrión-Grant, G., Murray, P. B., Clark, V. E., Ercan-Sencicek, A. G., Knight, J., Sencar, L., Altınok, S., … Günel, M. (2016). Integrated genomic characterization of IDH1-mutant glioma malignant progression. *Nature Genetics*, *48*(1), 59–66. https://doi.org/10.1038/ng.3457

Bao, Z. S., Chen, H.-M., Yang, M.-Y., Zhang, C. B., Yu, K., Ye, W. L., Hu, B. Q., Yan, W., Zhang, W., Akers, J., Ramakrishnan, V., Li, J., Carter, B., Liu, Y. W., Hu, H. M., Wang, Z., Li, M. Y., Yao, K., Qiu, X. G., … Jiang, T. (2014). RNA-seq of 272 gliomas revealed a novel, recurrent *PTPRZ1-MET* fusion transcript in secondary glioblastomas. *Genome*

Research, *24*(11), 1765–1773. https://doi.org/10.1101/gr.165126.113

Bassi, F. M., Brahmi, H., Sabraoui, A., Amri, A., Nsarellah, N., Nachit, M. M., Al-Abdallat, A., Chen, M. S., Lazraq, A., & El Bouhssini, M. (2019). Genetic identification of loci for Hessian fly resistance in durum wheat. *Molecular Breeding*, *39*(2). https://doi.org/10.1007/s11032-019-0927-1

Bassi, F. M., Brahmi, H., Sabraoui, A., Amri, A., Nsarellah, N., Nachit, M. M., Al-Abdallat, A., Chen, M. S., Lazraq, A., & El Bouhssini, M. (2019). Genetic identification of loci for Hessian fly resistance in durum wheat. *Molecular Breeding*, *39*(2), 24. https://doi.org/10.1007/s11032-019-0927-1

Behura, S. K., Nair, S., Sahu, S. C., & Mohan, M. (2000). An AFLP marker that differentiates biotypes of the Asian rice gall midge (*Orseolia oryzae*, Wood-Mason) is sex-linked and also linked to avirulence. *Molecular and General Genetics*, *263*(2), 328–334. https://doi.org/10.1007/s004380051175

Bent, A. F., & Mackey, D. (2007). Elicitors, effectors, and *R* genes: the new paradigm and a lifetime supply of questions. *Annual Review of Phytopathology*, *45*(1), 399–436. https://doi.org/10.1146/annurev.phyto.45.062806.094427

Bini, K., Saha, T., Radhakrishnan, S., Ravindran, M., & Uthup, T. K. (2022). Development of novel markers for yield in *Hevea brasiliensis* muell. Arg. Based on candidate genes from biosynthetic pathways associated with latex production. *Biochemical Genetics*. https://doi.org/10.1007/s10528-022-10211-w

Bos, J. I. B., Prince, D., Pitino, M., Maffei, M. E., Win, J., & Hogenhout, S. A. (2010). A functional genomics approach identifies candidate effectors from the Aphid species *Myzus persicae* (Green peach Aphid). *PLoS Genetics*, *6*(11), e1001216. https://doi.org/10.1371/journal.pgen.1001216

Buerstmayr, M., Steiner, B., & Buerstmayr, H. (2020). Breeding for fusarium head blight resistance in wheat progress and challenges. *Plant Breeding*, *139*(3), 429–454. https://doi.org/10.1111/pbr.12797

Buntin, D. (2007). Hessian fly and aphid management in wheat in Georgia. *University of Georgia College of Agricultural and Environmental Sciences. 14; 1–12*

Buntin, D. G., & Chapin, J. W. (1990). Biology of Hessian fly (Diptera: Cecidomyiidae) in the Southeastern United States: Geographic variation and temperature-dependent phenology. *Journal of Economic Entomology*, *83*(3), 1015–1024. https://doi.org/10.1093/jee/83.3.1015

Buntin, David G., & Hudson, R. D. (1991). Spring control of the Hessian fly (Diptera: Cecidomyiidae) in winter wheat usingi insecticides. *Journal of Economic Entomology*, *84*(6), 1913–1919. https://doi.org/10.1093/jee/84.6.1913

Buntin, G. D., Ott, S. L., & Johnson, J. W. (1992). Integration of plant resistance, insecticides, and planting date for management of the Hessian fly (diptera: cecidomyiidae) in Winter wheat. *Journal of Economic Entomology*, *85*(2), 530–538. https://doi.org/10.1093/jee/85.2.530

Buntin, G. David. (1999). Hessian fly (diptera: cecidomyiidae) injury and loss of winter wheat grain yield and quality. *Journal of Economic Entomology*, *92*(5), 1190–1197. https://doi.org/10.1093/jee/92.5.1190

Byers, R. A., & Gallun, R. L. (1972). Ability of the Hessian fly to stunt winter wheat. Effect of larval feeding on elongation of leaves. *Journal of Economic Entomology*, *65*(4), 955–958. https://doi.org/10.1093/jee/65.4.955

Cambron, S. E., Buntin, G. D., Weisz, R., Holland, J. D., Flanders, K. L., Schemerhorn, B. J., & Shukle, R. H. (2010). Virulence in Hessian fly (diptera: cecidomyiidae) field collections from the Southeastern United States to 21 resistance genes in wheat. *Journal of Economic Entomology*, *103*(6), 2229–2235. https://doi.org/10.1603/EC10219

Chen, B. H., Foster, J. E., Araya, J. E., & Taylor, P. L. (1991). Parasitism of *Mayetiola destructor* (diptera: cecidomyiidae) by *Platygaster hiemalis* (hymenoptera: platygasteridae) on Hessian fly resistant wheats. *Journal of Entomological Science*, *26*(2), 237–243. https://doi.org/10.18474/0749-8004-26.2.237

Chen, M. S. (2008). Inducible direct plant defense against insect herbivores: A review. *Insect*

Science, *15*(2), 101–114. https://doi.org/10.1111/j.1744-7917.2008.00190.x

Chen, M. S., Echegaray, E., Whitworth, R. J., Wang, H., Sloderbeck, P. E., Knutson, A., Giles, K. L., & Royer, T. A. (2009). Virulence analysis of Hessian fly populations from Texas, Oklahoma, and Kansas. *Journal of Economic Entomology*, *102*(2), 774–780. https://doi.org/10.1603/029.102.0239

Chen, M. S., Liu, X., Wang, H., & El-Bouhssini, M. (2009). Hessian fly (Diptera: Cecidomyiidae) interactions with barley, rice, and wheat seedlings. *Journal of Economic Entomology*, *102*(4), 1663–1672. https://doi.org/10.1603/029.102.0434

Chen, M. S., Liu, X., Yang, Z., Zhao, H., Shukle, R. H., Stuart, J. J., & Hulbert, S. (2010). Unusual conservation among genes encoding small secreted salivary gland proteins from a gall midge. *BMC Evolutionary Biology*, *10*(1), 296. https://doi.org/10.1186/1471-2148-10-296

Chen, M.-S., Wheeler, S., Wang, H., & Whitworth, R. J. (2014). Impact of temperatures on Hessian fly (diptera: cecidomyiidae) resistance in selected wheat cultivars (poales: poaceae) in the Great Plains region. *Journal of Economic Entomology*, *107*(3), 1266–1273. https://doi.org/10.1603/EC13357

Chen, S., Rouse, M. N., Zhang, W., Zhang, X., Guo, Y., Briggs, J., & Dubcovsky, J. (2020). Wheat gene *Sr60* encodes a protein with two putative kinase domains that confers resistance to stem rust. *New Phytologist*, *225*(2), 948–959. https://doi.org/10.1111/nph.16169

Chen, Z., Tang, D., Ni, J., Li, P., Wang, L., Zhou, J., Li, C., Lan, H., Li, L., & Liu, J. (2021). Development of genic KASP SNP markers from RNA-seq data for map-based cloning and marker-assisted selection in maize. *BMC Plant Biology*, *21*(1), 157. https://doi.org/10.1186/s12870-021-02932-8

Cokus, S. J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C. D., Pradhan, S., Nelson, S. F., Pellegrini, M., & Jacobsen, S. E. (2008). Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature*, *452*(7184), 215–219. https://doi.org/10.1038/nature06745

Collard, B. C. Y., Jahufer, M. Z. Z., Brouwer, J. B., & Pang, E. C. K. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica*, *142*(1–2), 169–196. https://doi.org/10.1007/s10681-005-1681-5

Costa-Silva, J., Domingues, D., & Lopes, F. M. (2017). RNA-seq differential expression analysis: an extended review and a software tool. *PLOS ONE*, *12*(12), e0190152. https://doi.org/10.1371/journal.pone.0190152

Delaney, D. E., Friebe, B. R., Gill, B. S., Hulbert, S. H., & Hatchett, J. H. (1995). Targeted mapping of rye chromatin in wheat by representational difference analysis. *Genome*, *38*(3), 458–466. https://doi.org/10.1139/g95-060

Delibes, A., Moral, J. Del, Martín-Sanchez, J. A., Mejias, A., Gallego, M., Casado, D., Sin, E., & López-Braña, I. (1997). Hessian fly resistance gene transferred from chromosome *4Mv* of *Aegilops ventricosa* to *Triticum aestivum*. *Theoretical and Applied Genetics*, *94*(6–7), 858–864. https://doi.org/10.1007/s001220050487

Dubcovsky, J., & Dvorak, J. (2007). Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science*, *316*(5833), 1862–1866. https://doi.org/10.1126/science.1143986

Dvorak, J., Wang, L., Zhu, T., Jorgensen, C. M., Deal, K. R., Dai, X., Dawson, M. W., Müller, H., Luo, M., Ramasamy, R. K., Dehghani, H., Gu, Y. Q., Gill, B. S., Distelfeld, A., Devos, K. M., Qi, P., You, F. M., Gulick, P. J., & McGuire, P. E. (2018). Structural variation and rates of genome evolution in the grass family seen through comparison of sequences of genomes greatly differing in size. *The Plant Journal*, *95*(3), 487–503. https://doi.org/10.1111/tpj.13964

Erayman, M. (2004). Demarcating the gene-rich regions of the wheat genome. *Nucleic Acids Research*, *32*(12), 3546–3565. https://doi.org/10.1093/nar/gkh639

Evans, L. T., & Fischer, R. A. (1999). Yield potential: its definition, measurement, and significance. *Crop Science*, *39*(6), 1544–1551.

https://doi.org/10.2135/cropsci1999.3961544x

Food and Agriculture organization of the United States. (2010). http://www.fao.org/news/archive/news-by-date/2010/en/

Food and Agriculture organization of the United States. (2020). http://www.fao.org/news/archive/news-by-date/2020/en/

Fu, D., Uauy, C., Distelfeld, A., Blechl, A., Epstein, L., Chen, X., Sela, H., Fahima, T., & Dubcovsky, J. (2009). A Kinase-*START* gene confers temperature-dependent resistance to wheat stripe rust. *Science*, *323*(5919), 1357–1360. https://doi.org/10.1126/science.1166289

Gallun, R. L., & Patterson, F. L. (1977). Monosomic analysis of wheat for resistance to Hessian fly. *Journal of Heredity*, *68*(4), 223–226. https://doi.org/10.1093/oxfordjournals.jhered.a108818

Gechev, T. S., Van Breusegem, F., Stone, J. M., Denev, I., & Laloi, C. (2006). Reactive oxygen species as signals that modulate plant stress responses and programmed cell death. *BioEssays*, *28*(11), 1091–1101. https://doi.org/10.1002/bies.20493

Goodman, R. N., & Novacky, A. J. (1994). The hypersensitive reaction in plants to pathogens. *APS Press*, 256.

Gould, F. (1986). Simulation models for predicting durability of insect-resistant germ plasm: Hessian fly (Diptera: Cecidomyiidae)-resistant winter wheat. *Environmental Entomology*, *15*(1), 11–23. https://doi.org/10.1093/ee/15.1.11

Grant, S. R., Fisher, E. J., Chang, J. H., Mole, B. M., & Dangl, J. L. (2006). Subterfuge and manipulation: Type III effector proteins of phytopathogenic bacteria. *Annual Review of Microbiology*, *60*(1), 425–449. https://doi.org/10.1146/annurev.micro.60.080805.142251

Graybosch, R. A., & Peterson, C. J. (2010). Genetic improvement in winter wheat yields in the Great Plains of North America, 1959-2008. *Crop Science*, *50*(5), 1882–1890. https://doi.org/10.2135/cropsci2009.11.0685

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

Han, Y. Y., Li, A. X., Li, F., Zhao, M. R, & Wang, W. (2012). Characterization of a wheat (*Triticum aestivum* L.) expansin gene, *TaEXPB23*, involved in the abiotic stress response and phytohormone regulation. *Plant Physiology and Biochemistry*, *54*, 49–58. https://doi.org/10.1016/j.plaphy.2012.02.007

Harris, M. O., Freeman, T. P., Moore, J. A., Anderson, K. G., Payne, S. A., Anderson, K. M., & Rohfritsch, O. (2010). *H*-gene-mediated resistance to Hessian fly exhibits features of penetration resistance to fungi. *Phytopathology*, *100*, 279–289. https://doi.org/10.1094/PHYTO-100-3-0279

Harris, M. O., Friesen, T. L., Xu, S. S., Chen, M. S., Giron, D., & Stuart, J. J. (2015). Pivoting from *Arabidopsis* to wheat to understand how agricultural plants integrate responses to biotic stress. *Journal of Experimental Botany*, *66*(2), 513–531. https://doi.org/10.1093/jxb/eru465

Harris, M. O., Rose, S., & Malsch, P. (1993). The role of vision in the host plant-finding behaviour of the Hessian fly. *Physiological Entomology*, *18*(1), 31–42. https://doi.org/10.1111/j.1365-3032.1993.tb00446.x

Harris, M. O., Stuart, J. J., Mohan, M., Nair, S., Lamb, R. J., & Rohfritsch, O. (2003). Plant defense and insect adaptation. *Annual Review of Entomology*, *48*(1), 549–577. https://doi.org/10.1146/annurev.ento.48.091801.112559

Hasan, N., Choudhary, S., Naaz, N., Sharma, N., & Laskar, R. A. (2021). Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes. *Journal of Genetic Engineering and Biotechnology*, *19*(1), 128. https://doi.org/10.1186/s43141-021-00231-1

Heath, M. C. (2000). Hypersensitive response related death. In *Programmed Cell Death in Higher Plants* (pp. 77–90). Springer Netherlands. https://doi.org/10.1007/978-94-010-0934-8_6

Hill, C. C. (1933). Parasites of the Hessian fly in the North Central States. *USDA Circular*, 1–15.

Hilliker, A. J., Harauz, G., Reaume, A. G., Gray, M., Clark, S. H., & Chovnick, A. (1994). Meiotic gene conversion tract length distribution within the rosy locus of *Drosophila melanogaster*. *Genetics*, *137*(4), 1019–1026. https://doi.org/10.1093/genetics/137.4.1019

Hussain, S. S. & Rivandi. A. (2007). Molecular breeding for drought tolerance in plants: wheat perspective. *Proceeding of Pakistan Academy of Sciences*. 38(30), 35-62.

IWGSC, International wheat genome sequencing consortium, Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., Pozniak, C.J., Choulet, F., Distelfeld, A., Poland, J. & Ronen, G. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science, 361(6403), eaar7191.

Jeffreys, A. J., & May, C. A. (2004). Intense and highly localized gene conversion activity in human meiotic crossover hot spots. *Nature Genetics*, *36*(2), 151–156. https://doi.org/10.1038/ng1287

Johnston, J. S., Ross, L. D., Beani, L., Hughes, D. P., & Kathirithamby, J. (2004). Tiny genomes and endoreduplication in *Strepsiptera*. *Insect Molecular Biology*, *13*(6), 581–585. https://doi.org/10.1111/j.0962-1075.2004.00514.x

Jones, J. D. G., & Dangl, J. L. (2006). The plant immune system. *Nature*, *444*(7117), 323–329. https://doi.org/10.1038/nature05286

Kämper, J., Kahmann, R., Bölker, M., Ma, L.-J., Brefort, T., Saville, B. J., Banuett, F., Kronstad, J. W., Gold, S. E., Müller, O., Perlin, M. H., Wösten, H. A. B., de Vries, R., Ruiz-Herrera, J., Reynaga-Peña, C. G., Snetselaar, K., McCann, M., Pérez-Martín, J., Feldbrügge, M., … Birren, B. W. (2006). Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature*, *444*(7115), 97–101. https://doi.org/10.1038/nature05248

Kanno, H., & Harris, M. O. (2000). Physical features of grass leaves influence the placement of eggs within the plant by the Hessian fly. *Entomologia Experimentalis et Applicata*, *96*(1), 69–80. https://doi.org/10.1046/j.1570-7458.2000.00680.x

Khajuria, C., Wang, H., Liu, X., Wheeler, S., Reese, J. C., El Bouhssini, M., Whitworth, R. J., & Chen, M.-S. (2013). Mobilization of lipids and fortification of cell wall and cuticle are important in host defense against Hessian fly. *BMC Genomics*, *14*(1), 423. https://doi.org/10.1186/1471-2164-14-423

Klymiuk, V., Yaniv, E., Huang, L., Raats, D., Fatiukha, A., Chen, S., Feng, L., Frenkel, Z., Krugman, T., Lidzbarsky, G., Chang, W., Jääskeläinen, M. J., Schudoma, C., Paulin, L., Laine, P., Bariana, H., Sela, H., Saleem, K., Sørensen, C. K., … Fahima, T. (2018). Cloning of the wheat *Yr15* resistance gene sheds light on the plant tandem kinase-pseudokinase family. *Nature Communications*, *9*(1), 3735. https://doi.org/10.1038/s41467-018-06138-9

Koebner, R., & Summers, R. (2002). The impact of molecular markers on the wheat breeding paradigm. *Cellular & Molecular Biology Letters*, *7*(2B), 295–702.

Kolodziej, M. C., Singla, J., Sánchez-Martín, J., Zbinden, H., Šimková, H., Karafiátová, M., Doležel, J., Gronnier, J., Poretti, M., Glauser, G., Zhu, W., Köster, P., Zipfel, C., Wicker, T., Krattinger, S. G., & Keller, B. (2021). A membrane-bound ankyrin repeat protein confers race-specific leaf rust disease resistance in wheat. *Nature Communications*, *12*(1), 956. https://doi.org/10.1038/s41467-020-20777-x

Kong, L., Ohm, H. W., Cambron, S. E., & Williams, C. E. (2005). Molecular mapping determines that Hessian fly resistance gene *H9* is located on chromosome 1A of wheat. *Plant Breeding*, *124*(6), 525–531. https://doi.org/10.1111/j.1439-0523.2005.01139.x

Kong, Lingrang, Cambron, S. E., & Ohm, H. W. (2008). Hessian fly resistance genes *H16* and *H17* are mapped to a resistance gene cluster in the distal region of chromosome 1AS in wheat. *Molecular Breeding*, *21*(2), 183–194. https://doi.org/10.1007/s11032-007-9119-5

Krasileva, K. V., Vasquez-Gross, H. A., Howell, T., Bailey, P., Paraiso, F., Clissold, L., Simmonds, J., Ramirez-Gonzalez, R. H., Wang, X., Borrill, P., Fosker, C., Ayling, S., Phillips, A. L., Uauy, C., & Dubcovsky, J. (2017). Uncovering hidden variation in polyploid wheat. *Proceedings of the National Academy of Sciences*, *114*(6). https://doi.org/10.1073/pnas.1619268114

Lamb, C., & Dixon, R. A. (1997). The oxidative burst in plant disease resistance. *Annual Review of Plant Physiology and Plant Molecular Biology*, *48*(1), 251–275. https://doi.org/10.1146/annurev.arplant.48.1.251

Lehmann, A. R. (2006). New functions for y family polymerases. *Molecular Cell*, *24*(4), 493–495. https://doi.org/10.1016/j.molcel.2006.10.021

Li, C., Chen, M. S., Chao, S., Yu, J., & Bai, G. (2013). Identification of a novel gene, *H34*, in wheat using recombinant inbred lines and single nucleotide polymorphism markers. *Theoretical and Applied Genetics*, *126*(8), 2065–2071. https://doi.org/10.1007/s00122-013-2118-5

Li, G., Wang, Y., Chen, M.-S., Edae, E., Poland, J., Akhunov, E., Chao, S., Bai, G., Carver, B. F., & Yan, L. (2015). Precisely mapping a major gene conferring resistance to Hessian fly in bread wheat using genotyping-by-sequencing. *BMC Genomics*, *16*(1), 108. https://doi.org/10.1186/s12864-015-1297-7

Li, H., Peng, Z., Yang, X., Wang, W., Fu, J., Wang, J., Han, Y., Chai, Y., Guo, T., Yang, N., Liu, J., Warburton, M. L., Cheng, Y., Hao, X., Zhang, P., Zhao, J., Liu, Y., Wang, G., Li, J., & Yan, J. (2013). Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nature Genetics*, *45*(1), 43–50. https://doi.org/10.1038/ng.2484

Li, Y., Zhang, S., Zhang, N., Zhang, W., Li, M., Liu, B., & Shi, Z. (2019). MYB-CC transcription factor, *TaMYBsm3*, cloned from wheat is involved in drought tolerance. *BMC Plant Biology*, *19*(1), 143. https://doi.org/10.1186/s12870-019-1751-9

Liu, G., Liu, X., Xu, Y., Bernardo, A., Chen, M., Li, Y., Niu, F., Zhao, L., & Bai, G. (2020). Reassigning Hessian fly resistance genes *H7* and *H8* to chromosomes 6A and 2B of the wheat cultivar ‘Seneca’ using genotyping-by-sequencing. *Crop Science*, *60*(3), 1488–1498. https://doi.org/10.1002/csc2.20148

Liu, W., Frick, M., Huel, R., Nykiforuk, C. L., Wang, X., Gaudet, D. A., Eudes, F., Conner, R. L., Kuzyk, A., Chen, Q., Kang, Z., & Laroche, A. (2014). The stripe rust resistance gene

Yr10 encodes an evolutionary-conserved and unique CC–NBS–LRR sequence in wheat. *Molecular Plant*, *7*(12), 1740–1755. https://doi.org/10.1093/mp/ssu112

Liu, X. M., Brown-Guedira, G. L., Hatchett, J. H., Owuoche, J. O., & Chen, M. S. (2005). Genetic characterization and molecular mapping of a Hessian fly-resistance gene transferred from *T. turgidum* ssp. *dicoccum* to common wheat. *Theoretical and Applied Genetics*, *111*(7), 1308–1315. https://doi.org/10.1007/s00122-005-0059-3

Liu, X. M., Fritz, A. K., Reese, J. C., Wilde, G. E., Gill, B. S., & Chen, M. S. (2005). *H9, H10,* and *H11* compose a cluster of Hessian fly resistance genes in the distal gene-rich region of wheat chromosome 1AS. *Theoretical and Applied Genetics*, *110*(8), 1473–1480. https://doi.org/10.1007/s00122-005-1982-z

Liu, X., Williams, C. E., Nemacheck, J. A., Wang, H., Subramanyam, S., Zheng, C., & Chen, M. S. (2010). Reactive oxygen species are involved in plant defense against a gall midge . *Plant Physiology*, *152*(2), 985–999. https://doi.org/10.1104/pp.109.150656

Lord, J., McMullan, D. J., Eberhardt, R. Y., Rinck, G., Hamilton, S. J., Quinlan-Jones, E., Prigmore, E., Keelagher, R., Best, S. K., Carey, G. K., Mellis, R., Robart, S., Berry, I. R., Chandler, K. E., Cilliers, D., Cresswell, L., Edwards, S. L., Gardiner, C., Henderson, A., … Wilson, E. (2019). Prenatal exome sequencing analysis in fetal structural anomalies detected by ultrasonography (PAGE): a cohort study. *The Lancet*, *393*(10173), 747–757. https://doi.org/10.1016/S0140-6736(18)31940-8

Ma, Z. Q., Gill, B. S., Sorrells, M. E., & Tanksley, S. D. (1993). RELP markers linked to two Hessian fly resistance genes in wheat (*Triticum aestivum* L.) from *Triticum tauschii* (Coss.) Schmal. *Theoretical and Applied Genetics*, *85–85*(6–7), 750–754. https://doi.org/10.1007/BF00225015

Martín-Sánchez, J. A., Gómez-Colmenarejo, M., Del Moral, J., Sin, E., Montes, M. J., González-Belinchón, C., López-Braña, I., & Delibes, A. (2003). A new Hessian fly resistance gene (*H30*) transferred from the wild grass *Aegilops triuncialis* to hexaploid wheat. *Theoretical and Applied Genetics*, *106*(7), 1248–1255. https://doi.org/10.1007/s00122-002-1182-z

Mass, F. B., Patterson, F. L., Foster, J. E., & Ohm, H. W. (1989). Expression and inheritance of resistance of els 6404-160 durum wheat to hessian fly. *Crop Science*, *29*(1), 23–28. https://doi.org/10.2135/cropsci1989.0011183X002900010005x

McDonald, M. J., Ohm, H. W., Rinehart, K. D., Giovanini, M. P., & Williams, C. E. (2014). *H33* : A wheat gene providing Hessian fly resistance for the Southeastern United States. *Crop Science*, *54*(5), 2045–2053. https://doi.org/10.2135/cropsci2013.12.0811

McMullen, M. D., Frey, M., & Degenhardt, J. (2009). Genetics and biochemistry of insect resistance in maize. In *Handbook of Maize: Its Biology* (pp. 271–289). Springer New York. https://doi.org/10.1007/978-0-387-79418-1_14

Merry, R., Butenhoff, K., Campbell, B. W., Michno, J., Wang, D., Orf, J. H., Lorenz, A. J., & Stupar, R. M. (2019). Identification and fine-mapping of a soybean quantitative trait locus on chromosome 5 conferring tolerance to iron deficiency chlorosis. *The Plant Genome*, *12*(3), 190007. https://doi.org/10.3835/plantgenome2019.01.0007

Mika, A., Minibayeva, F., Beckett, R., & Lüthje, S. (2004). Possible functions of extracellular peroxidases in stress-induced generation and detoxification of active oxygen species. *Phytochemistry Reviews*, *3*(1–2), 173–193. https://doi.org/10.1023/B:PHYT.0000047806.21626.49

Mittler, R., Vanderauwera, S., Gollery, M., & Van Breusegem, F. (2004). Reactive oxygen gene network of plants. *Trends in Plant Science*, *9*(10), 490–498. https://doi.org/10.1016/j.tplants.2004.08.009

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nature Methods*, *5*(7), 621–628. https://doi.org/10.1038/nmeth.1226

Niu, F., Xu, Y., Liu, X., Zhao, L., Bernardo, A., Li, Y., Liu, G., Chen, M.-S., Cao, L., Hu, Z., Xu, X., & Bai, G. (2020). The Hessian fly recessive resistance gene *h4* mapped to chromosome 1A of the wheat cultivar 'Java' using genotyping-by-sequencing. *Theoretical and Applied Genetics*, *133*(10), 2927–2935. https://doi.org/10.1007/s00122-020-03642-9

Ohm, H. W., Ratcliffe, R. H., Patterson, F. L., & Cambron, S. E. (1997). Resistance to Hessian fly conditioned by genes *H19* and proposed gene *H27* of durum wheat line PI 422297. *Crop Science*, *37*(1), 113–115. https://doi.org/10.2135/cropsci1997.0011183X003700010018x

Painter, R. H. (1951). Population biology and parasitism of Hessian fly (*Mayetiola destructor*) (Diptera: Cecidomyiidae) on *Bromus willdenowii* in New Zealand. In *Insect Resistance in Crop Plants* (6th ed., Vol. 72, pp. 481–495). Williams & Wilkins.

Parks, T. H. (1917). A county-wide survey to determine the effect of time of seeding and presence of volunteer wheat upon the extent of damage by the Hessian fly. *Journal of Economic Entomology*, 249–253.

Paterson, A. H., Bowers, J. E., Burow, M. D., Draye, X., Elsik, C. G., Jiang, C.-X., Katsar, C. S., Lan, T.-H., Lin, Y.-R., Ming, R., & Wright, R. J. (2000). Comparative genomics of plant chromosomes. *The Plant Cell*, *12*(9), 1523–1539. https://doi.org/10.1105/tpc.12.9.1523

Prodhomme, C., Vos, P. G., Paulo, M. J., Tammes, J. E., Visser, R. G. F., Vossen, J. H., & van Eck, H. J. (2020). Distribution of P1(D1) wart disease resistance in potato germplasm and GWAS identification of haplotype-specific SNP markers. *Theoretical and Applied Genetics*, *133*(6), 1859–1871. https://doi.org/10.1007/s00122-020-03559-3

Quinn, E. M., Cormican, P., Kenny, E. M., Hill, M., Anney, R., Gill, M., Corvin, A. P., & Morris, D. W. (2013). Development of strategies for SNP detection in RNA-seq data: application to lymphoblastoid cell lines and evaluation using 1000 genomes data. *PLoS ONE*, *8*(3), e58815. https://doi.org/10.1371/journal.pone.0058815

Rafalski, A. (2002). Applications of single nucleotide polymorphisms in crop genetics. *Current Opinion in Plant Biology*, *5*(2), 94–100. https://doi.org/10.1016/S1369-5266(02)00240-6

Ratcliffe, R. H., Cambron, S. E., Flanders, K. L., Bosque-Perez, N. A., Clement, S. L., & Ohm, H. W. (2000). Biotype composition of hessian fly (Diptera: Cecidomyiidae) populations from the Southeastern, Midwestern, and Northwestern United States and virulence to resistance genes in wheat. *Journal of Economic Entomology*, *93*(4), 1319–1328. https://doi.org/10.1603/0022-0493-93.4.1319

Razem, F. A., & Bernards, M. A. (2003). Reactive oxygen species production in association with suberization: evidence for an NADPH-dependent oxidase. *Journal of Experimental Botany*, *54*(384), 935–941. https://doi.org/10.1093/jxb/erg094

Rossi, M., Goggin, F. L., Milligan, S. B., Kaloshian, I., Ullman, D. E., & Williamson, V. M. (1998). The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. *Proceedings of the National Academy of Sciences*, *95*(17), 9750–9754. https://doi.org/10.1073/pnas.95.17.9750

Saintenac, C., Lee, W. S., Cambon, F., Rudd, J. J., King, R. C., Marande, W., Powers, S. J., Bergès, H., Phillips, A. L., Uauy, C., Hammond-Kosack, K. E., Langin, T., & Kanyuka, K. (2018). Wheat receptor-kinase-like protein *Stb6* controls gene-for-gene resistance to fungal pathogen *Zymoseptoria tritici*. *Nature Genetics*, *50*(3), 368–374. https://doi.org/10.1038/s41588-018-0051-x

Sandhu, D., & Gill, K. S. (2002). Gene-containing regions of wheat and the other grass genomes. *Plant Physiology*, *128*(3), 803–811. https://doi.org/10.1104/pp.010745

Sardesai, N., Nemacheck, J. A., Subramanyam, S., & Williams, C. E. (2005). Identification and mapping of *H32*, a new wheat gene conferring resistance to Hessian fly. *Theoretical and Applied Genetics*, *111*(6), 1167–1173. https://doi.org/10.1007/s00122-005-0048-6

Schliesky, S., Gowik, U., Weber, A. P. M., & Bräutigam, A. (2012). RNA-seq assembly – are we there yet? *Frontiers in Plant Science*, *3*. https://doi.org/10.3389/fpls.2012.00220

Schuster, M. F., & Lidell, M. C. (1990). Distribution and seasonal abundance of Hessian fly (Diptera: Cecidomyiidae) parasitoids in Texas. *Journal of Economic Entomology*, *83*(6), 2269–2273. https://doi.org/10.1093/jee/83.6.2269

Schweiger, W., Steiner, B., Vautrin, S., Nussbaumer, T., Siegwart, G., Zamini, M., Jungreithmeier, F., Gratl, V., Lemmens, M., Mayer, K. F. X., Bérgès, H., Adam, G., & Buerstmayr, H. (2016). Suppressed recombination and unique candidate genes in the divergent haplotype encoding *Fhb1*, a major Fusarium head blight resistance locus in wheat. *Theoretical and Applied Genetics*, *129*(8), 1607–1623. https://doi.org/10.1007/s00122-016-

2727-x

Seo, Y. W., Johnson, J. W., & Jarret, R. L. (1997). A molecular marker associated with the *H21* Hessian fly resistance gene in wheat. *Molecular Breeding*, *3*(3), 177–181. https://doi.org/10.1023/A:1009606304447

Shukle, R. H., Grover, P. B., & Foster, J. E. (1990). Feeding of Hessian fly (Diptera: Cecidomyiidae) larvae on resistant and susceptible wheat. *Environmental Entomology*, *19*(3), 494–500. https://doi.org/10.1093/ee/19.3.494

Sorrells, M. E., & Yu, J. (2009). Linkage disequilibrium and association mapping in the *Triticeae*. In *Genetics and Genomics of the Triticeae* (pp. 655–683). Springer US. https://doi.org/10.1007/978-0-387-77489-3_22

Sosa, O., & Foster, J. E. (1976). Temperature and the expression of resistance in wheat to the Hessian fly. *Environmental Entomology*, *5*(2), 333–336. https://doi.org/10.1093/ee/5.2.333

Spanu, P. D., Abbott, J. C., Amselem, J., Burgis, T. A., Soanes, D. M., Stüber, K., Loren, V., Themaat, E., Ver, Brown, J. K. M., Butcher, S. A., Gurr, S. J., Lebrun, M. H., Ridout, C. J., Schulze-Lefert, P., Talbot, N. J., Ahmadinejad, N., Ametz, C., Barton, G. R., Benjdia, M., Bidzinski, P., & Panstruga, R. (2010). Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science*, *330*(6010), 1543–1546. https://doi.org/10.1126/science.1194573

Statista 2021 https://www.statista.com/topics/6034/wheat-in-the-us/

Stewart, A. V. (1992). 'Grasslands Gala' grazing brome (*Bromus stamineus Desv.*)— a new dryland pasture grass. *New Zealand Journal of Agricultural Research*, *35*(3), 349–353. https://doi.org/10.1080/00288233.1992.10427513

Stuart, J. J., Chen, M. S., Shukle, R., & Harris, M. O. (2012). Gall midges (Hessian flies) as plant pathogens. *Annual Review of Phytopathology*, *50*(1), 339–357. https://doi.org/10.1146/annurev-phyto-072910-095255

Su, Z., Bernardo, A., Tian, B., Chen, H., Wang, S., Ma, H., Cai, S., Liu, D., Zhang, D., Li, T.,

Trick, H., St. Amand, P., Yu, J., Zhang, Z., & Bai, G. (2019). A deletion mutation in *TaHRC* confers *Fhb1* resistance to Fusarium head blight in wheat. *Nature Genetics*, *51*(7), 1099–1105. https://doi.org/10.1038/s41588-019-0425-8

Taagen, E., Tanaka, J., Gul, A., & Sorrells, M. E. (2021). Positional-based cloning 'fail-safe' approach is overpowered by wheat chromosome structural variation. *The Plant Genome*, *14*(2). https://doi.org/10.1002/tpg2.20106

Thind, A. K., Wicker, T., Šimková, H., Fossati, D., Moullet, O., Brabant, C., Vrána, J., Doležel, J., & Krattinger, S. G. (2017). Rapid cloning of genes in hexaploid wheat using cultivar-specific long-range chromosome assembly. *Nature Biotechnology*, *35*(8), 793–796. https://doi.org/10.1038/nbt.3877

Tjallingii, W. F. (2006). Salivary secretions by aphids interacting with proteins of phloem wound responses. *Journal of Experimental Botany*, *57*(4), 739–745. https://doi.org/10.1093/jxb/erj088

Tohge, T., & Fernie, A. R. (2012). Co-expression and co-responses: within and beyond transcription. *Frontiers in Plant Science*, *3*. https://doi.org/10.3389/fpls.2012.00248

Tooker, J. F. (2012, April 11). Hessian fly on wheat. *College of Agricultural Sciences, Cooperative Extension, Entomological Notes, Penn State University*, 1–8.

Tuteja, N. (2007). Abscisic Acid and Abiotic Stress Signaling. *Plant Signaling & Behavior*, *2*(3), 135–138. https://doi.org/10.4161/psb.2.3.4156

Tyler, J. M., & Hatchett, J. H. (1983). Temperature influence on expression of resistance to Hessian fly (Diptera: Cecidomyiidae) in wheat derived from *Triticum tauschii*. *Journal of Economic Entomology*, *76*(2), 323–326. https://doi.org/10.1093/jee/76.2.323

U.S. wheat associates, https://www.uswheat.org/working-with-buyers/wheat-classes/

VanDuyn, J., Bambara, S., & Weisz, R. (2003). The Hessian fly: a pest of wheat in North Carolina. *North Carolina Agricultural Extension Service*, 368–381.

Vuylsteke, M., Mank, R., Antonise, R., Bastiaans, E., Senior, M. L., Stuber, C. W., Melchinger, A. E., Lübberstedt, T., Xia, X. C., Stam, P., Zabeau, M., & Kuiper, M. (1999). Two high-density AFLP linkage maps of *Zea mays* L.: analysis of distribution of AFLP markers. *Theoretical and Applied Genetics*, *99*(6), 921–935. https://doi.org/10.1007/s001220051399

Wang, H., Sun, S., Ge, W., Zhao, L., Hou, B., Wang, K., Lyu, Z., Chen, L., Xu, S., Guo, J., Li, M., Su, P., Li, X., Wang, G., Bo, C., Fang, X., Zhuang, W., Cheng, X., Wu, J., … Kong, L. (2020). Horizontal gene transfer of *Fhb7* from fungus underlies Fusarium head blight resistance in wheat. *Science*, *368*(6493). https://doi.org/10.1126/science.aba5435

Wang, T., Xu, S. S., Harris, M. O., Hu, J., Liu, L., & Cai, X. (2006). Genetic characterization and molecular mapping of Hessian fly resistance genes derived from *Aegilops tauschii* in synthetic wheat. *Theoretical and Applied Genetics*, *113*(4), 611–618. https://doi.org/10.1007/s00122-006-0325-z

Wang, X., Zeng, J., Li, Y., Rong, X., Sun, J., Sun, T., Li, M., Wang, L., Feng, Y., Chai, R., Chen, M. S., Chang, J., Li, K., Yang, G., & He, G. (2015). Expression of *TaWRKY44*, a wheat WRKY gene, in transgenic tobacco confers multiple abiotic stress tolerances. *Frontiers in Plant Science*, *6*. https://doi.org/10.3389/fpls.2015.00615

Weech, M. H., Chapleau, M., Pan, L., Ide, C., & Bede, J. C. (2008). Caterpillar saliva interferes with induced *Arabidopsis thaliana* defence responses via the systemic acquired resistance pathway. *Journal of Experimental Botany*, *59*(9), 2437–2448. https://doi.org/10.1093/jxb/ern108

Wilde, G. E., Whitworth, R. J., Claassen, M., & Shufran, R. A. (2001). Seed treatment for control of wheat insects and its effect on yield. *Journal of Agriculture Urban Entomology*, *18*(1), 1–11.

Williams, C. E., Collier, C. C., Sardesai, N., Ohm, H. W., & Cambron, S. E. (2003). Phenotypic assessment and mapped markers for *H31*, a new wheat gene conferring resistance to Hessian fly (Diptera: Cecidomyiidae). *Theoretical and Applied Genetics*, *107*(8), 1516–1523. https://doi.org/10.1007/s00122-003-1393-y

Xie, J., Guo, G., Wang, Y., Hu, T., Wang, L., Li, J., Qiu, D., Li, Y., Wu, Q., Lu, P., Chen, Y., Dong, L., Li, M., Zhang, H., Zhang, P., Zhu, K., Li, B., Deal, K. R., Huo, N., & Liu, Z. (2020). A rare single nucleotide variant in *Pm5e* confers powdery mildew resistance in common wheat. *New Phytologist*, *228*(3), 1011–1026. https://doi.org/10.1111/nph.16762

Yahiaoui, N., Srichumpa, P., Dudler, R., & Keller, B. (2004). Genome analysis at different ploidy levels allows cloning of the powdery mildew resistance gene *Pm3b* from hexaploid wheat. *The Plant Journal*, *37*(4), 528–538. https://doi.org/10.1046/j.1365-313X.2003.01977.x

Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., Sanchez, A., Valarik, M., Yasuda, S., & Dubcovsky, J. (2006). The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*. *Proceedings of the National Academy of Sciences*, *103*(51), 19581–19586. https://doi.org/10.1073/pnas.0607142103

Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T., & Dubcovsky, J. (2003). Positional cloning of the wheat vernalization gene *VRN1*. *Proceedings of the National Academy of Sciences*, *100*(10), 6263–6268. https://doi.org/10.1073/pnas.0937399100

Yan, Liuling, Loukoianov, A., Blechl, A., Tranquilli, G., Ramakrishna, W., SanMiguel, P., Bennetzen, J. L., Echenique, V., & Dubcovsky, J. (2004). The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. *Science*, *303*(5664), 1640–1644. https://doi.org/10.1126/science.1094305

Yim, Y. S., Davis, G. L., Duru, N. A., Musket, T. A., Linton, E. W., Messing, J. W., McMullen, M. D., Soderlund, C. A., Polacco, M. L., Gardiner, J. M., & Coe, E. H. (2002). Characterization of three maize bacterial artificial chromosome libraries toward anchoring of the physical map to the genetic map using high-density bacterial artificial chromosome filter hybridization. *Plant Physiology*, *130*(4), 1686–1696. https://doi.org/10.1104/pp.013474

Yu, G. T., Cai, X., Harris, M. O., Gu, Y. Q., Luo, M. C., & Xu, S. S. (2009). Saturation and comparative mapping of the genomic region harboring Hessian fly resistance gene *H26* in wheat. *Theoretical and Applied Genetics*, *118*(8), 1589–1599.

https://doi.org/10.1007/s00122-009-1006-5

Yu, Y., Wei, J., Zhang, X., Liu, J., Liu, C., Li, F., & Xiang, J. (2014). SNP discovery in the transcriptome of white pacific shrimp *Litopenaeus vannamei* by next generation sequencing. *PLoS ONE*, *9*(1), e87218. https://doi.org/10.1371/journal.pone.0087218

Zhang, C., Liu, L., Zheng, Z., Sun, Y., Zhou, L., Yang, Y., Cheng, F., Zhang, Z., Wang, X., Huang, S., Xie, B., Du, Y., Bai, Y., & Li, J. (2013). Fine mapping of the *Ph-3* gene conferring resistance to late blight (*Phytophthora infestans*) in tomato. *Theoretical and Applied Genetics*, *126*(10), 2643–2653. https://doi.org/10.1007/s00122-013-2162-1

Zhang, H., Mao, X., Jing, R., Chang, X., & Xie, H. (2011). Characterization of a common wheat (*Triticum aestivum* L.) *TaSnRK2.7* gene involved in abiotic stress responses. *Journal of Experimental Botany*, *62*(3), 975–988. https://doi.org/10.1093/jxb/erq328

Zhang, L., Xu, Y., Chen, M.-S., Su, Z., Liu, Y., Xu, Y., La, G., & Bai, G. (2021). Identification of a major QTL for Hessian fly resistance in wheat cultivar 'Chokwang.' *The Crop Journal*. https://doi.org/10.1016/j.cj.2021.08.004

Zhao, C., Escalante, L. N., Chen, H., Benatti, T. R., Qu, J., Chellapilla, S., Waterhouse, R. M., Wheeler, D., Andersson, M. N., Bao, R., Batterton, M., Behura, S. K., Blankenburg, K. P., Caragea, D., Carolan, J. C., Coyle, M., El-Bouhssini, M., Francisco, L., Friedrich, M., … Richards, S. (2015a). A massive expansion of effector genes underlies gall-formation in the wheat pest *Mayetiola destructor*. *Current Biology*, *25*(5), 613–620. https://doi.org/10.1016/j.cub.2014.12.057

Zhao, C., Escalante, L. N., Chen, H., Benatti, T. R., Qu, J., Chellapilla, S., Waterhouse, R. M., Wheeler, D., Andersson, M. N., Bao, R., Batterton, M., Behura, S. K., Blankenburg, K. P., Caragea, D., Carolan, J. C., Coyle, M., El-Bouhssini, M., Francisco, L., Friedrich, M., … Richards, S. (2015b). A massive expansion of effector genes underlies gall-formation in the wheat pest *Mayetiola destructor*. *Current Biology*, *25*(5), 613–620. https://doi.org/10.1016/j.cub.2014.12.057

Zhao, H. X., Liu, X. M., & Chen, M.-S. (2006). *H22*, a major resistance gene to the Hessian fly

(*Mayetiola destructor*), is mapped to the distal region of wheat chromosome 1DS. *Theoretical and Applied Genetics*, *113*(8), 1491–1496. https://doi.org/10.1007/s00122-006-0396-x

Zhao, Lanfei, Abdelsalam, N. R., Xu, Y., Chen, M.-S., Feng, Y., Kong, L., & Bai, G. (2020). Identification of two novel Hessian fly resistance genes *H35* and *H36* in a hard winter wheat line SD06165. *Theoretical and Applied Genetics*, *133*(8), 2343–2353. https://doi.org/10.1007/s00122-020-03602-3

Zhao, L, Zhang, H., Kohnen, M. V., Prasad, K. V. S. K., Gu, L., & Reddy, A. S. N. (2019). Analysis of transcriptome and epitranscriptome in plants using Pacbio Iso-seq and Nanopore-based direct RNA sequencing. *Frontiers in Genetics*, *10*. https://doi.org/10.3389/fgene.2019.00253

Zhong, C., Sun, S., Li, Y., Duan, C., & Zhu, Z. (2018). Next-generation sequencing to identify candidate genes and develop diagnostic markers for a novel *Phytophthora* resistance gene, *RpsHC18*, in soybean. *Theoretical and Applied Genetics*, *131*(3), 525–538. https://doi.org/10.1007/s00122-017-3016-z

Zhu, L., Chen, M.-S., & Liu, X. (2011). Changes in phytohormones and fatty acids in wheat and rice seedlings in response to Hessian fly (Diptera: Cecidomyiidae) infestation. *Journal of Economic Entomology*, *104*(4), 1384–1392. https://doi.org/10.1603/EC10455

Zhu, L., Liu, X., Liu, X., Jeannotte, R., Reese, J. C., Harris, M., Stuart, J. J., & Chen, M. S. (2008). Hessian fly (*Mayetiola destructor*) attack causes a dramatic shift in carbon and nitrogen metabolism in wheat. *Molecular Plant-Microbe Interactions*, *21*(1), 70–78. https://doi.org/10.1094/MPMI-21-1-0070

Zhu, L., Liu, X., Wang, H., Khajuria, C., Reese, J. C., Whitworth, R. J., Welti, R., & Chen, M.-S. (2012). Rapid mobilization of membrane lipids in wheat leaf sheaths during incompatible interactions with Hessian fly. *Molecular Plant-Microbe Interactions*, *25*(7), 920–930. https://doi.org/10.1094/MPMI-01-12-0022-R

Zou, S., Wang, H., Li, Y., Kong, Z., & Tang, D. (2018). The NB-LRR gene *Pm60* confers

powdery mildew resistance in wheat. *New Phytologist*, *218*(1), 298–309. https://doi.org/10.1111/nph.14964

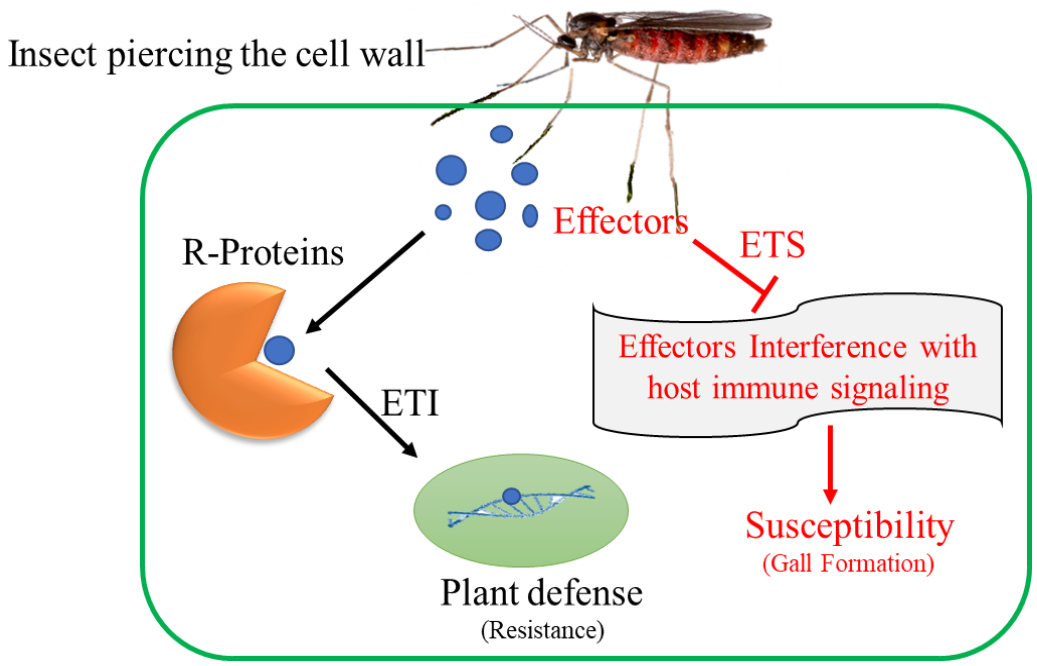

Figure 1.1. Diagrammatic illustration of effector triggered immunity (ETI) in the wheat-Hessian fly (HF) system. HF injects effector proteins through saliva just below the cell wall of a wheat plant to interfere plant immune signaling; in return, the effector recognition R-proteins produced in a host cell trigger plant resistance response to kill the insect larvae. However, if the effector recognition R-proteins are absent in wheat plants, effector proteins will interfere with host immune signaling to induce plant susceptibility, a phenomenon called effector triggered susceptibility (ETS).

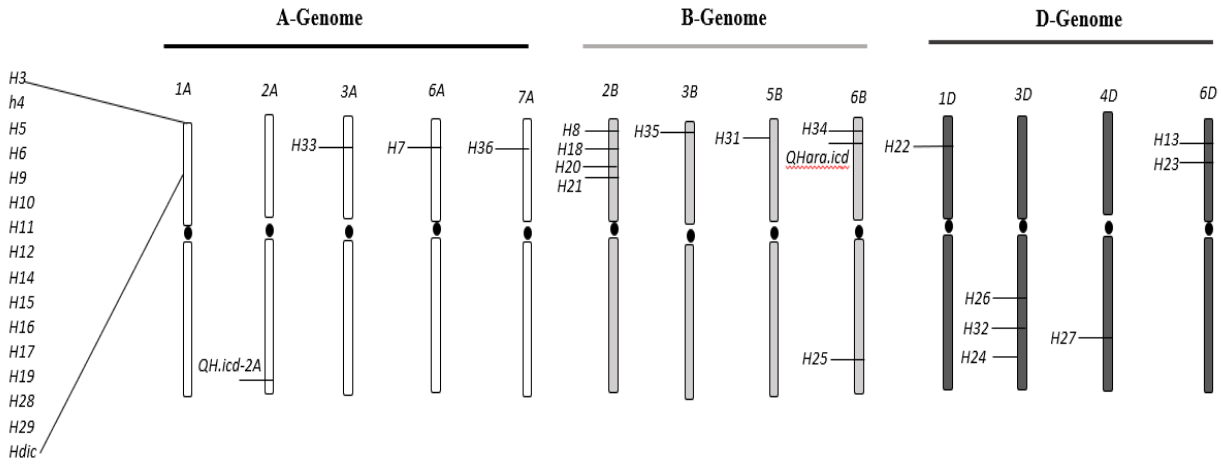

Figure 1.2. Chromosomal distribution of different Hessian fly genes (*H1-H37*, *Hdic*) in wheat.

Table A.1. Hessian fly genes sources, their location and marker types in wheat genome

Resistant Parent	**Gene name**	**Source**	**Location**	**Linked Marker**	**Reference**
	H1	*Triticum aestivum*			
	H2	*Triticum aestivum*			
Monon	*H3*	*Triticum aestivum*	1AS	RAPD	(Wang et al., 2006)
Java	*h4*	*Triticum aestivum*	1AS	KASP	(Niu et al., 2020)
Erin	*H5*	*Triticum aestivum*	1AS	AFLP	(Liu et al., 2005)
Caldwell	*H6*	*Triticum turgidum* ssp. *durum*	1AS	RAPD	(Gallun & Patterson, 1977)
Seneca	*H7*	*Triticum aestivum*	6AS	SNP	(Liu et al., 2020)
Seneca	*H8*	*Triticum aestivum*	2BS	SNP	(Liu et al., 2020)
Iris	*H9*	*Triticum turgidum* ssp. *durum*	1AS	SSR	(Liu et al., 2005)
Joy	*H10*	*Triticum turgidum* ssp. *durum*	1AS	SSR	(Liu et al., 2005)
Karen	*H11*	*Triticum turgidum* ssp. *durum*	1AS	SSR	(Liu et al., 2005)
Lola	*H12*	*Triticum aestivum*	1AS	RAPD	(Wang et al., 2006)
Molly	*H13*	*Aegilops tauschii*	6DS	RAPD	(Wang et al., 2006)
ELS 6404-160-5	*H14*	*Triticum turgidum* ssp. *durum*	1AS	RAPD	(Wang et al., 2006)
ELS 6404-160-5	*H15*	*Triticum turgidum* ssp. *durum*	1AS	RAPD	(Mass et al., 1989)
IN80164	*H16*	*Triticum turgidum* ssp. *durum*	1AS	RAPD	(Wang et al., 2006)
D6647-H17	*H17*	*Triticum turgidum* ssp. *durum*	1AS	RAPD	(Chen et al., 2009; Wang et al., 2006)
Brule	*H18*	*Triticum turgidum* ssp. *durum*	2BS	RAPD	(Martín-Sánchez et al., 2003)
PI422297	*H19*	*Triticum turgidum* ssp. *durum*	1AS	RAPD	(Ohm et al., 1997)
Jori	*H20*	*Triticum turgidum* ssp. *durum*	2BS	RAPD	(Amri et al., 1990)
KS86	*H21*	*Secale cereale*	2BS	RAPD	(Seo et al., 1997)

MHFSN91-44	*H22*	*Aegilops tauschii*	1DS	SSR and STS	(Zhao et al., 2006)
K89WGRC 3	*H23*	*Aegilops tauschii*	6DS	RFLP	(Ma et al., 1993)
K89WGRC 6	*H24*	*Aegilops tauschii*	3DL	RFLP	(Ma et al., 1993)
Cando	*H25*	*Secale cereale*	6BL	RFLP	(Delaney et al., 1995)
KS93 WGRC 26	*H26*	*Aegilops tauschii*	3DL	STS	(Yu et al., 2009)
Monon	*H27*	*Aegilops ventricosa*	4DL	Isozyme marker	(Delibes et al., 1997)
CI3146	*H28*	*Triticum turgidum* ssp. *durum*	1AS	RAPD	(Kong et al., 2005; Kong et al., 2008)
PI422297SIN CAPE 90	*H29*	*Triticum turgidum* ssp. *durum*	1AS	SSR	(Kong et al., 2005; Kong et al., 2008)
TR-3531	*H30*	*Aegilops triuncialis*	Non-allelic	Isozyme marker	(Martín-Sánchez et al., 2003)
PI921696	*H31*	*Triticum turgidum* ssp. *durum*	5BS	AFLP and STS	(Williams et al., 2003)
W-7984	*H32*	*Aegilops tauschii*	3DL	SSR	(Sardesai et al., 2005)
PI 134942	*H33*	*Triticum turgidum* ssp. *durum*	3AS	SSR	(McDonald et al., 2014)
Clark	*H34*	*Triticum aestivum*	6BS	SNP	(Li et al., 2013)
SD06165	*H35*	*Triticum aestivum*	3BS	SNP/SSR	(Zhao et al., 2020)
SD06165	*H36*	*Triticum aestivum*	7AS	SNP/SSR	(Zhao et al., 2020)
KS99WGRC4	*Hdic*	*Triticum turgidum* L. Subsp. *dicoccon* (Emmer wheat)	1AS	SSR	(Liu et al., 2005)
Faraj	*QHara.icd-6B*	*Triticum araraticum*	6BS	KASP	(Bassi et al., 2019)
Nassira	*QH.icd-2A*	*Triticum dicoccum*	2AL	KASP	(Bassi et al., 2019)

Chapter 2 - Fine mapping of the Hessian fly resistance gene *H34*

2.1 Introduction

Hessian fly is one of the most serious insect pests in wheat (*Triticum aestivum* L.) in most wheat-growing regions in the U.S. and many other countries and can cause significant economic losses to wheat farmers (Ratcliffe et al., 2000; Stuart et al., 2012). Deployment of HF resistance genes proven to be the most effective and economical strategy for controlling the pest (Berzonsky et al., 2003). Therefore, a better understanding of genes controlling HF resistance can accelerate wheat improvement for HF resistance (Zhang et al., 2017).

To date, 37 HF resistance genes have been identified (Bassi et al., 2019; Li et al., 2013; Zhao et al., 2020). Many of those genes, as well as several quantitative trait loci (QTLs), have been mapped to specific wheat chromosomes (Carter et al., 2014; Hao et al., 2013; Li et al., 2013). However, some of these were assigned to chromosomes using cytogenetic tools without linked markers available. For these genes mapped using molecular markers, most of the markers are still far from the genes and recombination between markers and target genes reduce the selection accuracy in marker-assisted breeding. In addition, most of gene linked markers identified by linkage mapping using biparental populations are usually population specific and may not be useful for diverse breeding populations. Therefore, diagnostic markers are urgently needed for precise selection of the target genes in breeding programs (Nair et al., 2015).

To date, only a few HF resistance genes have been finely mapped to relatively small physical intervals. *QHf.osu-1Ad* in ‘Duster’ was mapped to a relatively short genetic interval based on recombination frequency within the QTL interval (Li et al., 2015). *QHf.hwwg-3B* from ‘Overland’ was precisely mapped to a 2.32 Mb interval (2,479,314–4,799,538 bp) on

chromosome 3B using near-isogenic lines (NILs) that have recombination within the QTL interval (Xu et al., 2021). *QHf.hwwg-6BS* in 'Chokwang' was delimited to a 4.75 Mb physical interval between 6,028,601 bp and 10,779,424 bp on chromosome arm 6BS after evaluation homozygous recombinants within the QTL region (Zhang et al., 2021). The QTL *QHara.icd-6B* from durum wheat was aligned to an interval between 5.276 and 13.789 Mb (Bassi et al., 2019). However, none of those genes have been finely mapped to such a short interval that reliable candidates underlining the genes of interest can be identified, and diagnostic gene markers are not available for these genes.

H34 is a major gene for HF resistance derived from 'Clark'. *H34* was previously mapped to chromosome arm 6BS using wheat 90K SNP arrays and was located between 10.25 and 18.15 Mb based on International Wheat Genome Sequence Consortium (IWGSC) reference sequence version 2.1 (RefSeq. v2.1; Li et al., 2013; IWGSC, 2018). The objective of this study was to finely map *H34* using heterogeneous inbred families (HIFs) developed from a cross between the two RILs (RIL115-S and RsIL118-R) carrying the contrasting alleles at *H34.* The two RILs were selected from a F_{12} population of Ning7840 x Clark. We delimited the QTL interval to approximately 1.0 Mb to predict the candidate genes underlying this QTL and developed diagnostic markers using the candidate genes from the QTL interval for marker-assisted selection (MAS) in breeding.

2.2 Materials and Methods

2.2.1 Materials

The two RILs differing in presence (RIL118-R) and absence (RIL115-S) of *H34* were selected from the (Ning7840 × Clark) F_{12} RIL population for development of the fine mapping population. These RILs were crossed to generate 286 F_3 lines for screening of recombinants using the two flanking KASP markers HF2823341 and HF15879400 to develop HIFs for fine mapping. A total of 10,000 plants derived from the HIFs were screened in six greenhouse cycles (2017-2021) for fine mapping.

A diversity panel of 203 U.S. winter wheat accessions (AM203) (Zhang et al., 2010) was evaluated for HF resistance and genotyped using the selected KASP markers to validate the effectiveness of the newly developed markers in predicting presence of *H34*.

2.2.2 Evaluation of Hessian fly resistance

The F_3 mapping population and the progeny derived from the HIFs, two parents and four controls, Ike (*H3*), 'Caldwell' (*H6*), *H13*, and 'Karl 92' (susceptible control) were infested by HF biotype GP in fall 2017 at Kansas State University, Manhattan, KS. The HF data of F_3 progeny from corresponding F_2 plants in the mapping population were used to infer the phenotypes of each F_2 plants. Later, the selected recombinants with recombination within the QTL interval from different HIFs were evaluated for HF resistance from 2017 to 2021. In each experiment, 20 seeds of each recombinant lines were uniformly planted in spaced rows (24 half-rows per flat) in flats (52 x 36 x 10 cm) containing a mixture (1:1) of soil and vermiculite in greenhouse at 18 ± 1 ^{o}C with 14:10 h (light: dark) photoperiod. Seedlings at the one leaf stage were infested by confining

~200 newly mated HF females in each flat within a cheesecloth tent. Three weeks after infestation, the infested seedlings were scored for responses to HF infestation. Susceptible plants were stunted with dark green leaves and harbored live larvae, whereas resistant plants grew normally with light green leaves and had dead larvae between the leaf sheaths. In some cases, the plants with smaller sized live larvae than in susceptible plants were also scored as resistant (Fig. 2.1)

2.2.3 DNA extraction and KASP markers development

The seeds from F_3, selected heterozygous recombinants and NILs from different HIFs were planted in a greenhouse for genotyping. For DNA extraction, the leaf tissues of these plants at two-leaf stage were collected into 1.1- mL 96-deep-well plates with a 3-mm stainless bead in each well, freeze-dried for 48 h, and then ground for 3 min at 30 cycles per sec to a fine powder in a Mixer Mill (Retsch GmbH, Germany). DNA was isolated using the cetyl trimethyl ammonium bromide (CTAB) method (Bai et al., 1999).

In addition, four pairs of recombinant NILs from four different HIFs and parents were selected for GBS analysis to enrich SNPs in the *H34* region. In brief, CTAB-extracted genomic DNA was digested with *HF-PstI* and *MspI*, ligated with barcoded adaptor and Y common adaptor, and amplified by PCR (Poland et al. 2012). PCR product was sequenced using Ion Proton next generation sequencer (ThermoFisher Scientific, Waltham, MA) The GBS SNPs were called using the reference-based pipeline v2.0 in TASSEL (v5.2.63) by comparing the tags between parents in the *H34* interval.

The GBS SNPs in the QTL region were converted to KASP markers. Chromosome-specific KASP primers were designed based on the sequence from IWGSC RefSeq v1.0

(IWGSC 2018) using PolyMarker pipeline (http://www.polymarker.info/). Two tail sequences, GAAGGTGACCAAGTTCATGCT and GAAGGTCGGAGTCAACGGATT, were added to the 5'-end of both forward primers to match with the FAM- and HEX-fluorescence-dye-labeled sequences in the KASP reaction mix, respectively. KASP assays were performed in a Veriti thermal cycler (Applied Biosystems, CA) using a 4 µL reaction mix consisting of 1.94 µL of 2× PACE^{TM} Genotyping Master Mix (www.3crbio.com), 0.06 µL primer mix and 2 µL DNA. PCR products were scanned in a FLUOstar® Omega microplate reader (BMG Labtech Inc., Cary, NC), and data were analyzed using Klustercaller software v3.4.1.39 (LGC group, Teddington, UK).

KASP markers were first screened for polymorphism between the parents, and the polymorphic markers were then used to genotype the recombinants and *H34* NILs (Fig. 2.4). Selected KASP markers were validated in the AM203 panel to confirm their selection accuracy.

2.2.4 Recombinant isolation and high-resolution map construction

HF2823341 (2,823,341 bp) and HF15879400 (15,879,400 bp) were selected as the flanking markers for screening of the F_3 progeny to identify recombinants in the *H34* region and 14 HIFs were selected for further marker enrichment to shorten the QTL region (Table 2.1). Initially, 800 plants from different HIFs were screened with the flanking markers HF2823341 and HF15879400. Heterozygous plants showing recombination events within the target region were self-pollinated to generate seeds for next cycles of recombinant selection. In each of the subsequent cycles, new heterozygous plants were screened using newly developed markers inside the flanking markers. Total 159 homozygous plants from the HIFs with

different recombination within the target region were selected as the *H34* NILs and evaluated for HF resistance for fine mapping.

In each of following cycles, approximately 600 seeds from the newly selected heterozygous recombinants between new flanking markers HF5008058 and HF10779424 from the previous cycles were screened with the flanking markers to identify homozygous recombinants for phenotyping and new heterozygous recombinants for next cycle of recombinant screening. The selected heterozygous and homozygous recombinants with possible new recombination in *H34* region were genotyped with nine newly developed markers from GBS-SNPs identified by comparing of the GBS sequences between selected pairs of NILs and phenotyped to further delimit the *H34* region to a smaller interval.

2.2.5 Candidate gene prediction

To identify the putative candidate genes in the *H34* candidate interval, the flanking SNP-carrying sequences were blasted against IWGSC Refseq v2.1 to define possible physical interval of *H34*. The annotated high confidence genes in the flanking region of IWGSC RefSeq. v2.1 (IWGSC 2018) were selected as putative candidates for *H34*.

2.2.6 RNA sample preparation and sequencing

The four sets of NILs (which were used for GBS-SNP analysis) were used for RNA-seq analysis. The cheesecloth tent was removed about one week after HF infestation when larvae hatched from eggs, and they crawled down to the base of plants. Seedlings were checked for infestation status at 12 hours (12-h) and 48 hours (48-h) after HF larvae were first seen at the base of a plant. The inner leaf sheath of second leaf that were attacked by larvae were sampled at

12-h and 48-h. Five plants were sampled by cutting a 1-cm leaf sheath from the base of each larvae-attached plant for RNA extraction. RNA was extracted using Trizol reagent (ThermoFisher Scientific, Waltham, MA) following the manufacturer's protocol. All the RNA samples were treated with DNase1 to avoid possible DNA contamination. The RNA quality was first checked in a 1% denaturing agarose gel and further confirmed using ND-2000 (NanoDrop Technology). Only high-quality RNA samples with RNA integrity number (RIN) > 8.5 were used to generate the sequence library**.**

2.2.7 RNA-seq library preparation and data analysis

RNA-seq was carried out using Illumina HiSeq™ 2000 sequencer in Novogene Corporation Inc. (CA, USA). The sequence reads were aligned to the IWGSC RefSeq. v2.1 (IWGSC 2018) and the differentially expressed genes (DEGs) were analyzed by comparing the HF infested resistant samples with the HF infested susceptible samples (control) using DEGseq R package (Mortazavi et al. 2008). The difference in gene expression levels was considered statistically significant at a false discovery rate (FDR) < 0.05. The raw expression values were converted into log2 change ≤ -1 or ≥ 1.

2.2.8 Cloning of full-length genomic and coding sequences of DEGs

All the DEGs under the *H34* interval were selected for further sequencing to isolate full-length of these genes for marker development. The full-length genomic DNA sequences of the selected DEGs were amplified from the two parents with contracting alleles at *H34* by PCR using the gene-specific primers (Table 2.4). To amplify the full-length coding sequences (CDS) of the DEGs, total RNA was reverse-transcribed to cDNA using SuperScript IV First-Strand Synthesis System and an oligo$(dT)_{20}$ primer following the manufacturer's instruction

(Invitrogen, USA). The full-length CDS of the DEGs were amplified by PCR using the gene-specific primers (Table 2.4). The PCR products of the full-length genomic and coding sequences of the DEGs from the parents were cloned into the vector from NEB PCR cloning kit (New England Biolabs, Ipswich, MA, USA). The inserted fragments in the construct were verified by sequencing in an ABI 3730 DNA Sequencer (ThermoFisher Scientific). The sequence data were assembled and aligned using Chromas v.2.6.6 (https://technelysium.com.au/wp/) and DNAMAN v.10.0 (https://www.lynnon.com/). Gene structure was predicted using the SnapGene Viewer program (https://www.snapgene.com).

2.3 Results

2.3.1 Construction of a linkage map and development of KASP markers in *H34* region

A linkage map of chromosome 6B was reconstructed by using the 18 GBS-SNPs in the QTL region based on the IWGSC RefSeq. v2.1 (Fig. 2.3). Using this map, *H34* was remapped to a 0.68-3.11 cM interval between markers HF6028581 and HF7398225 and the two flanking markers were mapped to 6,028,581 bp and 7,398,225 bp on chromosome 6BS, respectively, based on IWGSC RefSeq. v2.1. Hence, *H34* is delimited to a 1,369,644 bp physical interval (Fig. 2.3).

2.3.2 Fine mapping to delimit candidate interval of *H34*

To verify the accuracy of SNPs data, 65 KASP were designed from the SNP data within the *H34* region generated by 90K-SNP chips, GBS and exome capture. Among them, 18 KASP assays (Table 2.2) showed polymorphisms between parents and among the recombinants selected

from different HIFs. KASP assays separated recombinants into three clusters, with one cluster on Y-axis representing plants carrying the resistance allele at *H34* from Clark, one cluster on X-axis representing the susceptibility allele from Ning 7840, and the cluster between X and Y-axis representing heterozygotes carrying both alleles (Fig. 2.4).

A physical map was constructed based on the physical positions of the SNPs in the *H34* candidate region in the IWGSC RefSeq v2.1 (Fig. 2.5). This map showed perfect co-linearity with the linkage map (Fig. 2.3) and was used to determine the physical location of *H34*. Based on this map, about 800 heterozygous plants selected from F_3 progeny were screened with the flanking markers HF2823341 and HF15879400, which identified 14 HIFs. Recombinant plants from those HIFs were self-pollinated and about 159 homozygous recombinant and non-recombinant sister lines were screened as NILs with seven other KASP markers and evaluated for HF resistance in the greenhouse experiments.

Two critical recombination from one HIF were identified between HF5008058 (homozygous Ning 7840 allele at 5,008,058 bp) and HF10252606 (homozygous Clark allele at 10,252,606 bp) (Fig. 2.5a). Recombinant in the NIL 58_02-15 was identified between HF5008058 (homozygous Ning 7840 allele at 5,008,058 bp) and HF5501616 (homozygous Clark allele at 5,501,616 bp) showed a high HF resistance score (90%), suggesting that KASP-HF5008058 is the left flanking marker for *H34*. Another NIL 58_02-17 carries another recombination between SNP HF5501616 (homozygous Ning 7840 allele at 5,501,616 bp) and HF10252606 (homozygous Clark allele at 10,252,606 bp) were identified. This line showed complete susceptibility to HF, suggesting that HF10252606 is the right flanking marker for *H34*. These data delimited *H34* to a 5.0 Mb region on chromosome 6B (Fig. 2.5a).

A second round of recombinant screening was conducted for 600 HIF's progeny from the first cycle of fine mapping using the flanking markers HF5008058 and HF10252606 to identify more heterozygous and homozygous recombinants in this interval. In the following three recombinant screen cycles, recombinant plants were self-pollinated and about 75 homozygous recombinant and non-recombinant sister lines were screened as NILs with nine new KASP markers that were developed from the SNP sequences of GBS from 4 pairs of NILs. These recombinant lines were also evaluated for HF resistance in two greenhouse experiments. Two new key recombination occurred between markers HF6028581 (6,028,581 bp on 6B) and HF7398225 (7,398,225 bp on 6B) were found in the lines 58_17-14-5 and 58_17-23-13. NIL 58_17-14-5 carries the homozygous Ning7840 allele at SNP HF7118763 and homozygous Clark allele at SNP HF7398225 and showed 100% susceptibility to HF, suggesting HF7398225 as the right flanking marker. Another recombinant NIL 58_17-23-13 contains homozygous Clark allele at HF6900547 (6,900,547 bp) and homozygous Ning 7840 allele at HF6028581 (6,028,581 bp) and showed 100% resistance to HF, suggesting HF6028581 is more likely the left flanking marker for *H34*. The new recombinants narrowed the physical interval of *H34* down to 1.3 Mb between 6,028,581 bp and 7,398,225 bp on 6B based on IWGSC RefSeq. v2.1 (Fig. 2.5b).

2.3.3 Putative annotated genes underlying *H34*

The fine mapping delimited *H34* to a 1.37 Mb interval between 6,028,581 to 7,398,225 bp on chromosome 6BS. In this interval, 16 high confidence (HC) genes were annotated (Fig. 2.6) and five of them including three encoding leucine-rich repeat receptor-like proteins (TraesCS6B03G0019400, TraesCS6B03G0019500 and TraesCS6B03G0020400), one encoding a kinase family protein (TraesCS6B03G0022400) and one encoding an ankyrin rich family protein (TraesCS6B03G0022700) could be the candidate genes for *H34* as they all have been

reported to be involved in resistance against insects and pathogens in cereal crops. In addition, six other genes (TraesCS6B03G0021800, TraesCS6B03G0022100, TraesCS6B03G0022200, TraesCS6B03G0022300, TraesCS6B03G0022500, and TraesCS6B03G0022600) encode F-Box protein (Table 2.3).

2.3.4 Identification of differentially expressed genes in *H34* interval

To identify possible candidate genes for marker development, RNA-sequencing was conducted using the parents and four pairs of NILs carrying the contrasting *H34* alleles. Three genes (TraesCS6B03G0019400, TraesCS6B03G0019500 and TraesCS6B03G0022700) in this interval showed differential at 12-h and 48-h of larval attack on the wheat seedlings (Fig. 2.6).

The three DEGs were sequenced to get the full-length genomic DNA from the parents using the gene specific primers that were designed based on RNA-seq sequencing data. (Table 2.4). Comparison of sequences between two parents identified SNPs in all three genes with 20 SNPs from TraesCS6B03G0019400 (5,965 bp), 26 SNPs from TraesCS6B03G0019500 (12,046 bp) and 11 SNPs from TraesCS6B03G0022700 (2,842 bp). These SNPs locates in 3'UTR, 5'UTR, exons, and intron regions of these DEGs (Fig. 2.7). All these SNPs were converted into KASP markers to develop gene specific functional markers.

2.3.5 KASP markers validation in diversity panel

The four flanking KASP markers (HF5501616, HF10252606, HF6028581 and HF7118763) generated from fine mapping were validated in the AM203 diversity panel. Among them, marker HF6028581 showed significant difference in HF resistance ($p = 0.00002$) between the two allelic groups with 11 out of 12 accessions that carry the Clark allele showing resistance

scores of more than 50%. For the other three markers, HF5501616, HF10252606 and HF7398225, the differences between the two allelic groups (resistant and susceptible) were not significant. When the data from all those markers were combined, the detection accuracy was not improved compared to the result from HF6028581 alone (Fig. 2.8).

The SNPs from the three differentially expressed genes (TraesCS6B03G0019400, TraesCS6B03G0019500 and TraesCS6B03G002270) were also converted to KASP markers. After 20 KASP SNPs from TraesCS6B03G0019400 were tested between the two parents, three were polymorphic and were further used to screen AM203 diversity panel. The three markers showed significant difference in HF resistance. The Clark alleles of G0019400-S3 and G0019400-S6 ($p = 0.00002$) presented in 11 accessions that all had the resistance score of more than 50%. This marker identified the same set of HF resistant lines as by marker HF6028581. While G0019400-S4 ($p = 0.00001$) identified 6 accessions with the Clark allele and they all had more than 91% resistance score. Among 26 SNPs in TraesCS6B03G0019500, only two were diagnostic with one from the 3'UTR region (G0019500-3UTRS1) and the other from the second exon (G0019500-Ex2). The two markers identified the same set of accessions as by the markers HF6028581, G0019400-S3 and G0019400-S6 in the AM203 diversity panel (Fig. 2.8).

Only one polymorphic KASP SNP (G0022700-S1) from the third exon of TraesCS6B03G002270 was nearly diagnostic in the AM203 U.S. winter wheat diversity panel. It identified the same set of 11 accessions with Clark allele at G0022700-S1 as these previously identified fine-mapping derived markers in the panel (Fig. 2.8).

2.4 Discussion

2.4.1 Fine mapping of *H34*

Host resistance is the most cost-effective strategy for managing HF in the U.S. and around the world. Resistance to HF is considered an important trait to be incorporated into the new varieties in most of the U.S. breeding programs (Subramanyam et al., 2021; Bassi et al. 2019; Li et al. 2013). Among the 37 reported HF genes, only four genes are mapped on B genome including *H20* on 2B, *H31* and *H35* on 5BS, and *H34* on 6BS. In addition, one QTL *QHara.icd-6B* has also been mapped on 6BS (Bassi et al., 2019). Based on the physical locations of the flanking markers in IWGSC RefSeq v1.0, *H34* was located between 8.409 and 15.134 Mb on 6BS in 'Clark' (Li et al., 2013), whereas *QHara.icd-6B* was delimited to the interval of 0.167–11.299 Mb in a durum wheat (Bassi et al., 2019), therefore, the two QTLs are most likely the same QTL for HF resistance. However, none of them has been delimited to a small interval.

In this study, a physical map for the *H34* candidate region was generated by aligning the SNPs identified from the 90K SNP-chips, exome capture and GBS sequences within and nearby the *H34* interval to the IWGSC RefSeq v2.1. The recombinant NILs were developed from HIFs which were used to narrow down target QTL region to the 1.3 Mb region on chromosome arm 6BS (Fig 2.5). Thus, fine mapping of *H34* to this small interval reduced the number of high confidence genes underlying the *H34* interval to a small set and made it possible for development of diagnostic and near-diagnostic markers, which brought us one step closer to identification of causal gene(s) for *H34* in this region. Marker-screening of the HIFs derived recombinant population to select heterozygous recombinants in the target region for developing NILs is a rapid and effective way to develop fine-mapping populations for gene cloning, which

proven to be effective in many crops (Hao et al., 2013; Merry et al., 2019). However, fine mapping for HF resistance QTLs or genes has not been reported to date in wheat.

2.4.2 Prediction of *H34* candidate genes

Extensive work has not been done to identify and clone insect resistance genes in wheat due to its big genome size, polyploidy, and gene sequence redundancy. However, many wheat databases are being recently updated with improved gene annotations, making it possible for genome-wide identification and comprehensive analysis of gene families in wheat (Jiang et al., 2020). In the present study, 16 high-confidence genes were found in the *H34* interval after blasting the flanking marker sequences to the IWGSC RefSeq v2.1 (Table 2.3). Among those genes, TraesCS6B03G0019400, TraesCS6B03G0019500 and TraesCS6B03G0020400 encode disease resistance protein belonging to NBS-LRR family. The NBS-LRR gene families represent the largest family of plant disease resistance (R) genes that play an important role in defending plants from numerous pathogens and pests. It was speculated that receptor-like genes as membrane-related R proteins are the first "guard" to monitor or physically interact with its native partners or the exotic avirulent proteins from environmental invaders, for example, the salivary proteins secreted by *GB-I* (Zhang et al., 2022). In the current study, among the three DEGs, TraesCS6B03G0019400 and TraesCS6B03G0019500 carrying NBS-LRR domain showed significant differential expression between susceptible and resistance NILs and parental lines in RNA-seq analysis. Plant resistance R-genes have been cloned through map-based cloning from several plant species, which confer resistance to different biotic challenges including fungi, bacteria, insects and viruses (Osuna-Cruz et al., 2018; Periyannan et al., 2013). More recently, several wheat stripe rust (*Puccinia striiformis* f. sp. *Tritici*) resistance genes have been cloned belonging to the NBS-LRR protein family (Hu et al., 2021; Zhang et al., 2021). Also some wheat

leaf rust resistance gene were been cloned in wheat carrying the NBS-LRR domain (Timothy et al., 2021). These data support that one of the differentially expressed R-genes in the finely mapped *H34* region is the most likely candidate gene for HF resistance.

TraesCS6B03G0022700 was another important DEG identified within the *H34* region. It encodes an ankyrin family protein that confers race-specific resistance against plant pathogens. TraesCS6B03G0022700 shares a similar structure and function with *Lr14a,* a recently cloned wheat leaf rust resistance gene encoding a membrane-localized protein containing 12 ankyrin (ANK) repeats (Kolodziej et al., 2021). In *Arabidopsis*, a gene for *Arabidopsis* cell death (*ACD6*) containing nine ankyrin repeats acts as an activator of the defense pathway against bacterial pathogens and plays a role in program cell death by regulating salicylic acid pathway (Dong, 2004). These studies suggest TraesCS6B03G0022700 could also be a putative candidate for *H34.* Further studies need to be conducted to validate the functions of these DEGs on HF resistance.

Six genes encoding F-box proteins are identified in the *H34* region. The genes for F-box proteins have been reported to play a fundamental role in response to phytohormones that regulate pollen fertility, plant root growth, wounding and healing, and also defense against pathogens and insects (Wang et al., 2005). Furthermore, another major class of disease resistant genes are protein kinases. Protein kinases have been demonstrated to regulate plant development, adaptation to diverse environmental conditions and biotic stresses through transmembrane signaling. In this study, the gene TraesCS6B03G0022400 carries kinase domain and belongs to this family, however, RNA-seq data did not show differential expression of these genes between the two parents at 12-h and 48-h after HF infestation.

2.4.3 Development of tightly linked diagnostic markers for *H34*

Although 37 genes and some QTLs for HF resistance have been reported, but DNA markers are unavailable for selecting these genes or QTLs in diverse genetic backgrounds in breeding. Recently, KASP markers have been widely used in wheat gene tagging and breeding, and developed for several genes or QTLs, including *H32* (Tan et al., 2017), *QHara.icd-6B* (Bassi et al., 2019), *h4* (Niu et al., 2020), *H7* (Liu et al., 2020), and *H35* and *H36* (Zhao et al., 2020). However, most of those markers were developed based on the sequence tags loosely linked to target genes and they are population-specific and have not been validated in a diversity panel. Development and use of the diagnostic markers directly from HF resistance genes or tightly linked markers may provide more accurate selection and facilitate deployment of these genes in new wheat cultivars (Bassi et al., 2019; Pernet et al., 1999).

In this study, several SNPs in the *H34* interval were successfully converted into KASP markers and four markers (Table 2.2) were validated on a panel of 203 U.S. winter wheat elite breeding lines. Among them, HF6028581 showed significant difference in HF resistance between the two groups carrying the contrast alleles at HF6028581 (Fig. 2.3) and showed higher prediction accuracy than other markers because 11 of 12 accessions in the group carrying the positive allele were highly HF resistant (>50% resistance). The higher prediction accuracy of HF6028581 also suggests that the *H34* causal gene is closer to HF6028581 than to another marker HF7118763 (Fig 2.2).

Besides development of diagnostic markers from DNA sequence tags generated through genome-wide sequencing, we also explored diagnostic markers using sequence information from RNA-seq. RNA-seq represents a powerful tool for transcriptomic profiling, which provides rapid

access to a collection of expressed gene sequences (Hao et al., 2013; Wang et al., 2009). Moreover, RNA-seq data can also provide a rich resource for development of potential functional markers. For instance, 5,344 In-Del markers were developed from the transcriptomic data of both Sudan grass S722 and sorghum *Tx623B* (Li et al., 2016), and 1,276 polymorphic SSRs and 261,000 SNPs were identified from transcriptomic data of large yellow croaker (Xiao et al., 2015). In the current study, three DEGs were identified in the *H34* interval and considered candidate genes for *H34*. By comparing RNA-seq data between resistant and susceptible NILs, we designed gene-specific primers to develop the functional markers for *H34.* Six KASP SNPs were diagnostic in the U.S. winter wheat diversity panel. Among them, G0019400-S6 showed 99% prediction accuracy and five of the six accessions with the Clark allele had resistance scores >91%, while the five other markers (G0019400-S3, G0019400-S4, G0019500-Ex2, G0019500-3UTRS1 and G00227700-S1) identified the same set of resistant genotypes as by the marker HF6028581 with resistance scores >50% (Fig. 2.8). The six markers can be used in wheat breeding to increase the selection efficiency of *H34*.

Taken together, we finely mapped the H34 to the 1.3 Mb region on 6BS, identified three putative candidate genes based on expressed annotated high confidence genes in the region and RNA-seq data. We developed seven markers based on the difference in genomic and RNA sequence from one of the putative *H34* candidates between NILs. These markers have been validated to be diagnostic for *H43* in the U.S. winter wheat diversity panel and proved to be superior to previously reported markers with improved accuracy and high flexibility. Therefore, these markers should be useful for marker-assisted selection of *H34* to improve HF resistance in wheat breeding programs.

2.5 References

Ashikari, M., Sakakibara, H., Lin, S., Yamamoto, T., Takashi, T., Nishimura, A., Angeles, E. R., Qian, Q., Kitano, H., & Matsuoka, M. (2005). Cytokinin oxidase regulates rice grain production. *Science*, *309* (5735), 741–745. https://doi.org/10.1126/science.1113373

Aslam, M. S., Xue, P. H., Bashir, S., Alfakhri, Y., Nurunnabi, M., & Nguyen, V. C. (2021). Assessment of rice and wheat production efficiency based on data envelopment analysis. *Environmental Science and Pollution Research*, *28*(29), 38522–38534. https://doi.org/10.1007/s11356-021-12892-z

Bai, G., Kolb, F. L., Shaner, G., & Domier, L. L. (1999). Amplified fragment length polymorphism markers linked to a major quantitative trait locus controlling scab resistance in wheat. *Phytopathology*, *89*(4), 343–348. https://doi.org/10.1094/PHYTO.1999.89.4.343

Bassi, F. M., Brahmi, H., Sabraoui, A., Amri, A., Nsarellah, N., Nachit, M. M., Al-Abdallat, A., Chen, M. S., Lazraq, A., & El Bouhssini, M. (2019). Genetic identification of loci for Hessian fly resistance in durum wheat. *Molecular Breeding*, *39*(2), 24. https://doi.org/10.1007/s11032-019-0927-1

Carter, A. H., Cambron, S. E., Ohm, H. W., Bosque-Pérez, N. A., & Kidwell, K. K. (2014). Hessian fly (*Mayetiola destructor* [Say]) resistance identified on chromosome 1AS in the spring wheat (*Triticum aestivum* L.) cultivar 'Louise.' *Crop Science*, *54*(3), 971–981. https://doi.org/10.2135/cropsci2012.11.0636

Dong, X. (2004). The role of membrane-bound ankyrin-repeat protein *ACD6* in programmed cell death and plant defense. *Science*. *2004*(221). https://doi.org/10.1126/stke.2212004pe6

Hao, Y., Cambron, S. E., Chen, Z., Wang, Y., Bland, D. E., Buntin, G. D., & Johnson, J. W. (2013). Characterization of new loci for Hessian fly resistance in common wheat. *Theoretical and Applied Genetics*, *126*(4), 1067–1076. https://doi.org/10.1007/s00122-012-2037-x

Timothy, H., Zhang, J., Li, H., Narayana, U., Li, J., Park, R., Hoxha, S., McIntosh, R., Lagudah, E., & Zhang. P. (2021). Wheat leaf rust resistance gene *Lr13* is a specific *Ne2* allele for

hybrid necrosis. *Molecular Plant*, *14*(7), 1025-1028.

Hossain, A., Skalicky, M., Brestic, M., Maitra, S., Ashraful Alam, M., Syed, M. A., Hossain, J., Sarkar, S., Saha, S., Bhadra, P., Shankar, T., Bhatt, R., Kumar Chaki, A., EL Sabagh, A., & Islam, T. (2021). Consequences and mitigation strategies of abiotic stresses in wheat (*Triticum aestivum* L.) under the changing climate. *Agronomy*, *11*(2), 241. https://doi.org/10.3390/agronomy11020241

Hu, Y., Tao, F., Su, C., Zhang, Y., Li, J., Wang, J., Xu, X., Chen, X., Shang, H., & Hu, X. (2021). NBS-LRR gene *TaRPS2* is positively associated with the high-temperature seedling plant resistance of wheat against *Puccinia striiformis* f. sp. *tritici*. *Phytopathology*, *111*(8), 1449–1458. https://doi.org/10.1094/PHYTO-03-20-0063-R

Jiang, W., Geng, Y., Liu, Y., Chen, S., Cao, S., Li, W., Chen, H., Ma, D., & Yin, J. (2020). Genome-wide identification and characterization of *SRO* gene family in wheat: Molecular evolution and expression profiles during different stresses. *Plant Physiology and Biochemistry*, *154*, 590–611. https://doi.org/10.1016/j.plaphy.2020.07.006

Kolodziej, M. C., Singla, J., Sánchez-Martín, J., Zbinden, H., Šimková, H., Karafiátová, M., Doležel, J., Gronnier, J., Poretti, M., Glauser, G., Zhu, W., Köster, P., Zipfel, C., Wicker, T., Krattinger, S. G., & Keller, B. (2021). A membrane-bound ankyrin repeat protein confers race-specific leaf rust disease resistance in wheat. *Nature Communications*, *12*(1), 956. https://doi.org/10.1038/s41467-020-20777-x

Ladejobi, O., Salaudeen, M. T., Lava Kumar, P., Menkir, A., Adesoye, A., Atiri, G., & Gedil, M. (2018). Mapping of QTLs associated with recovery resistance to streak virus disease in maize. *Annals of Agricultural Sciences*, *63*(1), 115–121. https://doi.org/10.1016/j.aoas.2018.05.006

Li, C., Chen, M., Chao, S., Yu, J., & Bai, G. (2013). Identification of a novel gene, *H34*, in wheat using recombinant inbred lines and single nucleotide polymorphism markers. *Theoretical and Applied Genetics*, *126*(8), 2065–2071. https://doi.org/10.1007/s00122-013-2118-5

Li, G., Wang, Y., Chen, M.-S., Edae, E., Poland, J., Akhunov, E., Chao, S., Bai, G., Carver, B.

F., & Yan, L. (2015). Precisely mapping a major gene conferring resistance to Hessian fly in bread wheat using genotyping-by-sequencing. *BMC Genomics*, *16*(1), 108. https://doi.org/10.1186/s12864-015-1297-7

Li, J., Wang, L., Zhan, Q., Liu, Y., & Yang, X. (2016). Transcriptome characterization and functional marker development in sorghum sudanense. *PLOS ONE*, *11*(5), e0154947. https://doi.org/10.1371/journal.pone.0154947

Liu, G., Liu, X., Xu, Y., Bernardo, A., Chen, M., Li, Y., Niu, F., Zhao, L., & Bai, G. (2020). Reassigning Hessian fly resistance genes *H7* and *H8* to chromosomes 6A and 2B of the wheat cultivar 'Seneca' using genotyping-by-sequencing. *Crop Science*, *60*(3), 1488–1498. https://doi.org/10.1002/csc2.20148

Merry, R., Butenhoff, K., Campbell, B. W., Michno, J., Wang, D., Orf, J. H., Lorenz, A. J., & Stupar, R. M. (2019). Identification and fine-mapping of a soybean quantitative trait locus on chromosome 5 conferring tolerance to iron deficiency chlorosis. The Plant Genome, 12, 190007. https://doi.org/10.3835/plantgenome2019.01.0007

Nair, S. K., Babu, R., Magorokosho, C., Mahuku, G., Semagn, K., Beyene, Y., Das, B., Makumbi, D., Lava Kumar, P., Olsen, M., & Boddupalli, P. M. (2015). Fine mapping of *Msv1*, a major QTL for resistance to Maize streak virus leads to development of production markers for breeding pipelines. *Theoretical and Applied Genetics*, *128*(9), 1839–1854. https://doi.org/10.1007/s00122-015-2551-8

Niu, F., Xu, Y., Liu, X., Zhao, L., Bernardo, A., Li, Y., Liu, G., Chen, M.-S., Cao, L., Hu, Z., Xu, X., & Bai, G. (2020). The Hessian fly recessive resistance gene *h4* mapped to chromosome 1A of the wheat cultivar 'Java' using genotyping-by-sequencing. *Theoretical and Applied Genetics*, *133*(10), 2927–2935. https://doi.org/10.1007/s00122-020-03642-9

Osuna-Cruz, C. M., Paytuvi-Gallart, A., Di Donato, A., Sundesha, V., Andolfo, G., Aiese Cigliano, R., Sanseverino, W., & Ercolano, M. R. (2018). PRGdb 3.0: a comprehensive platform for prediction and analysis of plant disease resistance genes. *Nucleic Acids Research*, *46*(D1), D1197–D1201. https://doi.org/10.1093/nar/gkx1119

Periyannan, S., Moore, J., Ayliffe, M., Bansal, U., Wang, X., Huang, L., Deal, K., Luo, M.,

Kong, X., Bariana, H., Mago, R., McIntosh, R., Dodds, P., Dvorak, J., & Lagudah, E. (2013). The gene *Sr33*, an ortholog of barley *Mla* genes, encodes resistance to wheat stem rust race Ug99. *Science*, *341*(6147), 786–788. https://doi.org/10.1126/science.1239028

Pernet, A., Hoisington, D., Dintinger, J., Jewell, D., Jiang, C., Khairallah, M., Letourmy, P., Marchand, J. L., Glaszmann, J. C., & González de León, D. (1999). Genetic mapping of maize streak virus resistance from the Mascarene source. II. Resistance in line CIRAD390 and stability across germplasm. *Theoretical and Applied Genetics*, *99*(3–4), 540–553. https://doi.org/10.1007/s001220051267

Ratcliffe, R. H., Cambron, S. E., Flanders, K. L., Bosque-Perez, N. A., Clement, S. L., & Ohm, H. W. (2000). Biotype composition of Hessian Fly (Diptera: Cecidomyiidae) populations from the Southeastern, Midwestern, and Northwestern United States and virulence to resistance genes in wheat. *Journal of Economic Entomology*, *93*(4), 1319–1328. https://doi.org/10.1603/0022-0493-93.4.1319

Schmalenbach, I., March, T. J., Bringezu, T., Waugh, R., & Pillen, K. (2011). High-resolution genotyping of wild barley introgression lines and fine-mapping of the threshability locus *Thresh-1* using the Illumina golden gate assay. *G3 Genes|Genomes|Genetics*, *1*(3), 187–196. https://doi.org/10.1534/g3.111.000182

Schwarting, H. N., Whitworth, R. J., Chen, M. S., Cramer, G., & Maxwell, T. (2016). Impact of Hessian fly, *Mayetiola destructor* , on developmental aspects of hard red winter wheat in Kansas. *Southwestern Entomologist*, *41*(2), 321–330. https://doi.org/10.3958/059.041.0208

Shi, Z., Ren, W., Zhao, Y., Wang, X., Zhang, R., Su, A., Wang, S., Li, C., Wang, J., Wang, S., Zhang, Y., Ji, Y., Song, W., & Zhao, J. (2021). Identification of a locus associated with genic male sterility in maize via EMS mutagenesis and bulked-segregant RNA-seq. *The Crop Journal*, *9*(6), 1263–1269. https://doi.org/10.1016/j.cj.2020.09.007

Sime, S. S., Menkir, A., Adetimirin, V. O., Gedil, M., & Kumar, P. L. (2021). Validation of diagnostic markers for streak virus disease resistance in maize. *Agriculture*, *11*(2), 130. https://doi.org/10.3390/agriculture11020130

Smiley, R. W., Gourlie, J. A., Whittaker, R. G., Easley, S. A., & Kidwell, K. K. (2004).

Economic impact of Hessian fly (Diptera: Cecidomyiidae) on spring wheat in Oregon and additive yield losses with Fusarium crown rot and lesion nematode. *Journal of Economic Entomology*, *97*(2), 397–408. https://doi.org/10.1093/jee/97.2.397

Stuart, J. J., Chen, M.-S., Shukle, R., & Harris, M. O. (2012). Gall midges (Hessian flies) as plant pathogens. *Annual Review of Phytopathology*, *50*(1), 339–357. https://doi.org/10.1146/annurev-phyto-072910-095255

Subramanyam, S., Nemacheck, J. A., Bernal-Crespo, V., & Sardesai, N. (2021). Insect derived extra oral *GH32* plays a role in susceptibility of wheat to Hessian fly. *Scientific Reports*, *11*(1), 2081. https://doi.org/10.1038/s41598-021-81481-4

Tan, C.-T., Yu, H., Yang, Y., Xu, X., Chen, M., Rudd, J. C., Xue, Q., Ibrahim, A. M. H., Garza, L., Wang, S., Sorrells, M. E., & Liu, S. (2017). Development and validation of KASP markers for the greenbug resistance gene *Gb7* and the Hessian fly resistance gene *H32* in wheat. *Theoretical and Applied Genetics*, *130*(9), 1867–1884. https://doi.org/10.1007/s00122-017-2930-4

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, *10*(1), 57–63. https://doi.org/10.1038/nrg2484

William A. B, Ohm H. W., Patterson, F. L., Ding, H., Porter, D. R., & Shanower, T. G. (2003). Breeding wheat for resistance to insects. In W. A. Berzonsky (Ed.), *Plant Breeding Reviews* (Vol. 22, pp. 221–296).

Xiao, S., Han, Z., Wang, P., Han, F., Liu, Y., Li, J., & Wang, Z. Y. (2015). Functional marker detection and analysis on a comprehensive transcriptome of large yellow croaker by next generation sequencing. *PLOS ONE*, *10*(4), e0124432. https://doi.org/10.1371/journal.pone.0124432

Xu, Y., La, G., Fatima, N., Liu, Z., Zhang, L., Zhao, L., Chen, M.-S., & Bai, G. (2021). Precise mapping of QTL for Hessian fly resistance in the hard winter wheat cultivar 'Overland.' *Theoretical and Applied Genetics*, *134*(12), 3951–3962. https://doi.org/10.1007/s00122-021-03940-w

Yang, Q., Zhang, D., & Xu, M. (2012). A sequential quantitative trait locus fine-mapping

strategy using recombinant-derived progeny. *Journal of Integrative Plant Biology*, *54*(4), 228–237. https://doi.org/10.1111/j.1744-7909.2012.01108.x

Zhang, D., Bai, G., Zhu, C., Yu, J., & Carver, B. F. (2010). Genetic diversity, population structure, and linkage disequilibrium in U.S. elite winter wheat. *The Plant Genome*, *3*(2), plantgenome2010.03.0004. https://doi.org/10.3835/plantgenome2010.03.0004

Zhang, H., Huang, J., & Huang, Y. (2022). Identification and characterization of plant resistance genes (*R* genes) in sorghum and their involvement in plant defense against aphids. *Plant Growth Regulation*, *96*(3), 443–461. https://doi.org/10.1007/s10725-022-00797-x

Zhang, H., Mittal, N., Leamy, L. J., Barazani, O., & Song, B. H. (2017). Back into the wild-apply untapped genetic diversity of wild relatives for crop improvement. *Evolutionary Applications*, *10*(1), 5–24. https://doi.org/10.1111/eva.12434

Zhang, J., Hewitt, T. C., Boshoff, W. H. P., Dundas, I., Upadhyaya, N., Li, J., Patpour, M., Chandramohan, S., Pretorius, Z. A., Hovmøller, M., Schnippenkoetter, W., Park, R. F., Mago, R., Periyannan, S., Bhatt, D., Hoxha, S., Chakraborty, S., Luo, M., Dodds, P., … Lagudah, E. S. (2021). A recombined *Sr26* and *Sr61* disease resistance gene stack in wheat encodes unrelated NLR genes. *Nature Communications*, *12*(1), 3378. https://doi.org/10.1038/s41467-021-23738-0

Zhang, L., Xu, Y., Chen, M.-S., Su, Z., Liu, Y., Xu, Y., La, G., & Bai, G. (2021). Identification of a major QTL for Hessian fly resistance in wheat cultivar ‘Chokwang.’ *The Crop Journal*. https://doi.org/10.1016/j.cj.2021.08.004

Zhao, L., Abdelsalam, N. R., Xu, Y., Chen, M.-S., Feng, Y., Kong, L., & Bai, G. (2020). Identification of two novel Hessian fly resistance genes *H35* and *H36* in a hard winter wheat line SD06165. *Theoretical and Applied Genetics*, *133*(8), 2343–2353. https://doi.org/10.1007/s00122-020-03602-3

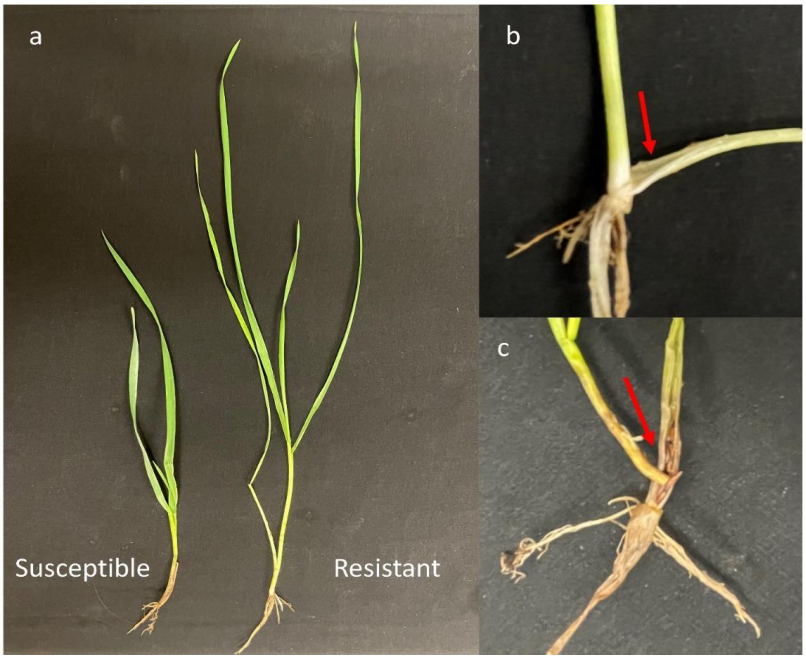

Figure 2.1. Hessian fly (HF)-infested wheat seedlings. (a) A susceptible plant is stunted without newly grown leaves (left) but resistant plant shows normal growth three weeks after infestation; b) HF resistant plant does not carry living larvae; c) HF susceptible plant carries living larvae between leaf sheaths.

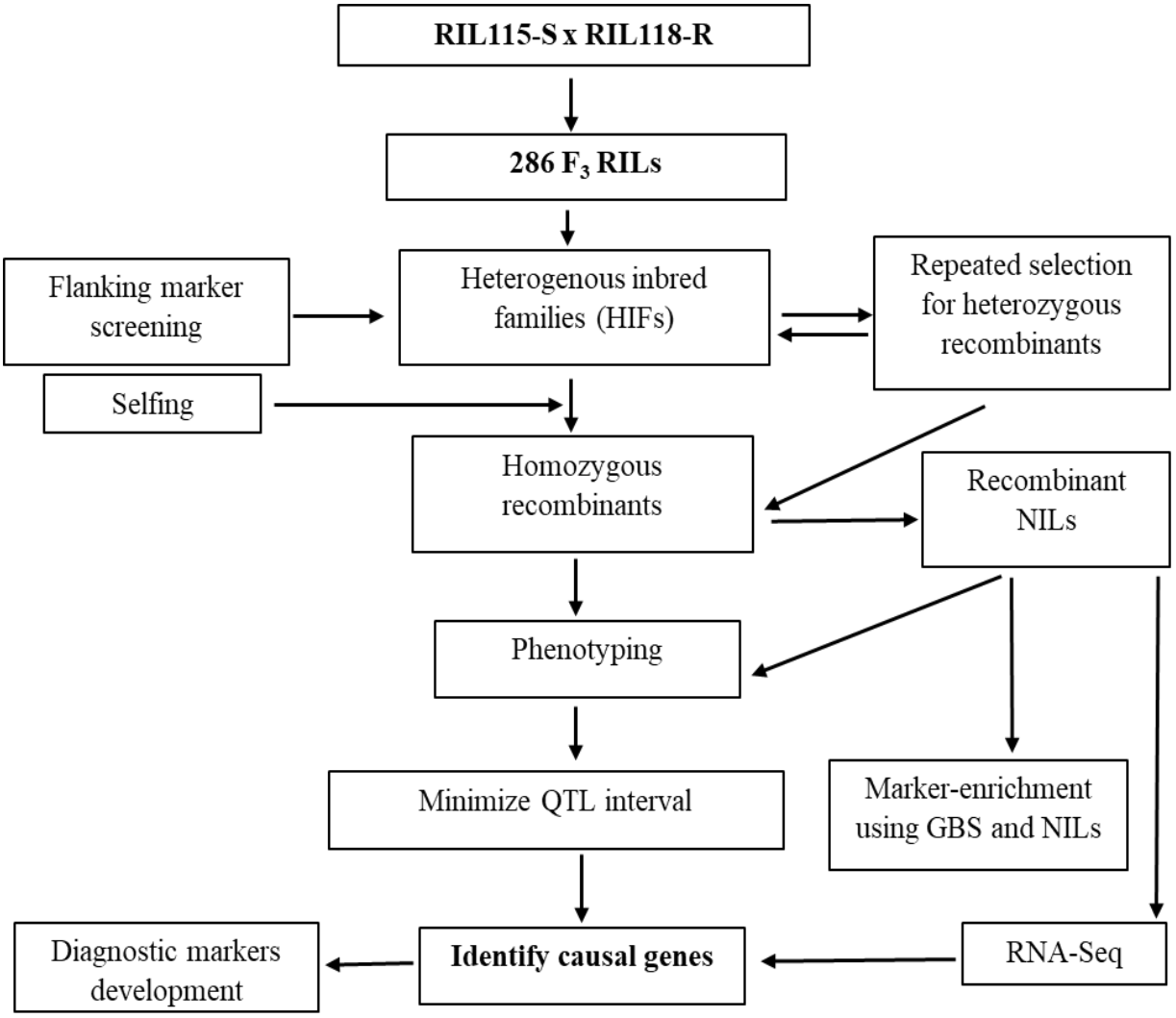

Figure 2.2. The overall procedure of fine mapping and marker development for *H34*. RILs = recombinant inbred lines, NILs = near isogenic lines, GBS = genotyping by sequencing, QTL= quantitative trait locus.

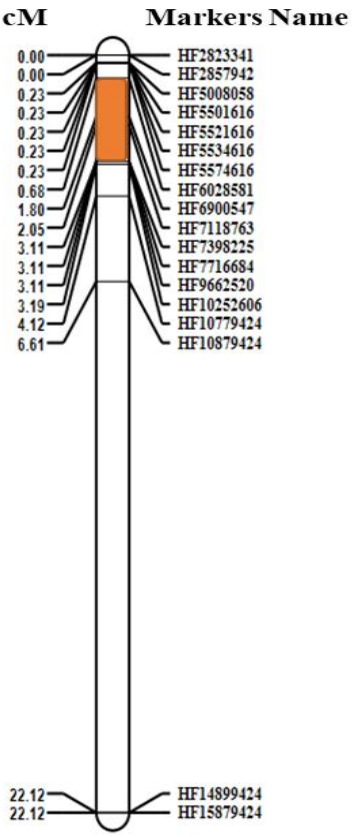

Figure 2.3. A portion of chromosome 6BS linkage map showing the position of *H34* interval and physical position based on IWGSC RefSeq. v2.1. The markers on the right were named by combination of HF (Hessian fly) with the physical position of respective SNPs in IWGSC RefSeq v2.1.

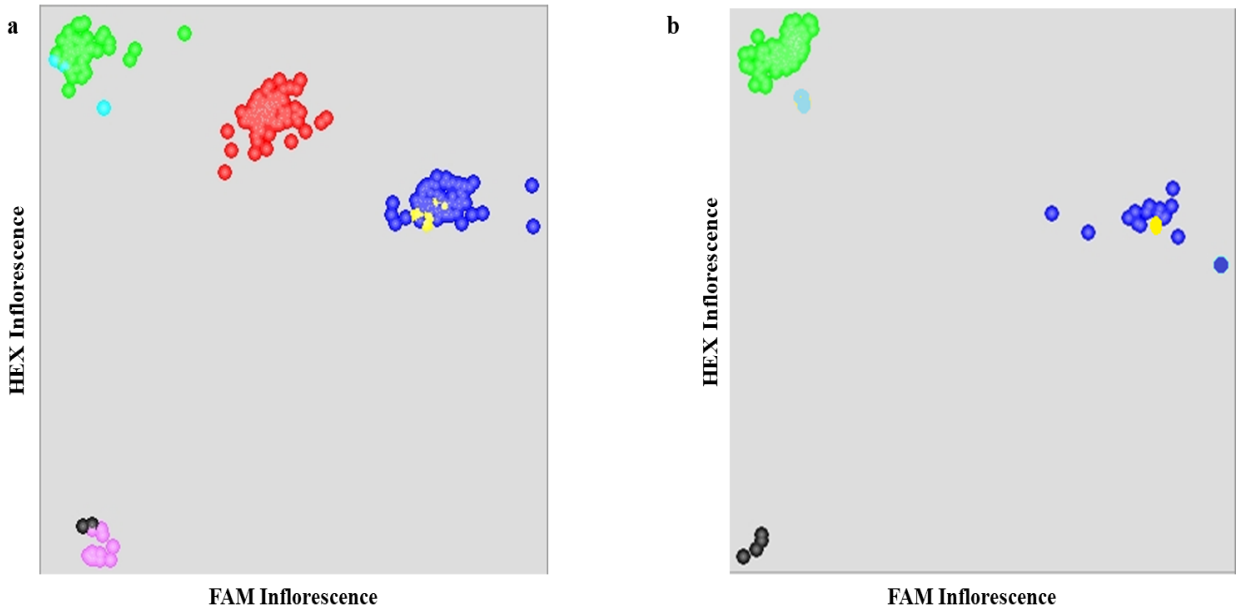

Figure 2.4. KASP assay profiles of selected markers on a) selected homozygous and heterozygous recombinants and b) near-isogenic lines (NILs) selected from the HIFs. Blue dots represent the Ning (susceptibility) allele, green dots represent the Clark (resistance) allele, and red dots represent the heterozygotes with both alleles. Light blue dots in the green cluster are for positive control (RIL118-R) and yellow dots in blue cluster are for negative control (RIL115-S) parent. Pink dots represent the genotypes with poor PCR amplification and black dots represent the water control.

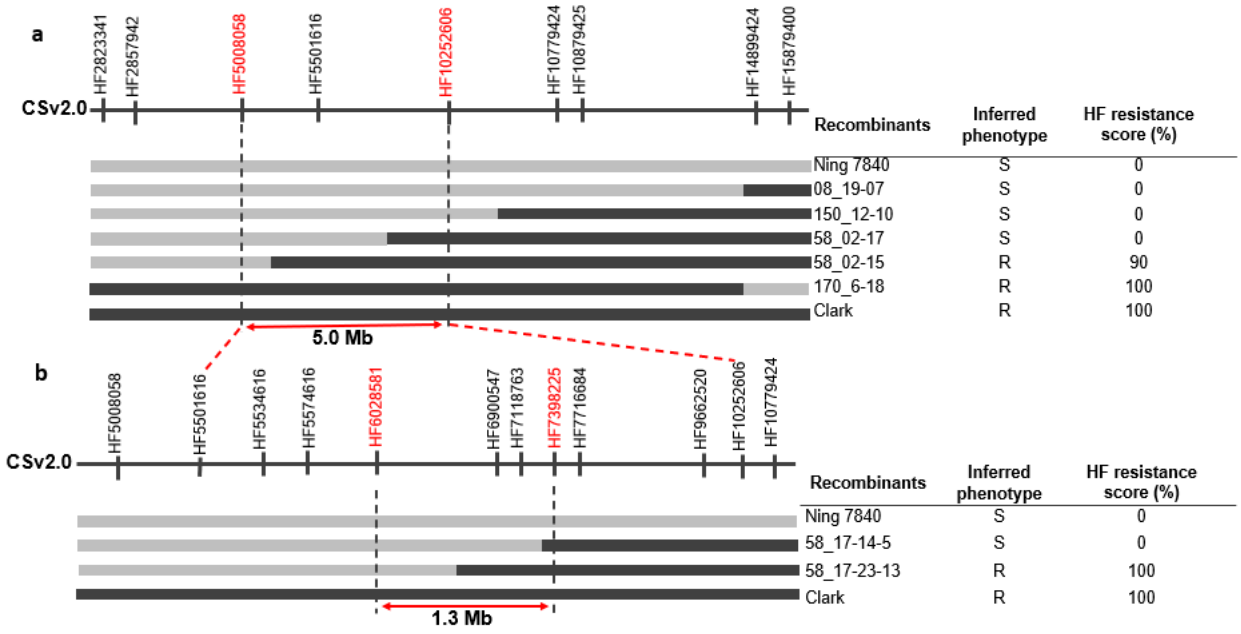

Figure 2.5. Chromosome walking of *H34* using recombinants identified from heterogeneous inbred families (HIFs). a) Recombinants identified in the *H34* region from the first round of screening using nine new markers and NILs selected from different HIFs; b) Recombinants in the *H34* region identified from the second round of screening progeny from HIF-58 using nine new markers identified by genotype-by-sequencing (GBS). Black and grey bars represent chromosome segments from Clark and Ning 7840 respectively. R and S represent resistant and susceptible phenotypes of each line, respectively. The four markers highlighted in red color are the flanking markers.

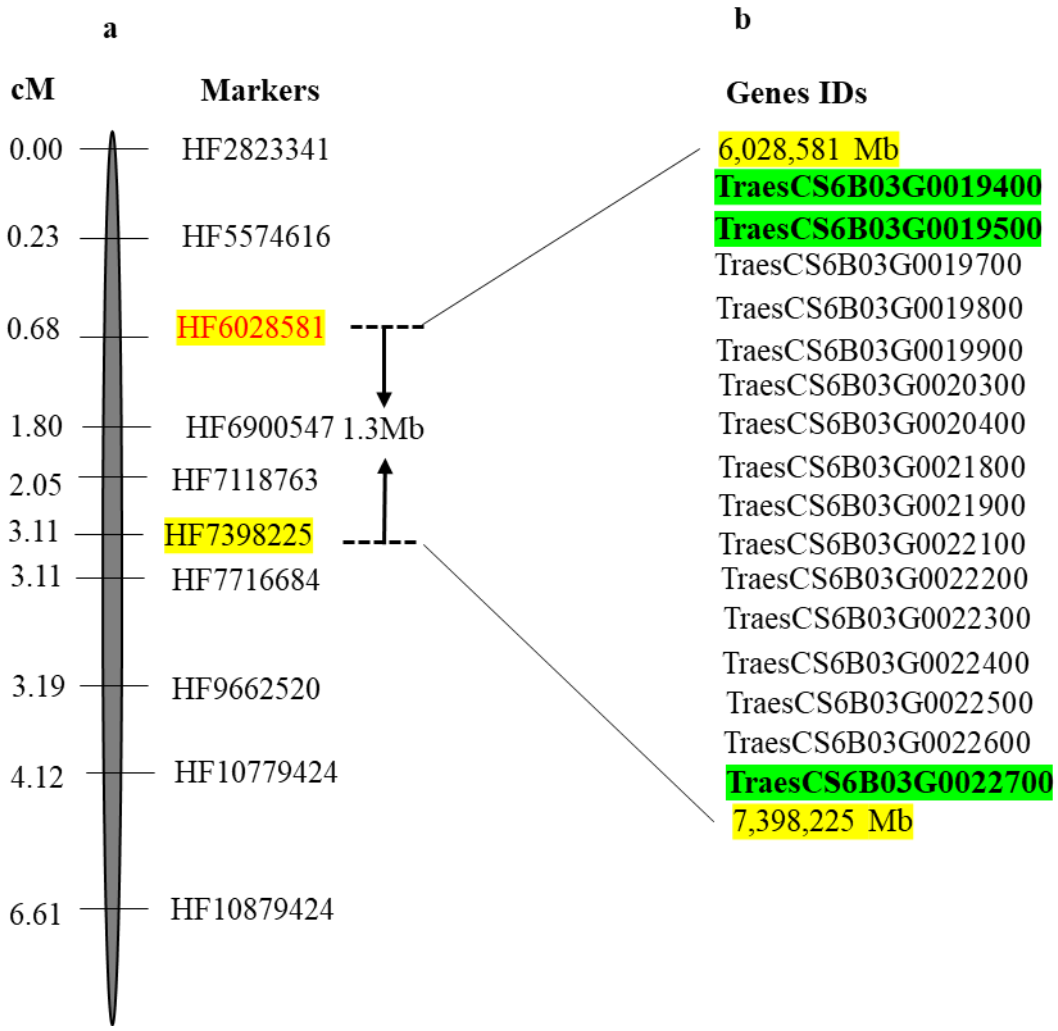

Figure 2.6. High-confidence (HC) genes and differentially expressed genes (DEGs) in the 1.3 Mb *H34* interval on chromosome arm 6BS based on IWGSC RefSeq v2.1 (IWGSC, 2018). (a) The genetic (left) and physical (right, indicated by the number after HF) map locations of the markers in the recombinant lines (RIL) selected from different heterogeneous inbred families (HIFs). Flanking markers are highlighted in yellow color. (b) Gene identity (IDs) of HC genes and DEGs (highlighted in green) in the *H34* interval as annotated in IWGSC RefSeq v2.1.

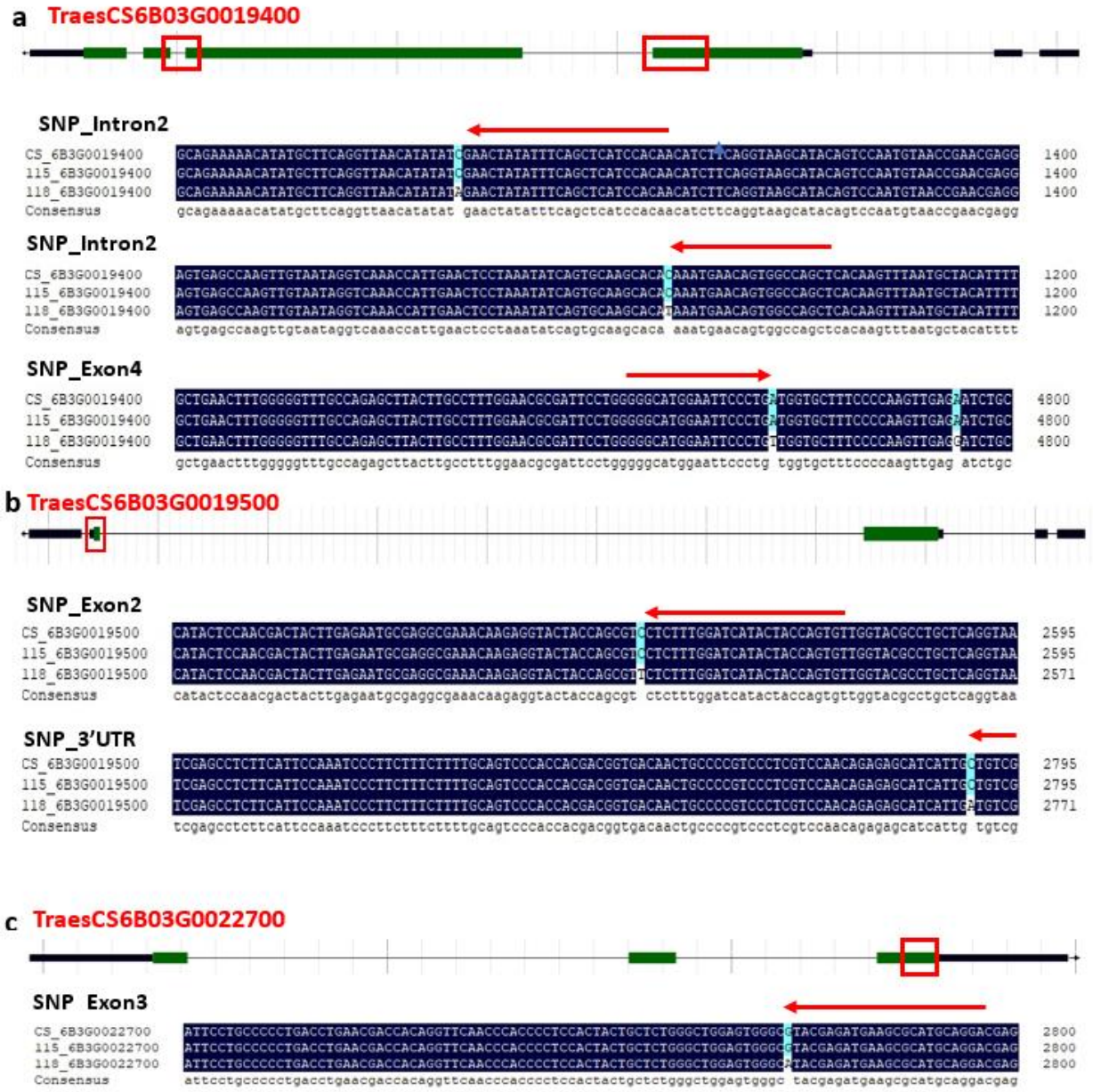

Figure 2.7. Sequence alignment of DEGs in *H34* interval of RIL115-S and RIL118-R against CS reference genome IWGSC RefSeq v2.1 to identify the SNPs for development of diagnostic markers. (a-c) KASP primers sequences and locations for gene TraesCS6B03G0019400, TraesCS6B03G0019500 and TraesCS6B03G0022700, respectively. In each gene structure black boxes represent 3' and 5' UTR regions, green boxes represent exons and black lines represent introns. Red boxes showed the SNP locations in the DEGs. Red arrows labeled the primer sequences of KASP marker in the DEGs.

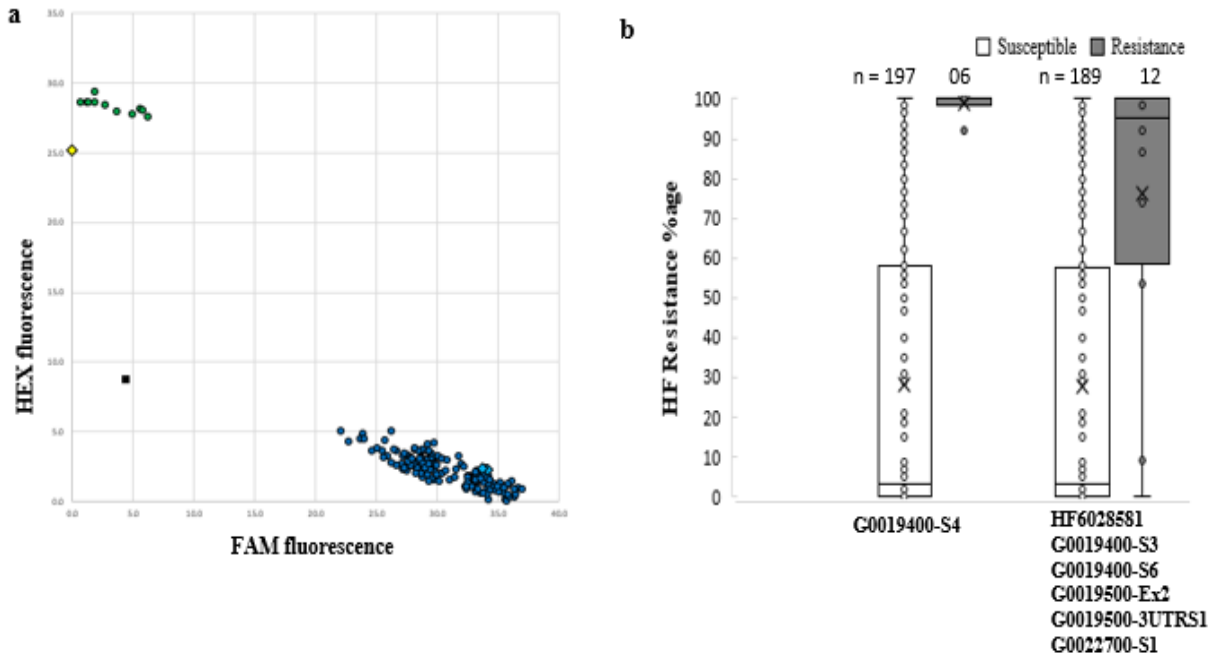

Figure 2.8. Allele separation of KASP markers in the diversity panel of AM203. (a) Allele separation of KASP marker HF6028581. The green dots represent the accessions with the Clark allele, the blue dots represent the accessions with the Ning7840 allele, the yellow dot is Clark (control), the light blue dot is Ning7840 (control) and the black dots are NTC (Non-template control). (b) HF resistance score (%) of the accessions from the 203 U.S. diversity panel. Grey bar (dots) represents the accessions carrying Clark (R) marker allele and white boxes represent the accessions carrying Ning7840 (S) allele.

Table A.1. Marker haplotypes of the 14 recombinants selected from F_3 for heterogenous inbred families (HIFs) development

Sr#	RILs for HIF construction	HF2823341	HF2857942	HF14899424	HF15879400	Resistance %
1	115/118_F_3_8	H	H	H	R	52
2	115/118_F_3_43	H	H	S	S	33
3	115/118_F_3_58	H	H	H	H	76
4	115/118_F_3_102	S	H	H	H	100
5	115/118_F_3_122	H	H	S	S	100
6	115/118_F_3_134	H	H	H	S	44
7	115/118_F_3_141	H	H	H	S	95
8	115/118_F_3_150	H	H	R	R	66
9	115/118_F_3_162	H	H	H	S	33
10	115/118_F_3_170	R	H	H	H	64
11	115/118_F_3_181	H	H	R	R	70
12	115/118_F_3_215	H	H	R	R	50
13	115/118_F_3_245	R	R	H	H	100
14	115/118_F_3_254	S	S	H	H	20

"**R**" is for resistance allele coming from Clark and "**S**" is for susceptible allele from Ning7840.

Table A.2. A list of KASP primers developed from 90K SNP arrays, RNA-sequencing and genotyping-by-sequencing of near-isogenic lines contrasting in *H34* alleles.

Primer Name	Source	Sequence
HF2823341_FAM	GBS_SNP	TCTGGGTCTGAAACACCGTCG
HF2823341_HEX		TCTGGGTCTGAAACACCGTCA
HF2823341_R		GATTAGGGATTCGGGAGGAG
HF2857942_FAM	GBS_SNP	CATGGGCAGTCATAGCCTCC
HF2857942_HEX		CATGGGCAGTCATAGCCTCCT
HF2857942_R		CAAGCTCCCCCAATTGTATC
HF5008058_FAM	GBS_SNP	GATATCCCAGAGGCTGAT
HF5008058_HEX		GATATCCCAGAGGCTGAG
HF5008058_R		GATGGTGGTGACGTTCTG
HF5501616_FAM	90k_SNP	ATCGGGATAATCTTTCCCTGA
HF5501616_HEX		ATCGGGATAATCTTTCCCTGG
HF5501616_R		CAAAGAACTGCATATCGAGG
HF5521616_FAM	GBS_SNP	CGAATTGTTTGGTACAGTGA
HF5521616_HEX		CGAATTGTTTGGTACAGTGT
HF5521616_R		CAACGGATTGCCATTGCA
HF5534616_FAM	GBS_SNP	GCAGCATGTGTGAGCAAAG
HF5534616_HEX		GCAGCATGTGTGAGCAAAC
HF5534616_R		GCACGCAGTATGGGC
HF5574616_FAM	GBS_SNP	CGAGCCGAGTCCCCTC
HF5574616_HEX		CGAGCCGAGTCCCCTT
HF5574616_R		GGCAGAGTTGACGACC
HF6028581_FAM	GBS_SNP	GGCATGGAATTCCCTGT
HF6028581_HEX		GGCATGGAATTCCCTGA
HF6028581_R		AGAAACCTGAATGGCG
HF6900547_FAM	GBS_SNP	TCGTCCACTCTTGTGTTGCT
HF6900547_HEX		TCGTCCACTCTTGTGTTGCC
HF6900547_R		GCATAGGAGCATGGATTGGA
HF7118763_FAM	GBS_SNP	TTAAACCAGTCGATCGCA
HF7118763_HEX		TTAAACCAGTCGATCGCC
HF7118763_R		ATACCTGCAGCAAGTTAC
HF7398225_FAM	GBS_SNP	ACCCGGTGAGAGAGAACGA
HF7398225_HEX		ACCCGGTGAGAGAGAACGG
HF7398225_R		GATCATTGCACATGCATCAA
HF7716684_FAM	GBS_SNP	CGGAAACATTAACCACCTCAAT
HF7716684_HEX		CGGAAACATTAACCACCTCAAC
HF7716684_R		GCCACGGGTCATGGTATTAT
HF9662520_FAM	GBS_SNP	GTTTTCGAATGGGTGCTT

HF9662520_HEX		GTTTTCGAATGGGTGCTG
HF9662520_R		CTCCTTGGAACCGTACCA
HF10252606_F	90k_SNP	ACGACGTTGTAAAACGACACAAAAGTTCGAGGGACAACAT
HF10252606_R		CAAAACACAGGCACGACTTCTA
HF10779424_FAM	Exome Capture_ SNP	TTGACGACATGCTCGGGCAT
HF10779424_HEX		TTGACGACATGCTCGGGCAC
HF10779424_R		AGCAGGTGGACGATGTTGAT
HF10879424_FAM	Exome Capture_ SNP	GTACAATACAGGGTCATCAA
HF10879424_HEX		GTACAATACAGGGTCATCAT
HF10879424_R		ATCGTTTGTCAACGGGATTG
HF14899424_FAM	90k_SNP	TACTGCATGCTTCCTCTGTCGA
HF14899424_HEX		TACTGCATGCTTCCTCTGTCGC
HF14899424_R		GACCCTTTTATCATCTAGTGC
HF15879424_FAM	90k_SNP	GAAACAACATTTCTCCTCTCTCTA
HF15879424_HEX		GAAACAACATTTCTCCTCTCTCTG
HF15879424_R		CATGCACAGATTGATGCGTC
G00227700-S1_FAM	Candidate gene sequence	GCATGCGCTTCATCTCGTAC
G00227700-S1_HEX		GCATGCGCTTCATCTCGTAT
G00227700-S1_R		TCAGGTTCCTTGTCGAGAAC
G0019400-S3_FAM	Candidate gene sequence	TGTGGATGAGCTGAAATATAGTTCG
G0019400-S3_HEX		TGTGGATGAGCTGAAATATAGTTCT
G0019400-S3_R		GCCGAAAATGCTGCATATGC
G0019400-S4_FAM	Candidate gene sequence	AGCTGGCCACTGTTCATTTG
G0019400-S4_HEX		AGCTGGCCACTGTTCATTTA
G0019400-S4_R		TCAGGAGTGAGCCAAGTTGT
G0019400-S6_FAM	Candidate gene sequence	GGGGCATGGAATTCCCTGA
G0019400-S6_HEX		GGGGCATGGAATTCCCTGT
G0019400-S6_R		CTGCAGAAACCTGAATGGCG
G0019500-Ex1_FAM	Candidate gene sequence	ACACTGGTAGTATGATCCAAAGAGG
G0019500-Ex1_HEX		ACACTGGTAGTATGATCCAAAGAGA
G0019500-Ex1_R		GAGCAGGCGAAACAAGAGGT
G0019500-3UTRS1_FAM	Candidate gene sequence	AGGAGATGTACTGGCGACAG
G0019500-3UTRS1_HEX		AGGAGATGTACTGGCGACAT
G0019500-3UTRS1_R		CTTTTGCAGTCCCACCACGA

Tail sequences are added before the actual FAM and Hex sequence of each primer. *FAM_ gaaggtgaccaagttcatgct. HEX_ gaaggtcggagtcaacggatt

Table A.3. Annotated genes present within the finely-mapped *H34* interval based on IWGSC RefSeq v2.1

Gene ID in v2.1	Gene ID in v1.0	Description
TraesCS6B03G0019400	TraesCS6B02G000100	NBS-LRR-like resistance protein
TraesCS6B03G0019500	TraesCS6B02G000200	Disease resistance protein (NBS-LRR class) family
TraesCS6B03G0019700	TraesCS6B02G000300	E3 ubiquitin-protein ligase
TraesCS6B03G0019800	TraesCS6B02G000400	Cytochrome/ubiquinol oxidase subunit 1
TraesCS6B03G0019900	TraesCS6B02G000500	Casein kinase I
TraesCS6B03G0020300	TraesCS6B02G000600	50S ribosomal protein L14
TraesCS6B03G0020400	TraesCS6B02G000700	Leucine-rich repeat receptor-like protein kinase family protein
TraesCS6B03G0021800	TraesCS6B02G000800	F-box protein
TraesCS6B03G0021900	TraesCS6B02G000900	Pollen Ole e 1 allergen
TraesCS6B03G0022100	TraesCS6B02G001000	F-box protein
TraesCS6B03G0022200	TraesCS6B02G001100	F-box protein
TraesCS6B03G0022300	TraesCS6B02G001200	F-box protein
TraesCS6B03G0022400	TraesCS6B02G001300	Receptor-like protein kinase
TraesCS6B03G0022500	TraesCS6B02G009400	F-box protein
TraesCS6B03G0022600	TraesCS6B02G009500	F-box protein
TraesCS6B03G0022700	TraesCS6B02G009600	Ankyrin repeat family protein

Table A.4. A list of gene-specific primers used to amplify full-length sequence of deferentially expressed genes

Primer Name	Gene Name	Sequence
G0019400_F1	TraesCS6B03G0019400	ACAGCAGAGCCTGCCCAT
G0019400_R2	TraesCS6B03G0019400	ACCGGTCACCACCTCCAT
G0019400_F2	TraesCS6B03G0019400	GACTGCTTCCTAGCTAAAGGATCC
G0019400_R3	TraesCS6B03G0019400	CTCGCATTCTCAAGTAGTCGTTGG
G0019400_F4	TraesCS6B03G0019400	GTACTTAATTGTCATTGATGACATATGG
G0019400_R4	TraesCS6B03G0019400	CTGTGGTAGACCATATGCATCG
G0019400_F5	TraesCS6B03G0019400	GTGCCGAATTGGATAGGGAA
G0019400_R5	TraesCS6B03G0019400	ATAATACAGAAGAAGTATTAACAAGG
G0019500_F1	TraesCS6B03G0019500	AGGTAGCAGCAGAGACCGT
G0019500_R2	TraesCS6B03G0019500	CTCGCATTCTCAAGTAGTCGTTGG
G0019500_F2	TraesCS6B03G0019500	GCGAAACAAGAGGTACTACCA
G0019500_R2	TraesCS6B03G0019500	GAATCGATTTTTTTAAAGGAGAGCA
G0022700_F1	TraesCS6B03G0022700	ATGAATGAAGGAGTTATCGA
G0022700_R1	TraesCS6B03G0022700	AAGCAAATATTGTTTTGCATCTACC
G0022700_R2	TraesCS6B03G0022700	CTTCCTTCAGCTCGTCCAAC
G0022700F2	TraesCS6B03G0022700	AAGCAAATATTGTTTTGCATCTACC

*F represent forward and R represent reverse primer sequence in this table.

Chapter 3 - Transcriptomics to unravel mechanisms of Hessian fly resistance in wheat

3.1 Introduction

Many pest management strategies can be deployed to reduce HF population, but host plant resistance is an economic and eco-friendly approach for HF control (Harris et al., 2015; Chen et al., 2010; Ratcliffe et al., 2000; Zhao et al., 2015). Continuous monitoring the virulent HF biotypes can avoid area-wide crop losses. When a virulent HF biotype starts increasing, it is time to deploy the new resistance genes to the emerging HF biotypes (Chen et al., 2009). Most importantly, understanding of the molecular mechanisms of interactions between HF and wheat may facilitate improvement of wheat cultivars for durability of HF resistance (Chen et al., 2009; Liu et al., 2005).

The identification of differentially expressed genes (DEGs) is a key step for characterizing transcriptome sequence variation and serves as an important resource for researchers to identify key genes involved in HF resistance in wheat. RNA sequencing (RNA-seq) is a next-generation-sequencing based technology and has become a popular tool for evaluating the gene expression levels in various organisms (Costa-Silva et al., 2017; Mortazavi et al., 2008) due to its high throughput with affordable cost (Bai et al., 2016). Gene ontology (GO) enrichment analysis can evaluate if the candidate genes are critical components of gene co-expression functional networks (Tohge & Fernie, 2012). Furthermore, comparisons of RNA-seq data between contrasting genotypes not only can identify SNPs, but also simultaneously evaluate the expression levels of the functional genes and their sequence variations. In addition, the

locations of causal variations in coding regions of the genes of interests can be identified and used to predict the phenotypes (Yu et al., 2014). Recently, RNA-seq has been successfully utilized in wheat to analyze gene expression during flowering (Pearce et al., 2016), and in responses to heat (Kumar et al., 2015), salt (Jiang et al., 2017; Xiong et al., 2017) and drought stresses (Liu et al., 2015). Particularly, with the availability of 'Chinese Spring' wheat genome reference (IWGSC 2014) and pan-genome sequences (www.10wheatgenomes.com) from ten wheat cultivars, physical positions of the genes identified through RNA-seq can be estimated.

Although 37 genes have been reported, none of the causal genes for HF resistance has been cloned. Therefore, mechanisms that cause HF larvae to die in resistant plants remain unknown. Different types of defensive, toxic chemicals including protease inhibitors, reactive oxygen species, toxic lectins, and secondary metabolites were specifically induced in resistant plants upon HF larval attack (Hargarten et al., 2017). In the current study, transcriptomic profiling of susceptible and resistant wheat lines infested with Hessian fly larvae revealed what genes involved in the wheat defense against HF infestation. Hundreds of differentially expressed genes (DEGs) at two different time points (12-h and 48-h) after larval attack on wheat seedlings were identified. To elucidate defense related pathways associated with the HF infestation at different infection stages, we performed the Gene Ontology (GO), and constructed heatmaps after 12-h and 48-h of HF infestation in four pairs of near isogenic lines (NILs) carrying contrasting alleles at *H34*. Our findings will help to identify the gene regulatory networks in the resistant wheat plants compared to their susceptible control.

3.2 Materials and methods

3.2.1 Evaluation of HF resistance

Four pairs of NILs derived from four different HIFs (chapter 2) were used for transcriptomic analysis to understand the HF resistance mechanism. Gene expression levels of four pairs of the resistant NILs (NIL-R) were compared with the susceptible NILs (NIL-S). These NILs were evaluated for HF resistance together with the four control cultivars, Carol (*H3*), Caldwell (*H6*), Molly (*H13*), and Danby (susceptible control) for reactions to HF biotype GP in the greenhouse experiments at Kansas State University using the same procedure described in Chapter 2.

3.2.2 RNA sample preparation and sequencing

RNA was isolated for all four NIL pairs after 12 hours (12-h) and 48 hours (48-h) of the time when HF larvae were first seen at the base of a plant after initial HF larval attack as explained in chapter 2.

3.2.3 RNA-seq library preparation and data analysis

RNA-eq was carried out using IIlumina HiSeq™ 2000 sequencer in Novogene Corporation Inc. (CA, USA). Data were analyzed for identification of DEGs as described in chapter 2.

3.2.4 Function and pathway analyses of DEGs

All the identified DEGs at 12-h and 48-h after initial HF larval attack, gene annotation was done using the GO database, and GO enrichment analysis of DEGs was performed using

cluster profile (4.0) (Wu et al., 2021). MapMan (v3.5.1R2) covering plant-specific pathways and processes (https://mapman.gabipd.org/mapman) was used to determine HF-resistance related pathways using the Log2-fold changes of common DEGs induced in NIL-R compared to NIL-S (control) at 12-h and 48-h hours after initial HF attack. A custom-specific mapping file for the MapMan was created using the Mercator pipeline (http://mapman.gabipd.org/web/guest/mercator) based on the wheat sequence output (Lohse et al. 2014). MapMan (v3.5.1R2) was also used to visualize the expression changes of DEGs to multiple MapMan functional categories (Thimm et al., 2004).

3.2.5 Validation of RNA-seq data using qRT-PCR

Three DEGs were analyzed by a quantitative real-time polymerase chain reaction (qRT-PCR) assay with the gene-specific primers to validate the RNA-seq data. Primers were designed using the sequences from RNA-seq data for three DEGs in the transcriptomic library (Table 3.1). The qRT-PCR analysis was completed in a CFX96 Real-Time PCR system (Bio-Rad, USA). Relative gene expression levels were calculated according to the $2^{-\Delta\Delta Ct}$ method, with the endogenous *Actin* gene as the control. Mean values and standard errors were calculated for the three independent experiments with three biological replications and four technical replications.

3.2.6 Neutral red staining to assess cell wall permeability

To determine whether the HF larvae disrupt the integrity of the epidermal cell layer, resistant and susceptible wheat plants were stained with neutral red (NR) stain (Sigma-Aldrich, MO, USA) to assess permeability as described by Williams et al. 2011. Resistant and susceptible NILs were grown and infested with HF in a greenhouse as described above. Wheat seedlings after 12-h and 48-h of initial larval attack were dissected below the root junction. After peeling

back to remove fist leaf, the bottom 2 cm portion of seedling was stained with 0.1 % NR solution for 10 min and then the stained tissues were rinsed thoroughly with sterile water. Uninfested control plants were also dissected and stained after piercing them with a 0.2-mm minute pin as the negative controls. Following staining, photomicrographs were taken with a Motic digital microscope (Kowloon, Hong Kong).

3.3 Results

To confirm the contrast responses to HF infestation between the resistant and susceptible NILs, they were phenotyped for HF resistance in greenhouse. These plants were planted and infested with HF together with the samples for the RNA-seq experiment. After RNA samples were collected, the rest of NILs in the same tray were rated for HF resistance three weeks after infestation. At that time, resistant NILs grew normally with light green leaves and dead larvae at the bottom of the leaf sheath, although some tiny larva survived on some normally grown plants, whereas susceptible NILs were stunted with dark green leaves and larvae grown at the bottom of the plants between leaf sheaths (Fig. 3.1).

3.3.1 Differentially expressed genes in response to Hessian fly attack identified by RNA-seq analysis

RNA-sequencing generated about 1.1 billion clean reads after removing low quality reads and they were mapped to IWGSC RefSeq v2.1 with 42 to 90 million reads per line for each treatment. Among these mapped reads, 88.3%, ranging from 84 to 90%, were mapped to unique positions and approximately 5% were mapped to multiple positions (Table S1). Moreover, 913 and 1,157 DEGs were significant (FDR < 0.05 and fold changes ≥ 1 or ≤ -1) at 12-h and 48-h of

HF infestation, respectively, in the NIL-R compared to the NIL-S (Fig. 3.2a). Only 119 DEGs were common between the two time-points. At 12-h after infestation, 257 genes were up-regulated and 656 were down-regulated, whereas 848 were up-regulated and 309 were down-regulated at 48-h after infestation in the resistant NILs compared to their susceptible NILs (Fig 3.2b). Overall, fewer DEGs between the NILs were identified at 12-h than 48-h timepoint. However, significant fewer up-regulated DEGs were observed than down-regulated DEGs at 12-h timepoint, and the opposite trend was observed at 48-h timepoint (Fig 3.2b).

3.3.2 Gene ontology (GO) classification

Gene ontology classification assigned all significantly DEGs in resistant NILs into three functional classes: biological process, cellular components, and molecular function. Majority of the DEGs at 12-h and 48-h timepoint treatments were grouped into cellular component domain followed by biological process and molecular function. However, the highest number of DEGs at the12-h treatment fell under the peptide metabolic process (GO:0003674) class followed by cytoplasmic membrane (GO:0016740), while the highest number of DEGs at the 48-h treatment belongs to cytoplasmic membrane class (GO:0065007) followed by cellular organelles membrane (GO:0033554) group in the cellular component domain. Overall, from 12-h to 48-h after infestation by HF, the numbers of DEGs in GO domains showed a considerably increased trend in cellular membranes, metabolic pathways, and biological regulation processes in response to HF larval attack (Fig. 3.3).

3.3.3 Distribution of DEGs across the wheat genome

The DEGs in the NIL-R were relatively evenly distributed across the three wheat genomes in comparison to the NIL-S at both 12-h and 48-h timepoints. The A-genome had the

maximum number of DEGs (382) at 12-h timepoint, followed by D-genome (300) and B-genome (285). While almost equal number of DEGs were found at the 48-h timepoint among the three genomes, with 382 genes on B-genome, followed by A (381) and D-genome (379) respectively (Fig. 3.4).

3.3.4 Biotic stress and defense adaptation related DEGs induced after Hessian fly attack

The key DEGs in biotic stress pathways including hormonal signaling, enzymatic activity, redox activity, transcription factors, heat shock proteins and production of secondary metabolites in the NIL-R lines were observed at both 12-h and 48-h of HF attack on wheat (Fig. 3.5). Fewer DEGs in the resistant NIL-R were observed at 12-h timepoint but most of those are documented to play a role in active stress responses such as hormonal signaling, respiratory burst, heat shock proteins and production of secondary metabolites to protect the wheat plant against HF (Fig. 3.5a); most of the DEGs involved in different defense related pathways were up-regulated at the 48-h in the NIL-R compared to the NIL-S (Fig 3.5b). Moreover, the DEGs involved in redox reaction were down-regulated at 12-h, then the expression level increased, and they were up-regulated at 48-h in the NIL-R (Fig 3.5).

3.3.5 DEGs involved in hormonal homeostasis, photosynthesis, and secondary metabolite production

Significant changes in the transcriptional profiles of the DEGs encoding hormone-signaling were observed at 12-h and 48-h of HF larval infestation in NIL-R compared to the NIL-S. The expression level of the DEGs involving in ethylene (ET) and jasmonic acid (JA)

biosynthesis were significantly increased from 12-h to 48-h after larval attack on wheat seedlings. On the contrary, most of the DEGs in auxins (AUX) pathways, except for one DEG at 48-h, were down-regulated (Fig. 3.6a). Only three DEGs were identified in brassinosteroids (BA) pathway, with one down-regulated and two up-regulated at 48-h compared to 12-h after infestation in the NIL-R compared to the NIL-S (Fig. 3.6a). The DEGs for secondary metabolite production in the NIL-R showed up-regulation at the 48-h. The transcriptome level of the photosynthesis related DEGs in the NIL-R were also up-regulated at 48-h after initial down regulation at the 12-h (Fig 3.6b).

3.3.6 DEGs involved in redox reaction and cell wall integrity

A notable boost in expression level of differentially expressed R-genes can be seen at 12-h with decrease in their activity at 48-h of HF attack in the NIL-R. Moreover, the resistant NILs showed differential regulation of the genes encoding ROS generating and scavenging enzymes at both 12-h and 48-h in the NIL-R when compared with the NIL-S (Fig. 3.7).

After initial attack by HF larvae, key DEGs which are involved in cell wall and cuticle metabolism were up-regulated at 12-h and down-regulated 48-h after Hessian fly infestation in the NIL-R. Along with that, for other DEGs which were involved in acquiring resistance against pest and pathogens such as F-box and heat shock protein (HSP), an increase in expression level of three F-box and two HSPs was observed at 48-h while they were down-regulated at 12-h (Fig. 3.7).

To examine if elevated expression of cell-wall-related DEGs correlated with cell wall permeability in the resistant and susceptible NILs following Hessian fly attack, neutral red staining of cell wall penetration was conducted. Before staining, living larvae can be seen on the

susceptible lines, while larvae were not observed on the resistant lines (Fig. 3.8 a-b). In the uninfested plants, the stain was localized to the manually punctured spots and did not spread to adjacent areas in the leaf (Fig. 3.8 c-d). In the susceptible NILs, more intense stain was observed as dark streaks covering the entire length of crown tissue at 48-h compared to 12-h after infestation (Fig. 3.8 e, g). On the contrary, in the resistant NILs, limited distribution of NR in form of localized patches and broken lines were seen at 48-h compared with 12-h after HF attack (Fig. 3.8 f, h).

3.3.7 Differential expression of transcription factors involved in HF resistance

A total 25 differentially expressed TFs belonging to four different families; 17 from WRKY (60 amino acid long four-stranded β-sheet WRKY DNA binding domain/s) family, six from Zinc finger protein family (ZnF), three MYB (DNA-binding, helix-turn-helix (HTH) domain of approximately 55 amino acids) related DEGs and two from NAC (NAM, ATAF1/2, and CUC2) family showed differential expression at 12-h and 48-h after larval attack in the NIL-R. Twelve out of the17 deferentially expressed WRKY TFs were up-regulated, five were down-regulated after 48-h in the NIL-R. Four out of six DEGs encoding ZnF TFs were down-regulated at 12-h, but significantly increased after 48-h in the NIL-R compared to NIL-S. On the contrary in the NIL-R, the down-regulation of DEGs encoding MYB and NAC TFs at 48-h were observed compared to 12-h of HF infestation attack. While DEG encoding ERF TF was up-regulated at 48-h and down-regulated at 12-h in the NIL-R (Fig. 3.9).

3.3.8 Verification of differentially expressed genes via qRT-PCR

Three highly DEGs from our transcriptomic data which were found to be involved in plant defense mechanism upon HF attack on the resistant NILs were selected for the verification

via qRT-PCR. Among them, TraesCS1B02G245800 was associated with regulation of systemic acquired resistance; TraesCS4B02G086600 involved in regulation of membrane components (Figure 3.6) and TraesCS4B02G070300 is an auxin repressor protein (Figure 3.7). From qRT-PCR data, the expression levels of TraesCS1B02G245800 and TraesCS4B02G086600 were significantly increased in the NIL-R compared to NIL-S after 12-h and 48-h of larval attack. On the contrary, significantly decreased in expression of TraesCS4B02G070300 has been observed, compared to NIL-S both at 12-h and 48-h of initial larval attack on wheat seedlings (Fig. 3.10). All these results were consistent with the RNA-seq data, which verified the accuracy of our RNA-Seq data.

3.4 Discussion

Previous studies revealed that *Brachypodium distachyon,* a model grass genome (https://www.nature.com/articles/s41598-019-39615-2 - ref-CR1), exhibited nonhost resistance to HF and shared similarities to wheat HF resistance at the physical and metabolic levels (Hargarten et al., 2017; Subramanyam et al., 2019). On the contrary, wheat is a polyploid crop with a big genome, so using *Brachypodium distachyon* as a model plant may facilitate identification of mechanisms of wheat resistance to HF. Moreover, utilizing the well-annotated wheat reference genome, and NILs with the same background may help functionally characterize the candidate genes for wheat resistance to Hessian fly and understand molecular mechanisms of HF resistance in wheat. In this study, we analyzed changes in gene expression levels at two time points (12-h and 48-h) following HF infestation in four pair of contrasting NILs. These two time points were selected for the reason, 12-h is the first critical time when HF larvae start feeding on wheat plants while 48-h is the time when a significant change in expression of genes involved in

plant defense can be seen (Liu et al., 2007). In-depth analysis revealed that the DEGs early response at 12-h were different from late response at 48-h in NIL-R compared to their respective susceptible control (NIL-S), as wheat resistant plants showed an incompatible reaction to HF by up-regulating genes involved in strengthening cell wall and some defense-related genes after initial larval attack (Fig. 3.5).

3.4.1 Fitness cost of defense regulated by phytohormones and their interaction with photosynthesis, and secondary metabolites production

Many plants growth regulatory phytohormones act as a sensor to respond to both biotic and abiotic stresses and specify a strong inter-connection between two physiological processes: plant development and adjustment to environmental cues. Hormonal regulatory potential allows plants to quickly respond to biotic stresses and to utilize their limited nutrients in a cost-efficient manner by down-regulating the DEGs involved in nutrient production (Denancé et al., 2013). While enhanced secondary metabolism undoubtedly influences primary metabolism, accumulating evidence suggests that insect herbivory reduces photosynthetic carbon fixation through a re-programming gene expression rather than stimulating photosynthesis (Kerchev et al., 2012). In this study, the DEGs encoding different hormones and later the DEGs involved in photosynthesis and production of secondary metabolites significantly increased in the NIL-R (Fig. 3.6), suggesting HF triggered the activation of these key pathways that lead to secondary metabolite biosynthesis in wheat (Liu et al., 2015). Moreover, in response to HF infestation, the NIL-R plants exhibited inconsistent photosynthetic responses, starting with down-regulation of the DEGs at 12-h followed by up-regulation of most of the DEGs at 48-h after HF infestation in comparison with the NIL-S (Fig 3.6). Down-regulation of photosynthesis-related DEGs in

resistant plants is a common plant response to insect feeding, representing a host defense mechanism that deprives HF larvae from taking food and starves them to death. It is possible that wheat plants channel energy and carbon resources towards production of defense related metabolites to bolster a defense response (Bilgin et al., 2010; Zhao et al., 2015). The same results were observed in this study with higher production of secondary metabolites in NIL-R at 12-h compared to 48-h of HF attack (Fig. 3.6).

3.4.2 Resistant wheat plants strengthen cell wall and use HR-mediated key defense strategy

Cellular defense often starts with the recognition of pathogen and insect associated molecular patterns (PAMPs) present on the surface of plant cells. PAMPs-triggered immunity induces reactive oxygen species (ROS) in the extra cellular spaces which increase the activity of cell membrane related genes to increase the cross linking of phenolics in the cell wall and fortify them (Keinath et al., 2010; Liu et al., 2010; Tuteja, 2007). This phenomenon eventually stops the further spread of HF damage to neighboring cells by starving them to death in 3 to 5 days, a similar process as to pathogens (Gechev et al., 2006; Heath, 2000). ROS molecules activate several other defense-related genes such as genes encoding F-box protein and heat shock proteins in response to insect attack. These proteins contribute to cellular homeostasis, perform critical chaperone functions, and assist in the protein refolding within the stress-damaged cells (Park & Seo, 2015). Liu et al. (2010) provided more insight into pathways contributing to accumulation of ROS molecules in resistant plants during HF larval attack. A two-fold increase in peroxidases activity was observed in the resistant plants carrying *H13* and *H9* during HF larval attack, while no change has been detected in the infested susceptible control plants (Das &

Roychoudhury, 2014; Liu et al., 2010). Increased transcripts for ROS-generating and scavenging enzymes (Fig. 3.7), and presence of lesions on resistant plants clearly support the involvement of ROS in wheat defense against HF from our results (Fig. 3.1).

Furthermore, the cell wall can be promptly strengthened in the HF resistant plants in response to HF larval attack, and absence of the nutritive cells at the feeding site in these plants, indicating the lack of nutrients in the resistant plants may result in the death of HF larvae (Harris et al., 2010; Khajuria et al., 2013). Similarly, analyses of the dynamic differences in cell wall strengthening gene expression in the resistant NILs compared to susceptible NILs found that rapid mobilization and re-utilization of nutrient resources and enhanced fortification of cell walls were coordinated defense processes that may be crucial for HF resistance in wheat (Fig. 3.5). Consistent with above observations, the data from neutral red staining demonstrated that Hessian fly larvae were unable to increase cell wall permeability in the infested resistant plants, suggesting that cell wall and cuticle were indeed strengthened in these resistant plants (Fig. 3.8). Fortified cell walls could prevent HF larvae from delivering effectors into wheat cells, preventing the insect from manipulating normal host tissue into nutritive tissues (Harris et al., 2015). Therefore, lack of nutritive tissues prevents HF larvae from obtaining nutrition from host cells in resistant plants, resulting in their death.

3.4.3 Wheat transcription factors mediated plant defense against HF

Transcription factors (TFs) are crucial upstream regulators of stress-related genes resulting in activation or inhibition of these genes, thereby imparting tolerance to a stress including plant pests (Alves et al., 2014). Investigating how TFs regulating stress-related gene expressions may provide insight on how the wheat plants respond to insect infestation resulting

in resistance or susceptibility. Genes for secondary metabolites were up-regulated in the resistant wheat line and TFs are the key regulators of those genes, supporting the roles of TFs in plant defense against pests (Satapathy et al., 2014). TFs are also involved in the ROS pathway. In tobacco, mitogen-activated protein kinases (MAPK) WRKY pathway activated ROS bursts by activating respiratory burst oxidative homologs *RbohB*, and a *Nicotiana benthamiana* WRKY TF *NbWRKY2* to enhance resistance to *Phytophthora nicotianae* by increasing the H_2O_2 accumulation and inducing defense against pest (Adachi et al., 2015). The largest TF WRKY family is a key player in plant responses to biotic stress. At least nine *Oryza sativa OsWRKYs* have been identified to regulate positive rice response to *M. oryzae*. For example, overexpression of different WRKY TF such as *OsWRKY31*, *OsWRKY45*, *OsWRKY47*, *OsWRKY53*, or *OsWRKY67* in rice plants enhanced resistance to *M. oryzae* (Vo et al., 2018). In this study, the increased WRKY transcript levels after 48-h of HF infestation in the NIL-R compared to the NIL-S suggested their involvement in HF resistance. *NAC* and *MYB* genes also play an important role in the response to abiotic stresses and negatively regulated tolerance to salt and osmotic stress by reducing ROS scavenging capability and proline biosynthesis. In wheat, the overexpression of root-expressed NAC transcription factor *TaRNAC1* enhanced drought tolerance and the overexpression of a NAC transcription factor *TaNAC29* increased salt tolerance by enhancing the antioxidant system to reduce H_2O_2 accumulation and membrane damage (Chen et al., 2018; He et al., 2017). We also observed down regulation of *NAC* genes after the activity of ROS genes increased, indicating their roles in regulating plant defense genes. The ZnF protein family is another TF superfamily which is involved in regulation of resistance mechanisms for various biotic and abiotic stresses. Zinc finger TFs were induced by *Spodoptera littoralis* feeding in *Arabidopsis* and played a JA-independent role in plant defense (Schweizer et al., 2013).

Similarly, a zinc finger TF, *StZFP2*, were induced in potato (*Solanum tuberosum*) in response to infestation by tobacco hornworm (*Manduca sexta*) and Colorado potato beetle (*Leptinotarsa decemlineata*), implicating their roles in plant defense against chewing insects (Lawrence et al., 2014). Same trend of up regulation of few ZnF DEGs in response to phytohormones at both 12-h and 48-h after HF attack was also observed in the NIL-R in the current study, demonstrated their roles in HF resistance.

3.4.5 Conclusion

Our results from transcriptomic analysis revealed the mechanisms of resistant plants to survive after HF larval attack. When HF attacks on HF resistant wheat plants (NIL-R), it triggers the interaction of insect secretions with R-gene receptors to initiate a hypersensitive response (HR). This HR response results in production of ROS by showing differential regulation of genes encoding ROS-generating and -scavenging enzymes which lead to up-regulation of downstream DEGs involved in cell wall strengthening to prevent HF to access the nutrients in the resistant plants and eventually to result in the death of HF larvae. At the same time, several other defense strategies are mounted, including that plant phytohormones are induced to up-regulate the genes involved in secondary metabolite production and initially down-regulation of genes involved in photosynthesis to preserve the plant nutrients against HF larvae and up-regulation after 48-h. Activation and suppression of different TFs up regulate several other defense genes including F-box and HSPs and ultimately bring resistance by incompatible interaction between resistant wheat plants and HF (Fig. 3.11). Moreover, in this study three highly important DEGs were verified using qRT-PCR analysis and they are involved in bringing HF resistance by interacting with other DEGs. Future characterization of functions of the candidate HF responsive genes will aid in crop improvement to increase wheat resistance against HF.

3.5 References

Adachi, H., Nakano, T., Miyagawa, N., Ishihama, N., Yoshioka, M., Katou, Y., Yaeno, T., Shirasu, K., & Yoshioka, H. (2015). WRKY transcription factors phosphorylated by MAPK regulate a plant immune NADPH oxidase in *Nicotiana benthamiana*. *The Plant Cell*, *27*(9), 2645–2663. https://doi.org/10.1105/tpc.15.00213

Alves, M., Dadalto, S., Gonçalves, A., de Souza, G., Barros, V., & Fietto, L. (2014). Transcription factor functional protein-protein interactions in plant defense responses. *Proteomes*, *2*(1), 85–106. https://doi.org/10.3390/proteomes2010085

Bai, H., Harmancı, A. S., Erson-Omay, E. Z., Li, J., Coşkun, S., Simon, M., Krischek, B., Özduman, K., Omay, S. B., Sorensen, E. A., Turcan, Ş., Bakırcığlu, M., Carrión-Grant, G., Murray, P. B., Clark, V. E., Ercan-Sencicek, A. G., Knight, J., Sencar, L., Altınok, S., … Günel, M. (2016). Integrated genomic characterization of IDH1-mutant glioma malignant progression. *Nature Genetics*, *48*(1), 59–66. https://doi.org/10.1038/ng.3457

Bilgin, D. D., Zavala, J. A., Zhu, J., Clough, S. J., Ort, D. R., & Delucia, E. H. (2010). Biotic stress globally downregulates photosynthesis genes. *Plant, Cell & Environment*, *33*(10), 1597–1613. https://doi.org/10.1111/j.1365-3040.2010.02167.x

Byers, R. A., & Gallun, R. L. (1972). Ability of the Hessian fly to stunt winter wheat. 1. Effect of larval feeding on elongation of leaves. *Journal of Economic Entomology*, *65*(4), 955–958. https://doi.org/10.1093/jee/65.4.955

Chen, D., Chai, S., McIntyre, C. L., & Xue, G.P. (2018). Overexpression of a predominantly root-expressed NAC transcription factor in wheat roots enhances root length, biomass and drought tolerance. *Plant Cell Reports*, *37*(2), 225–237. https://doi.org/10.1007/s00299-017-2224-y

Chen, M.-S., Liu, X., Wang, H., & El-Bouhssini, M. (2009). Hessian fly (Diptera: Cecidomyiidae) interactions with barley, rice, and wheat seedlings. *Journal of Economic Entomology*, *102*(4), 1663–1672. https://doi.org/10.1603/029.102.0434

Chen, M.-S., Liu, X., Yang, Z., Zhao, H., Shukle, R. H., Stuart, J. J., & Hulbert, S. (2010). Unusual conservation among genes encoding small secreted salivary gland proteins from a gall midge. *BMC Evolutionary Biology*, *10*(1), 296. https://doi.org/10.1186/1471-2148-10-296

Cokus, S. J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C. D., Pradhan, S., Nelson, S. F., Pellegrini, M., & Jacobsen, S. E. (2008). Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature*, *452*(7184), 215–219. https://doi.org/10.1038/nature06745

Costa-Silva, J., Domingues, D., & Lopes, F. M. (2017). RNA-seq differential expression analysis: An extended review and a software tool. *PLOS ONE*, *12*(12), e0190152. https://doi.org/10.1371/journal.pone.0190152

Das, K., & Roychoudhury, A. (2014). Reactive oxygen species (ROS) and response of antioxidants as ROS-scavengers during environmental stress in plants. *Frontiers in Environmental Science*, *2*. https://doi.org/10.3389/fenvs.2014.00053

Denancé, N., Sánchez-Vallet, A., Goffner, D., & Molina, A. (2013). Disease resistance or growth: the role of plant hormones in balancing immune responses and fitness costs. *Frontiers in Plant Science*, *4*. https://doi.org/10.3389/fpls.2013.00155

Gechev, T. S., Van Breusegem, F., Stone, J. M., Denev, I., & Laloi, C. (2006). Reactive oxygen species as signals that modulate plant stress responses and programmed cell death. *BioEssays*, *28*(11), 1091–1101. https://doi.org/10.1002/bies.20493

Hargarten, A. M., Nemacheck, J. A., Subramanyam, S., Xiao, X., Schemerhorn, B. J., & Williams, C. E. (2017). Physical and metabolic consequences of Hessian fly infestation are more severe on nonhost *Brachypodium distachyon* than on host-plant resistant wheat. *Arthropod-Plant Interactions*, *11*(6), 767–783. https://doi.org/10.1007/s11829-017-9542-4

Harris, M. O., Freeman, T. P., Moore, J. A., Anderson, K. G., Payne, S. A., Anderson, K. M., & Rohfritsch, O. (2010). *H* -gene-mediated resistance to Hessian fly exhibits features of penetration resistance to fungi. *Phytopathology*, *100*(3), 279–289. https://doi.org/10.1094/PHYTO-100-3-0279

Harris, M. O., Friesen, T. L., Xu, S. S., Chen, M. S., Giron, D., & Stuart, J. J. (2015). Pivoting from *Arabidopsis* to wheat to understand how agricultural plants integrate responses to biotic stress. *Journal of Experimental Botany*, *66*(2), 513–531. https://doi.org/10.1093/jxb/eru465

Harris, M. O., Stuart, J. J., Mohan, M., Nair, S., Lamb, R. J., & Rohfritsch, O. (2003). Plant defense and insect adaptation. *Annual Review of Entomology*, *48*(1), 549–577. https://doi.org/10.1146/annurev.ento.48.091801.112559

He, L., Shi, X., Wang, Y., Guo, Y., Yang, K., & Wang, Y. (2017). Arabidopsis *ANAC069* binds to C[A/G]CG[T/G] sequences to negatively regulate salt and osmotic stress tolerance. *Plant Molecular Biology*, *93*(4–5), 369–387. https://doi.org/10.1007/s11103-016-0567-3

Heath, M. C. (2000). Hypersensitive response-related death. In *Programmed Cell Death in Higher Plants* (pp. 77–90). Springer Netherlands. https://doi.org/10.1007/978-94-010-0934-8_6

Jiang, Q., Niu, F., Sun, X., Hu, Z., Li, X., Ma, Y., & Zhang, H. (2017). RNA-seq analysis of unintended effects in transgenic wheat overexpressing the transcription factor *GmDREB1*. *The Crop Journal*, *5*(3), 207–218. https://doi.org/10.1016/j.cj.2016.12.001

Keinath, N. F., Kierszniowska, S., Lorek, J., Bourdais, G., Kessler, S. A., Shimosato-Asano, H., Grossniklaus, U., Schulze, W. X., Robatzek, S., & Panstruga, R. (2010). PAMP (Pathogen-associated Molecular Pattern) induced changes in plasma membrane compartmentalization reveal novel components of plant immunity. *Journal of Biological Chemistry*, *285*(50), 39140–39149. https://doi.org/10.1074/jbc.M110.160531

Kerchev, P. I., Fenton, B., Foyer, C. H., & Hancock, R. D. (2012). Plant responses to insect herbivory: interactions between photosynthesis, reactive oxygen species and hormonal signalling pathways. *Plant, Cell & Environment*, *35*(2), 441–453. https://doi.org/10.1111/j.1365-3040.2011.02399.x

Khajuria, C., Wang, H., Liu, X., Wheeler, S., Reese, J. C., El Bouhssini, M., Whitworth, R. J., & Chen, M. S. (2013). Mobilization of lipids and fortification of cell wall and cuticle are important in host defense against Hessian fly. *BMC Genomics*, *14*(1), 423. https://doi.org/10.1186/1471-2164-14-423

Kumar, R. R., Goswami, S., Sharma, S. K., Kala, Y. K., Rai, G. K., Mishra, D. C., Grover, M., Singh, G. P., Pathak, H., Rai, A., Chinnusamy, V., & Rai, R. D. (2015). Harnessing next generation sequencing in climate change: RNA-seq analysis of heat stress-responsive genes in wheat (*Triticum aestivum* L.). *OMICS: A Journal of Integrative Biology*, *19*(10), 632–647. https://doi.org/10.1089/omi.2015.0097

Lawrence, S. D., Novak, N. G., Jones, R. W., Farrar, R. R., & Blackburn, M. B. (2014). Herbivory responsive C2H2 zinc finger transcription factor protein *StZFP2* from potato. *Plant Physiology and Biochemistry*, *80*, 226–233. https://doi.org/10.1016/j.plaphy.2014.04.010

Liu, X. M., Fritz, A. K., Reese, J. C., Wilde, G. E., Gill, B. S., & Chen, M.-S. (2005). *H9*, *H10*, and *H11* compose a cluster of Hessian fly-resistance genes in the distal gene-rich region of wheat chromosome 1AS. *Theoretical and Applied Genetics*, *110*(8), 1473–1480. https://doi.org/10.1007/s00122-005-1982-z

Liu, X., Williams, C. E., Nemacheck, J. A., Wang, H., Subramanyam, S., Zheng, C., & Chen, M.-S. (2010). Reactive oxygen species are involved in plant defense against a gall midge. *Plant Physiology*, *152*(2), 985–999. https://doi.org/10.1104/pp.109.150656

Liu, Z., Xin, M., Qin, J., Peng, H., Ni, Z., Yao, Y., & Sun, Q. (2015). Temporal transcriptome profiling reveals expression partitioning of homeologous genes contributing to heat and drought acclimation in wheat (*Triticum aestivum* L.). *BMC Plant Biology*, *15*(1), 152. https://doi.org/10.1186/s12870-015-0511-8

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nature Methods*, *5*(7), 621–628. https://doi.org/10.1038/nmeth.1226

Park, C.-J., & Seo, Y.-S. (2015). Heat shock proteins: a review of the molecular chaperones for plant immunity. *The Plant Pathology Journal*, *31*(4), 323–333. https://doi.org/10.5423/PPJ.RW.08.2015.0150

Pearce, S., Kippes, N., Chen, A., Debernardi, J. M., & Dubcovsky, J. (2016). RNA-seq studies using wheat phytochrome B and phytochrome C mutants reveal shared and specific

functions in the regulation of flowering and shade-avoidance pathways. *BMC Plant Biology*, *16*(1), 141. https://doi.org/10.1186/s12870-016-0831-3

Ratcliffe, R. H., Cambron, S. E., Flanders, K. L., Bosque-Perez, N. A., Clement, S. L., & Ohm, H. W. (2000). Biotype composition of Hessian fly (Diptera: Cecidomyiidae) populations from the Southeastern, Midwestern, and Northwestern United States and virulence to resistance genes in wheat. *Journal of Economic Entomology*, *93*(4), 1319–1328. https://doi.org/10.1603/0022-0493-93.4.1319

Sardesai, N., Nemacheck, J. A., Subramanyam, S., & Williams, C. E. (2005). Identification and mapping of *H32*, a new wheat gene conferring resistance to Hessian fly. *Theoretical and Applied Genetics*, *111*(6), 1167–1173. https://doi.org/10.1007/s00122-005-0048-6

Satapathy, L., Singh, D., Ranjan, P., Kumar, D., Kumar, M., Prabhu, K. V., & Mukhopadhyay, K. (2014). Transcriptome-wide analysis of WRKY transcription factors in wheat and their leaf rust responsive expression profiling. *Molecular Genetics and Genomics*, *289*(6), 1289–1306. https://doi.org/10.1007/s00438-014-0890-9

Schweizer, F., Bodenhausen, N., Lassueur, S., Masclaux, F. G., & Reymond, P. (2013). Differential contribution of transcription factors to *Arabidopsis thaliana* defense against *Spodoptera littoralis*. *Frontiers in Plant Science*, *4*. https://doi.org/10.3389/fpls.2013.00013

Stuart, J. J., Chen, M.-S., Shukle, R., & Harris, M. O. (2012). Gall midges (Hessian flies) as plant pathogens. *Annual Review of Phytopathology*, *50*(1), 339–357. https://doi.org/10.1146/annurev-phyto-072910-095255

Subramanyam, S., Nemacheck, J. A., Bernal-Crespo, V., & Sardesai, N. (2021). Insect derived extra oral *GH32* plays a role in susceptibility of wheat to Hessian fly. *Scientific Reports*, *11*(1), 2081. https://doi.org/10.1038/s41598-021-81481-4

Subramanyam, S., Nemacheck, J. A., Hargarten, A. M., Sardesai, N., Schemerhorn, B. J., & Williams, C. E. (2019). Multiple molecular defense strategies in *Brachypodium distachyon* surmount Hessian fly (*Mayetiola destructor*) larvae-induced susceptibility for plant survival. *Scientific Reports*, *9*(1), 2596. https://doi.org/10.1038/s41598-019-39615-2

Subramanyam, S., Smith, D. F., Clemens, J. C., Webb, M. A., Sardesai, N., & Williams, C. E. (2008). Functional characterization of *HFR1*, a high-mannose N-glycan-specific wheat lectin induced by Hessian fly larvae. *Plant Physiology*, *147*(3), 1412–1426. https://doi.org/10.1104/pp.108.116145

Thimm, O., Bläsing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P., Selbig, J., Müller, L. A., Rhee, S. Y., & Stitt, M. (2004). Mapman: A user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *The Plant Journal*, *37*(6), 914–939. https://doi.org/10.1111/j.1365-313X.2004.02016.x

Tohge, T., & Fernie, A. R. (2012). Co-expression and co-responses: within and beyond transcription. *Frontiers in Plant Science*, *3*. https://doi.org/10.3389/fpls.2012.00248

Tuteja, N. (2007). Abscisic acid and abiotic stress signaling. *Plant Signaling & Behavior*, *2*(3), 135–138. https://doi.org/10.4161/psb.2.3.4156

Vo, K. T. X., Kim, C. Y., Hoang, T. V., Lee, S. K., Shirsekar, G., Seo, Y. S., Lee, S. W., Wang, G. L., & Jeon, J.-S. (2018). *OsWRKY67* plays a positive role in basal and *XA21*-mediated resistance in rice. *Frontiers in Plant Science*, *8*. https://doi.org/10.3389/fpls.2017.02220

Williams, C. E., Nemacheck, J. A., Shukle, J. T., Subramanyam, S., Saltzmann, K. D., & Shukle, R. H. (2011). Induced epidermal permeability modulates resistance and susceptibility of wheat seedlings to herbivory by Hessian fly larvae. *Journal of Experimental Botany*, *62*(13), 4521–4531. https://doi.org/10.1093/jxb/err160

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X., & Yu, G. (2021). ClusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation*, *2*(3). https://doi.org/10.1016/j.xinn.2021.100141

Xiong, H., Guo, H., Xie, Y., Zhao, L., Gu, J., Zhao, S., Li, J., & Liu, L. (2017). RNA-seq analysis reveals pathways and candidate genes associated with salinity tolerance in a spaceflight-induced wheat mutant. *Scientific Reports*, *7*(1), 2731. https://doi.org/10.1038/s41598-017-03024-0

Yu, Y., Wei, J., Zhang, X., Liu, J., Liu, C., Li, F., & Xiang, J. (2014). SNP discovery in the

transcriptome of white pacific shrimp *Litopenaeus vannamei* by next generation sequencing. *PLoS ONE*, *9*(1), e87218. https://doi.org/10.1371/journal.pone.0087218

Zhao, C., Escalante, L. N., Chen, H., Benatti, T. R., Qu, J., Chellapilla, S., Waterhouse, R. M., Wheeler, D., Andersson, M. N., Bao, R., Batterton, M., Behura, S. K., Blankenburg, K. P., Caragea, D., Carolan, J. C., Coyle, M., El-Bouhssini, M., Francisco, L., Friedrich, M., … Richards, S. (2015). A massive expansion of effector genes underlies gall-formation in the wheat pest *Mayetiola destructor*. *Current Biology*, *25*(5), 613–620. https://doi.org/10.1016/j.cub.2014.12.057

Zhao, L., Abdelsalam, N. R., Xu, Y., Chen, M.-S., Feng, Y., Kong, L., & Bai, G. (2020). Identification of two novel Hessian fly resistance genes *H35* and *H36* in a hard winter wheat line SD06165. *Theoretical and Applied Genetics*, *133*(8), 2343–2353. https://doi.org/10.1007/s00122-020-03602-3

Zhu, L., Liu, X., Liu, X., Jeannotte, R., Reese, J. C., Harris, M., Stuart, J. J., & Chen, M. S. (2008). Hessian Fly (*Mayetiola destructor*) attack causes a dramatic shift in carbon and nitrogen metabolism in wheat. *Molecular Plant-Microbe Interactions®*, *21*(1), 70–78. https://doi.org/10.1094/MPMI-21-1-0070

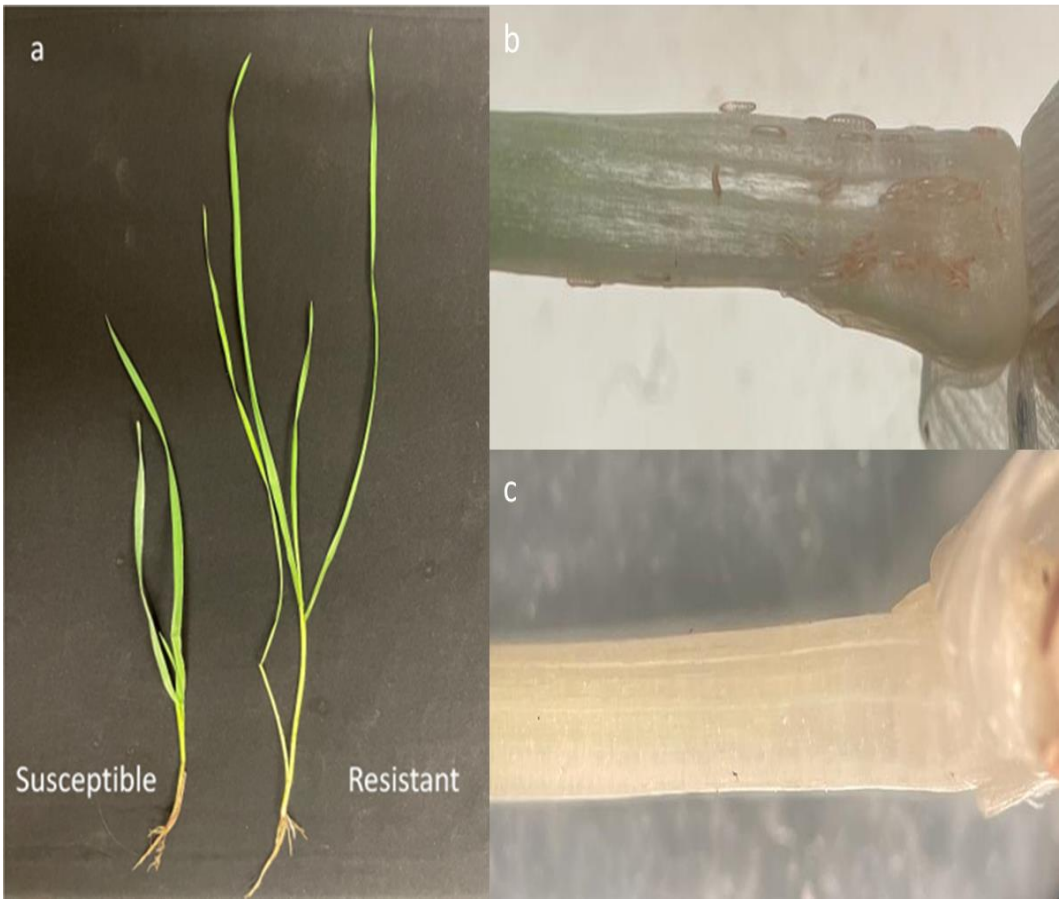

Figure 3.1. Responses of resistant and susceptible NILs to HF infestation two weeks after attack by HF larvae. A) Hessian fly-infested wheat seedlings showing normal growth in resistant NIL-R (right), but stunted plant in the susceptible NIL-S (left) (b) Living larvae lived on second leaf, close to the root crown of the susceptible NIL-S 12-h after initial attack by HF larvae. (c) Resistant line with no living larvae on NIL-R 12-h after initial attack by HF larvae.

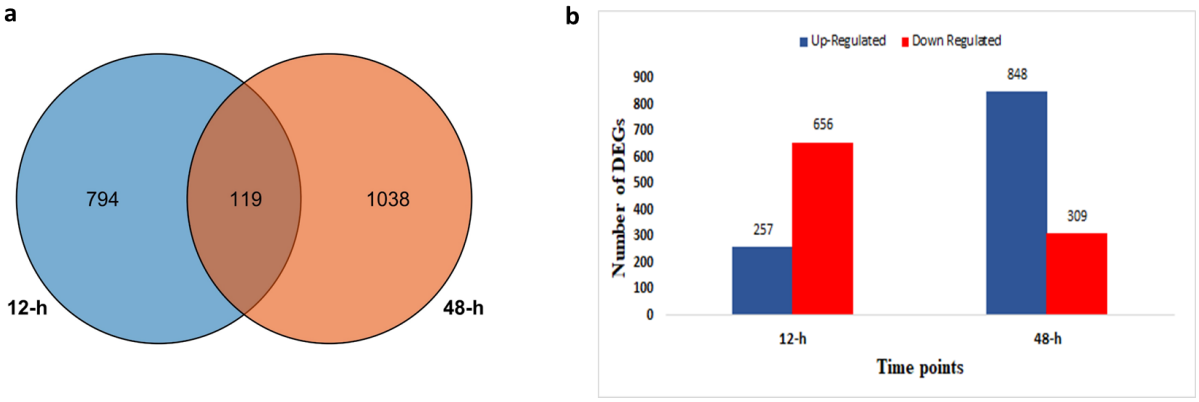

Figure 3.2. (a) Number of significantly (≥1 or ≤−1-fold change, $p < 0.05$) differentially expressed genes (DEGs) in the four pairs of resistant NIL-R compared to their susceptible NIL-S at 12-h (blue circle) and 48-h (orange circle) after HF larval attack. B) Bar chart comparison of total numbers of DEGs in NIL-R at 12-h and 48-h after infestation. The numbers of DEGs are marked on the top of each bar.

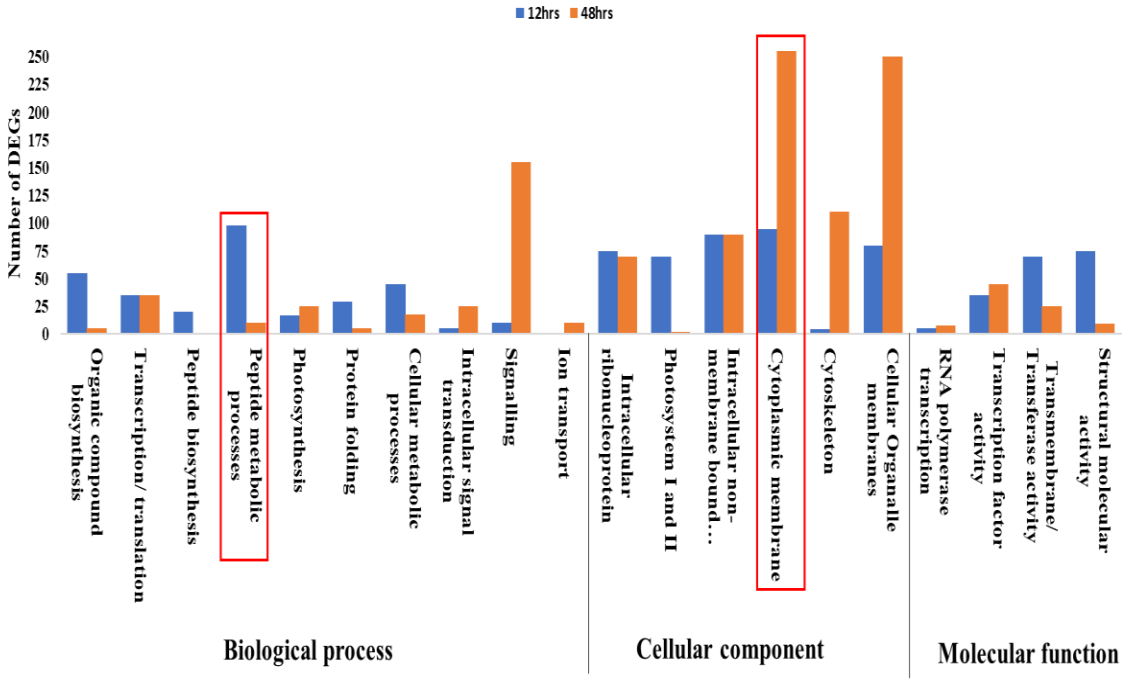

Figure 3.3. Gene ontology (GO) distribution of differentially expressed genes (DEGs) in the four pair of resistant NIL-R identified at 12-h and 48-h after larval attack on wheat. Blue bars represent DEGs at the 12-h and orange bars represent DEGs after the 48-h treatment.

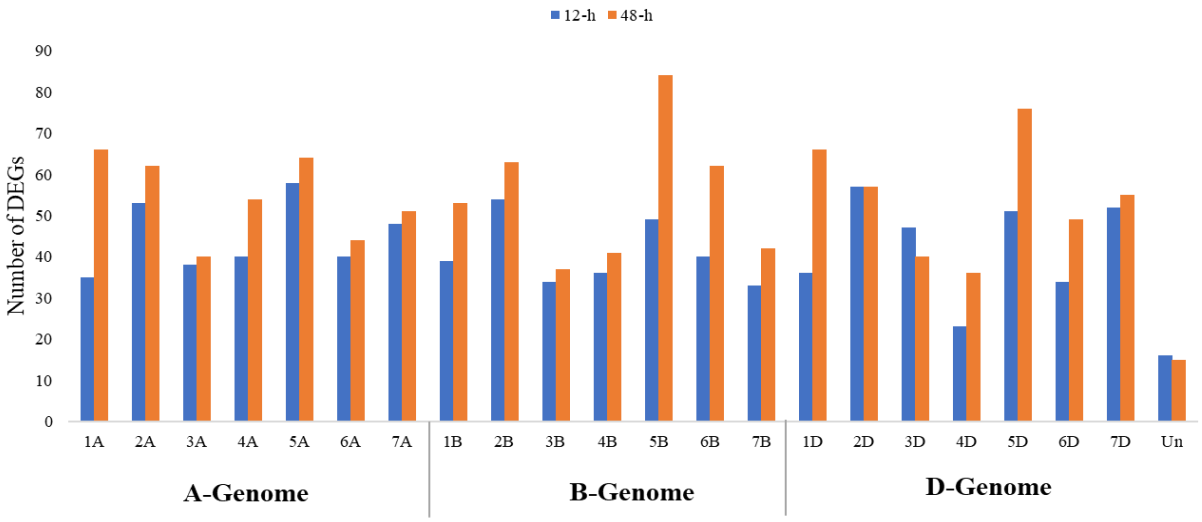

Figure 3.4. Chromosome distribution of the differentially expressed genes (DEGs) in the four pair of resistant NIL-R at 12-h and 48-h after larval attack on wheat. Blue bars represent DEGs after 12 hours and orange bars represent DEGs after 48 hours of Hessian fly larval attack.

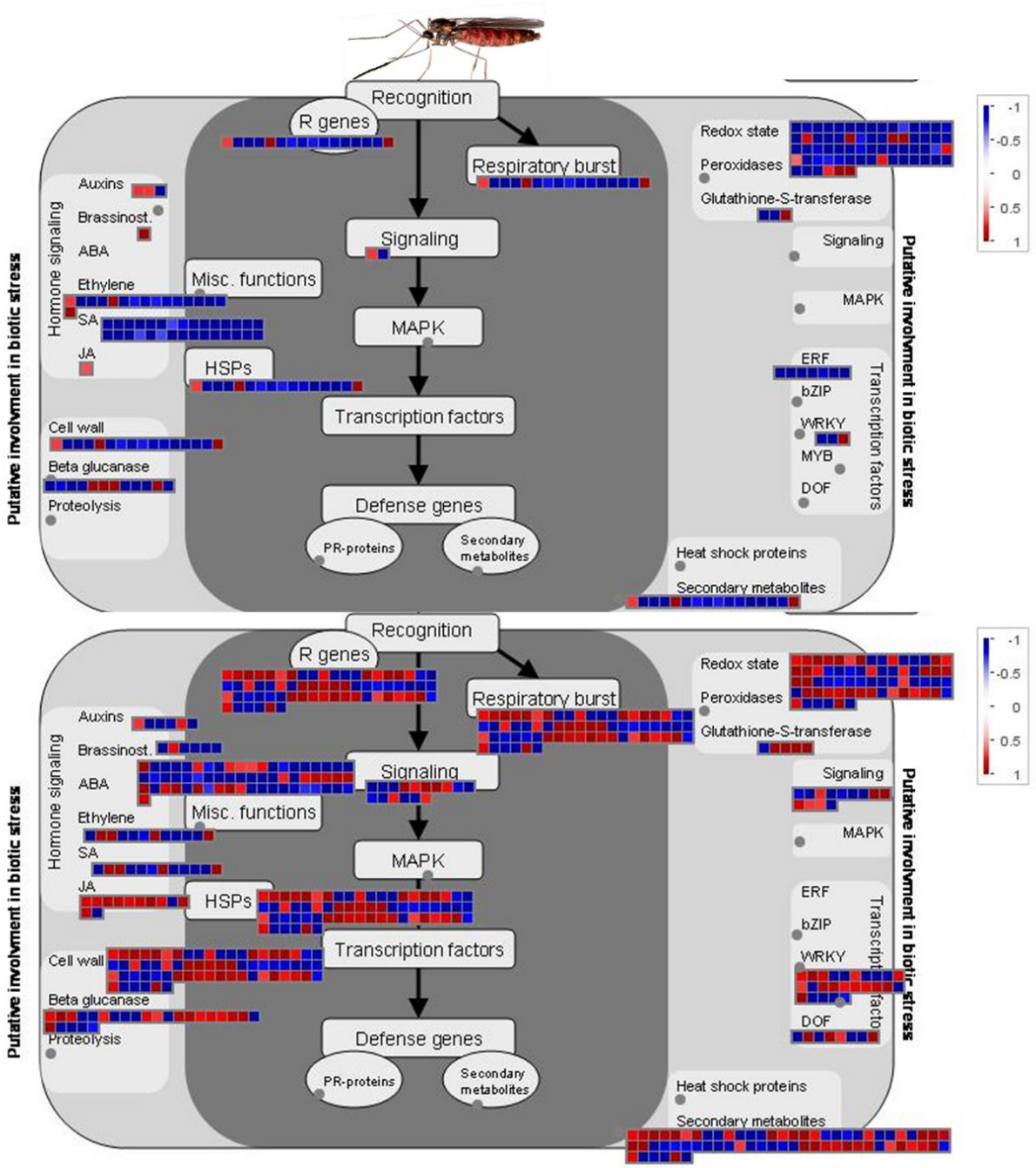

Figure 3.5. Pathway analysis highlighted significant differentially expressed genes (DEGs) ($p < 0.05$) involved in resistance against Hessian fly in the four pairs of resistant NILs at (a) 12-h and (b) 48-h after larval attack. Blue and red colors represent down and up-regulated DEGs, respectively.

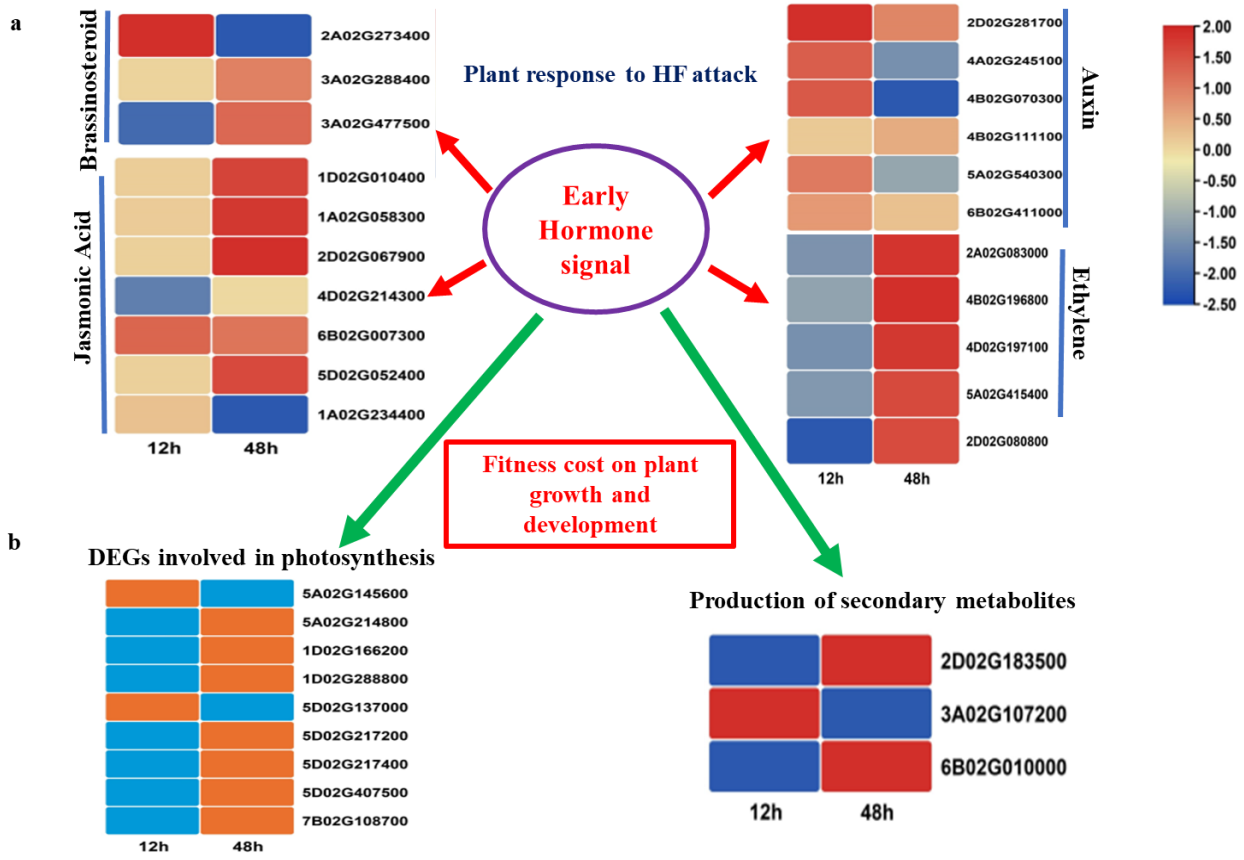

Figure 3.6. Differential expressed genes (DEGs) involving in hormonal homeostasis, secondary metabolite production and photosynthesis after 12-h and 48-h of HF larval attack in the four pairs of NIL-R. The DEGs are grouped within a heatmap based on their associated functions. Scales indicate log2-fold changes (≤−1 or ≥1). Gene identity (IDs) were named using their number only after removal of the prefix “TraesCS”.

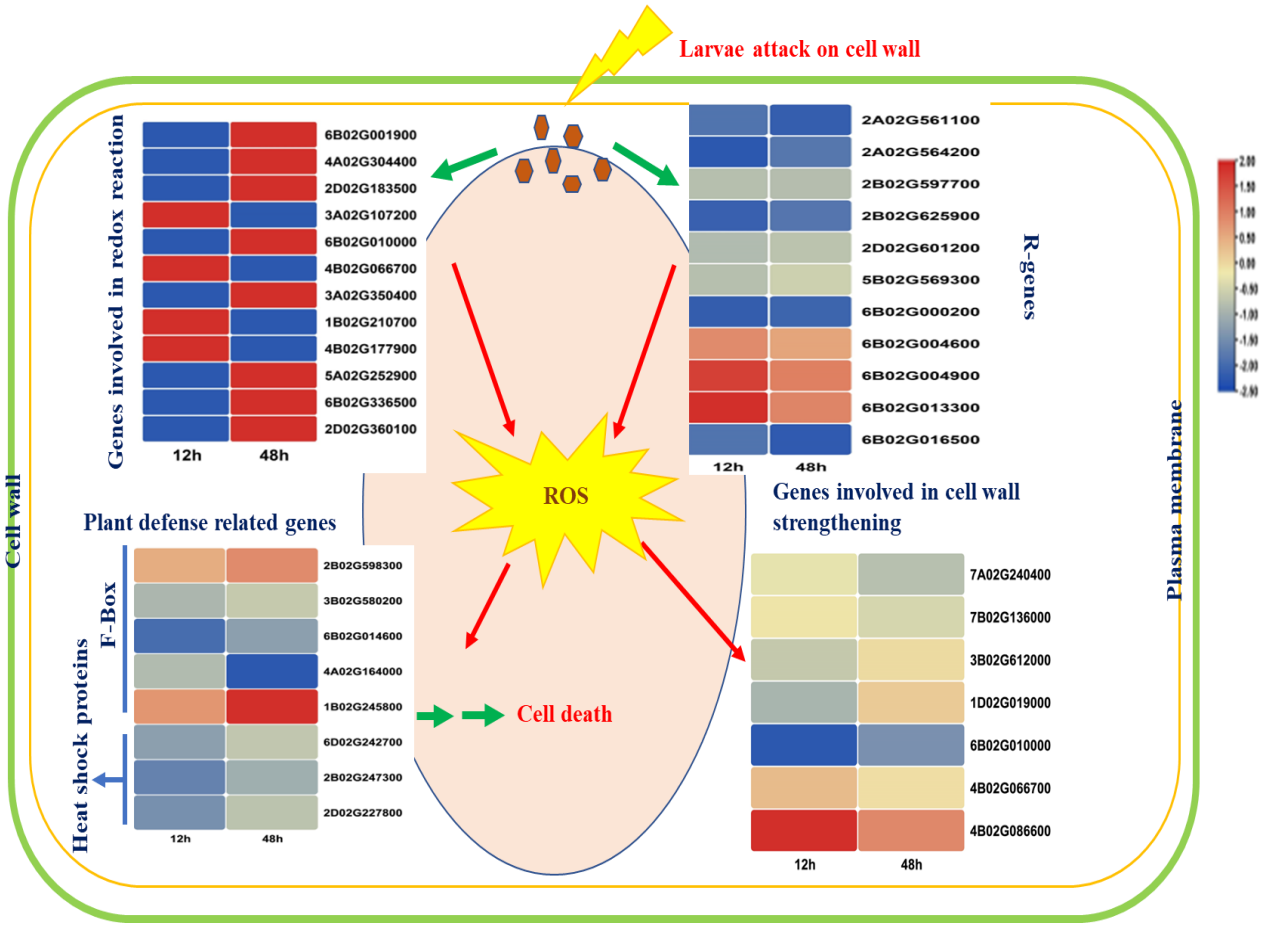

Figure 3.7. Expression profile of wheat differentially expressed R-genes and differentially expressed genes (DEGs) involved in redox reaction, defense signaling and cell wall strengthening 12-h and 48-h after HF infestation in the four pairs of resistant NIL-R. The DEGs are grouped within a heatmap based on their associated function. Red and blue represent up-regulated and down-regulated DEGs, respectively. Scales indicating log2 fold change (≤−1 or ≥1) shown on right side of the figure. Gene identification number (IDs) was written with their number only and prefix “TraesCS” is removed to make them short. ROS; Reactive oxygen species.

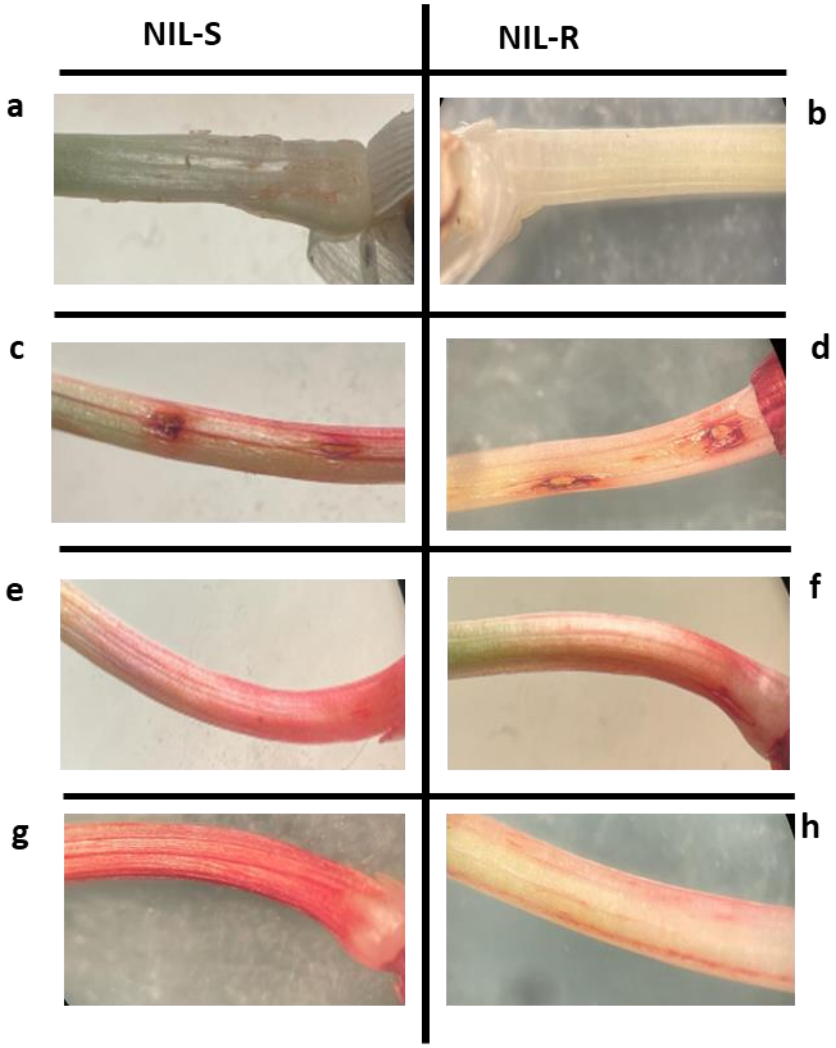

Figure 3.8. Neutral red stained wheat plant represents change in plant cell permeability in the wheat seedling at the bottom of second leaf right above the root crown harboring the HF larvae at 12-h and 48-h of larval attack. Left panel represent a susceptible (NIL-S) NIL-1; 115/118_F_3_43 and right panel is a resistant NIL-1; 115/118_F_3_43 (NIL-R). (a) NIL-S before staining; living larvae can be seen clearly (b) NIL-R before staining with no visible larvae. (c-d) Uninfested NIL-S and NIL-R punctured with a minute pin (diameter 200 µm) immediately before staining and taken as negative control (e-f) NIL-S and NIL-R after 12-h of HF infestation. (g-h) NIL-S and NIL-R after 48-h of HF infestation.

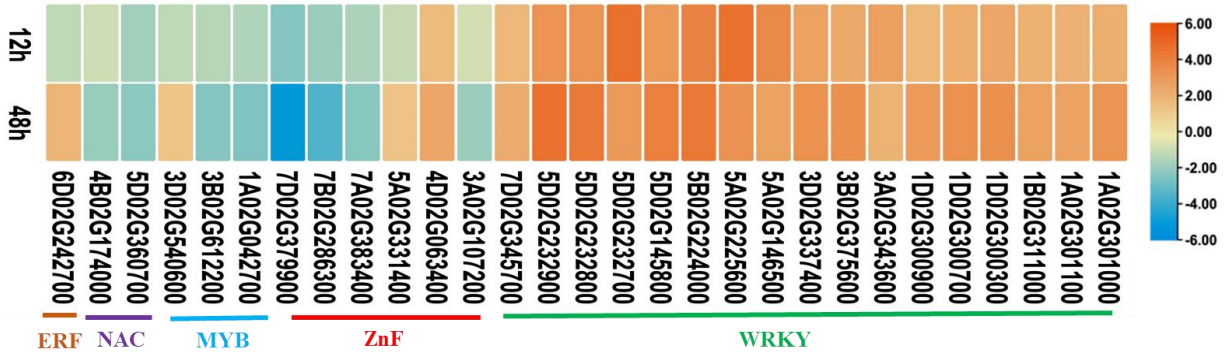

Figure 3.9. Differential expression of transcription factors in the four pairs of resistant NILs (NIL-R) attached by HF. Scales indicate log2-fold change (≤−1 or ≥1). shown on right side of the figure. Gene identification number (IDs) was written with their number only and prefix "TraesCS" is removed to make them short.

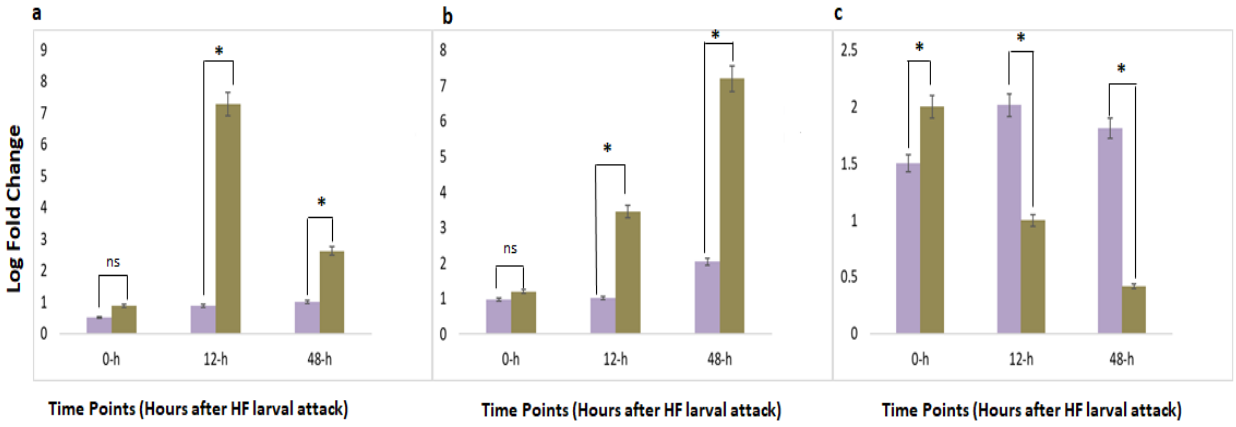

Figure 3.10. Verification of differentially expressed genes (DEGs) by qRT-PCR using NIL-1; 115/118_F_3_43 (a) TraesCS4B02G086600 (b) TraesCS1B02G245800 and (c) TraesCS4B02G070300 were verified. The x-axis in each chart representing the three time points 0-h: before attack of larvae; 12-h: 12 hours after attack of HF larvae; 48-h: 48 hours after attack of HF larvae and the y-axis representing log fold change value of expression determined by qRT-PCR. Purple bars represent susceptible NIL-1; 115/118_F_3_43 and green bars represent resistant NIL-1; 115/118_F_3_43. Asterisks representing significant differences between each comparison using a two-tailed Student's t-test. ns; non-significant difference between each comparison.

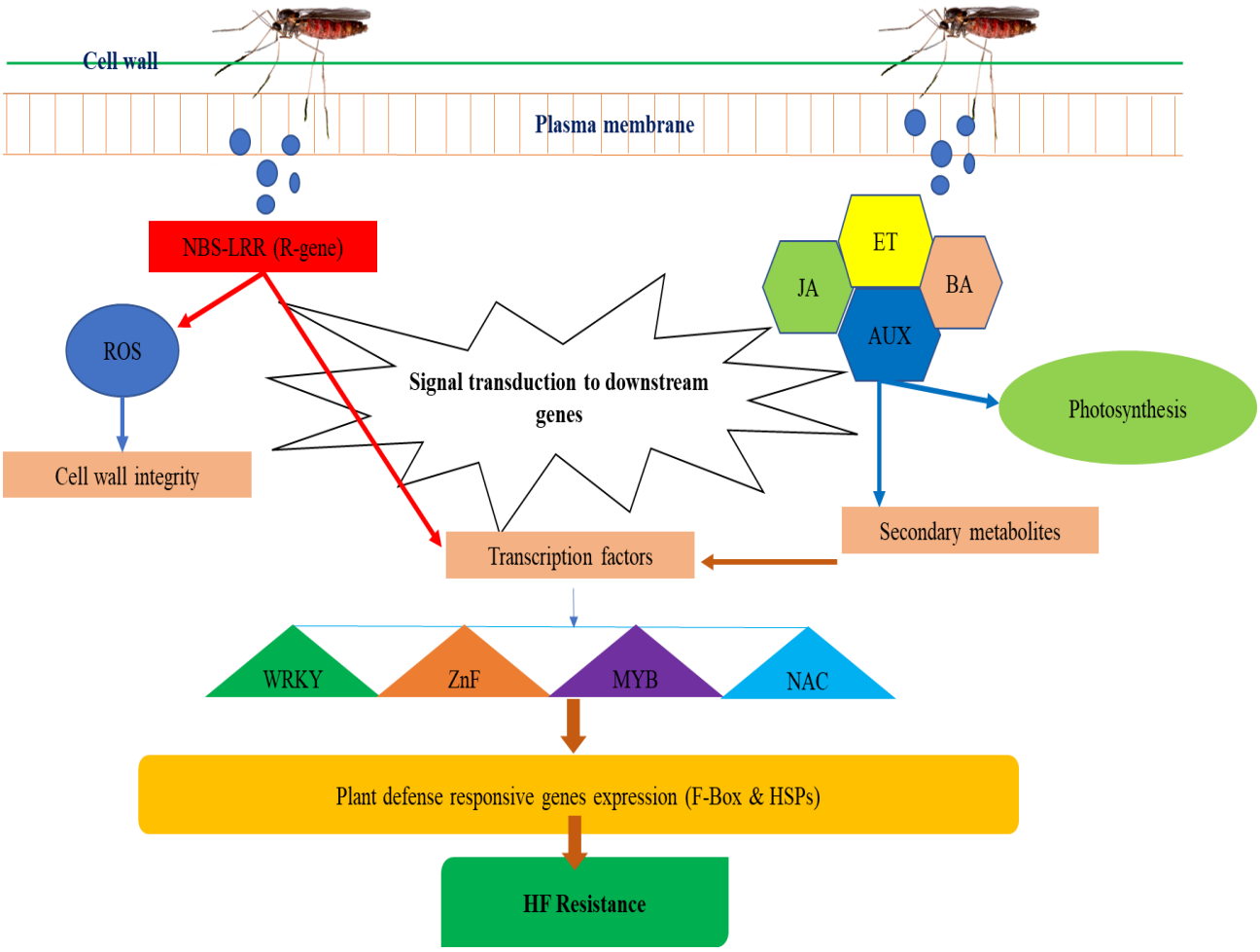

Figure 3.11. General figure showing mechanism of Hessian fly resistance in wheat. HSPs; Heat Shock proteins.

Table A.1. Primers used in this study for verification of the deferentially expressed genes by qRT-PCR

Primer Name	Gene Name	Sequence
4B02G086600F	TraesCS4B02G086600	GTGGGATCTGTGATTGTGATAA
4B02G086600R	TraesCS4B02G086600	TGGCTGCGACTGCAA
1B02G245800F	TraesCS1B02G245800	TGCGACGGCAACAACA
1B02G245800R	TraesCS1B02G245800	CCCGGATGTTGGCCAA
4B02G070300F	TraesCS4B02G070300	CTAGCGAGTGATGGGTGTGT
4B02G070300R	TraesCS4B02G070300	TTGAGAACACCACACCACAC

*F represent forward primer, R represent reverse gene primers

Appendix A - Supplementary material chapter 3

Table A.1: Summary of RNA-seq reads and percentage of reads mapped in the IWGSC RefSeq v2.1

Sample Names	Total Reads	Mapped	Uniquely Mapped (%)	Multi-mapped (%)
12-h_S_NIL1	65,727,678	62,399,403	89.97	4.97
12-h_S_NIL2	75,483,996	71,584,951	90.13	4.7
12-h _S_NIL3	80,231,046	75,403,019	89.18	4.8
12-h _S_NIL4	63,442,556	56,305,467	84.22	4.53
12-h_R_NIL1	52,193,376	49,112,994	89.49	4.61
12-h_R_NIL2	72,339,936	68,264,760	89.55	4.82
12-h_R_NIL3	96,373,446	90,279,949	88.55	5.12
12-h_R_NIL4	62,437,524	59,003,804	89.66	4.84
48-h_S_NIL1	59,470,816	55,963,080	88.89	5.21
48-h_S_NIL2	56,822,500	52,588,687	87.81	4.74
48-h_S_NIL3	52,583,576	48,356,117	86.71	5.26
48-h_S_NIL4	58,604,668	54,615,011	88.13	5.06
48-h_R_NIL1	46,417,584	42,267,327	86.36	4.7
48-h_R_NIL2	55,453,204	51,527,004	88.51	4.41
48-h_R_NIL3	52,902,632	48,741,656	87.53	4.6
48-h_R_NIL4	55,602,178	51,185,317	87.18	4.88

*12-h_S, 48-h_S represent data point 12 hours and 48 after initial attack of Hessian fly larvae on susceptible NILs. 12-h_R, 48-h_R represent data point 12 hours and 48 hours after initial attack of Hessian fly larvae on resistant NILs. The numbers 1 to 4 represent NILs from different HIFs. NIL1; 115/118_F_3_43, NIL2; 115/118_F_3_122, NIL3; 115/118_F_3_150, NIL4; 115/118_F_3_215.