

Identification of quantitative trait loci, candidate genes and diagnostic markers for Hessian fly  
resistance in wheat

by

Xiaoting Xu

B.S., Nanjing Agriculture University, 2016  
M.S., Chinese Academy of Agricultural Sciences, 2019

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

2024

## Abstract

Hessian fly (HF), *Mayetiola destructor* (Say), is a destructive insect pest of wheat (*Triticum aestivum* L.) worldwide. Growing resistant cultivars is the most effective approach for HF management. Although over 50 HF resistance (*H*) genes and quantitative trait loci (QTL) have been reported, only a few of them have been deployed in limited areas due to their loss of effectiveness against predominant HF biotypes, and lack of diagnostic markers for breeding selection. To discover new resistant sources, screening Pakistani wheat landraces and cultivars identified a highly HF resistant wheat cultivar ‘Parvaz-94’. Two QTL (*QHf.hwwg-1AS.2* and *QHf.hwwg-6BS.2*) were mapped using a high-density linkage map of single nucleotide polymorphisms (SNPs) developed from recombinant inbred lines (RILs) of Parvaz-94 × Cadenza. *QHf.hwwg-1AS.2* and *QHf.hwwg-6BS.2* were mapped to a 4.0 Mb interval (4.6–8.6 Mb) on 1AS and a 5.4 Mb interval (2.3–7.7 Mb) on 6BS, respectively, based on IWGSC RefSeq v2.1 reference. *QHf.hwwg-6BS.2* alone conferred complete HF resistance, whereas *QHf.hwwg-1AS.2* showed partial resistance. To facilitate cloning of *H35* on 3BS that was previously identified in the U.S. winter wheat cultivar ‘Overland’, a 14.57 Gb assembly of Overland (Overland v1.1) was generated using PacBio high-fidelity (HiFi) long reads. Screening of over 8,000 plants with the flanking markers identified 36 critical new recombinants in the *H35* interval. With 12 newly developed kompetitive allele specific PCR (KASP) markers, *H35* interval was finely mapped to a 38 kb interval in Chinese Spring (IWGSC RefSeq v2.1) and around 70 kb interval in Overland (v1.1), with only three and four putative genes, respectively, in the interval. Two of the genes significantly differentially expressed between two *H35* near isogenic lines using RNA-sequencing and were selected as the candidates for further function validation. To facilitate marker-assisted breeding for HF resistance, we developed and validated

breeder-friendly KASP markers for *QHf.hwwg-6BS.2*, *H35* and *H13*. *H13* on 6DS is a recently cloned resistance gene that originated from *Aegilops tauschii*; however, selectable markers were not available. In this study, KASP marker *K6B\_7697506* tightly linked to *QHf.hwwg-6BS.2* was developed and validated in three U.S. winter wheat diversity panels; *KB120-2*, a near diagnostic marker for *H35*, was validated in the AM203 panel and showed best performance in RGON2022 among four markers derived from *H35* candidate genes; *H13-1a*, converted from a nonsense mutation of *H13*, was validated in the AM203 panel and three soft winter wheat lines known to carry *H13*. These markers can be used for marker-assisted selection to facilitate quick deployment of these genes in U.S. wheat cultivars to enhance HF resistance level and durability in breeding programs.

Identification of quantitative trait loci, candidate genes and diagnostic markers for Hessian fly  
resistance in wheat

by

Xiaoting Xu

B.S., Nanjing Agriculture University, 2016  
M.S., Chinese Academy of Agricultural Sciences, 2019

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

2024

Approved by:

Co-Major Professor  
Dr. Allan K. Fritz

Approved by:

Co-Major Professor  
Dr. Guihua Bai

# Copyright

© Xiaoting Xu 2024.

## Abstract

Hessian fly (HF), *Mayetiola destructor* (Say), is a destructive insect pest of wheat (*Triticum aestivum* L.) worldwide. Growing resistant cultivars is the most effective approach for HF management. Although over 50 HF resistance (*H*) genes and quantitative trait loci (QTL) have been reported, only a few of them have been deployed in limited areas due to their loss of effectiveness against predominant HF biotypes, and lack of diagnostic markers for breeding selection. To discover new resistant sources, screening Pakistani wheat landraces and cultivars identified a highly HF resistant wheat cultivar ‘Parvaz-94’. Two QTL (*QHf.hwwg-1AS.2* and *QHf.hwwg-6BS.2*) were mapped using a high-density linkage map of single nucleotide polymorphisms (SNPs) developed from recombinant inbred lines (RILs) of Parvaz-94 × Cadenza. *QHf.hwwg-1AS.2* and *QHf.hwwg-6BS.2* were mapped to a 4.0 Mb interval (4.6–8.6 Mb) on 1AS and a 5.4 Mb interval (2.3–7.7 Mb) on 6BS, respectively, based on IWGSC RefSeq v2.1 reference. *QHf.hwwg-6BS.2* alone conferred complete HF resistance, whereas *QHf.hwwg-1AS.2* showed partial resistance. To facilitate cloning of *H35* on 3BS that was previously identified in the U.S. winter wheat cultivar ‘Overland’, a 14.57 Gb assembly of Overland (Overland v1.1) was generated using PacBio high-fidelity (HiFi) long reads. Screening of over 8,000 plants with the flanking markers identified 36 critical new recombinants in the *H35* interval. With 12 newly developed kompetitive allele specific PCR (KASP) markers, *H35* interval was finely mapped to a 38 kb interval in Chinese Spring (IWGSC RefSeq v2.1) and around 70 kb interval in Overland (v1.1), with only three and four putative genes, respectively, in the interval. Two of the genes significantly differentially expressed between two *H35* near isogenic lines using RNA-sequencing and were selected as the candidates for further function validation. To facilitate marker-assisted breeding for HF resistance, we developed and validated

breeder-friendly KASP markers for *QHf.hwwg-6BS.2*, *H35* and *H13*. *H13* on 6DS is a recently cloned resistance gene that originated from *Aegilops tauschii*; however, selectable markers were not available. In this study, KASP marker *K6B\_7697506* tightly linked to *QHf.hwwg-6BS.2* was developed and validated in three U.S. winter wheat diversity panels; *KB120-2*, a near diagnostic marker for *H35*, was validated in the AM203 panel and showed best performance in RGON2022 among four markers derived from *H35* candidate genes; *H13-1a*, converted from a nonsense mutation of *H13*, was validated in the AM203 panel and three soft winter wheat lines known to carry *H13*. These markers can be used for marker-assisted selection to facilitate quick deployment of these genes in U.S. wheat cultivars to enhance HF resistance level and durability in breeding programs.

# Table of Contents

|                                                                                                                        |      |
|------------------------------------------------------------------------------------------------------------------------|------|
| List of Figures .....                                                                                                  | xi   |
| List of Tables .....                                                                                                   | xiii |
| Acknowledgements .....                                                                                                 | xiv  |
| Dedication .....                                                                                                       | xv   |
| Chapter 1 - Literature Review .....                                                                                    | 1    |
| 1.1 Wheat .....                                                                                                        | 1    |
| 1.1.1 Origin and evolution of wheat .....                                                                              | 1    |
| 1.1.2 Biotic stress of wheat .....                                                                                     | 2    |
| 1.2 HF and its managements in wheat .....                                                                              | 3    |
| 1.2.1 HF life cycle .....                                                                                              | 3    |
| 1.2.2 HF biotypes .....                                                                                                | 4    |
| 1.2.3 HF management .....                                                                                              | 5    |
| 1.3 Wheat resistance to HF .....                                                                                       | 5    |
| 1.3.1 Mapping of <i>H</i> genes in wheat .....                                                                         | 6    |
| 1.3.1.1 Locations of <i>H</i> genes in wheat .....                                                                     | 7    |
| 1.3.1.2 Response of <i>H</i> genes in wheat to different HF biotypes .....                                             | 8    |
| 1.3.1.3 Markers for marker-assisted breeding .....                                                                     | 9    |
| 1.3.2 HF responsive gene identified by reverse genetics .....                                                          | 10   |
| 1.3.3 Mechanisms of wheat HF resistance .....                                                                          | 11   |
| 1.4 Genomic resources for improvement of HF resistance in wheat .....                                                  | 12   |
| 1.4.1 Genomic resources of wheat .....                                                                                 | 12   |
| 1.4.2 Genomic resources in wheat's wild relatives .....                                                                | 13   |
| 1.4.3 Long-read sequencing technology facilitates wheat genome assembly and functional genomics study .....            | 14   |
| 1.5 Perspectives .....                                                                                                 | 15   |
| Chapter 2 - Identification of two QTL for Hessian fly resistance in a Pakistan spring wheat cultivar 'Parvaz-94' ..... | 18   |
| 2.1 Introduction .....                                                                                                 | 18   |
| 2.2 Methods and materials .....                                                                                        | 18   |

|                                                                                                                           |    |
|---------------------------------------------------------------------------------------------------------------------------|----|
| 2.2.1 Plant materials.....                                                                                                | 18 |
| 2.2.2 Evaluation of HF resistance .....                                                                                   | 19 |
| 2.2.3 DNA extraction and genotyping-by-sequencing (GBS) analysis .....                                                    | 19 |
| 2.2.4 Linkage map construction and QTL analysis .....                                                                     | 20 |
| 2.2.5 KASP marker development and validation.....                                                                         | 20 |
| 2.3 Results.....                                                                                                          | 21 |
| 2.3.1 Segregation of HF resistance in the Parvaz-94 × Cadenza population.....                                             | 21 |
| 2.3.2 Linkage map construction.....                                                                                       | 22 |
| 2.3.3 QTL analysis and KASP marker development.....                                                                       | 22 |
| 2.3.4 QTL effects on HF resistance .....                                                                                  | 23 |
| 2.3.5 Validation of KASP markers .....                                                                                    | 24 |
| 2.4 Discussion.....                                                                                                       | 25 |
| Chapter 3 - Toward cloning of <i>H35</i> for Hessian fly resistance in the hard winter wheat cultivar<br>‘Overland’ ..... | 38 |
| 3.1 Introduction.....                                                                                                     | 38 |
| 3.2 Materials and methods .....                                                                                           | 39 |
| 3.2.1 Plant materials.....                                                                                                | 39 |
| 3.2.2 Evaluation of HF resistance .....                                                                                   | 39 |
| 3.2.3 HiFi sequencing, genome assembly and gene annotation of Overland genome.....                                        | 39 |
| 3.2.4 Fine mapping .....                                                                                                  | 40 |
| 3.2.5 RNA isolation, sequencing and data analysis .....                                                                   | 40 |
| 3.2.6 Putative candidate gene cloning.....                                                                                | 41 |
| 3.3 Results.....                                                                                                          | 42 |
| 3.3.1 Dissection of <i>H35</i> from <i>H36</i> .....                                                                      | 42 |
| 3.3.2 Overland genome assembly .....                                                                                      | 42 |
| 3.3.2 Fine mapping and identifying putative candidate genes of <i>H35</i> .....                                           | 43 |
| 3.3.3 Differentially expressed genes in the <i>H35</i> region.....                                                        | 44 |
| 3.3.4 Development and validation of gene markers.....                                                                     | 45 |
| 3.4 Discussion.....                                                                                                       | 46 |
| 3.4.1 <i>H35</i> fine mapping .....                                                                                       | 46 |
| 3.4.2 <i>H35</i> candidate genes prediction .....                                                                         | 47 |

|                                                                                                 |    |
|-------------------------------------------------------------------------------------------------|----|
| 3.4.3 Development of tightly linked markers .....                                               | 48 |
| 3.4.4 Conclusion .....                                                                          | 49 |
| Chapter 4 - Development of a diagnostic marker for Hessian fly resistance gene <i>H13</i> ..... | 62 |
| 4.1 Introduction.....                                                                           | 62 |
| 4.2 Materials and methods .....                                                                 | 63 |
| 4.2.1 Plant materials.....                                                                      | 63 |
| 4.2.2 Hessian fly infestation and evaluation .....                                              | 63 |
| 4.2.3 DNA extraction.....                                                                       | 63 |
| 4.2.4 KASP marker design and validation.....                                                    | 64 |
| 4.3 Results.....                                                                                | 64 |
| 4.3.1 Development of <i>H13</i> specific markers.....                                           | 64 |
| 4.3.2 Validation of KASP markers .....                                                          | 65 |
| 4.4 Discussion.....                                                                             | 66 |
| References.....                                                                                 | 72 |
| Appendix A - Copyright information .....                                                        | 83 |
| Appendix B - Supplementary data of Chapter 2.....                                               | 84 |
| Appendix C - Supplementary data of Chapter 3.....                                               | 98 |

## List of Figures

|                                                                                                                                                                                                                                                                                                                   |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1.1 Chromosomal distribution of genes and quantitative trait loci (QTL) for Hessian fly resistance in wheat .....                                                                                                                                                                                          | 16 |
| Figure 1.2 Responses of Hessian fly resistance genes to various Hessian fly populations .....                                                                                                                                                                                                                     | 17 |
| Figure 2.1 Frequency distribution of Hessian fly resistance scores in the recombinant inbred line population of Parvaz-94 × Cadenza .....                                                                                                                                                                         | 30 |
| Figure 2.2 The composite interval mapping of the <i>QHf.hwwg-1AS.2</i> and <i>QHf.hwwg-6BS.2</i> on chromosomes 1AS and 6BS, respectively .....                                                                                                                                                                   | 31 |
| Figure 2.3 Hessian fly resistance scores for homozygous recombinant inbred lines with different allelic combinations at <i>QHf.hwwg-1AS.2</i> and <i>QHf.hwwg-6BS.2</i> .....                                                                                                                                     | 32 |
| Figure 2.4 Phenotypic distribution of Hessian fly resistance scores (%) of the 203 accessions in AM203 panel grouped by the Kompetitive allele specific PCR (KASP) markers <i>K1A_7985775</i> linked to <i>QHf.hwwg-1AS.2</i> , and <i>K6B_6936780</i> and <i>K6B_7697506</i> linked to <i>QHf.hwwg-6BS</i> ..... | 33 |
| Figure 3.1 Fine mapping of <i>H35</i> .....                                                                                                                                                                                                                                                                       | 50 |
| Figure 3.2 Alignment of sequences in <i>H35</i> region from IWGSC RefSeq v2.1 and Overland v1.1 .....                                                                                                                                                                                                             | 52 |
| Figure 3.3 Heatmap of normalized read counts of putative annotated genes in the <i>H35</i> region in the near iso-genetic lines NIL52 (Resistant) and NIL53 (Susceptible) .....                                                                                                                                   | 53 |
| Figure 3.4 Alignment of genomic DNA sequences of <i>R300</i> from Overland and Overlay .....                                                                                                                                                                                                                      | 54 |
| Figure 3.5 Alignment of coding sequences of <i>R300</i> from Overland and Overlay .....                                                                                                                                                                                                                           | 55 |
| Figure 3.6 Alignment of coding sequences of <i>B120</i> predicted in Overland v1.1 and amplified from Overland and Overlay .....                                                                                                                                                                                  | 56 |
| Figure 3.7 Alignment of proteins sequences of B120 from Overland, Overlay and its homoeologous sequences from published wheat genomes.....                                                                                                                                                                        | 57 |
| Figure 3.8 Phenotypic distribution of Hessian fly resistance scores (%) of the 203 accessions in the AM203 panel grouped by eight Kompetitive allele specific PCR markers from <i>H35</i> candidate genes <i>R300</i> and <i>B120</i> .....                                                                       | 58 |

|                                                                                                                                                                                                                                                                                                                  |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 3.9 Phenotypic distribution of Hessian fly resistance scores (%) of the 327 accessions in the RGON2022 panel grouped by four Kompetitive allele specific PCR markers from <i>H35</i> candidate genes <i>R300</i> and <i>B120</i> .....                                                                    | 59 |
| Figure 4.1 Alignment of <i>H13</i> sequence from Molly to its homologous sequences in Chinese Spring RefSeq v2.1 reference genome to design 6D-specific KASP primers <i>H13-1a</i> and <i>H13-15</i> .....                                                                                                       | 69 |
| Figure 4.2 Validation of the Kompetitive Allele Specific PCR (KASP) markers <i>H13-1a</i> and <i>H13-15</i> for <i>H13</i> showing clear cluster separation in the F <sub>3:4</sub> population derived from Molly × Newton (a), the AM203 panel (b) and a set of 38 accessions with known genes or QTL (c) ..... | 70 |

## List of Tables

|           |                                                                                                                                                                                                                                                 |    |
|-----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Table 2.1 | Chi-square ( $\chi^2$ ) test of single gene segregation ratio for Hessian fly resistance in (Parvaz-94 $\times$ Cadenza) F <sub>3:4</sub> and F <sub>4:5</sub> population evaluated in two greenhouse experiments .                             | 34 |
| Table 2.2 | Primer sequences of Kompetitive allele specific PCR markers .....                                                                                                                                                                               | 35 |
| Table 2.3 | Chromosome positions, marker intervals and gene effects of two quantitative trait loci (QTL) for Hessian fly resistance detected in a recombinant inbred population of Parvaz-94 $\times$ Cadenza evaluated in two greenhouse experiments ..... | 37 |
| Table 3.1 | Statistics of the Overland genome assembly and annotation.....                                                                                                                                                                                  | 60 |
| Table 3.2 | Fourteen accessions with resistance allele using gene markers in the AM203 panel..                                                                                                                                                              | 61 |
| Table 4.1 | Sequences of KASP primers and single nucleotide flanking region .....                                                                                                                                                                           | 71 |

## Acknowledgements

I would like to thank my co-major professor Dr. Guihua Bai, for giving me the opportunity to join this lab and for his continuous support and guidance over these years. I would like to thank my co-major professor Dr. Allan Fritz, thesis committee members Dr. Ming-Shun Chen, Dr. Sanzhen Liu and Dr. Guorong Zhang for broadening my knowledge and providing valuable suggestions.

I would like to thank all the lab members, Dr. Yunfeng Xu, Dr. Paul St. Amand, Dr. Amy Bernardo, Dr. Ruolin Bian, Dr. Hui Chen, Dr. Liangliang Gao, Dr. Lanfei Zhao, Dr. Yuzhou Xu, Dr. Yuhui Pang, Dr. Nida Ghorri, Dr. Yujiao Gao, Dr. Tingting Zhou, Dr. Yi He, graduate students Ms. Wei Zhao, Ms. Julia Eilert, Ms. Jinan Park, undergraduate student Ms. Skylar Jost, collaborators Dr. Xiaomao Lin, Dr. Xuming Liu, Dr. Katherine Jordan and Dr. Hongliang Wang; without their supports, I would never finish all the projects.

I would like to thank my former mentors Dr. Zhonghu He, Dr. Yuanfeng Hao and Dr. Aizhong Cao, without their support in past research and the application process for PhD, I would never have had a chance to study here.

I would like to thank my friends, Dr. Yucheng Wang, Dr. Nenghan Wan, Ms. Yujie Qiu, Ms. Jin Liang, Ms. Olivia Baus, Dr. Zitong Yu, Dr. Hui Liu, Dr. Fan Liu, Ms. Zixian Wang, Ms. Yizhi Wang, Dr. Hao Deng and all other friends in the U.S. and China, for their support in life at Manhattan during these years.

I would like to thank my mother Ms. Kaiping Fan, my father Mr. Mingcai Xu, my cousin Ms. Juping Du and all other families; without them, I would never be able to go this far.

I truly appreciate everyone who loved and helped me. Thank you all for making this a wonderful journey!

## **Dedication**

*To Shujun Xia (夏淑君), 1944–2022.*

# Chapter 1 - Literature Review

## 1.1 Wheat

Wheat is one of the major staple food crops for the world population, with over 750 million tons of grain harvested annually from over 200 million ha. Wheat alone contributes approximately 20% of calorie and 25% of protein globally ([www.fao.org/faostat](http://www.fao.org/faostat)). Bread wheat (*Triticum aestivum* L.), also known as common wheat, occupies about 95% of the world wheat production while the remaining 5% is mainly durum wheat (*Triticum turgidum* ssp. *durum*). Bread wheat predominantly grows in temperate regions, but has spread to an unrivaled range of locations from 65° N in Norway and Russia to 45° S in Argentina (Dubcovsky & Dvorak, 2007).

### 1.1.1 Origin and evolution of wheat

Bread wheat is allohexaploid ( $2N = 6X = 42$ ), comprising of three subgenomes A, B and D. It formed through hybridization between domesticated free-threshing tetraploid wheat (BBAA) and *Aegilops tauschii* (DD) (Levy & Feldman, 2022). The domesticated tetraploid wheat can be traced to wild emmer (*T. turgidum* ssp. *dicoccoides*) that was formed ~0.8 million years ago (MYA) through hybridization between two diploid species (Marcussen et al., 2014). The B subgenome donor might be a distinct species related to *Ae. speltooides* that diverged from it around 4.5 MYA. The A subgenome originated from a species very similar to *T. urartu*, being diverged around 1.3 MYA (Marcussen et al., 2014). *Ae. tauschii* formed much earlier, around 5.5 MYA, via hybridization between ancestral donors of B and A subgenomes (Marcussen et al., 2014).

Domestication of wheat, einkorn (*T. monococcum*) and wild emmer in southeastern Turkey, part of the 'Fertile Crescent', marked the beginning of transition of human from hunting and gathering to agrarian lifestyles (Dubcovsky & Dvorak, 2007). During the domestication,

genetic traits, primarily non-shattering and free threshing, were selected by early farmers (Dubcovsky & Dvorak, 2007). Hexaploid wheat was formed around 8,500–9,000 years ago, then quickly spread from its cradle in the Fertile Crescent to new habitats and climates and eventually became a staple food globally (Feldman & Levy, 2023). The D subgenome in bread wheat contributed pest resistance and desirable grain qualities, promoting its wide spread and adoption over tetraploid wheat (Levy & Feldman, 2022). The most recent significant leap in wheat breeding was the Green Revolution of the mid-1960s, which dramatically increased wheat yield and adaptation through the selection of wheat cultivars for reduced photoperiod sensitivity and height (Venske et al., 2019).

### **1.1.2 Biotic stress of wheat**

Since the global human population may exceed nine billion by 2050, a 50% increase in wheat production is necessary to meet global demand (Figueroa et al., 2018). However, wheat production is threatened by both abiotic and biotic stresses and exacerbated by abrupt environmental changes and emerging new pathogens. Biotic stress alone causes up to 30% of annual yield loss in wheat globally (Savary et al., 2019). Among biotic stress, fungi, viruses and herbivorous insects have been the major pests affecting wheat grain yield and quality.

Wheat fungal diseases include stripe rust (*Puccinia striiformis* f. sp. *tritici*), leaf rust (*Puccinia triticina* f. sp. *tritici*), stem rust (*Puccinia graminis* f. sp. *tritici*), powdery mildew (*Blumeria graminis* f. sp. *tritici*), Fusarium head blight (*Fusarium graminearum*) and wheat blast (*Magnaporthe oryzae*) etc. Major viral threats include soil borne mosaic virus, soil-borne wheat streak mosaic virus and barley yellow dwarf virus (Afzal et al., 2015). Among the insect pests, aphids and Hessian fly [HF, *Mayetiola destructor* (Say)] pose significant threats, with regional variation in yield losses (Shukle, 2005). HF is a major destructive pest of bread wheat in

Southwest Asia, New Zealand, Europe, North Africa and North America (Harris et al., 2003; Stuart et al., 2012). To ensure food security, the development of high-yielding and high grain quality cultivars with resistance to multiple pathogens and pests, including HF, is necessary.

## **1.2 HF and its managements in wheat**

HF belongs to a large group of gall-inducing insects called gall midges (Diptera: Cecidomyiidae) (Harris et al., 2003). It is believed to have originated together with wheat in the Fertile Crescent region of the Middle East and was introduced to North America from Europe in the 1700s during the American Revolution (Harris et al., 2003). HF is now present in all major wheat-growing areas of the United States and cause significant economic losses (Chen et al., 2009a). Historically, HF has caused ~\$100 million losses in a single year (Cartwright & Jones, 1953). The severity of HF infestation varies among different locations and years. The losses were estimated around \$4 million in 1986 and around \$28 million in 1989 in Georgia, \$5 million in Texas in 1984 (Shukle, 2005), and 386 kg/ha in Oklahoma (Alvey, 2009). Although wheat is the preferred host for HF, other species such as rye, barley and some wild grass species can also serve as alternative hosts for HF (Chen et al., 2009b; Harris et al., 2003).

### **1.2.1 HF life cycle**

Many previous reports have detailed the life cycle of HF (Harris et al., 2003; Schmid et al., 2018; Shukle, 2005; Stuart et al., 2012), which includes egg, larva, puparium and adult stages. Adult flies resemble small mosquitoes and are short-lived, weak fliers. The female fly's abdomen is reddish due to developing eggs inside. It lays reddish, elongated eggs in rows on the upper surfaces of wheat leaves. A single female lays 50–400 eggs, which hatch in three to 10 days depending on temperatures. After hatching, the first instar-larva crawls toward the base of a plant to establish a permanent feeding site on a leaf sheath or a stem. First-instar larvae gradually

turn into white second-instar larvae as they continue feeding. After about two to three weeks, the larvae prepare to form the puparia and then turn to pupae, also known as flaxseeds (*Linum usitatissimum* L.), because of their similarity in appearance. Encased within the shiny brown, hardened skin of the flaxseed is the nonfeeding third-instar larva. A larva passes the summer and winter time within the puparium via diapause. HF can complete its life cycle in as short as a month (Schmid et al., 2018). Depending on temperature and moisture conditions, HF completes at least two generations a year in the northern states and three to six generations in the southern states of the United States (Buntin & Bruckner, 1990).

### **1.2.2 HF biotypes**

HF biotypes were historically classified based on their reaction to three resistance genes (*H3*, *H5*, *H6*) and a gene combination (*H7H8*). A total of 16 HF biotypes, A to O plus Great Plains (GP) biotypes, have been identified (Gallun, 1977; Ratcliffe et al., 1994). Among them, biotype GP, a predominant biotype in the U.S. Great Plains, is the least virulent biotype and remains avirulent to all the four genes or gene combination, therefore, has been used routinely to screen HF resistance in hard winter wheat in the Great Plains area. In contrast, biotype L is considered the most virulent, as it is virulent to all genes in the differential set. The other 14 biotypes were in between (Chen et al., 2009a). Although these known biotypes remain useful for studying HF genetics and wheat-HF interactions, they may not represent current predominant HF biotypes in farmers' fields due to quick evolution of HF (Nemacheck et al., 2023). Also, with many new HF resistance (*H*) genes being continuously discovered, the original 16 HF biotypes may not be sufficient to provide useful insight into the virulence status of predominant HF biotypes in relationship to newly identified resistance genes. Therefore, using field HF collections to evaluate the resistance of the *H* genes offers a more accurate assessment of the

distribution of current virulent HF biotypes and responses of the *H* genes in modern wheat cultivars to those biotypes (Chen et al., 2009a).

### **1.2.3 HF management**

Evaluating HF distribution and severity is crucial for devising effective management strategies. Cultural practices can play a significant role in managing HF. For example, delay planting after fly-free date can effectively reduce HF damage, although this strategy is only suitable for certain northern states and might lead to decreased wheat yields (Buntin et al., 1992). Given that adult HF are weak fliers with a limited host range, crop rotation with non-host crops can help prevent infestations to break its life cycle. Similarly, controlling volunteer wheat can inhibit the feeding of early emerging HF (Schmid et al., 2018). In addition, insecticide application can be effective when used for seed treatment as a preemptive application, while foliar insecticides cannot reach larvae inside the plants, rendering them ineffective and potentially cost prohibitive (Buntin, 1990). Planting resistant cultivars can reduce yield losses and provide better economic returns compared to other management options (Buntin et al., 1992; Chen et al., 2009a).

## **1.3 Wheat resistance to HF**

Wheat has developed resistance to HF through the long history of co-evolution with wheat and its relatives (Chen & Liu, 2023). To discover the HF resistance mechanisms, both forward and reverse genetics approaches have been employed, and have been crucial in advancing our understanding on HF resistance in wheat and facilitating the development of HF resistant wheat cultivars.

### 1.3.1 Mapping of *H* genes in wheat

Identification and characterization of new *H* genes in wheat and its relatives has been a continuous effort. To date, 37 *H* genes (*H1–H36* and *Hdic*) from wheat and its wild relatives have been formally named (McIntosh et al., 2013; Xu et al., 2021; Zhao et al., 2020), though *H* gene cloning work has not been reported in peer-reviewed journals. All the *H* genes are dominant except *h4*, a recessive gene from wheat cultivar ‘Java’ (Niu et al., 2020). Among those, only 11 genes (*H1*, *H2*, *H3*, *h4*, *H5*, *H7*, *H8*, *H12*, *H34*, *H35* and *H36*) originated from common wheat and others were from wheat relatives including 15 (*H6*, *H9*, *H10*, *H11*, *H14*, *H15*, *H16*, *H17*, *H18*, *H19*, *H20*, *H28*, *H29*, *H31* and *H33*) from durum wheat (*Triticum turgidum* ssp. *durum*), six (*H13*, *H22*, *H23*, *H24*, *H26* and *H32*) from *Aegilops tauschii*, two (*H21* and *H25*) from rye (*Secale cereale*). *H27*, *H30* and *Hdic* originated from *Aegilops ventricosa*, *Aegilops triuncialis* and cultivated primitive emmer wheat (*Triticum turgidum* ssp. *dicoccum*), respectively (McIntosh et al., 2013; Zhao et al., 2020).

In addition, about a dozen quantitative trait loci (QTL) with major effects on HF resistance have been reported. Those QTL include *QHf.hwwg-1AS* from a soft winter wheat cultivar ‘Clark’ (Li et al., 2013), *QHf.osu-1A* from a hard winter wheat cultivar ‘2174’ (Tan et al., 2013), *QHf.wak-1A* from a soft winter wheat cultivar ‘Louise’ (Carter et al., 2014), *QHf.osu-1A<sup>d</sup>* from a hard winter wheat cultivar ‘Duster’ (Li et al., 2015), *QH.icd-2A* identified in durum wheat cultivars ‘DWHF01’ and ‘Nassira’, originated from *Triticum turgidum* ssp. *dicoccum* (Bassi et al., 2019), *QHf.hwwg-3B* and *QHf.wwg-7A* from a hard winter wheat cultivar ‘Overland’ (Xu et al., 2021), *QHf.uga-6AL* from a soft winter wheat cultivar ‘26R61’ (Hao et al., 2013), *QHara.icd-6B* from *Triticum araraticum* (Bassi et al., 2019), *HKelse* from hard spring wheat cultivars ‘Kelse’ and ‘Seahawk’ (Prather et al., 2022), *HWGRC4* and *HNC09MDD14*

from *Aegilops tauschii* (Gill et al., 1991; Miranda et al., 2010), *QHf.uga-6AL* from a soft winter wheat cultivar ‘26R61’ (Hao et al., 2013), and *QHft.nc-7D* from soft winter wheat line “LA03136E71” (Winn et al., 2021).

#### **1.3.1.1 Locations of *H* genes in wheat**

To assign resistance genes to specific chromosomes, cytogenetic tools have been used in early studies. Using monosomic analysis, *H5* was assigned to chromosome 1A (Roberts & Gallun, 1984), *H6* was assigned to chromosome 5A (Gallun & Patterson, 1977), *H7* was assigned to chromosome 5D (Amri et al., 1990) and *H13* was assigned to chromosome 6D (Gill et al., 1987). However, monosomic analysis may introduce a high error rate in chromosome assignment, and it may not always locate the genes to correct chromosome locations. With advancement in DNA marker technologies such as restriction fragment length polymorphisms (RFLP), simple sequence repeats (SSR), and more recently, next-generation sequencing-based single nucleotide polymorphisms (SNPs) have been available for mapping genes in wheat. Using these markers, the locations of several previously mapped *H* genes have been reassigned, including *H6* being reassigned to chromosome 1A using SSR markers (Liu et al., 2005b), and *H5* and *H11* being remapped to 6B (unpublished data), *H7* being remapped to 6A (Liu et al., 2020), *H8* being remapped to 2B (Liu et al., 2020), *H12* being remapped to 3B (unpublished data).

Figure. 1.1 summarized the locations of previously reported 37 *H* genes and 14 major-effect QTL based on the physical position of gene/QTL-related markers in International Wheat Genome Sequencing Consortium (IWGSC) RefSeq v2.1 (Zhu et al., 2021). Among these genes, the chromosome locations of *H1*, *H2*, *H18*, *H19* and *H30* remain to be determined, while the rest of 46 genes/QTL were mapped on 15 wheat chromosomes excluding chromosomes 1B, 2D, 4B, 5A, 5D, 7B (Figure 1.1).

Three chromosome regions (1AS, 6BS and 6DS) harbor multiple genes clusters. Among them, 1AS has the biggest cluster of 16 genes/QTL including *H3*, *h4*, *H6*, *H9*, *H10*, *H14*, *H15*, *H16*, *H17*, *H28*, *H29*, *Hdic*, *QHf.hwwg-1AS*, *QHf.osu-1A*, *QHf.ous-1A<sup>d</sup>* and *QHf.wak-1A*, and half of them have been mapped to a region of less than 20 Mb. Five genes/QTL including *H5*, *H34*, *HKelse*, *QHara.icd-6B* and *QHf.hwwg-6BS* have been mapped to a ~20 Mb interval on chromosome arm 6BS. *H13*, *H23*, *HNC09MDD14* and *Hwgrc4*, all originated from *Ae. taushii*, were mapped to chromosome 6DS region. Investigating the relationships among these genes in each cluster would provide valuable insights to designing an effective strategy for their deployment in breeding to improve wheat HF resistance.

#### **1.3.1.2 Response of *H* genes in wheat to different HF biotypes**

Most of the *H* genes were identified decades ago and their effectiveness has changed since their initial deployment. Figure 1.2 summarized HF scores of reported *H* genes to HF populations collected from southern and central U.S. wheat growing regions including Alabama (AL), Georgia (GA), Kansas (KS), Louisiana (LA), North Carolina (NC), Oklahoma (OK), South Carolina (SC), and Texas (TX) in the United States (Cambron et al., 2010; Chen et al., 2009a; Schemerhorn et al., 2024; Shukle et al., 2015). Based on their average HF resistance scores, *H* genes can be classed into four categories: highly susceptible (<20%), moderate susceptible ( $\geq 20\%$  and < 50%), moderate resistant ( $\geq 50\%$  and <80%) and highly resistant ( $\geq 80\%$ ). Most of those *H* genes (*H1*–*H36* and *Hdic*) were evaluated at least once, except for *H27*, *H29*, *H30*, *H35* and *H36*. *H27*, *H29*, *H30* were identified from alien sources and might not be deployed in wheat cultivars yet. *H35* and *H36* have been identified recently and have not yet been included in these tests.

Many of previously reported genes or QTL either lost or reduced their resistance to current predominant biotypes. *H1H2*, *H3*, *h4*, *H5*, *H6*, *H7H8*, *H11*, *H14*, *H16* and *H34* have lost their effectiveness (< 50%) against most HF populations; *H9*, *H10*, *H12*, *H17*, *H22*, *H23* and *H24* showed susceptibility or reduced resistance to some collections although they exhibited broad effectiveness against most populations. Schemerhorn et al. (2024) tested *H15*, *H19*, *H20* and *H28* for resistance to an HF collection from Iberville Parish, LA, and observed that *H20* and *H28* remain highly resistant to the HF population, while *H15* and *H19* showed moderate resistance. *H13*, *H18*, *H21*, *H25*, *H26*, *H31*, *H32*, *H33* and *Hdic* remain resistant to all tested HF populations, therefore they are valuable sources of resistance for breeding (Figure 1.2).

Overall, the effectiveness of *H* genes varies against different HF populations and most resistance genes exhibit partial resistance. To extend the effectiveness and durability of the resistant cultivars, pyramiding of multiple resistance loci may be a wise strategy for sustainable HF management (Dweikat et al., 1997; Shukle et al., 2015). Development of DNA markers tightly linked to resistance genes will facilitate the deployment of the genes in breeding programs.

#### **1.3.1.3 Markers for marker-assisted breeding**

Linked or flanking markers have been developed for 25 *H* genes/QTL, and only a few of them are useful for marker-assisted breeding because many of the linked markers remain distant from the causal genes. Some markers co-segregated with HF resistance in bi-parental populations, but they may show high frequency false positive signal in diverse genetic backgrounds. Thus, it is necessary to validate these markers in a natural population. As the rapid development of next generation sequencing (NGS) technologies, generation of genome-wide, high-throughput SNPs becomes quick and affordable for routine breeding and genetics research.

Once the linked NGS-based SNP markers are identified, they can be converted to Kompetitive allele specific PCR (KASP) markers, which make it possible for genotyping of breeding materials with high throughput at affordable cost (Semagn et al., 2014).

To date, the linked markers have been validated only for six *H* genes/QTL. Tan et al. (2017) validated the KASP marker *IWB65911-KASP* and demonstrated its tight linkage to *H32* in a collection of 35 wheat lines consisting of 11 wheat lines and 14 synthetic lines. Ghori (2022) found the KASP markers *G0019400\_S4* and *HF6028581* were nearly diagnostic for *H34*, after validating them in a diversity panel of 203 U.S. winter wheat accessions (AM203). Xu et al. (2021) screened the AM203 panel with the *KASP-3B4525164* marker tightly linked with *QHf.hwwg-3B*, and found that 60.9% of 23 accessions carrying the resistance marker alleles showed HF resistance. Bassi et al. (2019) demonstrated that the KASP marker, *BS00072387*, showed 98% accuracy for predicting *QHara.icd-6B* in a population of 94 ICARDA's elite lines. Using haplotype analysis, Prather et al. (2022) identified a haplotype of three KASP markers, *kelse6BS\_167037*, *kelse6BS\_4554201* and *kelse6BS\_6196634* that tightly linked to *HKelse* in a panel of 250 North American spring wheat lines. Zhang et al. (2021) reported the left flanking marker *KASP-6B112698* was nearly diagnostic for *QHf.hwwg-6BS* in the AM203 panel. The limited availability of effective markers for breeding selection highlights the urgency for developing diagnostic markers for *H* genes with broad effectiveness.

### **1.3.2 HF responsive genes identified by reverse genetics**

Several studies have investigated the possible mechanisms of wheat resistance to HF by examining plant and HF interactions from the perspective of reverse genetics. *Hfr-1* (Hessian fly-response gene 1), *Hfr-2* and *Hfr-3* encoded insecticidal lectins were upregulated in wheat cultivar 'Iris' containing *H9* resistance gene during interaction with virulent or avirulent HF biotypes

(Giovanini et al., 2007; Puthoff et al., 2005; Williams et al., 2002). *HfrDrd*, a dirigent-like protein, increased transcript expression levels within a few hours of HF infestation in Iris (*H9*), which may protect the plants from HF attack by mediating the formation of lignins and lignans in the cell envelope to thicken the host plant cell wall as a barrier (Subramanyam et al., 2013). *Mds-1* (*Mayetiola destructor susceptibility-1*), a heat-shock protein cloned from a susceptible wheat ‘Newton’, is a major susceptibility gene for infestation of wheat by HF. One of the mechanisms for *Mds-1* may be the induction of virulent effectors from HF, which then activates susceptibility pathways by upregulation of nutrient metabolic pathways and the formation of nutritive cells (Liu et al., 2013).

### **1.3.3 Mechanisms of wheat HF resistance**

HF may share similarities with biotrophic or hemibiotrophic plant pathogens based on its feeding behavior (Stuart et al., 2012). The mandibles of first-instar larvae appear to act in a manner that is analogous to nematode stylets and the haustoria of filamentous plant pathogens (Stuart et al., 2012). Delivering salivary fluids that induce gall by turning plant cells to nutritive feeding cells resembles effectors from pathogens that manipulate host plant responses (Harris et al., 2010). Generally, two types of plant defense mechanisms have been reported: pathogen-associated molecular patterns (PAMPs)-triggered immunity (PTI) or innate immunity and effector-triggered immunity (ETI) or induced defense (Jones & Dangl, 2006).

PTI is present in all plants and is triggered by a general perception of PAMPs or damage-associated molecular patterns (DAMPs) (Zipfel, 2014). This recognition triggers responses such as hypersensitive response, reactive oxygen species (ROS) accumulation, cell wall modifications, and the release of insect-toxic molecules, including wheat protease inhibitors,

lectins, and enzymes that produce phenylpropanoids (Chen et al., 2016; Khajuria et al., 2013; Liu et al., 2010).

On the contrary, ETI is related to plant evolved *R* gene specific recognition of effectors released by pathogens, also known as gene-for-gene interactions. The gene-for-gene model applying to the genetic interaction between wheat and HF was first proposed by Hatchett and Gallun (1970). Over the past five decades, only one HF resistance genes (*H13*) have been cloned in wheat (not peer-reviewed), encoding a nucleotide binding (NB) and leucine-rich repeat (LRR) protein, a typical *R* gene for plant disease resistance (Joshi, 2018). Some progresses have been made in the study of effectors in HF. Six HF *Avr* genes (*vH5*, *vH6*, *vH9*, *vH13*, *vH24*, and *vHdic*) have been mapped (Behura et al., 2004; Harris et al., 2003; Schulte et al., 1999; Zhao et al., 2016) and a large number of genes encoding putative effector proteins or secreted salivary gland proteins (SSGPs) have been identified (Chen et al., 2016; Zhao et al., 2015). Moreover, one of the gene family SSGP-71 encoding proteins resemble ubiquitin E3 ligases in plants and E3-ligase-mimicking effectors in plant pathogenic bacteria was proved to have effector function by mutation (Zhao et al., 2015).

## **1.4 Genomic resources for improvement of HF resistance in wheat**

### **1.4.1 Genomic resources of wheat**

The whole genome assemblies have facilitated research in functional genomics including gene cloning, eventually accelerating crop improvement. However, wheat has lagged behind other cereal crops such as rice and maize in the application of the genomic tools for its improvement, because it has the largest genome size (~16 Gb) with 85% repeated sequences (IWGSC et al., 2018). To decode the bread wheat genome, scientists and breeders initiated the IWGSC in 2005. Thirteen years later, the first assembly of a reference-quality whole genome

sequence was released for bread wheat cultivar Chinese Spring (IWGSC RefSeq v1.1) in 2018 (IWGSC et al., 2018). Continuous efforts to improve the quality of the assembly using new technologies available have resulted in refined IWGSC RefSeq v2.1 (Zhu et al., 2021). The new sequence assembly has a total size of 14.58 Gb and includes annotations for 108,010 high confidence genes and 161,535 low confidence genes (Zhu et al., 2021). In addition, wheat pangenome project generated ten chromosome pseudomolecules and five scaffold assemblies (Walkowiak et al., 2020). More recently, chromosome-scale assembly of wheat cultivars ‘Fielder’, ‘Kariega’, ‘Renan’ and ‘Attraktion’ have been generated using long-read sequencing technologies (Athiyannan et al., 2022; Aury et al., 2022; Kale et al., 2022; Sato et al., 2021). These whole genome assemblies will be useful for gene cloning and functional study and for development of diagnostic markers for important traits identified from different resources.

#### **1.4.2 Genomic resources in wheat’s wild relatives**

Wheat’s wild relatives may preserve rich genetic diversity for disease and insect resistance genes. Among the 51 *H* genes/QTL, only 21 were from bread wheat, 18 were derived from tetraploid wheat, eight from *Ae. tauschii* and four from other wild relatives of wheat. The genetic resources in wheat and its relatives play a crucial role in breeding for pest resistance.

*Ae. taushii* (~4.3 Gb) is a diploid progenitor of bread wheat and has a relatively smaller and less complicated genome than tetraploid and hexaploid wheat, which made it the first to be decoded. The first draft genome of *Ae. taushii* cultivar AL8/78 was published in 2013, with 4.23 Gb scaffolds representing 83.4% of the genome (Jia et al., 2013). Later, a reference-quality genome sequence (Aet v4.0) was generated with 4.0 Gb pseudomolecules (Luo et al., 2017), and continuous improvement of the assembly led to Aet v5.0 featuring greater sequence contiguity

and improved annotation (Wang et al., 2021). Recently, four new reference genomes have been assembled from *Ae. tauschii* accessions AY17, AY61, T09 and XJ02 (Zhou et al., 2021).

Tetraploid wheat—another progenitor of AABB genomes in bread wheat—has a large genome (~12 Gb) with over 80% repeated sequences. The first reference-level genome of the tetraploid wheat was published in 2017, with a 10.1 Gb assembly (WEW\_v1.0) of wild emmer cultivar Zavitan, representing 87.5% of estimated tetraploid wheat genome (Avni et al., 2017). The assemble was improved to 10.4 Gb (WEW\_v2.0) using an optical map with Bionano Direct Label and Stain (DLS) technology (Zhu et al., 2019). Subsequently, the first durum wheat assembly of cultivar Svevo (10.45 Gb) was released in 2019 (Maccaferri et al., 2019). Another durum wheat cultivar, Kronos, has also been sequenced ([GrainGenes](#)).

### **1.4.3 Long-read sequencing technology facilitates wheat genome assembly and functional genomic studies**

Long-read sequencing technologies have resulted in establishment of several high-quality, chromosome-level genome assemblies for bread wheat cultivars (Athiyannan et al., 2022; Aury et al., 2022; Sato et al., 2021). For example, Sato et al. (2021) assembled 21 pseudomolecules with a cumulative size of 14.7 Gb for a U.S. wheat cultivar ‘Fieldier’ using Pacific Biosciences’ (PacBio) single-molecule real-time (SMRT) sequencing technology with the long high-fidelity (HiFi) reads. Athiyannan et al. (2022) generated a 14.7 Gb chromosome-scale assembly using the South African wheat cultivar ‘Kariega’ by combining HiFi reads, optical mapping and chromosome conformation capture. This assembly facilitated the cloning of the race-specific resistance gene *Yr27*, which encodes an intracellular immune receptor. Aury et al. (2022) assembled a 14.2 Gb genome of the French wheat cultivar ‘Renan’ using Oxford Nanopore Technologies’ (ONT) contigs that were polished with Illumina short reads.

Additionally, Kale et al. (2022) constructed a 14.7 Gb genome sequence assembly for a European wheat cultivar ‘Attraktion’ using HiFi reads and conducted reference base genome-wide associations scans for the yellow rust and leaf rust resistance genes. With advent of long-read sequencing technologies, the complex wheat genome has become more accessible, ultimately aiding in dissecting resistance genes and developing the next generation of modern wheat cultivars.

### **1.5 Perspectives**

Continuing to produce elite wheat cultivars with HF resistance is critical for maintaining and increasing wheat yield in the future to feed the growing global population. Wheat germplasm lines and its wild relatives provide a diverse pool for HF resistance genes. Exploring new sources of HF resistance and developing reliable markers will facilitate the breeding of cultivars with durable resistance. Advances in sequencing technologies will enhance genomic resources and aid in understanding the mechanisms of wheat resistance to HF.



| Hessian fly                                 | S <20%                        | 20 ≤ MS <50%                       | 50 ≤ MR <80%               | R ≥80%                                                                         |
|---------------------------------------------|-------------------------------|------------------------------------|----------------------------|--------------------------------------------------------------------------------|
| KS<br>(Chen et al, 2009)                    | H1H2                          | H3, h4, H10, H11,<br>H16, H22, H24 | H5, H6, H12,<br>H17, H18   | H7H8, H9, H13, H21,<br>H23, H25, H26, Hdic                                     |
| OK<br>(Chen et al, 2009)                    | h4, H5, H6,<br>H7H8, H9, H11  | H3, H12, H16,<br>H17, H24          | H10, H18                   | H13, H21, H22, H23,<br>H25, H26, Hdic                                          |
| TX<br>(Chen et al, 2009)                    | H6                            | H7H8, H9                           | h4, H16, H17,<br>H24, H31  | H3, H5, H10, H11, H12,<br>H13, H18, H21, H22,<br>H23, H25, H26, Hdic           |
| AL, GA, NC, SC, LA<br>(Cambron et al, 2010) | H3, H5, H6,<br>H7H8, H11      | H14, H16, H23                      | H9, H10, H17,<br>H22, H31  | H12, H13, H18, H21,<br>H24, H25, H26, H32                                      |
| AL, GA, NC<br>(Shukle et al, 2015)          | H3, H5, H6,<br>H7H8, H34      |                                    | H9                         | H12, H13, H18, H24, H25,<br>H26, H31, H32, H33, Hdic                           |
| LA<br>(Schemerhorn et al,<br>2024)          | H3, H5, H6,<br>H7H8, H11, H34 | H23                                | H10, H14, H15,<br>H19, H24 | H9, H12, H13, H16, H17,<br>H18, H20, H21, H22, H25,<br>H26, H28, H31, H32, H33 |

**Figure 1.2 Responses of Hessian fly resistance genes to various Hessian fly populations**

The Hessian fly populations were collected from the southern and central U.S. wheat growing states including Alabama (AL), Georgia (GA), Kansas (KS), Louisiana (LA), North Carolina (NC), Oklahoma (OK), South Carolina (SC), and Texas (TX).

## **Chapter 2 - Identification of Two QTL for Hessian Fly Resistance in a Pakistan Spring Wheat Cultivar ‘Parvaz-94’**

### **2.1 Introduction**

Pakistan is a country near the center of wheat origin. To discover new resistant sources from Pakistani wheat, we screened a panel of 184 Pakistani wheat landraces and cultivars for HF resistance (Hussain et al., 2022) and found that the cultivar ‘Parvaz-94’ is highly resistant to HF. The objectives of this study were to determine the QTL responsible for HF resistance in Parvaz-94 using a population of recombinant inbred lines (RILs) from Parvaz-94 × Cadenza and develop breeder-friendly markers tightly linked to the QTL for effective selection of these QTL in breeding.

### **2.2 Methods and materials**

#### **2.2.1 Plant materials**

A panel of 184 wheat cultivars and landraces from Pakistan was infested with HF biotype Great Plains (GP) (Hussain et al., 2022) and Parvaz-94 was identified as the sole accession with a high level of resistance (Table B.1). To identify QTL in Parvaz-94, a mapping population of 178 RILs was developed by crossing Parvaz-94 to Cadenza, a susceptible spring wheat cultivar from the United Kingdom, using single seed decent. The population was screened for HF resistance in two greenhouse experiments using ‘Carol’ (*H3*), ‘Caldwell’ (*H6*) and ‘Molly’ (*H13*) as the resistant controls and ‘Newton’ as the susceptible control. A diversity panel of 203 (AM203) U.S. winter wheat accessions (Zhang et al., 2010) was used to validate Kompetitive allele specific PCR (KASP) markers linked to the QTL identified in this study. Near diagnostic markers were selected to further screen two panels of 407 U.S. hard winter wheat germplasm from 2022 Regional Germplasm Observation Nursery and Regional Performance Nurseries (RGON2022 and

RPN2022; <https://www.ars.usda.gov/plains-area/lincoln-ne/wheat-sorghum-and-forage-research/docs/hard-winter-wheat-regional-nursery-program/research/>).

### **2.2.2 Evaluation of HF resistance**

To evaluate the RILs and their parents for HF resistance, HF biotype GP was used to infest the wheat seedlings as previously described (Li et al., 2013). The HF biotype GP is the predominant biotype in the U.S. Great Plains and maintained in the USDA Hard Winter Wheat Genetics Research Unit, Manhattan, Kansas (Chen et al., 2009a). The F<sub>3:4</sub> and F<sub>4:5</sub> RILs, together with their parents and controls were evaluated in the 2022 and 2023 fall greenhouse experiments. In each experiment, about 20 seeds per line were planted in each half row in plastic flats (56×36×10 cm) filled with a mixture of soil, sand and vermiculite (2:1:1). Each flat was divided into 24 15-cm half rows and four checks were planted in the middle two rows (four half-rows). All flats were arranged on greenhouse benches with a 14/10 h (light/dark) photoperiod under 20/18 °C day/night.

Wheat seedlings at 1.5-leaf stage were infested by placing HF stock trays with emerging flies within a mesh cloth tent on each bench. The HF stock trays were removed from the tent once egg density reached 5–10 per tested plant. Three weeks after infestation, wheat phenotypic responses to HF were scored as percent resistant plants per line. Plants that grew normally with a light green color and dead HF larvae were scored as resistant, whereas stunted plants with a dark green color and live HF larvae were scored as susceptible.

### **2.2.3 DNA extraction and genotyping-by-sequencing (GBS) analysis**

Leaf tissue collection, DNA isolation and quantification were carried out following the protocol developed by USDA Central Small Grain Genotyping Lab, Manhattan, KS (<https://hwwgenotyping.ksu.edu/protocols/>). GBS library construction and sequencing were conducted using previously described protocols (Xu et al., 2021). In brief, leaf tissues were

collected from a single plant in each F<sub>3</sub> line and parents into individual well of 1.1-mL 96-deep-well plates at three-leaf stage and freeze-dried for 48 h in a freeze-dryer (ThermoFisher, Waltham, MA). The tissue was then ground into a fine powder at 30 cycles per sec for 3 min with a 4.5-mm stainless bead in each well using a Mixer Mill (Retsch GmbH, Germany). Genomic DNA extraction used a cetyltrimethyl ammonium bromide method. DNA concentration was normalized using a Quant-iT<sup>TM</sup> PicoGreen® dsDNA Assay Kit (Life Technologies, Grand Island, NY). The GBS libraries were constructed using 200 ng DNA (20 ng/ul) per sample and sequenced on an Illumina NextSeq 2000 (Illumina, San Diego, CA). Single nucleotide polymorphisms (SNPs) were called from GBS reads (GBS-SNPs) using a reference-based pipeline in TASSEL 5.0 (Bradbury et al., 2007) and the IWGSC RefSeq v2.1 (Zhu et al., 2021) as the reference genome. GBS-SNPs with < 20% missing datapoints, < 10% heterozygotes and minor allele frequency >30% were used for linkage map construction.

#### **2.2.4 Linkage map construction and QTL analysis**

The GBS-SNPs were filtered to remove redundant SNPs using the BIN function in QTL IciMapping v4.1 (Meng et al., 2015) and all the remaining SNPs were grouped into 21 linkage groups based on anchored GBS-SNPs using MAP function in QTL IciMapping v4.1. Each linkage group was constructed using the regression mapping algorithm and Kosambi mapping function in JoinMap 4.0 (Van Ooijen, 2006). QTL analysis was performed using composite interval mapping (CIM) module in the Windows QTL Cartographer v2.5 (Wang et al., 2012). A logarithm of odds (LOD) threshold of 2.5 was set to claim significant QTL.

#### **2.2.5 KASP marker development and validation**

The GBS-SNPs significantly linked to the newly identified QTL in this study were converted to KASP markers. KASP primers were designed using PolyMarker

(<http://www.polymarker.info/>) and tested in the parents first. The polymorphic markers were then used for genotyping of the mapping population. The newly developed KASP markers were used to replace corresponding GBS-SNPs and reconstruct the linkage groups harboring the HF resistance QTL. KASP markers that were closest to the identified QTL were selected for validation in the AM203 panel. The KASP assays were performed in a GeneAmp™ PCR System 9700 Fast Thermal Cycler using a 5 µL reaction mix consisting of 1.94 µL of 2× PACE (PCR allelic competitive extension) Genotyping Master Mix (<https://3crbio.com/>), 0.06 µL primer mix, and 3 µL DNA at 10–20 ng/µL. The PCR profile started with denaturation of 94 °C for 15 min, followed by 10 touch-down PCR cycles at 94 °C for 20 s and 67 °C for 60 s with -1.0 /cycle, and then 32 cycles at 94 °C for 20 s and 57 °C for 60 s. MapChart v2.32 was used to draw the linkage maps (Voorrips, 2002). Data from combined QTL effects and analysis of linked markers were plotted using the ‘ggpubr’ and ‘ggplot2’ packages in R (Kassambara, 2023).

## **2.3 Results**

### **2.3.1 Segregation of HF resistance in the Parvaz-94 × Cadenza population**

Three HF resistant checks showed moderate (45.0% for Carol) to high levels (86.5% for Caldwell and 100.0% for Molly) of HF resistance with reddish dead larvae at the base of the resistant plants. In contrast, the susceptible check Newton was completely susceptible with the typical symptoms of stunted plants with dark green leaves, indicating the phenotyping conditions were adequate for evaluation of the mapping population for HF resistance. Parvaz-94 was 100% resistant like the resistant control Molly, while Cadenza was completely susceptible with a HF score of 0% for all tested plants, resembling Newton’s susceptibility (Figure 2.1). The (Parvaz-94 × Cadenza) F<sub>3:4</sub> RIL population segregated with 80 lines at resistance scores of 90% or higher, 65 lines with the resistance scores between 10%–90% and 33 lines with scores of 10% or lower (Table

2.1). The segregation ratio does not fit the single gene segregation. The same trend was observed for the F<sub>4:5</sub> generation, suggesting the existence of at least two QTL conferring the HF resistance in Parvaz-94. The correlation coefficient of HF scores was 0.71 between F<sub>3:4</sub> and F<sub>4:5</sub> RILs, indicating high repeatability of HF phenotypic data between the two experiments.

### **2.3.2 Linkage map construction**

GBS data were generated for 178 RILs from Parvaz-94 × ‘Cadenza’. After initial filtration, GBS analysis generated 4,690 high-quality SNPs. These SNPs were assigned into 3,600 bins, and the representative SNPs with the least missing data in each bin were used for linkage map construction. A total of 3,469 SNPs (96.4%) were mapped to 21 linkage groups, which cover 3,836.3 cM with an average marker density of 1.1 cM per SNP. The linkage groups were assigned to 21 chromosomes based on the IWGSC RefSeq v2.1 reference genome. SNP numbers and map lengths varied among 21 chromosomes, ranging from 11 SNPs at 43.8 cM on chromosome 4D to 306 SNPs on chromosome 6B and 277.1 cM in 5A. The marker order in the linkage map largely matched well with the physical map of the reference genome (Figure B.1), revealing extensive conservation between the two maps. Low SNP density was typically observed around the centromere regions of chromosomes.

### **2.3.3 QTL analysis and KASP marker development**

Using CIM, two QTL from Parvaz-94 were identified for HF resistance with the data from both greenhouse experiments and their mean values. The QTL were mapped on the distal regions of the short arms of chromosomes 1A and 6B, hereafter designated as *QHf.hwwg-1AS.2* and *QHf.hwwg-6BS.2*, respectively, to differentiate them from previous identified QTL *QHf.hwwg-1AS* (Li et al., 2013) and *QHf.hwwg-6BS* (Zhang et al., 2021).

The sequences of seven GBS-SNPs in the *QHf.hwwg-1AS.2* region and 11 GBS-SNPs in the *QHf.hwwg-6BS.2* region were used to design KASP primers, and six and eight of them were successfully converted to KASP markers, respectively (Table 2.2). We analyzed additional KASP markers on chromosome 6BS that were reported previously, including *KASP\_6B112698* linked to *QHf.hwwg-6BS* (Zhang et al., 2021), *G0019400\_S4* linked to *H34* (Ghori, 2022) and *kelse6BS\_167037*, *kelse6BS\_4554201* and *kelse6BS\_6196634* linked to *Hkelse* (Prather et al., 2022). *KASP\_6B112698* and *G0019400\_S4* acted as dominant markers in the population, while *kelse6BS\_167037*, *kelse6BS\_4554201* and *kelse6BS\_6196634* were not polymorphic between the parents (Table B.2). Thus, they were not used for the subsequent linkage map update.

After the GBS-SNPs were replaced by the corresponding KASP markers in the linkage map, *QHf.hwwg-1AS.2* was in the interval of 7.1–13.4 cM flanked by *S1A\_8619460* and *K1A\_7985775*, which corresponds to a 4.0 Mb physical interval between 4.6 to 8.6 Mb based on the IWGSC RefSeq v2.1 reference (Table 2.3, Figure 2.2). This QTL showed a minor effect that explained 5.2–9.0% of the phenotypic variance with LOD scores from 4.2 to 12.6. In contrast, *QHf.hwwg-6BS.2* showed a major effect that explained 30.1–48.6% of the phenotypic variance with LOD scores from 19.8 to 35.8. *QHf.hwwg-6BS.2* was finally mapped between 8.3 and 15.4 cM, corresponding to a 5.4 Mb (2.3–7.7 Mb) physical interval based on the IWGSC RefSeq v2.1 reference.

### **2.3.4 QTL effects on HF resistance**

Among the 14 newly developed KASP markers (Table 2.2), *K1A\_6140823*, *K1A\_7528908* and *K1A\_7985775* were identified as the most closely linked markers to *QHf.hwwg-1AS.2*, while *K6B\_6914795*, *K6B\_6936780* and *K6B\_7697506* were the most closely linked markers to the *QHf.hwwg-6BS.2*. Those markers were selected for analysis of the QTL effects on HF resistance.

In the F<sub>3:4</sub> RIL population, four allele combinations were identified from 91 lines that carry homozygous parental alleles at both QTL: AABB, aaBB, AAbb and aabb, where AA stands for the resistance allele of *QHf.hwwg-1AS.2* and BB stands for the resistance allele of *QHf.hwwg-6BS.2*. As a result, 43 lines with either AABB or aaBB that carry the resistance allele at *QHf.hwwg-6BS.2* all exhibited high HF resistance with an average resistance score of 99.7%. Among the 19 AAbb lines that carry the resistance allele at *QHf.hwwg-1AS.2* alone, HF resistance scores varied from 25.8% to 100% with an average of 64.3%. In contrast, 29 lines carrying the susceptibility alleles aabb at both QTL were all scored as 0 (Figure 2.3). Those data suggested that *QHf.hwwg-6BS.2* alone conferred complete resistance against HF, whereas *QHf.hwwg-1AS.2* showed partial resistance with large variation among the lines.

### 2.3.5 Validation of KASP markers

To assess the diagnostic value of the six most closely linked KASP markers to the two HF resistance QTL, they were evaluated in the AM203 panel consisting of 203 elite U.S. winter wheat breeding lines and released cultivars. Three KASP markers, *K1A\_7985775* linked to 1AS QTL, and *K6B\_6936780* and *K6B\_7697506* linked to 6BS, showed a significant correlation with HF resistance (Figure 2.4, Table B.2). Among 163 accessions successfully genotyped with *K1A\_7985775*, 112 carried the resistance allele and had an average HF resistance score of 39.7%, whereas other 51 accessions carried the susceptibility allele and had an average score of 15.1%. The marker *K6B\_6936780* showed amplification in 168 accessions but only 12 carried the resistance allele, two were heterogeneous, and the other 154 accessions carried the homogeneous susceptibility allele. For *K6B\_7697506*, only seven accessions carried the resistance allele, and they were all included in the 12 resistant accessions detected by *K6B\_6936780*. The rest of the accessions have the susceptibility allele (177 accessions) or did not show amplification (17

accessions). The average HF resistance scores of the seven resistant accessions (92.2% to 100%) detected by *K6B-7697506* were much higher than that of the 12 accessions (76.1%, ranging from 0% to 100%) detected by *K6B\_6936780*, which indicated *K6B-7697506* may be more tightly linked to *QHf.hwwg-6BS.2* than *K6B-6936780*, and can be used as a near-diagnostic marker for breeding selection. Among the seven resistant accessions detected by *K6B-7697506* in the AM203 panel, six were soft red winter wheat from Indiana or South Carolina. *K6B-7697506* was also screened in both RGON2022 and RPN2022 panels, but the resistance allele was not detected in either of the two panels.

## 2.4 Discussion

Pakistan is a country close to the Fertile Crescent region where wheat and HF originated. To find new sources of HF resistance, a set of popular wheat cultivars and germplasm lines from Pakistan (Hussain et al., 2022) were screened for HF resistance. Out of 186 accessions, only Parvaz-94 was found to be highly resistant to HF GP biotype, suggesting the source of HF resistance is rare in Pakistani wheat. Parvaz-94 (also called Parwaz-94) is a spring wheat cultivar with CIMMYT wheat germplasms in its pedigree ‘Ciano 67’//‘Lerma Rojo 64’\*2//‘Sonora 64’/3/Sonora 64/‘Parula’sib (<http://www.wheatpedigree.net>). It also has *Lr10*, *Lr26*, *Lr27* for leaf rust (*Puccinia recondita* f.sp. *tritici*) resistance (Mustafa et al., 2013), *Yr6* and *Yr7* for stripe rust (*Puccinia striiformis*. f. sp. *tritici*) resistance (Afzal et al., 2010), high thousand kernel weight, superior bread quality, and low phytic acid (Anjum et al., 2002). Therefore, Parvaz-94 is an invaluable source of HF resistance and can be useful for simultaneous improvement of wheat cultivars for high grain quality and multiple pest resistance.

*QHf.hwwg-1AS.2* and *QHf.hwwg-6BS.2* were identified in Parvaz-94 with major effects on HF resistance in both experiments. *QHf.hwwg-6BS.2* was located between 2.3 and 7.7 Mb on 6BS.

Among the previously named genes, only *H34* was found near *QHf.hwwg-6BS.2* (Li et al., 2013). Besides, three QTL (*QHara.icd-6B*, *QHf.hwwg-6BS* and *HKelse*) were also mapped in the vicinity of *QHf.hwwg-6BS.2* (Bassi et al., 2019; Prather et al., 2022; Zhang et al., 2021). All these genes or QTL were identified from modern common wheat or durum wheat cultivars. *H34* from a soft red winter wheat cultivar ‘Clark’ was finely mapped to a 1.4 Mb interval between *HF6028581* (6.0 Mb) and *HF7398225* (7.4 Mb) (Ghori, 2022; Li et al., 2013). *QHara.icd-6B* was identified in a durum cultivar ‘Faraj’, and closely linked to markers *BS00072387* (5.3 Mb) and *BA00462979* (13.1 Mb) (Bassi et al., 2019). *QHf.hwwg-6BS* from a Korean winter wheat cultivar ‘Chokwang’ was mapped between 6.0 and 10.8 Mb (Zhang et al., 2021). *HKelse* from hard red spring wheat cultivars ‘Kelse’ and ‘Seahawk’ was identified at an interval between 4.4 and 7.1 Mb (Prather et al., 2022). These results suggested that those gene/QTL are all located in a 9 Mb short region (4.4–13.1 Mb) on chromosome arm 6BS.

Among these gene/QTL on 6BS, *QHara.icd-6B* from a wild emmer wheat is likely a unique QTL different from the other QTL that originated from common wheat. To investigate the relationship among the remaining QTL, seven KASP markers, *K6B\_6936780* and *K6B\_7697506* linked to *QHf.hwwg-6BS.2*, *G0019400\_S4* linked to *H34*, *KASP\_6B112698* linked to *QHf.hwwg-6BS*, *kelse6BS\_167037*, *kelse6BS\_4554201* and *kelse6BS\_6196634* linked to *HKelse*, were tested in the cultivars carrying the HF resistance genes or QTL on 6BS including Parvaz-94 (*QHf.hwwg-6BS.2*), Clark (*H34*), Chokwang (*QHf.hwwg-6BS*), Kelse and Seahawk (*HKelse*) (Table 2.2; Table B.2). Parvaz-94, Clark and Chokwang all had the same resistance alleles at the first four markers, whereas Kelse and Seahawk had resistance alleles at the other three markers. The genotyping results suggest that the 6BS QTL in Kelse and Seahawk are likely different from those in Parvaz-94, Clark and Chokwang. In the AM203 panel, marker *K6B\_7697506*, *K6B\_6936780*,

*G0019400\_S4* and *KASP\_6B112698* consistently identified similar accessions (Table B.3). However, *K6B\_7697506* is more diagnostic for the 6BS QTL than the other three markers, as the latter three markers showed some false positive signals in the AM203 panel based on the phenotypic data.

We also conducted pedigree analysis of these accessions with 6BS QTL to investigate the possible sources of the resistance QTL. Parvaz-94 with the pedigree of Ciano 67//Lerma Rojo 64\*2//Sonora 64/3/Sonora 64/Parula sib might have inherited the resistance QTL from Sonora 64 (CItr 13930). Li et al. (2013) mapped *H34* in Clark with the pedigree of ‘Beau’//‘Caldwell’sib/67137B5-16/4/‘Sullivan’/3/‘Beau’//5517B8-5-3-3/‘Logan’ (Ohm et al., 1988). The parents Beau and Sullivan were reported to carry *H5* that was derived from ‘Ribeiro’ (PI 56206), a landrace from Lisboa, Portugal (Dweikat et al., 1997; Patterson et al., 1979). Dweikat et al. (1997) also reported that Clark might carry *H5* (6B) and *H6* (1A); therefore, *H34* might be *H5*. Chokwang was derived from the cross between ‘Jaekwang’ from Korea and ‘Norin-72’ Japan; however, there is no genetic connection between Chokwang and Clark. Kelse (‘WestBred 906R’/‘SD 2961’//‘Scholar’) might have inherited the 6BS QTL from WestBred 906R (Prather et al., 2022), whereas Seahawk (‘Whit’/‘Alpowa’//Whit/Alpowa) likely obtained the QTL from Whit. Whit has the pedigree of ‘Challis’/5/‘El Gaucho’/Sonora 64//‘Spr Luke Mutant’/3/‘Centennial’/4/Alpowa (Kidwell et al., 2009) and Sonora 64 might be the donor of the 6BS QTL in Whit. The pedigree analysis suggests that the 6BS QTL in Seahawk and Parvaz-94 might be from the same source. However, the marker data showed the same haplotype between Kelse and Seahawk, but a different haplotype in Parvaz-94. Therefore, additional studies to evaluate their allelic relationships and differential responses to various HF biotypes, or to finely map these genes, are needed to confirm their relationships.

*QHf.hwwg-1AS.2* was detected on the short arm of chromosome 1A, where 14 other HF resistance genes (*H3*, *H6*, *H9–H12*, *H14–H17*, *H19*, *H28*, *H29* and *Hdic*) and three major QTL (*QHf.hwwg-1AS*, *QHf.osu-1A* and *QHf.osu-1A<sup>d</sup>*) have been reported. Among them, *H9*, *H10*, *H11*, *H16* and *H17* were located to the distal 14% of chromosome 1AS flanked by *Xcfa2153* (8.6 Mb) and *Xbarc263* (13.2 Mb) (Kong et al., 2008; Kong et al., 2005; Liu et al., 2005b); *Hdic* was located between *Xcfa2153* and *Xgwm33* (11.0 Mb) (Liu et al., 2005a); *QHf.hwwg-1AS* closely linked to *Xgwm33* in Clark (Li et al., 2013); *QHf.osu-1A* in ‘2174’ was flanked by the markers *Xcfa2153* and *Xpsp3151* (23.1 Mb) (Tan et al., 2013); and *QHf.osu-1A<sup>d</sup>* in ‘Duster’ linked to marker *TaOPR-A1* at 9.6 Mb (Li et al., 2015). These genes or QTL were close to either *Xcfa2153* or *Xgwm33* and they were mapped in a 15 Mb gene cluster between 8.6 and 23.1 Mb on 1AS. Here, we mapped *QHf.hwwg-1AS.2* to an interval between 4.6 and 8.6 Mb, near the gene cluster. Patterson et al. (1994) previously demonstrated differential responses of *H3*, *H6*, *H9*, *H10*, *H11* and *H12* to HF biotypes B, C, D and L, suggesting that they might be either linked genes in the gene cluster or different alleles of the same gene, as demonstrated by multiple functional alleles at *Pm3* for powdery mildew resistance on the same chromosome arm (Bourras et al., 2019).

Historically, HF resistance has been viewed as a qualitative trait that is governed by monogenic inheritance (Harris et al., 2003). However recent studies have revealed the quantitative contribution to HF resistance. In wheat cultivar ‘Seneca’, *H7* and *H8* collectively imparted high resistance, but showed different effects (Liu et al., 2020). A similar pattern was observed in wheat cultivars ‘SD06165’, ‘Overland’, Clark and Parvaz-94 (Li et al., 2013; Xu et al., 2021; Zhao et al., 2020). *H7* and *H8* showed a pure additive effect in Seneca, while significant epistatic interactions were observed between *H35* and *H36* in SD06165 and between *QHf.hwwg-3B* and *QHf.hwwg-7A* in Overland. Here, the QTL interaction in Parvaz-94 appears to be similar to that in Clark (Li et

al., 2013), where 6B QTL alone conferred high resistance, while 1A QTL provided partial resistance. Although QTL analysis showed *QHf.hwwg-1AS.2* as a minor QTL that explained less than 10% phenotypic variation, the RILs carrying only *QHf.hwwg-1AS.2* showed a high mean resistance score of 64.3% with 100% for some RILs (Figure 2.3). However, the marker identified in this study appears to not be diagnostic to *QHf.hwwg-1AS.2* because false positive signals were identified in the AM203 panel. Further fine mapping is essential to develop diagnostic markers for the QTL.

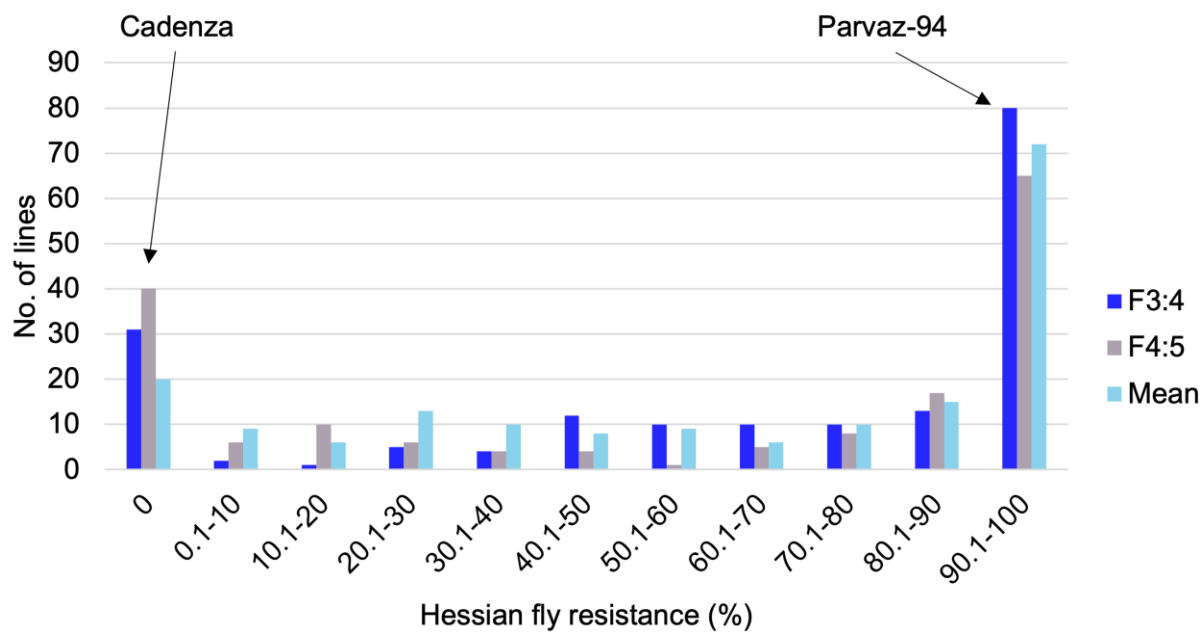


**Figure 2.1 Frequency distribution of Hessian fly resistance scores in the recombinant inbred line population of Parvaz-94 × Cadenza**

Resistance scores are in 0–100% scale where 0 is fully susceptible and 100% is fully resistant.



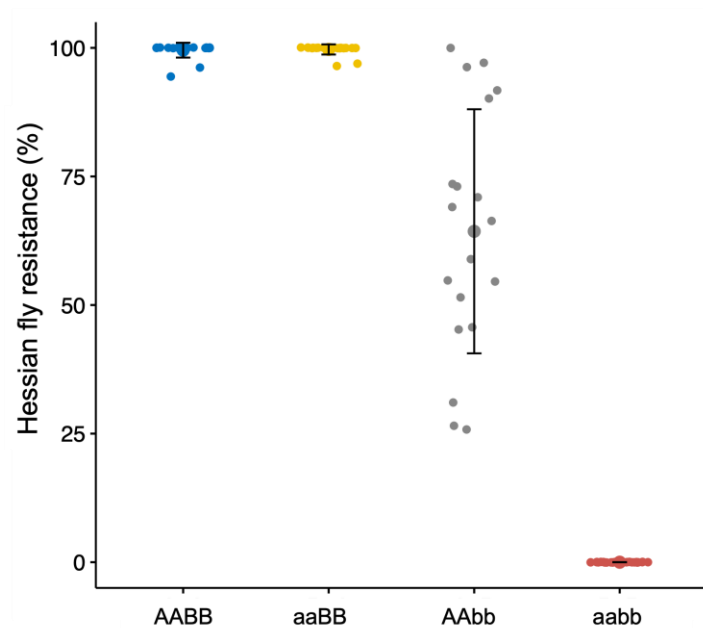


**Figure 2.3 Hessian fly resistance scores for homozygous recombinant inbred lines with different allelic combinations at *QHf.hwwg-1AS.2* and *QHf.hwwg-6BS.2***

Y-axis is percentage of Hessian fly resistance scores and x-axis is allelic combinations at the two quantitative trait loci, where A/a and B/b represent resistance/susceptibility alleles in *QHf.hwwg-1AS.2* and *QHf.hwwg-6BS.2*, respectively.

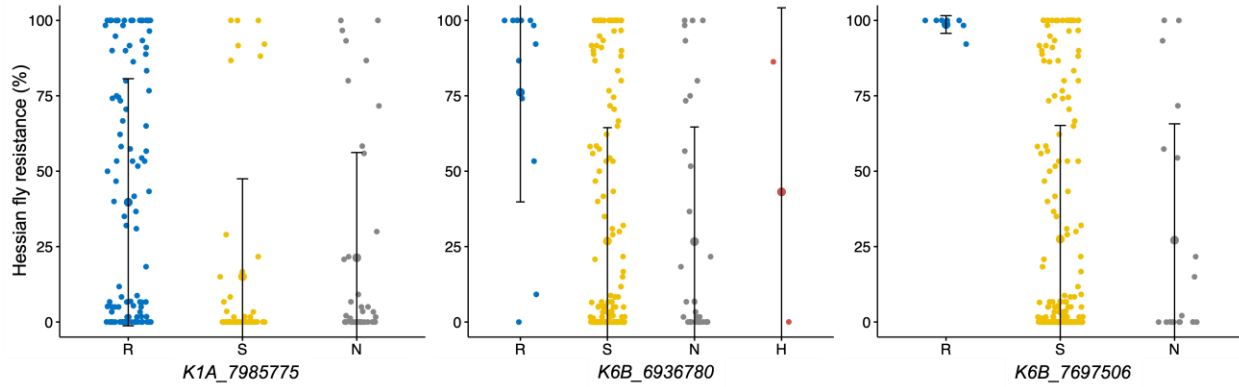


**Figure 2.4 Phenotypic distribution of Hessian fly resistance scores (%) of the 203 accessions in AM203 panel grouped by the Kompetitive allele specific PCR (KASP) markers *K1A\_7985775* linked to *QHf.hwwg-1AS.2*, and *K6B\_6936780* and *K6B\_7697506* linked to *QHf.hwwg-6BS.2***

Y-axis is percentage of Hessian fly resistance scores and x-axis showed different marker alleles where R is marker allele for resistance as in Parvaz-94, S is the allele for susceptibility as in Cadenza, N is missing data and H is heterozygote.

**Table 2.1 Chi-square ( $\chi^2$ ) test of single gene segregation ratio for Hessian fly resistance in (Parvaz-94  $\times$  Cadenza) F<sub>3:4</sub> and F<sub>4:5</sub> population evaluated in two greenhouse experiments**

|                        | F <sub>3:4</sub> (R:H:S = 3:2:3) |          |      |                 | F <sub>4:5</sub> (R:H:S = 7:2:7) |          |      |                 |
|------------------------|----------------------------------|----------|------|-----------------|----------------------------------|----------|------|-----------------|
|                        | Observed                         | Expected |      | <i>P</i> -value | Observed                         | Expected |      | <i>P</i> -value |
| Resistant (> 90%)      | 80                               | 66.8     | 29.1 | < 0.01          | 67                               | 73.5     | 65.9 | < 0.01          |
| Heterozygotes (10–90%) | 65                               | 44.5     |      |                 | 55                               | 21       |      |                 |
| Susceptible (< 10%)    | 33                               | 66.8     |      |                 | 46                               | 73.5     |      |                 |

**Table 2.2 Primer sequences of Kompetitive allele specific PCR markers**

| Marker name <sup>a</sup>        | SNP name     | Position (bp) | Primer sequence (5'– 3') <sup>a</sup> |
|---------------------------------|--------------|---------------|---------------------------------------|
| <i>K1A_4608940F</i>             | S1A_4608940  | 4,608,940     | CTTCTGACCTACGCAAGAAGATA/t             |
| <i>K1A_4608940R</i>             |              |               | CACGAGAATTGAGCGCATCG                  |
| <i>K1A_6140823F</i>             | S1A_6140823  | 6,140,823     | TGTAGAAGTAATGACAGTTTGCTCG/a           |
| <i>K1A_6140823R</i>             |              |               | GCTGCAGCCTAGGTTTTTTCAC                |
| <i>K1A_7528908F</i>             | S1A_7528908  | 7,528,908     | CAGCTCCGCCATTGGAACA/t                 |
| <i>K1A_7528908R</i>             |              |               | GTTACCTGTGTTGGGGGAG                   |
| <i>K1A_7985775F</i>             | S1A_7985775  | 7,985,775     | AGAGGAAGGATTGCCATGGTC/t               |
| <i>K1A_7985775R</i>             |              |               | CACTCGATCCCACCAGTAGC                  |
| <i>K1A_9634576F</i>             | S1A_9634576  | 9,634,576     | ACGAGGCAGCGCTCCTAC/t                  |
| <i>K1A_9634576R</i>             |              |               | CACGCTGGGAGTAGTACACG                  |
| <i>K1A_10935582F</i>            | S1A_10935582 | 10,935,582    | AGAAGTTTCAGCTCTTCCTTGA/g              |
| <i>K1A_10935582R</i>            |              |               | GGCCATCTGGACGTGAAGAA                  |
| <i>K6B_2311494F</i>             | S6B_2311494  | 2,311,494     | GCAGGCGCATTTGTCATATTTTG/t             |
| <i>K6B_2311494R</i>             |              |               | CGCGGGATAAGGAGTCTTCC                  |
| <i>K6B_4088590F</i>             | S6B_4088590  | 4,088,590     | CGGTGAATGGTGGGAGACG/c                 |
| <i>K6B_4088590R</i>             |              |               | GATCGTCACGGACCAAGGTG                  |
| <i>K6B_4351810F</i>             | S6B_4351810  | 4,351,810     | ACTTGGAATAAGCCTGCAGTTA/g              |
| <i>K6B_4351810R</i>             |              |               | AGTTTATTCGTGCGGACAAGT                 |
| <i>K6B_5588047F</i>             | S6B_5588047  | 5,588,047     | CTGCTTCTGTTTCTACTCGGTAA/t             |
| <i>K6B_5588047R</i>             |              |               | AGAATGTGCACTGCAGGGAT                  |
| <i>K6B_6914795F</i>             | S6B_6914795  | 6,914,795     | CAAAGGTTTCGACAACGTATGTTT/c            |
| <i>K6B_6914795R</i>             |              |               | CTGCAGTGATCTCAGATTGGA                 |
| <i>K6B_6936780F</i>             | S6B_6936780  | 6,936,780     | TGCATCGCCGCAATGACC/g                  |
| <i>K6B_6936780R</i>             |              |               | ACTGCGTACATCACCTTCCT                  |
| <i>K6B_7697506F</i>             | S6B_7697506  | 7,697,506     | TACCCGGCCATGAGAAGC/a                  |
| <i>K6B_7697506R</i>             |              |               | AGATAAGTGTTACCTACGTGTCC               |
| <i>K6B_22327187F</i>            | S6B_22327187 | 22,327,187    | TCATGGCATGCACCGCTG/c                  |
| <i>K6B_22327187R</i>            |              |               | ATGTTCCAGACGGCAGTAGC                  |
| <i>G0019400_S4F<sup>b</sup></i> | /            | 5,640,743     | AGCTGGCCACTGTTTCATTTG/a               |
| <i>G0019400_S4R</i>             |              |               | TCAGGAGTGAGCCAAGTTGT                  |

|                                         |           |                         |
|-----------------------------------------|-----------|-------------------------|
| <i>KASP_6B112698F</i> <sup>c</sup> /    | 6,026,607 | CAACTTGGGGAAAGCACCAT/a  |
| <i>KASP_6B112698R</i>                   |           | GGGGTTTGCCAGAGCTTACT    |
| <i>kelse6BS_4554201F</i> <sup>d</sup> / | 4,397,562 | CTCTGGAGTGAATGAGCATT/c  |
| <i>kelse6BS_4554201R</i>                |           | CTGGGTACGCCATAAGATTT    |
| <i>kelse6BS_167037F</i> <sup>d</sup> /  | 6,080,945 | GCACCCGCACTTCGAAATTCT/a |
| <i>kelse6BS_167037R</i>                 |           | GCACCCGCACTTCGAAATTCA   |
| <i>kelse6BS_6196634F</i> <sup>d</sup> / | 7,091,965 | CCAACAAGGTTGTTCTGCTA/g  |
| <i>kelse6BS_6196634R</i>                |           | GCCAAAGGCTCTCTTCAACT    |

---

<sup>a</sup> The marker name with F is forward primer, the last two letters in the forward primer sequence separated by a backslash are the polymorphic nucleotides in Cadenza (before the slash) and Parvaz-94 (after the slash); the name with R is reverse primer

<sup>b</sup> The marker is from Ghori (2022)

<sup>c</sup> The marker is from Zhang et al. (2021)

<sup>d</sup> The markers are from Prather et al. (2022)

**Table 2.3 Chromosome positions, marker intervals and gene effects of two quantitative trait loci (QTL) for Hessian fly resistance detected in a recombinant inbred population of Parvaz-94 × Cadenza evaluated in two greenhouse experiments**

| Experiments      | Chr. | Marker interval                | Genetic distance (cM) | Physical interval (Mb) | LOD <sup>a</sup> | ADD <sup>b</sup> | PVE <sup>c</sup> |
|------------------|------|--------------------------------|-----------------------|------------------------|------------------|------------------|------------------|
| F <sub>3:4</sub> | 1A   | <i>S1A_8619460–K1A_7985775</i> | 7.1–13.4              | 4.6–8.6                | 12.6             | 12.1             | 9.0              |
| F <sub>4:5</sub> | 1A   | <i>S1A_8619460–K1A_7985775</i> | 7.1–13.4              | 4.6–8.6                | 4.2              | 10.4             | 5.2              |
| Mean             | 1A   | <i>S1A_8619460–K1A_7985775</i> | 7.1–13.4              | 4.6–8.6                | 11.2             | 11.9             | 8.9              |
| F <sub>3:4</sub> | 6B   | <i>S6B_6158719–S6B_7044180</i> | 8.3–15.4              | 2.3–7.7                | 32.5             | 26.3             | 44.1             |
| F <sub>4:5</sub> | 6B   | <i>S6B_6158719–S6B_7044180</i> | 8.3–15.4              | 2.3–7.7                | 19.8             | 25.6             | 30.1             |
| Mean             | 6B   | <i>S6B_6158719–S6B_7044180</i> | 8.3–15.4              | 2.3–7.7                | 35.8             | 27.7             | 48.6             |

<sup>a</sup> LOD, log of odd value

<sup>b</sup> ADD, additive effect in which a positive value indicates increased resistance to Hessian fly contributed by the allele from Parvaz-94

<sup>c</sup> PVE, percentage of phenotypic variation explained by a QTL

## Chapter 3 - Toward Cloning of *H35* for Hessian Fly Resistance in the Hard Winter Wheat Cultivar ‘Overland’

### 3.1 Introduction

Previously, HF resistance genes *H35* and *H36* were identified in a U.S. hard winter wheat cultivar ‘SD06165’ (Zhao et al. 2020). Later, we identified *QHf.hwwg-3B* and *QHf.hwwg-7A* for HF resistance in another U.S. hard winter wheat cultivar ‘Overland’ (Xu et al. 2021). *QHf.hwwg-3B* and *QHf.hwwg-7A* located in regions overlapping with *H35* and *H36*, respectively.

Significant epistatic effects were detected between the QTL in both studies (Xu et al., 2021; Zhao et al., 2020). Although the common donor of HF resistance QTL was not identified in the pedigrees of SD06165 and Overland, the chromosome locations and epistatic effect pattern of the two QTL in the two cultivars suggest that Overland and SD06105 most likely carry the same set of QTL (Xu et al., 2021).

*H35* on chromosome 3BS showed a larger effect than *H36* on chromosome 7AS (Xu et al., 2021; Zhao et al., 2020). In the RIL population of Overland × Overley, *H35* alone increased 16.9–47.7% HF resistance scores in the three experiments, while *H36* alone increased only 9.3–20.9% HF resistance scores. Together with the epistatic effect, they increased up to 80.2% HF resistance score (Xu et al., 2021). *H35* was mapped to a 2.32 Mb physical interval between *KASP-3B2479314* (2,479,314 bp) and *KASP-3B4799538* (4,799,538 bp) on the distal end of chromosome arm 3BS (Xu et al., 2021) based on IWGSC RefSeq v1.1 (IWGSC et al., 2018). Since these two QTL are in U.S. adapted hard winter wheat cultivars, they are valuable sources of HF resistance for breeding programs and warrant of further study.

Due to major effect of *H35*, we dissected *H35* toward cloning of the gene using Overland × Overley derived secondary mapping population. The objectives of this study were to 1) finely

map *H35* to identify its putative causal gene for HF resistance, and 2) develop diagnostic markers for *H35* that can be used in MAS breeding.

## **3.2 Materials and methods**

### **3.2.1 Plant materials**

Overland is a hard red winter wheat cultivar selected from the cross ‘Millennium’ sib/ND8974 (Baenziger et al., 2008) and carries *H35* and *H36* for moderate resistance to HF biotype GP. Overley is a HF susceptible hard red winter wheat cultivar with the pedigree of U1275-1-4-2-2/‘Heyne sib’/‘Jagger’ (<https://bookstore.ksre.ksu.edu/pubs/L924.pdf>). Previously, a population of 186 F<sub>5:6</sub> RILs was developed from Overland × Overley (Xu et al., 2021) and this population was screened to dissect *H35* and identify heterogeneous inbred families (HIFs) for fine mapping and marker development. A diversity panel of 203 (AM203) U.S. winter wheat accessions (Zhang et al., 2010) was used to validate newly developed KASP markers for *H35*.

### **3.2.2 Evaluation of HF resistance**

New homozygous recombinants, selected by continuous screening of the secondary population derived from (Overland × Overley)F<sub>5</sub> using the flanking markers, were evaluated for HF resistance in the greenhouses from 2019 fall to 2021 spring. HF biotype GP was used to infest the wheat seedlings as described in Chapter 2.2.2.

### **3.2.3 HiFi sequencing, genome assembly and gene annotation of Overland genome**

DNA of Overland was extracted using CTAB. DNA concentration was quantified using a Qubit™ dsDNA Quantification Assay Kit (ThermoFisher Scientific) and quality was assessed by 0.8% agarose gel electrophoresis. The HiFi sequencing library was constructed following the manufacturer’s protocol using a SMRTbell® Express Template Prep Kit 2.0 (Novogene Corporation Inc., CA, USA). Sequencing was performed using circular consensus sequencing

(CCS) mode on a PacBio Sequel II platform and the PacBio HiFi reads were assembled by Novogene Corporation Inc using default parameters in hifiasm (Cheng et al., 2021). Following the HiFi assembly, RAGTAG (<https://github.com/malonge/RagTag>) was used to construct reference-guided pseudomolecules. The parameters used for minimap2 (Li, 2018) in RAGTAG were --mm2-params “-I 10G -K2G”. Genome was annotated at the Plant Genome and Systems Biology, German Research Center for Environmental Health, Neuherberg, Germany using a custom pipeline.

### **3.2.4 Fine mapping**

Leaf tissue collection, DNA isolation and development and assays of KASP markers used the same protocols as described in Chapter 2.2.5. Exome capture data of Overland and Overlay was used to identify SNPs in the target region. The exome capture data were obtained from the T3/Wheat website (<https://triticeaetoolbox.org/wheat/>). Dot plots of target region sequences from IWGSC RefSeq v2.1 and Overland v1.1 were made using D-GENIES with default parameters (Cabanettes & Klopp, 2018).

### **3.2.5 RNA isolation, sequencing and data analysis**

Overland and two near-isogenic lines (NILs), NIL52 (resistant NIL) and NIL53 (susceptible NIL), derived from the RHL DY157 were infested with HF Biotype GP in a greenhouse to produce leaf tissues for RNA-sequencing (RNA-seq). Plant growth and HF infestation conditions were described in Chapter 2.2.2. Tissues were collected at 24 and 48 hours past infestation (hpi) for RNA sequencing. Around 1 cm tissue was collected from bottom of second leaf sheath of a plant where HF larvae established their feeding sites. Each sample consisted of three 1-cm tissues from three separated plants. Samples were frozen immediately in liquid

nitrogen after larvae were washed away and water was removed from the tissues. The collected tissues were transferred to a -80 °C freezer till RNA isolation.

Total RNA was extracted from the leaf tissues using the Trizol RNA Reagent (ThermoFisher Scientific, Thousand Oaks, CA) and then treated with TURBO DNase (ThermoFisher Scientific) to avoid DNA contamination. RNA quality was evaluated in a 1% agarose gel and RNA concentration was measured in a ND-2000 NanoDrop Microvolume Spectrophotometer (ThermoFisher Scientific). RNA quality control, library construction and sequencing using Illumina HiSeq<sup>TM</sup> 2000 sequencer (Illumina, San Diego, CA) were carried out in Novogene Corporation Inc. Raw sequence reads were filtered by removing those with low quality or containing adapters using Trimmomatic (Bolger et al., 2014), and clean reads were aligned to the indexed genome of *Overland* v1.1 using STAR (Dobin et al., 2013). Differential expression analysis was performed using DESeq2 R package (Love et al., 2014).

### **3.2.6 Putative candidate gene cloning**

Genomic DNA (gDNA) sequences of putative candidate genes were amplified from the parents by PCR using gene specific primers, which were designed after aligning homoeologous sequences from chromosomes 3A, 3B and 3D. To amplify the full-length coding sequences (CDS) of candidate genes, the RNA of both parents were extracted using the methods of HF infestation, tissue collection and RNA extraction described for RNA-Seq experiment above. About 1.0 mg of total RNA was used for synthesis of the first-strand complementary DNA (cDNA) using a Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). The full-length CDS of candidate genes were amplified by PCR using gene specific primers (Table C.1).

The PCR products of the selected putative candidate genes were cloned into a pMiniT2.0 vector (New England Biolabs, Ipswich, MA, USA); however, they were cloned into a pCR-XL-2-

TOPO vector (Thermo Fisher Scientific) if they were longer than 3 kb. Plasmids were extracted using GenCatch™ Plasmid DNA Mini-Prep Kits (Epoch Life Science, Fort Bend County, TX, USA) and were sequenced using ABI 3730 sequencing platform (Thermo Fisher Scientific). Sequence reactions used a 10-μL reaction mix consisting of 1.0 μL of plasmid DNA, 0.5 μL of primer at 10 mM, 0.5 μL of BigDye™ terminator v3.1 ready reaction mix (Thermo Fisher Scientific), 2.0 μL 5X sequencing buffer, 0.5 μL of dimethylsulfoxide (DMSO, ≥ 99.9%) and 5.5 μL of sterile double-distilled water (ddH<sub>2</sub>O). PCR started at 94 °C for 60 s and followed by 25 cycles at 94 °C for 30 s, 50–60 °C (depending on primers) for 60 s and 60 °C for 4 min. PCR products were purified to remove dye terminators using Optima DTR Ultra 96-Well Plates (EdgeBio, San Jose, CA) before being loaded on sequencing platform.

### **3.3 Results**

#### **3.3.1 Dissection of *H35* from *H36***

(Overland × Overlay)<sub>F5</sub> RILs were screened using flanking KASP markers for both *H35* and *H36* (Table C.1) and three heterogeneous inbred lines (HIFs) of *H35*, DY33, DY95 and DY157, carrying heterogeneous alleles at the flanking markers for *H35* and homozygous susceptibility alleles at the flanking markers for *H36* were selected. Those lines were used as starting materials to generate a secondary population for identification of new recombinants in the *H35* flanking region for fine mapping of *H35*.

#### **3.3.2 Overland genome assembly**

To facilitate the cloning of *H35* from Overland, we generated a *de novo* genome assembly using PacBio CCS (HiFi) sequencing. The cumulative length of subreads from each of the 12 PacBio cells ranged from 14.35 to 28.13 Gb and N50 for HiFi reads ranged from 12.92 to 17.13 kb (Table C.2). In total, over 254 Gb of HiFi reads were generated with ~16x genome

coverage. A total 14.57 Gb (Overland v1.0) was assembled to pseudomolecules of 21 chromosomes using 1,827 contigs, with N50 of 18.01 Mb (Table 3.1). There were 1,680 contigs with a total length of 0.14 Gb that were not anchored to chromosomes. Overland assembly v1.0 was annotated (Overland v1.1) with the support from transcriptomes of Overland using RNA from the leaf sheath collected at 24 and 48 hpi and the Overland v1.1 defined 126,534 high-confidence genes with 93.5% (118,361 genes) anchored to pseudomolecules.

### **3.3.2 Fine mapping and identification of putative candidate genes of *H35***

*H35* was originally mapped between 2.5 and 4.8 Mb on IWGSC RefSeq v1.1 (Xu et al., 2021) and between 2.6 and 11.3 Mb on IWGSC RefSeq v2.1 by the flanking markers, *KASP-3B2479314* and *KASP-3B4799538*, respectively (Figure 3.1b). Over 8,000 plants derived by selfing of heterozygous plants from the selected HIFs were screened for recombination in the *H35* interval using the flanking KASP markers, which resulted in identification of a total of 36 critical new recombinants (Figure 3.1b). Further genotypic and phenotypic screening of the homozygous plants from these new recombinants with more markers, identified in the *H35* interval, reduced the flanking region to ~38 kb (3,492,790–3,531,272) between *K3492790* and *K3531272* on chromosome arm 3BS based on IWGSC RefSeq v2.1. This interval corresponded to ~71 kb (6,450,478–6,521,189) and was located in a single contig in Overland v1.1. In the *H35* flanking region, Overland v1.1 shows two large insertions that extended the *H35* interval by 33 kb compared with that in IWGSC RefSeq v2.1 (Figure 3.2).

In the *H35* interval, both IWGSC RefSeq v2.1 and Overland v1.1 harbor the genes *TraesCS3B03G0011900* (*B119*), *TraesCS3B03G0012000* (*B120*), and *TraesCS3B03G0012100* (*B121*). In addition, Overland v1.1 has *RagTagG00737300* (*R300*) that is absent in IWGSC RefSeq v2.1 (Figure 3.1c). These annotated genes encode a glycoside hydrolase, a protein with

an isopropylmalate dehydrogenase-like domain, a metal dependent phosphohydrolase and a calmodulin-binding kinase protein, respectively.

### **3.3.3 Differentially expressed genes in the *H35* region**

To identify possible candidate genes for *H35*, RNA-seq analysis using a pair of NILs contrasting in *H35* alleles identified two differentially expressed genes (*R300* and *B120*) in the *H35* interval based on Overland v1.1 (Figure 3.3). Both genes upregulated in the resistant NIL52 at 24 and 48 hpi, therefore, were selected as the putative candidates for further function analysis of *H35*.

The full length gDNA of *R300* was cloned from both parents, which isolated a 3825 bp fragment from the resistant parent Overland and a 3786 bp fragment from the susceptible parent Overlay (Figure 3.4). After sequence alignment between the two DNA fragments, 43 SNPs and 8 InDels were identified. A 2406 bp CDS of *R300* was isolated from Overland, which was identical to the predicted CDS of *R300* in Overland v1.1. However, amplification of CDS in Overlay failed, and thus, the 2421 bp CDS was predicted using the gDNA sequence of *R300* from Overlay. Comparison of *R300* CDS identified 26 SNPs and 3 InDels between the two parents (Figure 3.5).

CDS of *B120* were also cloned from both parents. They were 966 bp in Overland and 1203 bp in Overlay (Figure 3.6). Compared to Overlay, Overland lacks an extra 237 bp segment at the beginning of the gene due to a SNP that induced premature stop codon (Figure 3.6), and a new code starting at 238 bp reinitiated transcription to form *B120* CDS in Overland. In addition, alignment of *B120* protein sequence with pan-genome cultivars found that *B120* is a conserved protein across pan-genomes (Figure 3.7). Besides the initial 79 amino acids deletions at the

beginning of the protein sequence, variations were found only in five amino acids at 118<sup>th</sup>, 144<sup>th</sup>, 203<sup>rd</sup>, 272<sup>nd</sup> and 277<sup>th</sup> where Overland carries unique amino acids in B120.

### 3.3.4 Development and validation of gene markers

KASP markers were developed for both candidate genes. Among SNPs and InDels in the *R300* between the two parents, five were successfully converted to KASP markers and were polymorphic between the parents. They were *KR300-4*, *KR300-26*, *KR300-32*, *KR300-40*, and *KR300-41* (Table C.1). For *B120*, an attempt to convert the SNP that induced premature stop codon to a KASP marker failed due to a high GC content. However, four other SNPs, *KB120-1*, *KB120-2*, *KB120-5* and *KB120-6*, were successfully converted to KASP markers. Three of the four SNPs, except for *KB120-6*, caused amino acid changes in the protein at positions 118<sup>th</sup>, 144<sup>th</sup> and 277<sup>th</sup>. Two markers, *KB120-2* and *KB120-5*, were co-dominant, while the other two *KB120-1* and *KB120-6*, were dominant markers that only amplified in the resistant genotypes and heterozygotes, and did not amplify in susceptible genotypes.

All nine gene markers from the two candidate genes were used to screen the AM203 panel (Table 3.2; Figure 3.8). *KB120-2* amplified the resistance allele in 14 accessions that had an average HF resistance at 81.4%, varying from 30.9% to 100.0%, whereas the markers *KR300-4*, *KR300-32*, *KR300-40*, *KR300-41*, *KB120-1*, *KB120-5* and *KB120-6* amplified the marker alleles associated with resistance in a different set of 14 accessions that showed an mean HF resistance score of 73.5%, ranging from 6.5% to 100.0% (Figure 3.8). Comparing the two sets of the 14 accessions found that they differed in only two accessions with 12 accessions identical between the sets. In addition, *KR300-26* amplified the resistance allele in 88 accessions with an average HF of 31.8%. Therefore, *KB120-2* was the only diagnostic marker for *H35* in the AM203 panel because the accessions that amplified the resistance allele of *KB120-2* showed at

least moderate resistance for each individual accession. The 14 accessions harboring the resistance allele at *KB120-2* were all hard winter wheat from CO (1), KS (1), NE (7), SD (5).

To track the distribution of the two *H35* candidate genes in current hard winter wheat breeding programs, two KASP markers per gene, *KR300-4*, *KR300-40*, *KB120-2* and *KB120-6*, were further screened in the RGON2022 panel of 327 winter wheat accessions (Figure 3.9). The data showed that 49, 7, 25 and 31 accessions amplified the resistance alleles with average HF resistance scores of 52.1%, 48.8%, 69.9% and 52.1% at *KR300-4*, *KR300-40*, *KB120-2* and *KB120-6*, respectively. Among the four markers, *KB120-2* identified accessions with the highest HF resistance score of 69.9% and was a near diagnostic marker for *H35* in the panel, although none of them were diagnostic for *H35* because some accessions with the resistance alleles showed susceptible phenotypes.

## **3.4 Discussion**

### **3.4.1 *H35* fine mapping**

To date, 37 *H* genes have been reported (McIntosh et al., 2013; Zhao et al., 2020); however, studies on fine mapping or cloning of these genes are limited with only two fine mapping and cloning studies (*H13* and *H34*) being documented (Ghori, 2022; Joshi, 2018). *H13* was finely mapped to a 1.4 cM interval between SSR markers *Xcfd132* and *Xgdm36* on chromosome 6D, and an *R* gene *CNL-1* was cloned and validated through an EMS-mutagenized targeting induced local lesions in genomes (TILLING) population developed from the *H13* donor ‘Molly’ (Joshi, 2018). *H34* was initially mapped between 8.4 and 15.1 Mb on the chromosome arm 6BS based on IWGSC RefSeq v1.0 (Li et al., 2013). After fine mapping, the region was delimited to a 1.3 Mb interval based on IWGSC RefSeq v2.0 (Ghori, 2022). By leveraging the draft genome of Overland, utilizing KASP markers, and taking advantage of the high

recombination rate at the distal end of chromosome 3BS, we successfully narrowed *H35* region down to approximately 70 kb interval with only 4 putative genes, which laid a solid foundation for further identification of candidate genes for *H35* and development of its diagnostic markers for breeding.

### **3.4.2 *H35* candidate genes prediction**

After fine mapping, four putative genes were annotated in the *H35* interval (Figure 3.1c). Among them, only two genes (*R300* and *B120*) were differentially expressed between *H35* NILs. *R300* encodes a calmodulin-binding kinase protein, which has been reported to confer resistance to herbivore by regulating jasmonic or salicylic pathway through interacting with calmodulin and  $\text{Ca}^{2+}$  (Lv et al., 2019; Scholz et al., 2014). *B120* has an isopropylmalate dehydrogenase-like domain, which involves in leucine and glucosinolate biosynthesis pathways (Lee et al., 2016), but its role in pest resistance remains unclear. KASP markers were developed from both candidate genes, the KASP *B120*-2 identified accessions that exhibited the highest mean HF resistance score in the AM203 and RGON2022 panels among the gene markers screened, therefore it is the best diagnostic marker for *H35*. The data also suggests that *B120* is more likely the candidate gene of *H35*. Further transformation is necessary to validate its function.

Historically, the genetic interactions between wheat and HF have been reported to fit the gene-for-gene model (Hatchett & Gallun, 1970; Stuart et al., 2012). However, *H35* in this study is unlikely to be an *R* gene, because *R* gene was not annotated in the delimited *H35* region. Besides, *H35* conferred partial resistance up to 40%, whereas an *R* gene typically provides complete resistance (Kaur et al., 2021). Thus, *H35* may be more likely related to innate immunity.

Molecular responses of plants to herbivores are strongly correlated with the mode of feeding and the degree of tissue damage at the feeding sites (Du et al., 2009). Although plant defenses against sucking insects such as HF, aphids, whiteflies, are similar to those against pathogens (Du et al., 2009; Stuart et al., 2012), only a few insect resistance genes have been cloned in plants so far. Most cloned resistance genes are *R* genes, including *Mi* resistance gene against potato aphids (*Macrosiphum euphorbiae*) and whiteflies (*Bemisia tabaci*) (Nombela et al., 2003; Rossi et al., 1998), *gm3* (Sama et al., 2014) and *Gm4* (Divya et al., 2015) resistance to Asian rice gall midge (*Orseolia oryzae*), *Bph14* resistance to brown planthopper (*Nilaparvata lugens* Stål) in rice (Du et al., 2009), *H13* resistance to HF (Joshi, 2018), while *Gm8* resistance to Asian rice gall midge encodes a proline-rich protein (Divya et al., 2018). *R300* and *B120* are neither *R* genes nor genes encoding proline-rich proteins, and thus *H35* is probably a new type of resistance gene.

### **3.4.3 Development of tightly linked markers**

Although 37 *H* genes and some QTL for HF resistance have been reported, their effective markers are limited for breeding selection in diverse genetic backgrounds. Only a few genes have linked markers that were validated in diversity panels, including *H32* (Tan et al., 2017), *H34* (Ghori, 2022), *QHara.icd-6B* (Bassi et al., 2019), *HKelse* (Prather et al., 2022) and *QHf.hwwg-6BS* (Zhang et al., 2021), therefore, development of breeder-friendly DNA markers tightly linked to resistance genes will facilitate their deployment in breeding programs.

In this study, eight KASP markers were developed from two candidate genes of *H35* and they were validated in the AM203 panel. *B120-2* identified resistant accessions with the highest levels of resistance among the markers tested, with the lowest HF resistance score of 30.9% in only one accession (Figure 3.8). *B120-2* also showed the best performance in RGON2022 among

four gene markers tested although four accessions with the resistance alleles showed complete susceptibility (Figure 3.9). Detection of susceptible genotypes by the markers could be due to heterogeneity of seeds used for marker analysis and phenotyping or to the SNP is not the causal SNP of the candidate gene. Overexpression of the candidate gene in transgenic plants will confirm their function of the candidate gene and provide perfect marker for *H35* to be used in HF resistance breeding programs.

#### **3.4.4 Conclusion**

In summary, we sequenced Overland's genome, the doner of *H35*, using PacBio HiFi long-read technology and generated a 14.57 Gb assembly (Overland v1.1) for cloning of *H35*. Using fine mapping, *H35* was delimited to a ~70 kb interval in a single contig of Overland v1.1 where four putative genes were annotated. RAN-seq of NILs identified two differentially expressed genes, *R300* and *B120*, in the *H35* interval. Eight KASP markers were developed from the two genes. Those markers were used to screen both panels of AM203 and RGON2022, and *B120-2* showed the best performance and was validated as a near diagnostic marker for *H35*. The function validation of the candidate genes is in progress using gene transformation and other genomic approaches.

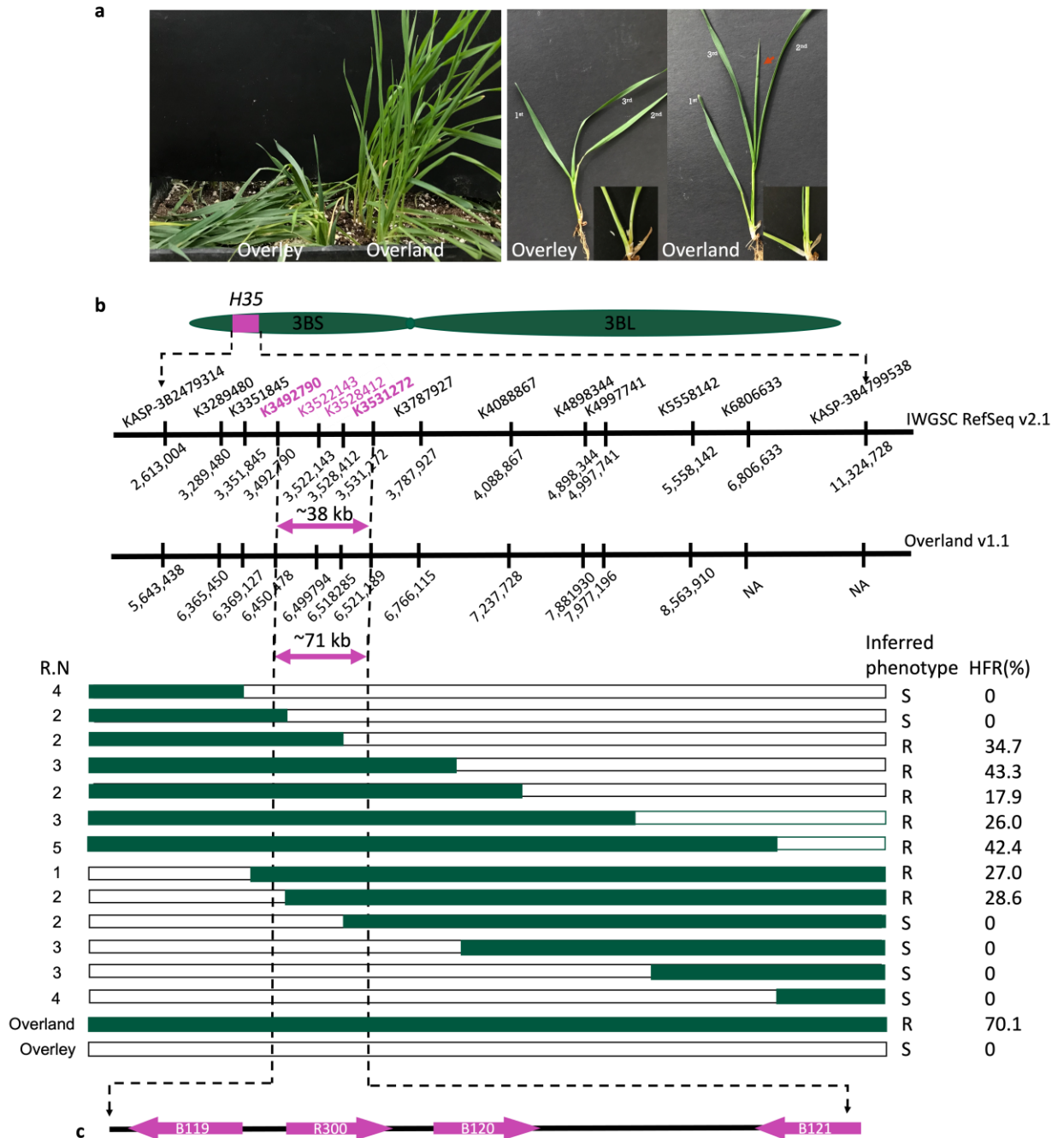


**Figure 3.1 Fine mapping of *H35***

a, Hessian fly infested seedlings of Overley and Overland. b, Wheat chromosome 3B showing the original QTL mapping of *H35* region and fine mapping delimited *H35* region to a 38 kb region. The name of Kompetitive allele specific PCR markers are above the physical map and their positions are below the physical map based on International Wheat Genome Sequencing Consortium (IWGSC) RefSeq v2.1 reference genome. The bars represent the selected lines with recombination within *H35* interval from the progeny of the residual heterozygotes, with the number of lines (R.N) for each type of recombination on the left and the inferred phenotype (R =

resistant and S = susceptible) and average Hessian fly resistance (HFR%) for each type of recombinants on the right. Solid and open bars represent chromosomal segments from Overland and Overlay, respectively. c, All putative genes annotated in the *H35* region from IWGSC RefSeq v2.1 and Overland v1.1.

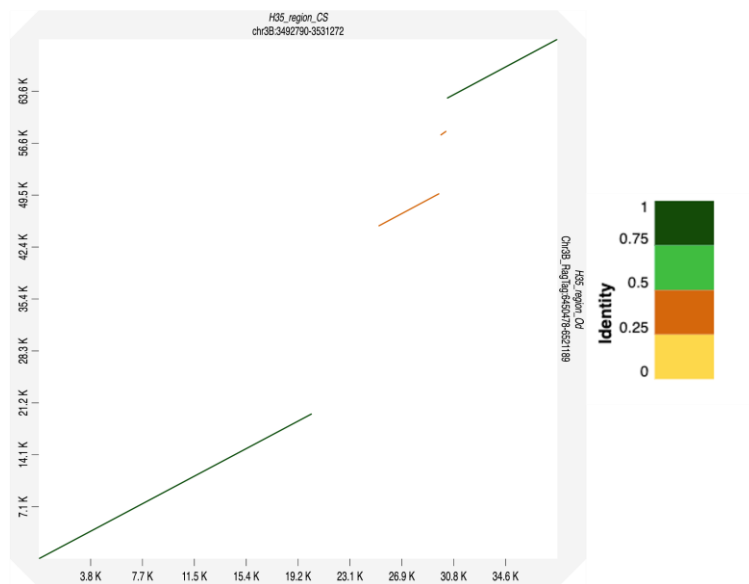


**Figure 3.2** Alignment of sequences in *H35* region from IWGSC RefSeq v2.1 and Overland v1.1

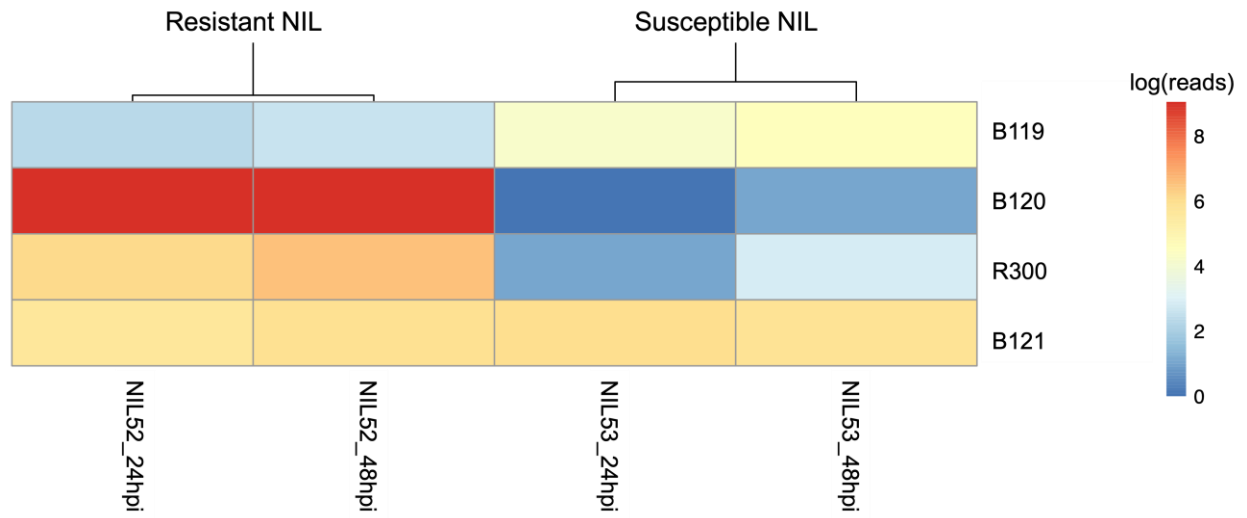


**Figure 3.3** Heatmap of normalized read counts of putative annotated genes in the *H35* region in the near iso-genetic lines NIL52 (Resistant) and NIL53 (Susceptible)

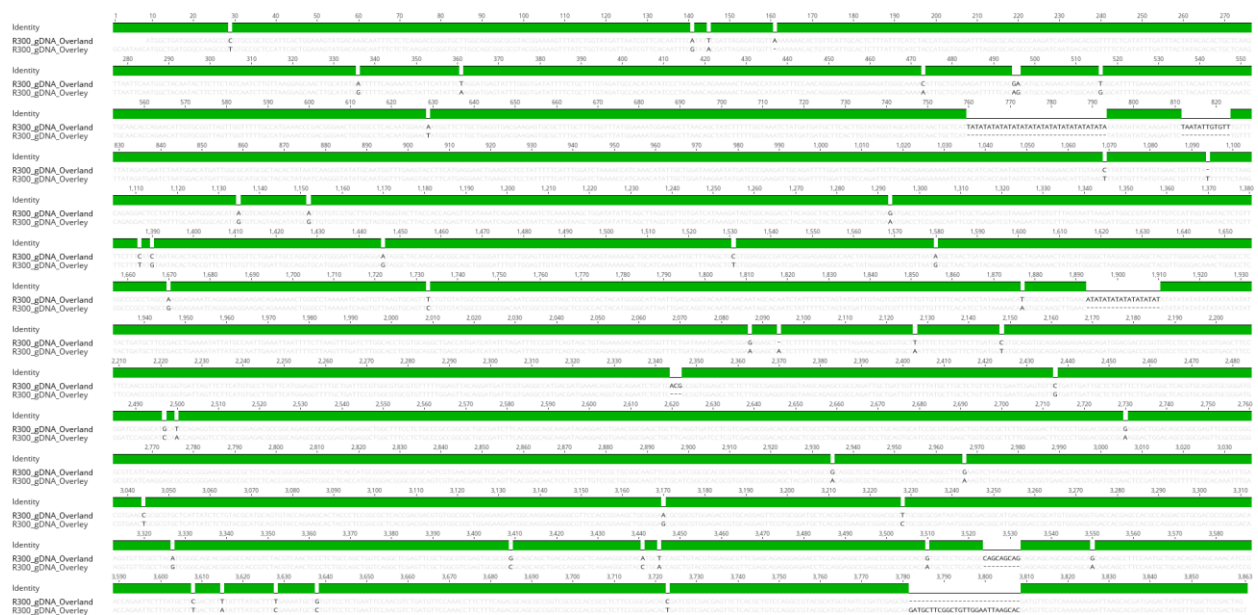


**Figure 3.4 Alignment of genomic DNA sequences of R300 from Overland and Overlay**

Open bar and dark nucleotides were disagreement regions.

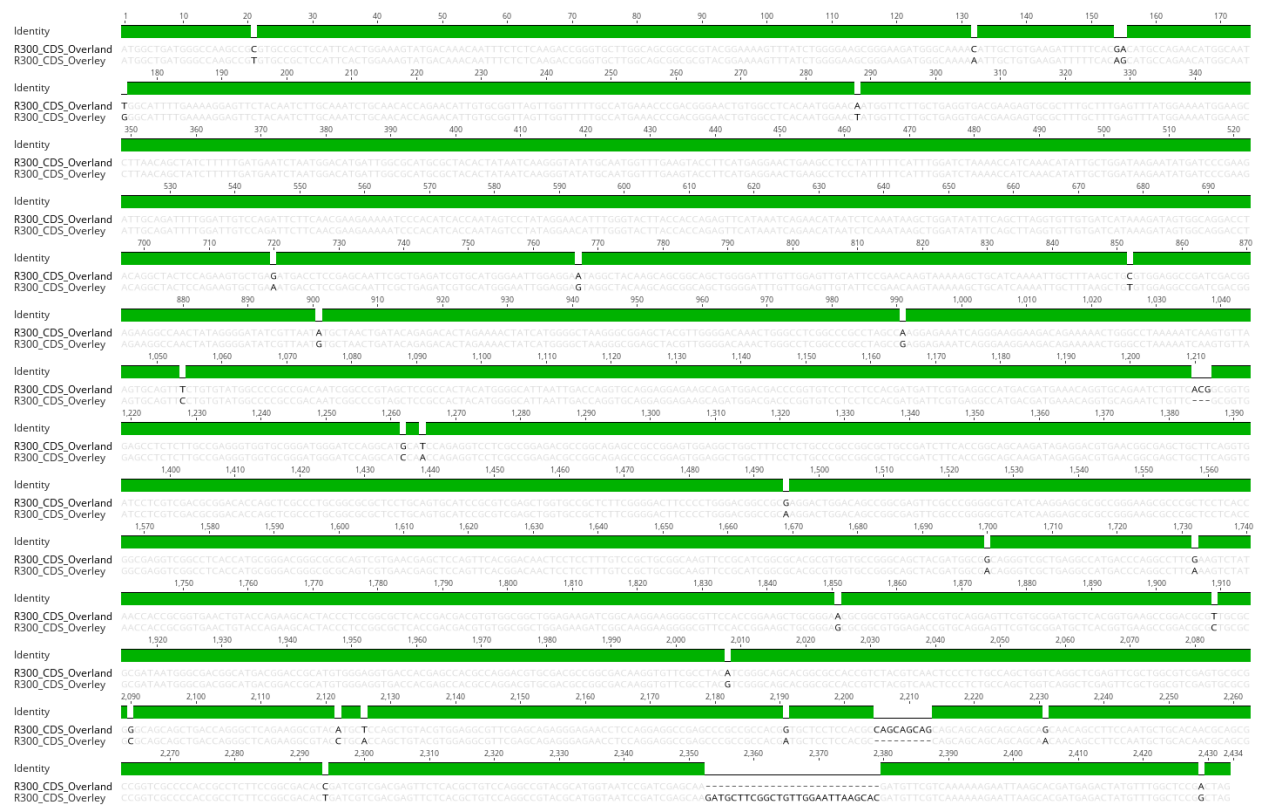


**Figure 3.5 Alignment of coding sequences of R300 from Overland and Overlay**

Green bar and grey nucleotides were identical regions; open bar and dark nucleotides were disagreement regions.

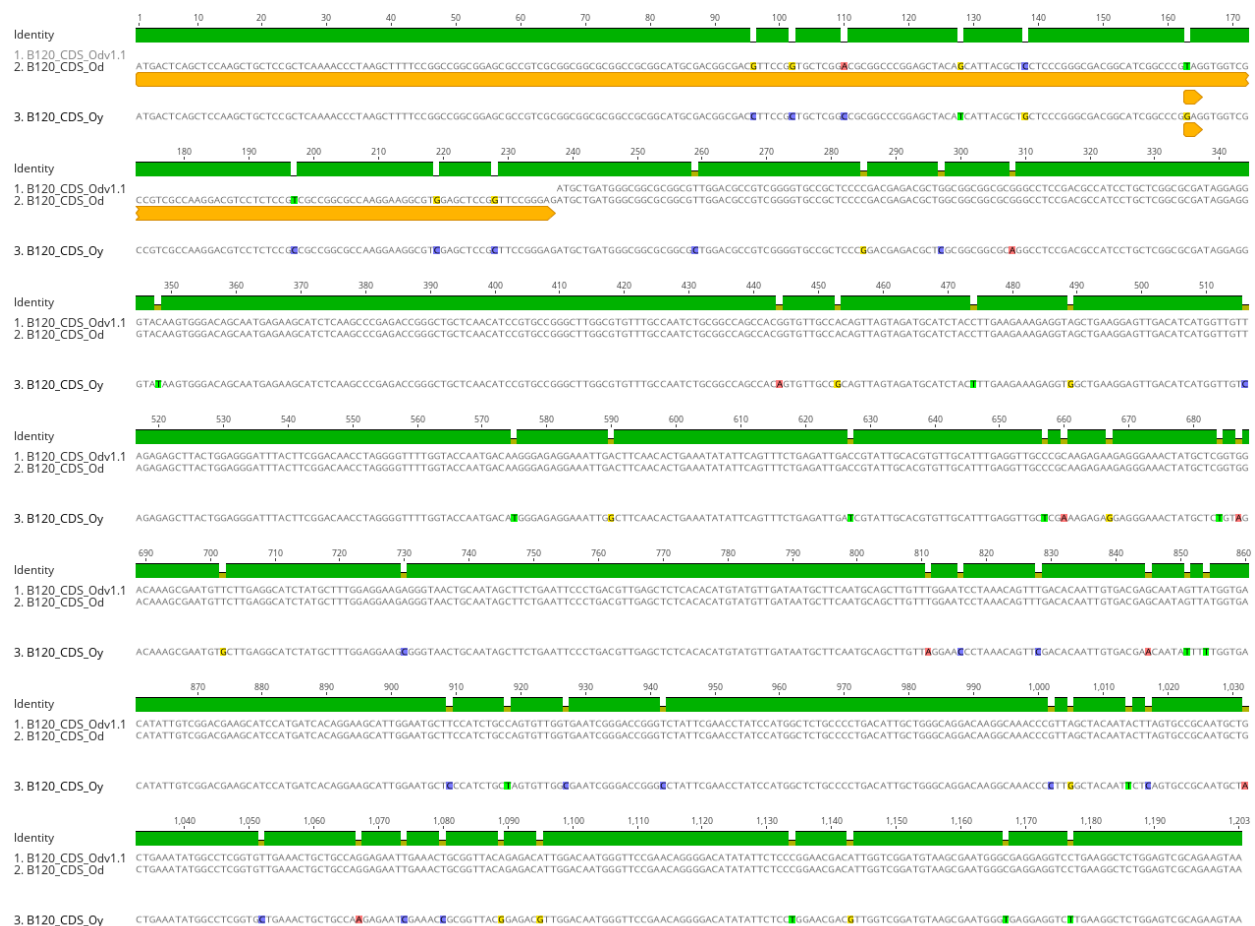


**Figure 3.6 Alignment of coding sequences of *B120* predicted in Overland v1.1 and amplified from Overland and Overlay**

Green bar and grey nucleotides were identical regions; open bar and dark nucleotides were disagreement regions; yellow annotations are the additional 237 bp in Overlay and the SNP caused premature stop codon.



**Figure 3.7 Alignment of proteins sequences of B120 from Overland, Overlay and its homoeologous sequences from published wheat genomes**

The protein sequence of B120 from Overland was set as the reference, the amino acids that disagree with the reference were in black.

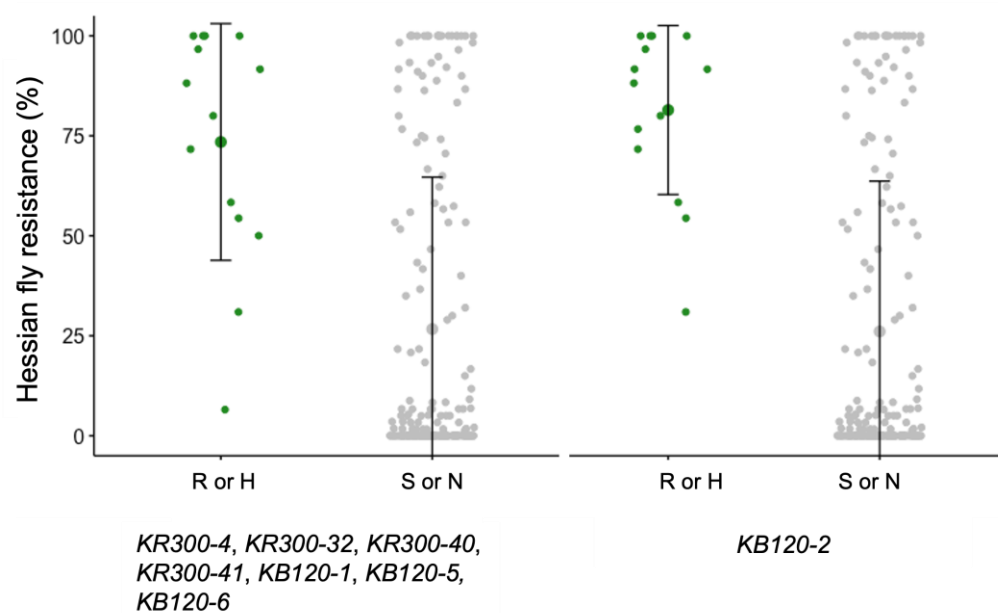


**Figure 3.8 Phenotypic distribution of Hessian fly resistance scores (%) of the 203 accessions in the AM203 panel grouped by eight Kompetitive allele specific PCR markers from *H35* candidate genes *R300* and *B120***

Y-axis is percentage of Hessian fly resistance scores and x-axis showed different marker alleles where R is marker allele for resistance as in Overland, H is heterozygote, S is the allele for susceptibility as in Overley and N is missing data.

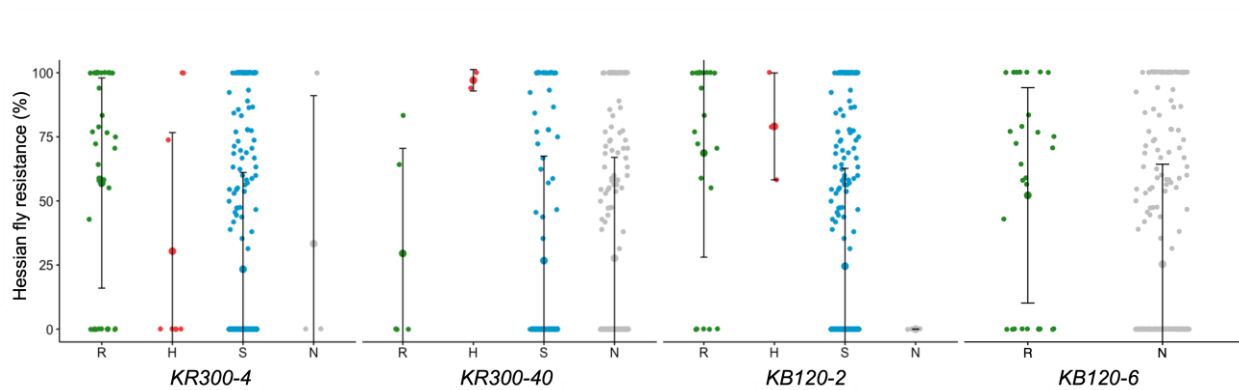


**Figure 3.9 Phenotypic distribution of Hessian fly resistance scores (%) of the 327 accessions in the RGON2022 panel grouped by four Kompetitive allele specific PCR markers from *H35* candidate genes *R300* and *B120***

Y-axis is percentage of Hessian fly resistance scores and x-axis showed different marker alleles where R is marker allele for resistance as in Overland, H is heterozygote, S is the allele for susceptibility as in Overley and N is missing data.

**Table 3.1 Statistics of the Overland genome assembly and annotation**

| Genomic feature                          | Value    |
|------------------------------------------|----------|
| Length of HiFi assembly                  | 14.57 Gb |
| Number of contigs                        | 3,507    |
| Length of contig N50                     | 18.10 Mb |
| Length of contig N90                     | 4.04 Mb  |
| Length of pseudomolecule assembly        | 14.57 Gb |
| Total length of anchored pseudomolecules | 14.43 Gb |
| Number of anchored contigs               | 1,827    |
| Number of high-confidence genes          | 118,361  |
| Total length of unanchored chromosome    | 0.14 Gb  |
| Number of unanchored contigs             | 1,680    |
| Number of high-confidence genes          | 8,173    |

**Table 3.2 Fourteen accessions with resistance allele using gene markers in the AM203 panel**

| Name                                | Loc. | HFR%  | KR300<br>-4 | KR300<br>-32 | KR300<br>-40 | KR300<br>-41 | KB12<br>0-1 | KB120<br>-5 | KB120<br>-6 | KB120<br>-2 |
|-------------------------------------|------|-------|-------------|--------------|--------------|--------------|-------------|-------------|-------------|-------------|
| SD06173                             | SD   | 100   | R           | H            | H            | H            | R           | H           | R           | R           |
| NE06619                             | NE   | 100   | R           | H            | H            | H            | R           | H           | R           | R           |
| SD07204                             | SD   | 100   | R           | H            | R            | R            | R           | R           | R           | H           |
| KS07HW117                           | KS   | 100   | R           | H            | H            | H            | R           | H           | R           | R           |
| NW04Y2188                           | NE   | 96.67 | R           | R            | R            | R            | R           | R           | R           | H           |
| NE04424                             | NE   | 91.67 | N           | N            | S            | N            | N           | N           | N           | H           |
| NE05548                             | NE   | 91.67 | R           | R            | R            | R            | R           | R           | R           | H           |
| SD05118                             | SD   | 88.16 | R           | H            | R            | R            | R           | R           | R           | R           |
| NE06549                             | NE   | 80    | R           | R            | R            | R            | N           | R           | R           | H           |
| NE05549                             | NE   | 76.67 | S           | N            | S            | N            | N           | N           | N           | R           |
| SD05W148-1                          | SD   | 71.67 | R           | H            | R            | R            | R           | R           | R           | H           |
| N02Y5117                            | NE   | 58.33 | R           | R            | R            | R            | R           | R           | R           | H           |
| SD05210                             | SD   | 54.39 | R           | H            | H            | H            | R           | H           | R           | R           |
| SD07W041                            | SD   | 50    | R           | H            | R            | R            | H           | R           | R           | S           |
| CO03064                             | CO   | 30.93 | R           | H            | R            | R            | R           | R           | R           | H           |
| TXHT001F8-<br>CS06/325-<br>PRE07/75 | TX   | 6.51  | R           | R            | R            | R            | H           | R           | R           | S           |

\* R is marker allele for resistance as in Overland, H is heterozygote, S is the allele for susceptibility as in Overley and N is missing data.

## Chapter 4 - Development of a Diagnostic Marker for Hessian Fly

### Resistance Gene *H13*

#### 4.1 Introduction

Diploid and tetraploid wheat relatives have been important sources of HF resistance. Among them, the diploid species *Aegilops tauschii* is the donor of eight *H* genes including *H13*, *H22*, *H23*, *H24*, *H26*, *H32*, *Hwgrc4* and *HNC09MDD14* (Bagwell et al., 2023; McIntosh et al., 2013). *H13* originated from an *Ae. tauschii* accession ‘KU276’, and it was introgressed into wheat through a synthetic hexaploid wheat by interspecific hybridization between tetraploid *Triticum persicum* L. var. *stramineum* Zhuk. (AABB) and *Ae. tauschii* KU276 (Hatchett et al., 1981). Two synthetic wheat lines were derived ‘KU221-14’ and ‘KU221-19’ exhibited 100% resistance to HF biotype D, characterized by dead first-instar larvae at the base of wheat plants. Monosomic and ditelosomic analyses located *H13* to the long arm of chromosome 6D in wheat line ‘KS81H164OHF’ [KU221-14 / ‘Eagle’ // ‘NE73640’ / ‘Cheney’] (Gill et al., 1987). *H13* was introgressed into ‘Molly’ (PI 562619) by backcrossing a wheat cultivar ‘Newton’ to KU221-19 (Patterson et al., 1994). Gene mapping using an F<sub>2:3</sub> families derived from PI 372129 × Molly reassigned *H13* to distal end of chromosome arm 6DS based on a closely linked SSR marker *Xcfd132* (Liu et al., 2005c). More recently, Joshi (2018) identified a candidate gene of *H13* to encode a typical resistance (*R*) gene *CNL-1* with a central nucleotide-binding adapter and a C-terminal leucine-rich repeat domain. The candidate gene was validated through Ethyl methanesulfonate (EMS) mutants.

Despite being discovered for four decades, *H13* continues to exhibit a high level of resistance against most HF populations in the major U.S. wheat-growing areas (Cambron et al., 2010; Chen et al., 2009a; Schemerhorn et al., 2024; Shukle et al., 2015). However, *H13* has not

been widely deployed in U.S. wheat due to the lack of diagnostic markers for *H13*. The objective of this study is to develop breeding friendly diagnostic KASP markers for *H13* to facilitate its identification and deployment in breeding programs. This work will provide a more effective strategy to manage HF outbreaks across diverse agricultural environments.

## **4.2 Materials and methods**

### **4.2.1 Plant materials**

A population of 87 F<sub>3:4</sub> families from Molly × Newton was used for marker evaluation. An association mapping population (AM203) consisting of 203 U.S. winter wheat breeding lines was used to validate the KASP markers developed in this study. In addition, three soft red winter wheat lines from Georgia, ‘GoWheat 2032’ (‘GA 07353-14E19’), ‘PGX 20-15’ and ‘GA101298-17LE11’ known to carry *H13* provided by Dr. Mohamed Mergoum (Mergoum et al., 2021) and a set of 38 accessions containing known *H* genes were also screened for further marker validation. The newly validated diagnostic markers were also used to screen wheat breeding lines in two 2022 hard winter regional nurseries (RGON2022 and RPN2022) to determine the distribution of *H13* in U.S. hard winter wheat breeding programs.

### **4.2.2 Hessian fly infestation and evaluation**

The three Georgia’s lines were evaluated for phenotypic reactions to infestation by HF biotype GP as previously described (Xu et al., 2021). More details on HF phenotyping were described in Chapter 2.2.2.

### **4.2.3 DNA extraction**

Methods for tissue collection and genomic DNA extraction were described in Chapter 2.2.3.

#### 4.2.4 KASP marker design and validation

SNPs due to missense and nonsense mutations in *H13* between Molly and Newton were selected after comparing protein sequences from Joshi (2018). The flanking sequences of the SNP were aligned to IWGSC RefSeq v2.1 using a Basic Local Alignment Search Tool (BLAST) to identify homoeologous sequences from chromosomes 6A, 6B and 6D in Chinese Spring for designing 6D-specific KASP markers. The KASP primers were first screened for polymorphisms between Molly and Newton, then the polymorphic markers were tested for segregation in the (Molly  $\times$  Newton) $F_{3,4}$  families, and were validated in the panels of AM203, RGON2022, RPN 2022 and 38 lines containing known *H* genes. The method for KASP assay was described in Chapter 2.2.5.

### 4.3 Results

#### 4.3.1 Development of *H13* specific markers

Joshi (2018) reported *H13* protein sequences in both Molly and Newton, but only reported the *H13* coding sequence in Molly. The *H13* protein sequence contains 1063 amino acids in Molly and 1499 amino acids in Newton with an identity of 92.5%. The protein length difference between the two parents was due to a SNP-induced premature stop codon at the 1064<sup>th</sup> position of *H13* protein in Molly, and the reverse translated DNA sequence was used to design KASP marker *H13-1a*. Besides, 15 aa mutations from the 117<sup>th</sup> to the 975<sup>th</sup> positions across the whole protein sequence were randomly selected for KASP primers design. Among them, 12 mutation sites were able to design 6D specific KASP primers and the other three failed due to high GC content or high secondary structure scores. Among the 12 sets of KASP primers, only *H13-1a* at 1064<sup>th</sup> amino acid and *H13-15* at the 217<sup>th</sup> amino acid showed polymorphisms

between Molly and Newton (Table 4.1, Figure 4.1) and both markers showed the same segregation pattern in the (Molly  $\times$  Newton) F<sub>3:4</sub> population, except missing data (Figure 4.2).

#### **4.3.2 Validation of KASP markers**

To evaluate their effectiveness in marker-assisted breeding, *H13-1a* and *H13-15* were used to screen the AM203 panel. Only one line, P03207A1-7, from Purdue University, IN, showed Molly's (resistance) alleles at both *H13-1a* and *H13-15*, and it had a 100% resistance score to HF biotype GP. Another line, 'OK05312' carried Molly's allele only at *H13-15* but it was susceptible in the HF evaluation. These results suggest that *H13-1a* is diagnostic for *H13*, while *H13-15* may produce false positive signals in U.S. hard winter wheat.

Due to the false positive signals detected by *H13-15*, only *H13-1a* was used to do further screening. To examine the distribution of *H13* in current U.S. hard winter wheat (HWW) breeding programs, *H13-1a* was used to screen the RGON2022, RPN2022, which did not find Molly's allele in either panel, indicating *H13* has not been deployed in the U.S. hard winter wheat breeding programs.

To further validate its diagnostic value, *H13-1a* was used to screen the three breeding lines from Georgia. Initially four plants per line were genotyped individually. All plants of PGX 20-15 and GA101298-17LE11 showed Molly's allele, while all four plants in GoWheat 2032 amplified Newton's allele. The three Georgia lines were subjected to phenotyping for resistance to HF GP biotype using a large sample size for each line, and the data revealed that PGX 20-15 and GA101298-17LE11 exhibited 100% resistance with 43 and 56 seedlings tested for each line, respectively, whereas only 26 out of 54 tested seedlings of GoWheat2032 were resistant. The DNA was then collected from the phenotyped plants and screened with *H13-1a*. The genotyping result showed a perfect agreement between genotypic and phenotypic data for the three lines. The data

indicated that GoWheat2032 is heterogeneous for *H13* and confirmed *H13-1a* is diagnostic for *H13*.

Among the 38 accessions containing known *H* genes, only Molly (*H13*), KS89WGRC03 (*H23*) and KS89WGRC04 (*Hwgrc4*) amplified the *H13* resistance allele at *H13-1a* and the *H* genes in those accessions all originated from *Ae. tauschii*. Those results imply that either the three *H* genes are the same gene or that the *H13-1a* is not the causal SNP of *H13*, but an *Ae. tauschii* genome-specific SNP.

#### 4.4 Discussion

Among the reported *H* genes, only a few confer consistent high resistance against HF over years, including *H13*, *H18*, *H21*, *H25*, *H26*, *H31*, *H32*, *H33* and *Hdic* (Cambron et al., 2010; Chen et al., 2009a; Schemerhorn et al., 2024; Shukle et al., 2015). *H13* is the first *H* gene that has been cloned although it has not yet been published in a peer-reviewed journal (Joshi, 2018). Previously, an SSR marker *Xcfd132* showed co-segregation with *H13* in a F<sub>2:3</sub> mapping population (Liu et al., 2005c). However, that marker has not been validated and is a gel-based marker which is not suitable for high-throughput screening in modern breeding programs. SNP markers, especially KASP, have significantly improved throughput for genotyping of breeding materials and reduced assay cost and complexity (Semagn et al., 2014). Our study developed two KASP markers, *H13-1a* and *H13-15*, for *H13*. Of them, *H13-1a* showed an agreement between marker alleles and HF phenotypes in the AM203 panel; however, *H13-15* had false positive signals in a susceptible accession. Therefore, *H13-1a* is diagnostic for *H13*. Screening of *H13-1a* in the AM203 panel identified the *H13* resistance allele in only one line, P03207A1-7 ('INW0304'\*2/'RSI5'/'981281'/3/'INW0315'/'99794') (Jin et al., 2013). Based on its pedigree, P03207A1-7 might inherit *H13* from INW0304 (Anderson et al., 2004). Besides, P03207A1-7 is

an elite breeding line with multiple disease resistance, including type II resistance to Fusarium head blight and moderate to high resistance to leaf rust (Jin et al., 2013; Li et al., 2017).

Therefore, it is a valuable *H13* donor for simultaneous improvement of wheat resistance to multiple pests. Further screening of RGON2022 and RPN2022 did not find any accession with the *H13* resistance marker allele, suggesting *H13* has not been deployed in U.S. HWW breeding programs.

*H13* was not detected in HWW, demonstrating *H13-1a* did not amplify false positive signals. To confirm the marker does not fail in detection of real positive genotypes, we requested three soft winter wheat lines from Georgia that were developed using *H13* as the HF resistant source. Among the three lines, PGX 20-15 and GA101298-17LE11 showed complete resistance and were validated to carry the resistance marker allele in all tested plants. Initial marker testing did not find the resistance marker allele in four tested plants of GoWheat2032. However, further genotyping a large sample of plants found heterogeneity in GoWheat2032 seeds and that the plants with the resistance marker allele were all resistant to the biotype GP in a greenhouse experiment. This result further confirmed robustness of *H13-1a* in selecting *H13*.

In the set of 38 accessions with known genes or QTL, KS89WGRC03 (*H23*) and KS89WGRC04 (*Hwgrc4*) also carry the resistance allele of Molly at *H13-1a*. In addition, the QTL *NC09MDD14* was also mapped on chromosome 6DS overlapping with *H13* (Liu et al., 2005c; Ma et al., 1993; Miranda et al., 2010). These genes all originated from different *Ae. Taushii* accessions. *H23* was from the ‘TA1642’/2\*‘Wichita’ (AABBDD), while *Hwgrc4* was from ‘TA1695’/3\*Wichita (Liu et al., 2005c). The QTL *NC09MDD14* was found from wheat line NC09MDD14 with pedigree of ‘Saluda’(AABBDD)\*2/‘TA2492’(DD)//Saluda\*2/‘TA2377’(DD), and therefore, it most likely

originated from TA 2492 or TA2377 (Miranda et al., 2010). The current study found that *H13*, *H23* and *Hwgrc4* all carry the resistance marker allele of *H13-1a*. The line NC09MDD14 was not available for genotyping. However, *H23*, *Hwgrc4* and *NC09MDD14* have all been reported to confer resistance to biotype *vH13* (virulent to *H13*), indicating that they are likely genes or alleles different from *H13* (Miranda et al., 2010). Since the marker can clearly separate resistant plants from susceptible plants in different populations, it is a diagnostic marker for *H13* and can be used for selection of *H13* in breeding programs.

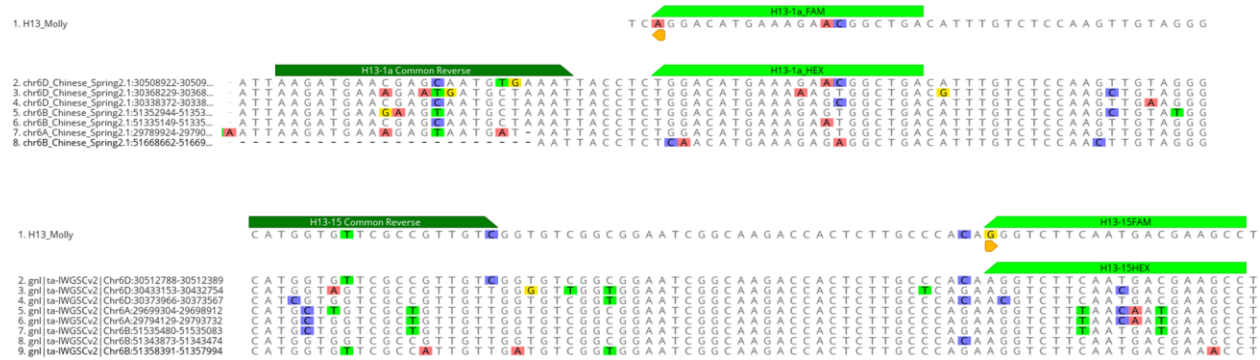


**Figure 4.1 Alignment of *H13* sequence from Molly to its homologous sequences in Chinese Spring RefSeq v2.1 reference genome to design 6D-specific KASP primers *H13-1a* and *H13-15***

The yellow arrows indicate the target single nucleotide polymorphisms; colored nucleotides denote sequence discrepancies.

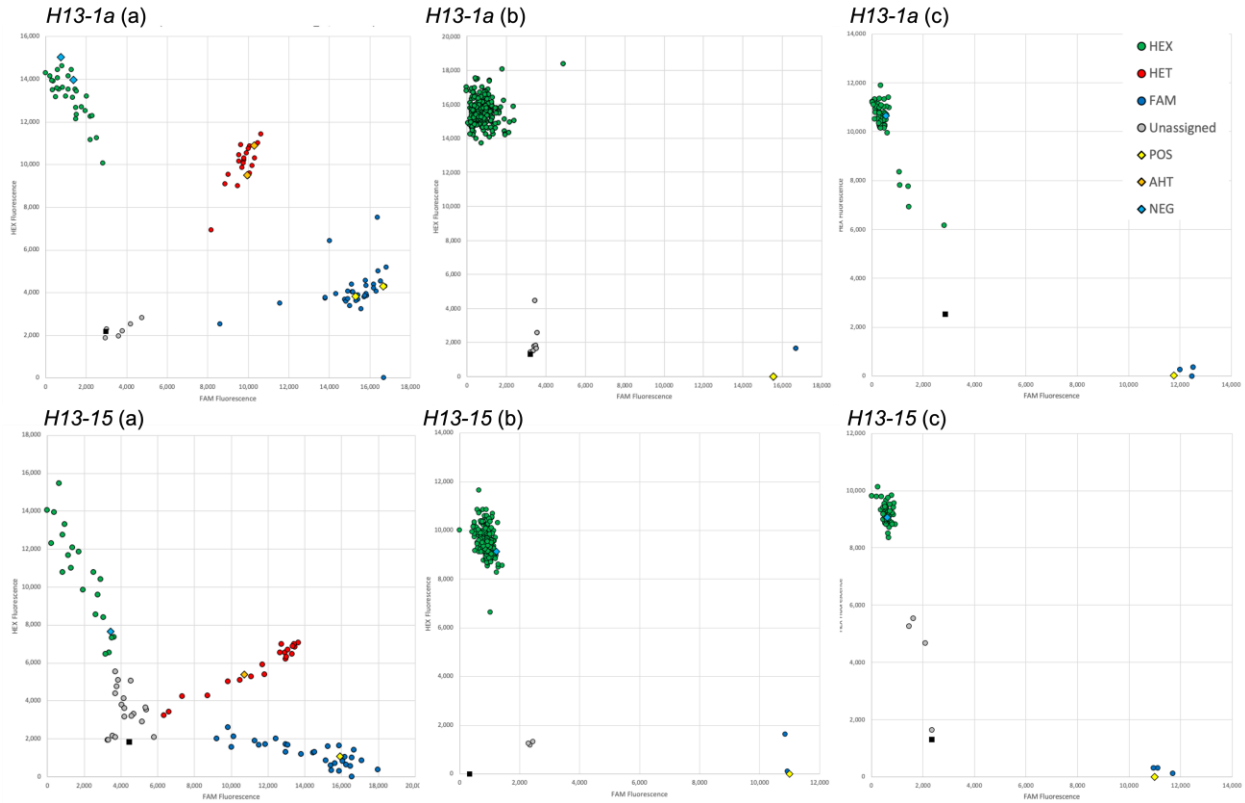


**Figure 4.2 Validation of the Kompetitive Allele Specific PCR (KASP) markers *H13-1a* and *H13-15* for *H13* showing clear cluster separation in the F<sub>3:4</sub> population derived from Molly × Newton (a), the AM203 panel (b) and a set of 38 accessions with known genes or QTL (c)**

The yellow diamonds are Molly as positive controls; blue dots are tested accessions with resistance alleles (Molly type); orange diamonds are artificial heterozygotes generated by PCR amplification of mixed DNA from both parents (Molly and Newton); red dots are accessions tested with heterozygotes; light blue diamonds are Newton as negative controls; green dots are the accessions tested with susceptibility alleles (Newton type); black dots are water controls, while grey dots are the undetermined accessions. FAM, resistance allele of *H13*; HEX, susceptibility allele of *H13*.

**Table 4.1 Sequences of KASP primers and single nucleotide flanking region**

| Name                                | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                          | Note                        | Position (bp)* |
|-------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------|----------------|
| H13-1aFAM                           | GAAGGTGACCAAGTTCATGCTTCAGCC<br>GTTCTTTTCATGTCCT                                                                                                                                                                                                                                                                                                                                                                                                   | <i>H13</i> -Molly type      | 6D:30,509,124  |
| H13-1aHEX                           | GAAGGTGCGGAGTCAACGGATTTCAGCC<br>GTTCTTTTCATGTCCA                                                                                                                                                                                                                                                                                                                                                                                                  | non <i>H13</i> -Newton type |                |
| H13-1aC                             | AAGATGAACGAGCAATGTGAAAT                                                                                                                                                                                                                                                                                                                                                                                                                           | Reverse                     |                |
| H13-15FAM                           | GAAGGTGACCAAGTTCATGCTGGCTTC<br>GTCATTGAAGACCC                                                                                                                                                                                                                                                                                                                                                                                                     | <i>H13</i> -Molly type      | 6D:30,512,589  |
| H13-15HEX                           | GAAGGTGCGGAGTCAACGGATTGGCTTC<br>GTCATTGAAGACCT                                                                                                                                                                                                                                                                                                                                                                                                    | non <i>H13</i> -Newton type |                |
| H13-15C                             | CATGGTGTTTCGCCGTTGTC                                                                                                                                                                                                                                                                                                                                                                                                                              | Reverse                     |                |
| Flanking<br>sequencing<br>of H13-1a | GAGAAAAAGCTAGTAATACATCTATGCAGACATTACTAGCATATTATTGTTGT<br>CCTACTTTAGAAGGGTCTATTCAAAAGGAATAGTTGGCAATATGCAAACCTCCA<br>GCATACTGACAAAATGTACAATACAGACAGAATAGTTCAGATAACCAGAATA<br>GATAGACAAATTAAGATGAACGAGCAATGTGAAATTACCTC[A/t]GGACATGA<br>AAGAACGGCTGACATTTGTCTCCAAGTTGTAGGGATCAGCTGTGTACAAGAC<br>GAACCATTTCTTTGTATTATCTCCTTCAGATGCATATGCTTTGACATGCTTAAC<br>ATGGCTGAACTTGTCCCACTCAGGGCCAGACTGTGCAGCGGCTATGGAAGTG<br>AGCAGTGCAAGGCTACAATATAGCTCCAAATGC  |                             |                |
| Flanking<br>sequencing<br>of H13-15 | GTGCAAGCCTCTCGTTCTGGCAATAAGAGACGTGAGACGTCAGGGGTGTTTG<br>ACCGGTCGGCTGTAGTCGGAGAGAAGATTGAACAAGACACAAGAAAGCTGGT<br>TGAGATCATGCTGAGTGAAAAGGAGGGCAACCCCAACATCATGGTGTTCGCC<br>GTTGTCTGGTGTCTGGCGGAATCGGCAAGACCACTCTTGCCCACA[G/a]GGTCTTC<br>AATGACGAAGCCTTGAACAACAAGTTCCAGAAGACGATATGGTTGAGCATCA<br>ACAAGGACTTTGATGAGGTTGAGTTGCTCAGGACAATCATCACGCAAGCGGG<br>GGGAGTACATTGTGGTCATAAGGCGTTGGCTGTGCTTCAGCCAATCCTTGTCA<br>CTACCTTGAAAGGGAAGAAGCTATTCCTGGTGCTG |                             |                |

\* The physical positions of single nucleotide polymorphisms were determined based on the IWGSC RefSeq v2.1 reference.

## References

- Afzal, F., Chaudhari, S. K., Gul, A., Farooq, A., Ali, H., et al. (2015). Bread wheat (*Triticum aestivum* L.) under biotic and abiotic stresses: an overview. In Hakeem, K. R. (Ed.), *Crop Production and Global Environmental Issues* (pp. 293–317). Springer Charm. [https://doi.org/10.1007/978-3-319-23162-4\\_13](https://doi.org/10.1007/978-3-319-23162-4_13)
- Afzal, S. N., Haque, M., Rauf, A., Ahmad, I., & Firdous, S. S. (2010). Vulnerability of Pakistani wheat (*Triticum aestivum* L.) varieties against stripe rust under rain-fed climate of the northern Punjab and NWFP. *Pakistan Journal of Botany*, 42, 2029–2042.
- Alvey, D.-P. R. (2009). *Evaluation of the impact of Hessian fly (Mayetiola destructor) on Oklahoma winter wheat systems* [Master's thesis, Oklahoma State University]. SHAREOK. <https://shareok.org/handle/11244/8985>
- Amri, A., Cox, T. S., Hatchett, J. H., & Gill, B. S. (1990). Complementary action of genes for Hessian fly resistance in the wheat cultivar 'Seneca'. *Journal of Heredity*, 81, 224–227. <https://doi.org/10.1093/oxfordjournals.jhered.a110977>
- Anderson, J. M., Cambron, S. E., Collier, C. C., Crane, C., Goodwin, S. B., et al. (2004). New cultivars. *Annual Wheat Newsletter*, 50, 207.
- Anjum, F. M., Butt, M. S., Ahmad, N., & Ahmad, I. (2002). Phytate and mineral content in different milling fractions of some Pakistani spring wheats. *International Journal of Food Science & Technology*, 37, 13–17. <https://doi.org/10.1046/j.1365-2621.2002.00540.x>
- Athiyannan, N., Abrouk, M., Boshoff, W. H. P., Cauet, S., Rodde, N., et al. (2022). Long-read genome sequencing of bread wheat facilitates disease resistance gene cloning. *Nature Genetics*, 54, 227–231. <https://doi.org/10.1038/s41588-022-01022-1>
- Aury, J. M., Engelen, S., Istace, B., Monat, C., Lasserre-Zuber, P., et al. (2022). Long-read and chromosome-scale assembly of the hexaploid wheat genome achieves high resolution for research and breeding. *Gigascience*, 11. <https://doi.org/10.1093/gigascience/giac034>
- Avni, R., Nave, M., Barad, O., Baruch, K., Twardziok, S. O., et al. (2017). Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science*, 357, 93–97. <https://doi.org/doi:10.1126/science.aan0032>
- Baenziger, P. S., Beecher, B., Graybosch, R. A., Ibrahim, A. M. H., Baltensperger, D. D., et al. (2008). Registration of 'NE01643' wheat. *Journal of Plant Registrations*, 2, 36–42. <https://doi.org/10.3198/jpr2007.06.0327crc>
- Bagwell, J. W., Subedi, M., Sapkota, S., Lopez, B., Ghimire, B., et al. (2023). Quantitative trait locus analysis of Hessian fly resistance in soft red winter wheat. *Genes*, 14, 1812.

- Bassi, F. M., Brahmi, H., Sabraoui, A., Amri, A., Nsarellah, N., et al. (2019). Genetic identification of loci for Hessian fly resistance in durum wheat. *Molecular Breeding*, 39, 24. <https://doi.org/10.1007/s11032-019-0927-1>
- Behura, S. K., Valicente, F. H., Rider Jr., S. D., Shun-Chen, M., Jackson, S., et al. (2004). A physically anchored genetic map and linkage to avirulence reveals recombination suppression over the proximal region of Hessian fly chromosome A2. *Genetics*, 167, 343–355. <https://doi.org/10.1534/genetics.167.1.343>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Bourras, S., Kunz, L., Xue, M., Praz, C. R., Müller, M. C., et al. (2019). The *AvrPm3-Pm3* effector-NLR interactions control both race-specific resistance and host-specificity of cereal mildews on wheat. *Nature Communications*, 10, 2292. <https://doi.org/10.1038/s41467-019-10274-1>
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., et al. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, 23, 2633–2635. <https://doi.org/10.1093/bioinformatics/btm308>
- Buntin, G., & Bruckner, P. (1990). Effect of planting date on Hessian fly infestation and production of triticale. *Applied Agricultural Research*, 5, 82–88.
- Buntin, G., Ott, S., & Johnson, J. (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, 530–538.
- Cabanettes, F., & Klopp, C. (2018). D-GENIES: dot plot large genomes in an interactive, efficient and simple way. *PeerJ*, 6, e4958. <https://doi.org/10.7717/peerj.4958>
- Cambron, S. E., Buntin, G. D., Weisz, R., Holland, J. D., Flanders, K. L., et al. (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, 2229–2235. <https://doi.org/10.1603/EC10219>
- 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, 971–981. <https://doi.org/10.2135/cropsci2012.11.0636>
- Cartwright, W., & Jones, E. T. (1953). The Hessian fly and how losses from it can be avoided. *USDA Farmers' Bulletin*, 162, 1–9.
- Chen, M.-S., Echegaray, E., Jeffrey Whitworth, R., Wang, H., Sloderbeck, P. E., et al. (2009a). Virulence analysis of Hessian fly populations from Texas, Oklahoma, and Kansas. *Journal of Economic Entomology*, 102, 774–780. <https://doi.org/10.1603/029.102.0239>

- Chen, M.-S., Liu, S., Wang, H., Cheng, X., El Bouhssini, M., et al. (2016). Genes expressed differentially in Hessian fly larvae feeding in resistant and susceptible plants. *International Journal of Molecular Sciences*, 17, 1324. <https://doi.org/10.3390/ijms17081324>
- Chen, M.-S., Liu, X., Wang, H., & El-Bouhssini, M. (2009b). Hessian fly (Diptera: Cecidomyiidae) interactions with barley, rice, and wheat seedlings. *Journal of Economic Entomology*, 102, 1663–1672. <https://doi.org/10.1603/029.102.0434>
- Chen, M.-S., & Liu, X.-M. (2023). Achievements, challenges and future perspective in managing the Hessian fly, *Mayetiola destructor*, in the USA. *Plant Health Cases*, 2023, 8. <https://doi.org/10.1079/planthealthcases.2023.0009>
- Cheng, H., Concepcion, G. T., Feng, X., Zhang, H., & Li, H. (2021). Haplotype-resolved *de novo* assembly using phased assembly graphs with hifiasm. *Nature Methods*, 18, 170–175. <https://doi.org/10.1038/s41592-020-01056-5>
- Divya, D., Himabindu, K., Nair, S., & Bentur, J. S. (2015). Cloning of a gene encoding LRR protein and its validation as candidate gall midge resistance gene, *Gm4*, in rice. *Euphytica*, 203, 185–195. <https://doi.org/10.1007/s10681-014-1302-2>
- Divya, D., Sahu, N., Nair, S., & Bentur, J. S. (2018). Map-based cloning and validation of a gall midge resistance gene, *Gm8*, encoding a proline-rich protein in the rice variety Aganni. *Molecular Biology Reports*, 45, 2075–2086. <https://doi.org/10.1007/s11033-018-4364-8>
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29, 15–21. <https://doi.org/10.1093/bioinformatics/bts635>
- Du, B., Zhang, W., Liu, B., Hu, J., Wei, Z., et al. (2009). Identification and characterization of *Bph14*, a gene conferring resistance to brown planthopper in rice. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 22163–22168.
- Dubcovsky, J., & Dvorak, J. (2007). Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science*, 316, 1862–1866. <https://doi.org/10.1126/science.1143986>
- Dweikat, I., Ohm, H., Patterson, F., & Cambron, S. (1997). Identification of RAPD markers for 11 Hessian fly resistance genes in wheat. *Theoretical and Applied Genetics*, 94, 419–423. <https://doi.org/10.1007/s001220050431>
- Feldman, M., & Levy, A. A. (2023). Evolution of wheat under cultivation. In *Wheat evolution and domestication* (pp. 605–649). Springer Nature. <https://doi.org/10.1007/978-3-031-30175-9>
- Figuerola, M., Hammond-Kosack, K. E., & Solomon, P. S. (2018). A review of wheat diseases—a field perspective. *Molecular Plant Pathology*, 19, 1523–1536. <https://doi.org/10.1111/mpp.12618>

- Gallun, R. (1977). Genetic basis of Hessian fly epidemics. *Annals of the New York Academy of Sciences*, 287, 223–229.
- Gallun, R. L., & Patterson, F. L. (1977). Monosomic analysis of wheat for resistance to Hessian fly. *Journal of Heredity*, 68, 223–226.  
<https://doi.org/10.1093/oxfordjournals.jhered.a108818>
- Ghori, N. (2022). *Towards map-based cloning of a Hessian fly resistance gene H34 in wheat* [Doctoral dissertation, Kansas State University]. K-State Research Exchange (K-REx).  
<https://krex.k-state.edu/handle/2097/42347>
- Gill, B., Wilson, D., Raupp, W., Hatchett, J., Harvey, T., et al. (1991). Registration of KS89WGRC4 hard red winter wheat germplasm with resistance to Hessian fly, greenbug, and soil-borne mosaic virus. *Crop Science*, 31, 246.
- Gill, B. S., Hatchett, J. H., & Raupp, W. J. (1987). Chromosomal mapping of Hessian fly-resistance gene *H13* in the D genome of wheat. *Journal of Heredity*, 78, 97–100.  
<https://doi.org/10.1093/oxfordjournals.jhered.a110344>
- Giovanini, M. P., Saltzmann, K. D., Puthoff, D. P., Gonzalo, M., Ohm, H. W., et al. (2007). A novel wheat gene encoding a putative chitin-binding lectin is associated with resistance against Hessian fly. *Molecular Plant Pathology*, 8, 69–82. <https://doi.org/10.1111/j.1364-3703.2006.00371.x>
- Hao, Y., Cambron, S. E., Chen, Z., Wang, Y., Bland, D. E., et al. (2013). Characterization of new loci for Hessian fly resistance in common wheat. *Theoretical and Applied Genetics*, 126, 1067–1076. <https://doi.org/10.1007/s00122-012-2037-x>
- Harris, M. O., Freeman, T. P., Moore, J. A., Anderson, K. G., Payne, S. A., et al. (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., Stuart, J. J., Mohan, M., Nair, S., Lamb, R. J., et al. (2003). Grasses and gall midges: plant defense and insect adaptation. *Annual Review of Entomology*, 48, 549–577.  
<https://doi.org/10.1146/annurev.ento.48.091801.112559>
- Hatchett, J. H., & Gallun, R. L. (1970). Genetics of the ability of the Hessian fly, *Mayetiola destructor*, to survive on wheats having different genes for resistance. *Annals of the Entomological Society of America*, 63, 1400–1407.  
<https://doi.org/10.1093/aesa/63.5.1400>
- Hatchett, J. H., Martin, T. J., & Livers, R. W. (1981). Expression and inheritance of resistance to Hessian fly in synthetic hexaploid wheats derived from *Triticum tauschii* (Coss) Schmal. *Crop Science*, 21, crops1981.0011183X002100050025x.  
<https://doi.org/10.2135/crops1981.0011183X002100050025x>
- Hussain, S., Habib, M., Ahmed, Z., Sadia, B., Bernardo, A., et al. (2022). Genotyping-by-sequencing based molecular genetic diversity of Pakistani bread wheat (*Triticum*

- aestivum* L.) accessions. *Frontiers in Genetics*, 13, 772517.  
<https://doi.org/10.3389/fgene.2022.772517>
- International Wheat Genome Sequencing Consortium (IWGSC), Appels, R., Eversole, K., Stein, N., Feuillet, C., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, 361, eaar7191.  
<https://doi.org/10.1126/science.aar7191>
- Jia, J., Zhao, S., Kong, X., Li, Y., Zhao, G., et al. (2013). *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature*, 496, 91–95.  
<https://doi.org/10.1038/nature12028>
- Jin, F., Zhang, D., Bockus, W., Baenziger, P. S., Carver, B., et al. (2013). Fusarium head blight resistance in U.S. winter wheat cultivars and elite breeding lines. *Crop Science*, 53, 2006–2013. <https://doi.org/10.2135/cropsci2012.09.0531>
- Jones, J. D. G., & Dangl, J. L. (2006). The plant immune system. *Nature*, 444, 323–329.  
<https://doi.org/10.1038/nature05286>
- Joshi, A. (2018). *Map-based cloning of the Hessian fly resistance gene H13 in wheat* [Doctoral dissertation, Kansas State University]. K-State Research Exchange (K-REx).  
<http://hdl.handle.net/2097/39145>
- Kale, S. M., Schulthess, A. W., Padmarasu, S., Boeven, P. H. G., Schacht, J., et al. (2022). A catalogue of resistance gene homologs and a chromosome-scale reference sequence support resistance gene mapping in winter wheat. *Plant Biotechnology Journal*, 20, 1730–1742. <https://doi.org/10.1111/pbi.13843>
- Kassambara, A. (2023). ggpubr: ‘ggplot2’ based publication ready plots. *R package version 0.6.0*. <https://rpkgs.datanovia.com/ggpubr/>
- Kaur, B., Bhatia, D., & Mavi, G. S. (2021). Eighty years of gene-for-gene relationship and its applications in identification and utilization of *R* genes. *Journal of Genetics*, 100, 50.  
<https://doi.org/10.1007/s12041-021-01300-7>
- Khajuria, C., Wang, H., Liu, X., Wheeler, S., Reese, J. C., et al. (2013). Mobilization of lipids and fortification of cell wall and cuticle are important in host defense against Hessian fly. *BMC Genomics*, 14, 423. <https://doi.org/10.1186/1471-2164-14-423>
- Kidwell, K. K., Shelton, G. B., DeMacon, V. L., Kuehner, J. S., Baik, B., et al. (2009). Registration of ‘Whit’ wheat. *Journal of Plant Registrations*, 3, 279–282.  
<https://doi.org/10.3198/jpr2009.01.0006crc>
- Kong, L., 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, 183–194. <https://doi.org/10.1007/s11032-007-9119-5>

- 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, 525–531. <https://doi.org/10.1111/j.1439-0523.2005.01139.x>
- Lee, S. G., Nwumeh, R., & Jez, J. M. (2016). Structure and mechanism of isopropylmalate dehydrogenase from *Arabidopsis thaliana*: Insights on leucine and aliphatic glucosinolate biosynthesis. *Journal of Biological Chemistry*, 291, 13421–13430. <https://doi.org/10.1074/jbc.M116.730358>
- Levy, A. A., & Feldman, M. (2022). Evolution and origin of bread wheat. *The Plant Cell*, 34, 2549–2567. <https://doi.org/10.1093/plcell/koac130>
- 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, 2065–2071. <https://doi.org/10.1007/s00122-013-2118-5>
- Li, G., Wang, Y., Chen, M. S., Edae, E., Poland, J., et al. (2015). Precisely mapping a major gene conferring resistance to Hessian fly in bread wheat using genotyping-by-sequencing. *BMC Genomics*, 16, 108. <https://doi.org/10.1186/s12864-015-1297-7>
- Li, G., Xu, X., Bai, G., Carver, B. F., Hunger, R., et al. (2017). Novel sources of leaf rust resistance in winter wheat. *Crop Science*, 57, 865–876. <https://doi.org/10.2135/cropsci2016.08.0725>
- Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, 34, 3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>
- Liu, G., Liu, X., Xu, Y., Bernardo, A., Chen, M., et al. (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, 1488–1498. <https://doi.org/10.1002/csc2.20148>
- Liu, X., Khajuria, C., Li, J., Trick, H. N., Huang, L., et al. (2013). Wheat *Mds-1* encodes a heat-shock protein and governs susceptibility towards the Hessian fly gall midge. *Nature Communications*, 4, 2070. <https://doi.org/10.1038/ncomms3070>
- Liu, X., Williams, C. E., Nemacheck, J. A., Wang, H., Subramanyam, S., et al. (2010). Reactive oxygen species are involved in plant defense against a gall midge. *Plant Physiology*, 152, 985–999. <https://doi.org/10.1104/pp.109.150656>
- Liu, X. M., Brown-Guedira, G. L., Hatchett, J. H., Owuochi, J. O., & Chen, M. S. (2005a). 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, 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., et al. (2005b). *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, 1473–1480.  
<https://doi.org/10.1007/s00122-005-1982-z>
- Liu, X. M., Gill, B. S., & Chen, M. S. (2005c). Hessian fly resistance gene *H13* is mapped to a distal cluster of resistance genes in chromosome 6DS of wheat. *Theoretical and Applied Genetics*, 111, 243–249. <https://doi.org/10.1007/s00122-005-2009-5>
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15, 550.  
<https://doi.org/10.1186/s13059-014-0550-8>
- Luo, M.-C., Gu, Y. Q., Puiu, D., Wang, H., Twardziok, S. O., et al. (2017). Genome sequence of the progenitor of the wheat D genome *Aegilops tauschii*. *Nature*, 551, 498–502.  
<https://doi.org/10.1038/nature24486>
- Lv, T., Li, X., Fan, T., Luo, H., Xie, C., et al. (2019). The calmodulin-binding protein IQM1 interacts with CATALASE2 to affect pathogen defense. *Plant Physiology*, 181, 1314–1327. <https://doi.org/10.1104/pp.19.01060>
- 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, 750–754.  
<https://doi.org/10.1007/BF00225015>
- Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., et al. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. *Nature Genetics*, 51, 885–895. <https://doi.org/10.1038/s41588-019-0381-3>
- Marcussen, T., Sandve, S. R., Heier, L., Spannagl, M., Pfeifer, M., et al. (2014). Ancient hybridizations among the ancestral genomes of bread wheat. *Science*, 345, 1250092.  
<https://doi.org/10.1126/science.1250092>
- McIntosh, R. A., Dubcovsky, J., Rogers, W. J., Morris, C., Appels, R., et al. (2013). Catalogue of gene symbols for wheat. 150–152.  
<https://shigen.nig.ac.jp/wheat/komugi/genes/download.jsp>
- Meng, L., Li, H., Zhang, L., & Wang, J. (2015). QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *The Crop Journal*, 3, 269–283. <https://doi.org/10.1016/j.cj.2015.01.001>
- Miranda, L. M., Bland, D. E., Cambron, S. E., Lyerly, J. H., Johnson, J., et al. (2010). Genetic mapping of an *Aegilops tauschii*-derived Hessian fly resistance gene in common wheat. *Crop Science*, 50, 612–616. <https://doi.org/10.2135/cropsci2009.05.0278>
- Mustafa, G., Alam, M. M., Khan, S. U., Naveed, M., & Mumtaz, A. S. (2013). Leaf rust resistance in semi dwarf wheat cultivars: a conspectus of post green revolution period in Pakistan. *Pak. J. Bot*, 45, 415–422.

- Nemacheck, J. A., Flynn, R. D., Zapf, K., & Subramanyam, S. (2023). Identification of new sources of Hessian fly resistance in tetraploid wheat. *Crop Science*, 63, 1354–1363. <https://doi.org/10.1002/csc2.20917>
- Niu, F., Xu, Y., Liu, X., Zhao, L., Bernardo, A., et al. (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, 2927–2935. <https://doi.org/10.1007/s00122-020-03642-9>
- Nombela, G., Williamson, V. M., & Muñiz, M. (2003). The root-knot nematode resistance gene *Mi-1.2* of tomato is responsible for resistance against the whitefly *Bemisia tabaci*. *Molecular Plant-Microbe Interactions*, 16, 645–649. <https://doi.org/10.1094/mpmi.2003.16.7.645>
- Ohm, H. W., Shaner, G., Foster, J. E., Patterson, F. L., & Buechley, G. (1988). Registration of ‘Clark’ wheat. *Crop Science*, 28, 1032. <https://doi.org/10.2135/cropsci1988.0011183X002800060058x>
- Patterson, F. L., Maas, F. B., Foster, J. E., Ratcliffe, R. H., Cambron, S., et al. (1994). Registration of eight Hessian fly resistant common winter wheat germplasm lines (Carol, Erin, Flynn, Iris, Joy, Karen, Lola, and Molly). *Crop Science*, 34, 315–316. <https://doi.org/10.2135/cropsci1994.0011183X003400010084x>
- Patterson, F. L., Shaner, G. E., Huber, D. M., Ohm, H. W., Finney, R. E., et al. (1979). Registration of Sullivan wheat. *Crop Science*, 19, 297. <https://doi.org/10.2135/cropsci1979.0011183X001900020043x>
- Prather, S., Schneider, T., Gaham Godoy, J., Odubiyi, S., Bosque-Perez, N. A., et al. (2022). Reliable DNA markers for a previously unidentified, yet broadly deployed Hessian fly resistance gene on chromosome 6B in Pacific Northwest spring wheat varieties. *Frontiers in Plant Science*, 13, 779096. <https://doi.org/10.3389/fpls.2022.779096>
- Puthoff, D. P., Sardesai, N., Subramanyam, S., Nemacheck, J. A., & Williams, C. E. (2005). *Hfr-2*, a wheat cytolytic toxin-like gene, is up-regulated by virulent Hessian fly larval feeding. *Molecular Plant Pathology*, 6, 411–423. <https://doi.org/10.1111/j.1364-3703.2005.00289.x>
- Ratcliffe, R. H., Safranski, G. G., Patterson, F. L., Ohm, H. W., & Taylor, P. L. (1994). Biotype status of Hessian fly (Diptera: Cecidomyiidae) populations from the eastern United States and their response to 14 Hessian fly resistance genes. *Journal of Economic Entomology*, 87, 1113–1121. <https://doi.org/10.1093/jee/87.4.1113>
- Roberts, J. J., & Gillum, R. L. (1984). Chromosome location of the *H5* gene for resistance to the Hessian fly in wheat. *Journal of Heredity*, 75, 147–148. <https://doi.org/10.1093/oxfordjournals.jhered.a109893>
- Rossi, M., Goggin, F. L., Milligan, S. B., Kaloshian, I., Ullman, D. E., et al. (1998). The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid.

- Proceedings of the National Academy of Sciences of the United States of America*, 95, 9750–9754.
- Sama, V. S. A. K., Rawat, N., Sundaram, R. M., Himabindu, K., Naik, B. S., et al. (2014). A putative candidate for the recessive gall midge resistance gene *gm3* in rice identified and validated. *Theoretical and Applied Genetics*, 127, 113–124.  
<https://doi.org/10.1007/s00122-013-2205-7>
- Sato, K., Abe, F., Mascher, M., Haberer, G., Gundlach, H., et al. (2021). Chromosome-scale genome assembly of the transformation-amenable common wheat cultivar 'Fielder'. *DNA Research*, 28. <https://doi.org/10.1093/dnares/dsab008>
- Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., et al. (2019). The global burden of pathogens and pests on major food crops. *Nature Ecology & Evolution*, 3, 430–439. <https://doi.org/10.1038/s41559-018-0793-y>
- Schemerhorn, B. J., Cambron, S. E., & Subramanyam, S. (2024). Hessian fly (Diptera: Cecidomyiidae) virulence in Louisiana: assessment of field populations from 2023 to efficacy of 27 Hessian fly resistance genes in wheat. *Journal of Economic Entomology*. <https://doi.org/10.1093/jee/toae002>
- Schmid, R. B., Knutson, A., Giles, K. L., & McCornack, B. P. (2018). Hessian fly (Diptera: Cecidomyiidae) biology and management in wheat. *Journal of Integrated Pest Management*, 9, 1–12. <https://doi.org/10.1093/jipm/pmy008>
- Scholz, S. S., Vadassery, J., Heyer, M., Reichelt, M., Bender, K. W., et al. (2014). Mutation of the *Arabidopsis* calmodulin-like protein CML37 deregulates the jasmonate pathway and enhances susceptibility to herbivory. *Molecular plant*, 7, 1712–1726.  
<https://doi.org/10.1093/mp/ssu102>
- Schulte, S. J., Rider Jr., S. D., Hatchett, J. H., & Stuart, J. J. (1999). Molecular genetic mapping of three X-linked avirulence genes, *vH6*, *vH9* and *vH13*, in the Hessian fly. *Genome*, 42, 821–828. <https://doi.org/10.1139/g98-162>
- Semagn, K., Babu, R., Hearne, S., & Olsen, M. (2014). Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. *Molecular Breeding*, 33, 1–14.  
<https://doi.org/10.1007/s11032-013-9917-x>
- Shukle, R. H. (2005). Hessian fly, *Mayetiola Destructor* (Say) (Diptera: Cecidomyiidae). In Capinera, J. L. (Ed.), *Encyclopedia of Entomology* (pp. 1070–1073). Springer Dordrecht. [https://doi.org/10.1007/0-306-48380-7\\_2004](https://doi.org/10.1007/0-306-48380-7_2004)
- Shukle, R. H., Cambron, S. E., Moniem, H. A., Schemerhorn, B. J., Redding, J., et al. (2015). Effectiveness of genes for Hessian fly (Diptera: Cecidomyiidae) resistance in the southeastern United States. *Journal of Economic Entomology*, 109, 399–405.  
<https://doi.org/10.1093/jee/tov292>

- Stuart, J. J., Chen, M. S., Shukle, R., & Harris, M. O. (2012). Gall midges (Hessian flies) as plant pathogens. *Annual Review of Phytopathology*, 50, 339–357.  
<https://doi.org/10.1146/annurev-phyto-072910-095255>
- Subramanyam, S., Zheng, C., Shukle, J. T., & Williams, C. E. (2013). Hessian fly larval attack triggers elevated expression of disease resistance dirigent-like protein-encoding gene, *HfrDrd*, in resistant wheat. *Arthropod-Plant Interactions*, 7, 389–402.  
<https://doi.org/10.1007/s11829-013-9253-4>
- Tan, C. T., Carver, B. F., Chen, M. S., Gu, Y. Q., & Yan, L. (2013). Genetic association of *OPR* genes with resistance to Hessian fly in hexaploid wheat. *BMC Genomics*, 14, 369.  
<https://doi.org/10.1186/1471-2164-14-369>
- Tan, C. T., Yu, H., Yang, Y., Xu, X., Chen, M., et al. (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, 1867–1884.  
<https://doi.org/10.1007/s00122-017-2930-4>
- Van Ooijen, J. (2006). JoinMap 4. *Software for the calculation of genetic linkage maps in experimental populations*. Kyazma BV, Wageningen, Netherlands, 33.
- Venske, E., dos Santos, R. S., Busanello, C., Gustafson, P., & Costa de Oliveira, A. (2019). Bread wheat: a role model for plant domestication and breeding. *Hereditas*, 156, 16.  
<https://doi.org/10.1186/s41065-019-0093-9>
- Voorrips, R. E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *Journal of Heredity*, 93, 77–78. <https://doi.org/10.1093/jhered/93.1.77>
- Walkowiak, S., Gao, L., Monat, C., Haberer, G., Kassa, M. T., et al. (2020). Multiple wheat genomes reveal global variation in modern breeding. *Nature*, 588, 277–283.  
<https://doi.org/10.1038/s41586-020-2961-x>
- Wang, L., Zhu, T., Rodriguez, J. C., Deal, K. R., Dubcovsky, J., et al. (2021). *Aegilops tauschii* genome assembly Aet v5.0 features greater sequence contiguity and improved annotation. *G3: Genes|Genomes|Genetics*, 11. <https://doi.org/10.1093/g3journal/jkab325>
- Wang, S., Basten C. J., & Z.-B., Z. (2012). Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC.  
<http://statgen.ncsu.edu/qtlcart/WQTLCart.htm>
- Williams, C. E., Collier, C. C., Nemacheck, J. A., Liang, C., & Cambron, S. E. (2002). A lectin-like wheat gene responds systemically to attempted feeding by avirulent first-instar Hessian fly larvae. *Journal of Chemical Ecology*, 28, 1411–1428.  
<https://doi.org/10.1023/A:1016200619766>
- Winn, Z. J., Acharya, R., Merrill, K., Lyster, J., Brown-Guedira, G., et al. (2021). Mapping of a novel major effect Hessian fly field partial-resistance locus in southern soft red winter

- wheat line LA03136E71. *Theoretical and Applied Genetics*.  
<https://doi.org/10.1007/s00122-021-03936-6>
- Xu, Y., La, G., Fatima, N., Liu, Z., Zhang, L., et al. (2021). Precise mapping of QTL for Hessian fly resistance in the hard winter wheat cultivar ‘Overland’. *Theoretical and Applied Genetics*, 134, 3951–3962. <https://doi.org/10.1007/s00122-021-03940-w>
- 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, 117–127. <https://doi.org/10.3835/plantgenome2010.03.0004>
- Zhang, L., Xu, Y., Chen, M.-S., Su, Z., Liu, Y., et al. (2021). Identification of a major QTL for Hessian fly resistance in wheat cultivar ‘Chokwang’. *The Crop Journal*, 10, 775–782. <https://doi.org/10.1016/j.cj.2021.08.004>
- Zhao, C., Escalante, L. N., Chen, H., Benatti, T. R., Qu, J., et al. (2015). A massive expansion of effector genes underlies gall-formation in the wheat pest *Mayetiola destructor*. *Current Biology*, 25, 613–620. <https://doi.org/10.1016/j.cub.2014.12.057>
- Zhao, C., Shukle, R., Navarro-Escalante, L., Chen, M., Richards, S., et al. (2016). Avirulence gene mapping in the Hessian fly (*Mayetiola destructor*) reveals a protein phosphatase 2C effector gene family. *Journal of Insect Physiology*, 84, 22–31. <https://doi.org/10.1016/j.jinsphys.2015.10.001>
- Zhao, L., Abdelsalam, N. R., Xu, Y., Chen, M. S., Feng, Y., et al. (2020). Identification of two novel Hessian fly resistance genes *H35* and *H36* in a hard winter wheat line SD06165. *Theoretical and Applied Genetics*, 133, 2343–2353 <https://doi.org/10.1007/S00122-020-03602-3>
- Zhou, Y., Bai, S., Li, H., Sun, G., Zhang, D., et al. (2021). Introgressing the *Aegilops tauschii* genome into wheat as a basis for cereal improvement. *Nature Plants*, 7, 774–786. <https://doi.org/10.1038/s41477-021-00934-w>
- Zhu, T., Wang, L., Rimbart, H., Rodriguez, J. C., Deal, K. R., et al. (2021). Optical maps refine the bread wheat *Triticum aestivum* cv. Chinese Spring genome assembly. *The Plant Journal*, 107, 303–314. <https://doi.org/10.1111/tpj.15289>
- Zhu, T., Wang, L., Rodriguez, J. C., Deal, K. R., Avni, R., et al. (2019). Improved genome sequence of wild emmer wheat Zavitan with the aid of optical maps. *G3: Genes|Genomes|Genetics*, 9, 619–624. <https://doi.org/10.1534/g3.118.200902>
- Zipfel, C. (2014). Plant pattern-recognition receptors. *Trends in Immunology*, 35, 345–351. <https://doi.org/10.1016/j.it.2014.05.004>

## **Appendix A - Copyright information**

This section contains the copyright information in the dissertation.

The research work in chapters 2–4 has not been published yet. As the copyright holder, we reserve the right to publish it first. Permission is required from the authors before using this unpublished content from this dissertation.

## Appendix B - Supplementary data of Chapter 2

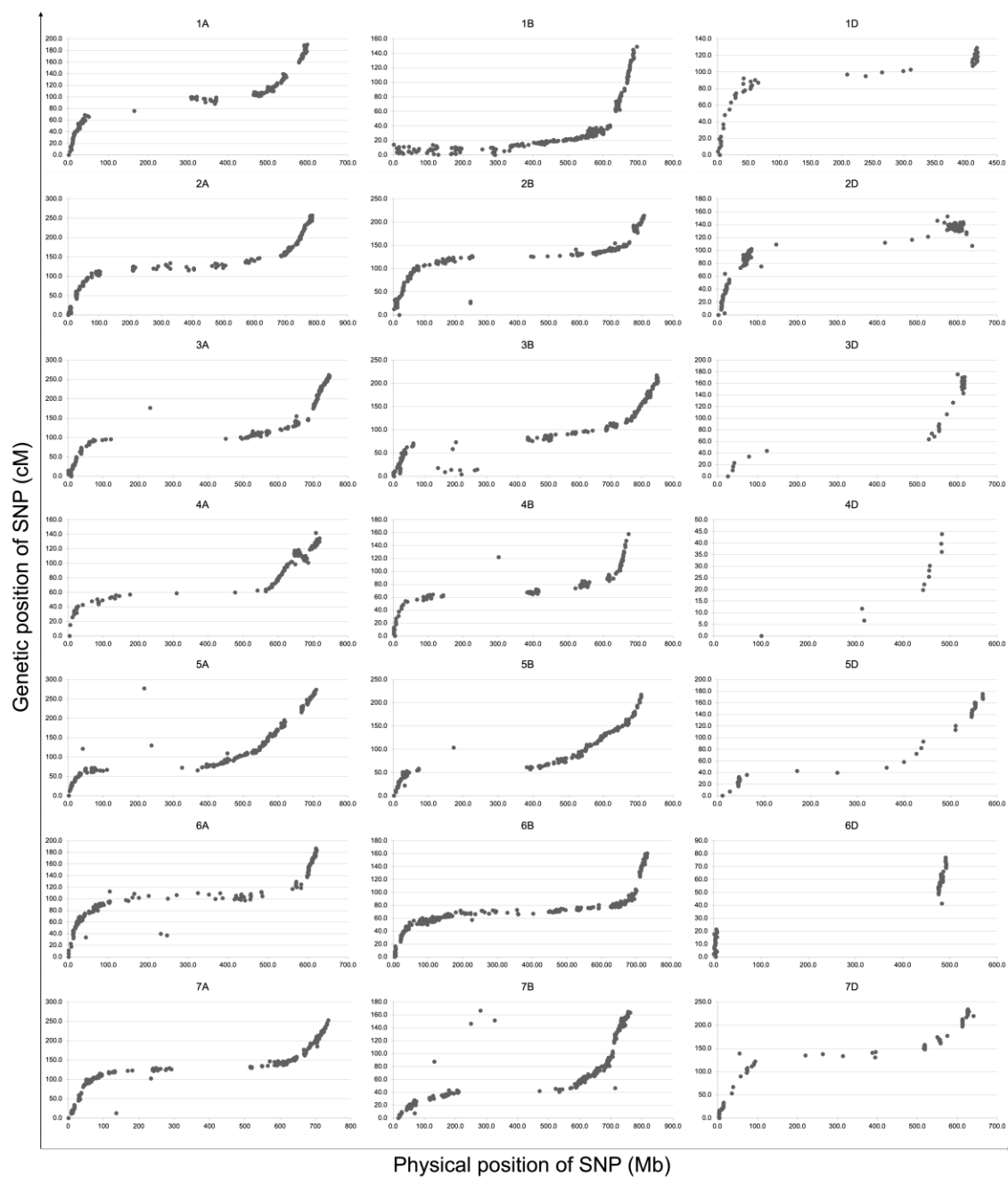


**Figure B.1 Collinearity of the genetic and physical maps of 21 wheat chromosomes based on Chinese Spring IWGSC RefSeq 2.1 reference genome**

X-axis is the physical positions (Mb) of the SNPs in the IWGSC RefSeq v2.1 reference genome, and y-axis is the genetic distance (cM) in the linkage map of Parvaz-94 × Cadenza.

**Table B.1 CB. No, name or pedigree, origin and Hessian fly resistance of the panel 184  
Pakistani wheat landraces and cultivars**

| No. | CB. No. | Name or pedigree   | Origin           | Hessian fly resistance (%) |
|-----|---------|--------------------|------------------|----------------------------|
| 184 | CB55    | PARVAZ-94          | Pakistan         | 100                        |
| 1   | CB3     | BHAKKAR-2000       | Pakistan         | 0                          |
| 2   | CB4     | CHAKWAL-50         | Pakistan         | 0                          |
| 3   | CB8     | FAREED-06          | Pakistan         | 0                          |
| 4   | CB11    | INQILAB 91         | Pakistan         | 0                          |
| 5   | CB14    | MANTHAR            | Pakistan         | 0                          |
| 6   | CB15    | MIRAJ-08           | Pakistan         | 0                          |
| 7   | CB17    | SHAFaq-06          | Pakistan         | 0                          |
| 8   | CB20    | BARS-09            | Pakistan         | 0                          |
| 9   | CB23    | BHITTAI            | Pakistan         | 0                          |
| 10  | CB30    | SASSI              | Pakistan         | 0                          |
| 11  | CB37    | PIRSABAK 2005      | Pakistan         | 0                          |
| 12  | CB42    | CHENAB-79          | Pakistan         | 0                          |
| 13  | CB46    | IQBAL2000          | Pakistan         | 0                          |
| 14  | CB47    | JAUHAR-78          | Pakistan         | 0                          |
| 15  | CB50    | KOHISTAN 97        | Pakistan         | 0                          |
| 16  | CB56    | PASINA 90          | Pakistan         | 0                          |
| 17  | CB57    | PUNJAB-76          | Pakistan         | 0                          |
| 18  | CB60    | PUNJAB 96          | Pakistan         | 0                          |
| 19  | CB61    | SA-42              | Pakistan         | 0                          |
| 20  | CB65    | SHAHKAR 95         | Pakistan         | 0                          |
| 21  | CB66    | SHALIMAR-88        | Pakistan         | 0                          |
| 22  | CB69    | ZA-77              | Pakistan         | 0                          |
| 23  | CB73    | KAUZ'S'            | Australia/CIMMYT | 0                          |
| 24  | CB74    | NACozARI F-76      | CIMMYT           | 0                          |
| 25  | CB75    | OASIS F-86         | CIMMYT           | 0                          |
| 26  | CB76    | PBW-343=ATTILA     | India/CIMMYT     | 0                          |
| 27  | CB80    | FRET-1             | CIMMYT           | 0                          |
| 28  | CB82    | WH-542             | India/CIMMYT     | 0                          |
| 29  | CB83    | HOOSAM-3           | ICARDA           | 0                          |
| 30  | CB85    | SAAR               | CIMMYT           | 0                          |
| 31  | CB94    | CHAM-4             | ICARDA           | 0                          |
| 32  | CB96    | CHILERO=CHIL'S'    | CIMMYT           | 0                          |
| 33  | CB99    | FRONTANA           | Brazil/CIMMYT    | 0                          |
| 34  | CB100   | HARTOG=HTG.(PAVON) | CIMMYT           | 0                          |

|    |       |                                                                   |                  |   |
|----|-------|-------------------------------------------------------------------|------------------|---|
| 35 | CB103 | OASIS/SKAUZ//4*BCN/3/2*PASTOR                                     | CIMMYT           | 0 |
| 36 | CB104 | BABAX/LR42//BABAX*2/3/VIVITSI                                     | CIMMYT           | 0 |
| 37 | CB106 | PBW 343*2/KUKUNA                                                  | CIMMYT           | 0 |
| 38 | CB107 | PBW 343*2/KURUKU                                                  | CIMMYT           | 0 |
| 39 | CB115 | PVN//CAR422/ANA/3/KAUZ*2/TRAP/<br>/KAUZ                           | CIMMYT           | 0 |
| 40 | CB117 | TRAP#1/PBW65/3/KAUZ*2/TRAP//KA<br>UZ                              | CIMMYT           | 0 |
| 41 | CB120 | PVN/PBW65/3/KAUZ*2/TRAP//KAUZ                                     | CIMMYT           | 0 |
| 42 | CB121 | PARULA=PRL                                                        | CIMMYT           | 0 |
| 43 | CB123 | NING-8319                                                         | China/CIMMYT     | 0 |
| 44 | CB124 | HARRIER 17.B                                                      | Australia/CIMMYT | 0 |
| 45 | CB126 | V-03007                                                           | Pakistan         | 0 |
| 46 | CB128 | CON.'S'/ANA 75//CON.'S'                                           | CIMMYT           | 0 |
| 47 | CB129 | HD2236//SA.42/HARRIER'S= V-97088                                  | Pakistan         | 0 |
| 48 | CB130 | PB81//F3.71/TRM/3/BULBUL// F3. 71/<br>TRM =V0005                  | CIMMYT           | 0 |
| 49 | CB131 | WEEBILL-1 = V-03158                                               | CIMMYT           | 0 |
| 50 | CB133 | WATAN/2*ERA                                                       | CIMMYT           | 0 |
| 51 | CB138 | TURACO/PRINIA                                                     | CIMMYT           | 0 |
| 52 | CB146 | PB-96/87094//MH-97                                                | Pakistan         | 0 |
| 53 | CB147 | MAYA/PVN                                                          | CIMMYT           | 0 |
| 54 | CB148 | WL 711/CROW "S"//ALD #1 /<br>CMH77A.917/3/HI 666/PVN 'S'          | CIMMYT           | 0 |
| 55 | CB150 | PRL/2*PASTOR//PARUS/5/NAC/TH.A<br>C//3*PVN/3/MIRLO/BUC/4/2*PASTOR | CIMMYT           | 0 |
| 56 | CB154 | FRET2*2/4/SNI/TRAP#1/3/KAUZ*2/T<br>RAP//KAUZ/5/ONIX               | CIMMYT           | 0 |
| 57 | CB163 | CROC-1/AE.SQ(224)//OPATA/3/FLAG-<br>7                             | CIMMYT           | 0 |
| 58 | CB165 | SERI.1B*2/3/KAUZ*2/BOW//K                                         | CIMMYT           | 0 |
| 59 | CB167 | SHUHA-4//NS732/HER/3/<br>MILAN/DUCULA                             | CIMMYT           | 0 |
| 60 | CB171 | BWP 122526                                                        | Pakistan         | 0 |
| 61 | CB174 | NW S-2001                                                         | Pakistan         | 0 |
| 62 | CB176 | PR-111                                                            | Pakistan         | 0 |
| 63 | CB179 | PR-106                                                            | Pakistan         | 0 |
| 64 | CB186 | 13248                                                             | Pakistan         | 0 |
| 65 | CB188 | LONG GRAIN                                                        | Unknown          | 0 |
| 66 | CB195 | NSW-14                                                            | Australia/CIMMYT | 0 |
| 67 | CB196 | V-12266                                                           | Pakistan         | 0 |
| 68 | CB198 | V-13270                                                           | Pakistan         | 0 |
| 69 | CB200 | 122557                                                            | Pakistan         | 0 |
| 70 | CB202 | V-02192                                                           | Pakistan         | 0 |

|     |       |                                                                                                                   |                  |   |
|-----|-------|-------------------------------------------------------------------------------------------------------------------|------------------|---|
| 71  | CB203 | V-02156                                                                                                           | Pakistan         | 0 |
| 72  | CB208 | V-04048                                                                                                           | Pakistan         | 0 |
| 73  | CB210 | V-05115                                                                                                           | Pakistan         | 0 |
| 74  | CB212 | V-06129                                                                                                           | Pakistan         | 0 |
| 75  | CB214 | V-056132                                                                                                          | Pakistan         | 0 |
| 76  | CB220 | V-06018                                                                                                           | Pakistan         | 0 |
| 77  | CB224 | V-06068                                                                                                           | Pakistan         | 0 |
| 78  | CB228 | KIRITATI//PBW65/2*SERI.1B                                                                                         | CIMMYT           | 0 |
| 79  | CB232 | KIRITATI/4/2*SERI.1B*2/3/KAUZ*2/BOW//KAUZ                                                                         | CIMMYT           | 0 |
| 80  | CB234 | WHEAR/VIVITSI//WHEAR                                                                                              | CIMMYT           | 0 |
| 81  | CB236 | WHEAR/CHAPIO//WHEAR                                                                                               | CIMMYT           | 0 |
| 82  | CB238 | WHEAR/KUKUNA/3/C80.1/3*BATAVIA//2*WBLL1                                                                           | CIMMYT           | 0 |
| 83  | CB239 | INQALAB91*2/KUKUNA//                                                                                              | CIMMYT           | 0 |
| 84  | CB240 | SUNCO//TNMU/TUI                                                                                                   | CIMMYT           | 0 |
| 85  | CB242 | SHARP/3/PRL/SARA//TSI/VEE#5/5/VEE/LIRA//BOW/3/BCN/4/KAUZ                                                          | CIMMYT           | 0 |
| 86  | CB244 | DOLLARBIRD                                                                                                        | Australia/CIMMYT | 0 |
| 87  | CB247 | KIRITATI                                                                                                          | CIMMYT           | 0 |
| 88  | CB248 | PFAU/WEAVER*2//KIRITATI                                                                                           | CIMMYT           | 0 |
| 89  | CB249 | PGO/SERI//BAV92                                                                                                   | CIMMYT           | 0 |
| 90  | CB251 | KINGBIRD#1                                                                                                        | CIMMYT           | 0 |
| 91  | CB253 | TAM200/TUI                                                                                                        | CIMMYT           | 0 |
| 92  | CB256 | CROC_1/AE.SQUARROSA (205)//FCT/3/PASTOR                                                                           | CIMMYT           | 0 |
| 93  | CB259 | HD 2169/C591//PBW343                                                                                              | Pakistan         | 0 |
| 94  | CB262 | V-86711TC/SH-88//CROW                                                                                             | Pakistan         | 0 |
| 95  | CB266 | AS2002/WL711//SHAFFAQ                                                                                             | Pakistan         | 0 |
| 96  | CB268 | INQ91/YR-31                                                                                                       | Pakistan         | 0 |
| 97  | CB269 | INQ91/YR-31                                                                                                       | Pakistan         | 0 |
| 98  | CB276 | V-04179/T7 (T.SPHAEROCOCCUM) – DROUGHT<br>WBLLI*2/VIVITSI/3/T.DICOCCOMP1                                          | Pakistan         | 0 |
| 99  | CB281 | 94624/AE.SQ(409)//BCN/4/WBLL1*2/VIVITSI/5/WBLLI<br>WBLLI*2/VIVITSI/3/T.DICOCCOMP1                                 | CIMMYT           | 0 |
| 100 | CB282 | 94624/AE.SQ(409)//BCN/4/WBLL1*2/VIVITSI/5/WBLLI<br>KAUZ//ALTAR84/AOS/3/MILAN/KAUZ/4/HUTES/5/T.SPELTAP1384764/6/2* | CIMMYT           | 0 |
| 101 | CB284 | KAUZ//ALTAR84/AOS                                                                                                 | CIMMYT           | 0 |
| 102 | CB288 | TOBA97/PASTOR*2//T.SPELTA P1348774                                                                                | CIMMYT           | 0 |

|     |       |                                                                                                        |          |   |
|-----|-------|--------------------------------------------------------------------------------------------------------|----------|---|
|     |       | T.SPELTA                                                                                               |          |   |
| 103 | CB290 | P1348764//INQ.91*2/TUKORU/3/WBL<br>L1*2/TUKURU                                                         | CIMMYT   | 0 |
| 104 | CB320 | V-11186                                                                                                | CIMMYT   | 0 |
| 105 | CB323 | MUNAL #1                                                                                               | CIMMYT   | 0 |
| 106 | CB326 | TACUPETO<br>F2001/BRAMBLING//KIRITATI                                                                  | CIMMYT   | 0 |
| 107 | CB328 | ATTILA/3*BCN//BAV92/3/TILHI/5/BA<br>V92/3/PRL/SARA//TSI/VEE#5/4/CROC<br>_1/AE.SQUARROSA (224)//2*OPATA | CIMMYT   | 0 |
| 108 | CB329 | ROLF07*2/KIRITATI                                                                                      | CIMMYT   | 0 |
| 109 | CB330 | FRET2/KUKUNA//FRET2/3/PARUS/5/<br>FRET2*2/4/SNI/TRAP#1/3/KAUZ*2/T<br>RAP//KAUZ                         | CIMMYT   | 0 |
| 110 | CB332 | PBW343*2/KUKUNA*2//YANAC                                                                               | CIMMYT   | 0 |
| 111 | CB334 | TRCH//PRINIA/PASTOR                                                                                    | CIMMYT   | 0 |
| 112 | CB336 | ACHTAR*3//KANZ/KS85-<br>85/4/MILAN/KAUZ//PRINIA/3/BAV92/<br>5/MILAN/KAUZ//PRINIA/3/BAV92               | CIMMYT   | 0 |
| 113 | CB339 | SOKOLL*2/TROST                                                                                         | CIMMYT   | 0 |
| 114 | CB341 | VILLA JUAREZ<br>F2009/SOLALA//WBLL1*2/BRAMBLI<br>NG                                                    | CIMMYT   | 0 |
| 115 | CB350 | NR 388                                                                                                 | CIMMYT   | 0 |
| 116 | CB351 | NR-371                                                                                                 | Pakistan | 0 |
| 117 | CB354 | NR-379                                                                                                 | Pakistan | 0 |
| 118 | CB357 | NR-390                                                                                                 | Pakistan | 0 |
| 119 | CB362 | NR-403                                                                                                 | Pakistan | 0 |
| 120 | CB369 | 76377                                                                                                  | Pakistan | 0 |
| 121 | CB372 | 99108                                                                                                  | Pakistan | 0 |
| 122 | CB373 | V-11183 –RF                                                                                            | Pakistan | 0 |
| 123 | CB375 | V-11365                                                                                                | Pakistan | 0 |
| 124 | CB378 | 11B2049                                                                                                | Pakistan | 0 |
| 125 | CB381 | 11BT004                                                                                                | Pakistan | 0 |
| 126 | CB383 | V-11143                                                                                                | Pakistan | 0 |
| 127 | CB385 | V-12284                                                                                                | Pakistan | 0 |
| 128 | CB387 | NW-10-1111-7                                                                                           | Pakistan | 0 |
| 129 | CB388 | V-11046                                                                                                | Pakistan | 0 |
| 130 | CB390 | NS-10                                                                                                  | Pakistan | 0 |
| 131 | CB392 | NR 411                                                                                                 | Pakistan | 0 |
| 132 | CB396 | V-11001                                                                                                | Pakistan | 0 |
| 133 | CB399 | 12257                                                                                                  | Pakistan | 0 |
| 134 | CB400 | 12292                                                                                                  | Pakistan | 0 |

|     |       |                                                          |               |   |
|-----|-------|----------------------------------------------------------|---------------|---|
| 135 | CB403 | D67.2/PARANA 66.270//AE.SQ<br>(320)/3/CUNNINGHAM/4/VORB  | Pakistan      | 0 |
| 136 | CB411 | ATTILA*2/PBW65*2//HAWFINCH #1                            | Pakistan      | 0 |
| 137 | CB413 | PFAU/SERI.1B//AMAD/3/WAXWING*<br>2/4/MUU                 | Pakistan      | 0 |
| 138 | CB420 | V-12253                                                  | Pakistan      | 0 |
| 139 | CB422 | V-11160                                                  | Pakistan      | 0 |
| 140 | CB423 | 12266                                                    | Pakistan      | 0 |
| 141 | CB425 | 11138                                                    | Pakistan      | 0 |
| 142 | CB427 | V-13005                                                  | Pakistan      | 0 |
| 143 | CB429 | V-13016                                                  | Pakistan      | 0 |
| 144 | CB433 | V-12130                                                  | Pakistan      | 0 |
| 145 | CB434 | V-12057                                                  | Pakistan      | 0 |
| 146 | CB435 | V-13241                                                  | Pakistan      | 0 |
| 147 | CB438 | V-13255                                                  | Pakistan      | 0 |
| 148 | CB440 | V-12066                                                  | Pakistan      | 0 |
| 149 | CB442 | V-13266                                                  | Pakistan      | 0 |
| 150 | CB443 | V-13270                                                  | Pakistan      | 0 |
| 151 | CB450 | 122557                                                   | Pakistan      | 0 |
| 152 | CB452 | 112095                                                   | Pakistan      | 0 |
| 153 | CB454 | 12BT012                                                  | Pakistan      | 0 |
| 154 | CB456 | 12C027                                                   | Pakistan      | 0 |
| 155 | CB462 | TW /424                                                  | Pakistan      | 0 |
| 156 | CB463 | TWS12268                                                 | Pakistan      | 0 |
| 157 | CB464 | MSW                                                      | Pakistan      | 0 |
| 158 | CB465 | NR-429                                                   | Pakistan      | 0 |
| 159 | CB466 | NR-449                                                   | Pakistan      | 0 |
| 160 | CB470 | TRITICUMPYRUM (V-2)                                      | Wild relative | 0 |
| 161 | CB471 | TRITICUMPYRUM (V-3)                                      | Wild relative | 0 |
| 162 | CB477 | F6 3013 (BWP)                                            | Pakistan      | 0 |
| 163 | CB484 | 088200 (MONO TILLER EARLY<br>MATURITY WITH LESS LODGING) | Pakistan      | 0 |
| 164 | CB488 | YOUNIS                                                   | Pakistan      | 0 |
| 165 | CB491 | 13B-3146                                                 | Pakistan      | 0 |
| 166 | CB493 | TWS-12464                                                | Pakistan      | 0 |
| 167 | CB497 | NR-443                                                   | Pakistan      | 0 |
| 168 | CB500 | NR-487                                                   | Pakistan      | 0 |
| 169 | CB505 | 14C036                                                   | Pakistan      | 0 |
| 170 | CB508 | NIBGE GANDUM N                                           | Pakistan      | 0 |
| 171 | CB512 | 13BT016                                                  | Pakistan      | 0 |
| 172 | CB516 | 14152                                                    | Pakistan      | 0 |
| 173 | CB520 | 14170                                                    | Pakistan      | 0 |

|     |       |            |          |   |
|-----|-------|------------|----------|---|
| 174 | CB521 | 14168      | Pakistan | 0 |
| 175 | CB522 | 13167      | Pakistan | 0 |
| 176 | CB523 | 14227      | Pakistan | 0 |
| 177 | CB524 | 13338      | Pakistan | 0 |
| 178 | CB2   | AUQAB-2000 | Pakistan | 0 |
| 179 | CB6   | CHAKWAL-97 | Pakistan | 0 |
| 180 | CB19  | UFAQ       | Pakistan | 0 |
| 181 | CB5   | CHAKWAL-86 | Pakistan | 0 |
| 182 | CB59  | PUNJAB-85  | Pakistan | 0 |
| 183 | CB12  | KOHISAR-95 | Pakistan | 0 |

**Table B.2 Comparison of the Kompetitive allele specific PCR marker alleles tightly linked to the gene/quantitative trait loci on chromosome 6B amplified in their different donors and susceptible controls.**

|                                        | <i>QHf.hwwg-6BS.2</i>  |                        | <i>H5 &amp; H34</i>    | <i>QHf.hwwg-6BS</i>      | <i>Hkelse</i>              |                             |                             |
|----------------------------------------|------------------------|------------------------|------------------------|--------------------------|----------------------------|-----------------------------|-----------------------------|
|                                        | <i>K6B_6</i><br>936780 | <i>K6B_7</i><br>697506 | <i>G00194</i><br>00_S4 | <i>KASP_6B</i><br>112698 | <i>kelse6BS</i><br>_167037 | <i>kelse6BS_</i><br>4554201 | <i>kelse6BS_</i><br>6196634 |
| Ning7840                               | UNA                    | FAM                    | FAM                    | FAM                      | FAM                        | UNA                         | FAM                         |
| Cadenza                                | FAM                    | FAM                    | UNA                    | FAM                      | FAM                        | FAM                         | FAM                         |
| AHT <sup>#</sup>                       | HET                    | HET                    | HEX                    | HEX                      | FAM                        | FAM                         | FAM                         |
| PARVAZ-94<br>( <i>QHf.hwwg-6BS.2</i> ) | HEX                    | HEX                    | HEX                    | HEX                      | FAM                        | FAM                         | FAM                         |
| Clark ( <i>H5 &amp; H34</i> )          | HEX                    | HEX                    | HEX                    | HEX                      | FAM                        | FAM                         | FAM                         |
| Chokwang<br>( <i>QHf.hwwg-6BS</i> )    | HEX                    | HEX                    | HEX                    | HEX                      | FAM                        | FAM                         | FAM                         |
| Kelse ( <i>Hkelse</i> )                | UNA                    | FAM                    | UNA                    | UNA                      | HEX                        | HEX                         | HEX                         |
| Seahawk ( <i>Hkelse</i> )              | UNA                    | FAM                    | UNA                    | UNA                      | HEX                        | HEX                         | HEX                         |

\* FAM, accession shows the FAM dye; HET, accession shows both FAM and HEX dyes; HEX, accession shows the HEX dye; UNA, accession has unknown allele

<sup>#</sup> AHT, artificial heterozygote by mixing DNA of Parvaz-94 and Cadenza

**Table B.3 Identification numbers, Hessian fly phenotypic scores of wheat accessions in the AM203 panel and their Kompetitive allele specific PCR marker alleles tightly linked to the gene/quantitative trait loci on chromosome arms 1AS and 6BS**

| ID    | Name            | Hessian fly resistance (%) | <i>QHf.hwwg</i> -1AS.2 | <i>QHf.hwwg</i> -6BS.2 |             | <i>H5 &amp; H34</i> | <i>QHf.hwwg</i> -6BS |
|-------|-----------------|----------------------------|------------------------|------------------------|-------------|---------------------|----------------------|
|       |                 |                            | 1A_798577 5*           | K6B-6936780            | K6B-7697506 | G0019 400_S4        | KASP-6B112698        |
| AM120 | P03207A1-7      | 100.0                      | R                      | R                      | R           | R                   | R                    |
| AM132 | INW0411         | 100.0                      | R                      | R                      | R           | R                   | R                    |
| AM150 | W98007V1        | 100.0                      | R                      | R                      | R           | R                   | R                    |
| AM155 | W98008J1        | 98.3                       | R                      | R                      | R           | R                   | R                    |
| AM082 | TX04M410164     | 92.2                       | S                      | R                      | R           | R                   | R                    |
| AM149 | M03-3616-C      | 100.0                      | N                      | R                      | R           | R                   | R                    |
| AM200 | P04287A1_10     | 100.0                      | N                      | R                      | R           | N                   | R                    |
| AM005 | KS980554-12~9   | 9.1                        | N                      | R                      | S           | R                   | R                    |
| AM098 | KS020304K~3     | 86.7                       | N                      | R                      | S           | R                   | R                    |
| AM059 | CO03W239        | 53.3                       | R                      | R                      | S           | R                   | R                    |
| AM101 | TX06A001376     | 74.1                       | R                      | R                      | S           | R                   | R                    |
| AM038 | Antelope        | 0.0                        | S                      | R                      | S           | R                   | R                    |
| AM179 | G61505          | 0.0                        | R                      | H                      | S           | R                   | S                    |
| AM197 | MO040152        | 86.3                       | R                      | H                      | S           | N                   | S                    |
| AM029 | SD06173         | 100.0                      | R                      | S                      | N           | N                   | S                    |
| AM048 | CO03W139        | 0.0                        | R                      | S                      | N           | N                   | S                    |
| AM058 | KS970093-8-9-#1 | 57.4                       | R                      | S                      | N           | N                   | S                    |
| AM073 | SD05210         | 54.4                       | R                      | S                      | N           | N                   | S                    |
| AM079 | MT0552          | 0.0                        | R                      | S                      | N           | N                   | S                    |
| AM177 | AR97124-4-3     | 0.0                        | R                      | S                      | N           | N                   | S                    |
| AM188 | OK06848W        | 100.0                      | R                      | S                      | N           | N                   | S                    |
| AM008 | N98L20040-44    | 0.0                        | R                      | N                      | S           | S                   | S                    |
| AM020 | NX03Y2489       | 6.8                        | R                      | N                      | S           | S                   | S                    |
| AM027 | TX02A0252       | 0.0                        | R                      | N                      | S           | S                   | S                    |
| AM039 | SD03164-1       | 73.3                       | R                      | N                      | S           | N                   | S                    |
| AM043 | OK03716W        | 18.3                       | R                      | N                      | S           | N                   | S                    |
| AM054 | OK05903C        | 51.7                       | R                      | N                      | S           | N                   | S                    |
| AM072 | SD06W117        | 1.7                        | R                      | N                      | S           | S                   | S                    |
| AM119 | KY96C-0769-7-3  | 100.0                      | R                      | N                      | S           | N                   | S                    |
| AM126 | Roane           | 100.0                      | R                      | N                      | S           | N                   | S                    |
| AM139 | IL00-8530       | 98.3                       | R                      | N                      | S           | N                   | S                    |
| AM152 | Coker 9553      | 6.7                        | R                      | N                      | S           | N                   | S                    |
| AM154 | B030543         | 0.0                        | R                      | N                      | S           | N                   | S                    |

|       |               |       |   |   |   |   |   |
|-------|---------------|-------|---|---|---|---|---|
| AM156 | OK05122       | 1.7   | R | N | S | S | S |
| AM164 | OK06345       | 75.0  | R | N | S | N | S |
| AM165 | OK06319       | 36.7  | R | N | S | S | S |
| AM166 | D04*5513      | 0.0   | R | N | S | N | S |
| AM170 | D04-5012      | 100.0 | R | N | S | N | S |
| AM171 | G59160        | 56.7  | R | N | S | N | S |
| AM174 | OK06518       | 0.0   | R | N | S | N | S |
| AM184 | VA04W-259     | 6.7   | R | N | S | N | S |
| AM003 | KS05HW136-3   | 74.5  | R | S | S | N | S |
| AM006 | KS980512-2-2  | 1.9   | R | S | S | N | S |
| AM009 | NI04420       | 88.8  | R | S | S | N | S |
| AM010 | Duster        | 100.0 | R | S | S | N | S |
| AM012 | Scout 66      | 5.2   | R | S | S | N | S |
| AM015 | NE04424       | 91.7  | R | S | S | N | S |
| AM019 | SD06165       | 91.0  | R | S | S | N | S |
| AM021 | NI04427       | 32.0  | R | S | S | N | S |
| AM022 | Endurance     | 0.0   | R | S | S | N | S |
| AM025 | HV9W03-539R   | 70.6  | R | S | S | N | S |
| AM026 | CO03064       | 30.9  | R | S | S | N | S |
| AM030 | NX04Y2107     | 1.8   | R | S | S | N | S |
| AM033 | Trego         | 53.3  | R | S | S | N | S |
| AM034 | HV9W03-696R-1 | 100.0 | R | S | S | N | S |
| AM041 | NE05549       | 76.7  | R | S | S | N | S |
| AM047 | NE05430       | 8.8   | R | S | S | N | S |
| AM051 | NE02533       | 0.0   | R | S | S | N | S |
| AM052 | NE05569       | 0.0   | R | S | S | N | S |
| AM061 | Jerry         | 5.0   | R | S | S | N | S |
| AM063 | NE02558       | 90.0  | R | S | S | N | S |
| AM067 | KS05HW121-2   | 96.5  | R | S | S | N | S |
| AM069 | KS970187-1-10 | 83.3  | R | S | S | N | S |
| AM075 | MTS0531       | 0.0   | R | S | S | N | S |
| AM081 | NE05496       | 3.3   | R | S | S | N | S |
| AM084 | SD05W030      | 0.0   | R | S | S | N | S |
| AM088 | OK02405       | 0.0   | R | S | S | N | S |
| AM090 | NE06619       | 100.0 | R | S | S | N | S |
| AM091 | MTS04120      | 5.0   | R | S | S | N | S |
| AM095 | OH02-7217     | 100.0 | R | S | S | N | S |
| AM096 | MD99W483-06-9 | 35.0  | R | S | S | N | S |
| AM099 | KS010143K-11  | 0.0   | R | S | S | N | S |
| AM102 | VA03W-412     | 3.3   | R | S | S | N | S |
| AM105 | HV9W05-881R   | 46.7  | R | S | S | N | S |

|       |                                     |       |   |   |   |   |   |
|-------|-------------------------------------|-------|---|---|---|---|---|
| AM107 | NW05M6011-6-1                       | 0.0   | R | S | S | N | S |
| AM108 | TX06A001431                         | 100.0 | R | S | S | N | S |
| AM113 | OK05511                             | 53.3  | R | S | S | N | S |
| AM114 | SD07W041                            | 50.0  | R | S | S | N | S |
| AM116 | NW05M6015-25-4                      | 0.0   | R | S | S | N | S |
| AM117 | TXHT001F8-<br>CS06/325-<br>PRE07/75 | 6.5   | R | S | S | S | S |
| AM125 | NE06472                             | 5.0   | R | S | S | N | S |
| AM130 | KS010514-9TM-<br>10                 | 100.0 | R | S | S | N | S |
| AM133 | MO040192                            | 0.0   | R | S | S | N | S |
| AM135 | KS07HW81                            | 0.0   | R | S | S | N | S |
| AM137 | TX05V5614                           | 0.0   | R | S | S | S | S |
| AM140 | IL02-18228                          | 100.0 | R | S | S | N | S |
| AM141 | KS07HW117                           | 100.0 | R | S | S | N | S |
| AM142 | NE06549                             | 80.0  | R | S | S | N | S |
| AM146 | Mocha exp.                          | 93.3  | R | S | S | N | S |
| AM148 | NC04-15533                          | 0.0   | R | S | S | N | S |
| AM151 | Arena exp.                          | 90.0  | R | S | S | N | S |
| AM153 | VA05W-258                           | 41.7  | R | S | S | N | S |
| AM157 | OK06210                             | 94.8  | R | S | S | N | S |
| AM159 | G69202                              | 66.7  | R | S | S | N | S |
| AM160 | USG 3555                            | 5.4   | R | S | S | N | S |
| AM161 | LA01138D-52                         | 8.3   | R | S | S | N | S |
| AM163 | OK05723W                            | 0.0   | R | S | S | N | S |
| AM167 | M04-4566                            | 0.0   | R | S | S | N | S |
| AM168 | NC03-6228                           | 0.0   | R | S | S | N | S |
| AM169 | AR96077-7-2                         | 0.0   | R | S | S | N | S |
| AM173 | OK06528                             | 6.7   | R | S | S | N | S |
| AM175 | KY97C-0321-02-<br>01                | 0.0   | R | S | S | N | S |
| AM178 | GA991336-6E9                        | 58.2  | R | S | S | N | S |
| AM180 | OK05134                             | 62.2  | R | S | S | S | S |
| AM181 | OK06313                             | 40.0  | R | S | S | N | S |
| AM183 | M04*5109                            | 5.0   | R | S | S | N | S |
| AM186 | GA991209-6E33                       | 6.7   | R | S | S | N | S |
| AM187 | G41732                              | 1.7   | R | S | S | N | S |
| AM189 | W06-202B                            | 0.0   | R | S | S | N | S |
| AM191 | LA99005UC-31-3-<br>C                | 0.0   | R | S | S | N | S |
| AM192 | P03112A1-7-14                       | 43.3  | R | S | S | N | S |

|       |                             |       |   |   |   |   |   |
|-------|-----------------------------|-------|---|---|---|---|---|
| AM193 | TN801                       | 0.0   | R | S | S | N | S |
| AM194 | GA991371_6E13               | 65.0  | R | S | S | N | S |
| AM196 | OK06336                     | 0.0   | R | S | S | S | S |
| AM198 | AGS 2000                    | 0.0   | R | S | S | N | S |
| AM199 | LA98214D_14_1_2_B           | 0.0   | R | S | S | N | S |
| AM201 | M04_4715                    | 100.0 | R | S | S | N | S |
| AM202 | GA991227_6A33               | 11.8  | R | S | S | N | N |
| AM109 | TXHT023F7-CS06/607-STA07/40 | 0.0   | H | S | S | N | S |
| AM032 | Deliver                     | 0.0   | S | S | N | N | S |
| AM070 | CO03W043                    | 0.0   | S | S | N | N | S |
| AM074 | NW03666                     | 0.0   | S | S | N | N | S |
| AM089 | KS010957K~4                 | 15.0  | S | S | N | N | S |
| AM110 | AR97044-10-2                | 0.0   | S | S | N | N | N |
| AM195 | OK05212                     | 0.0   | S | S | N | N | S |
| AM007 | TX04M410211                 | 0.0   | S | S | S | N | S |
| AM011 | OK02522W                    | 0.0   | S | S | S | N | S |
| AM013 | AP04T8211                   | 29.0  | S | S | S | N | S |
| AM016 | CO02W237                    | 1.8   | S | S | S | N | S |
| AM018 | TX04V075080                 | 0.0   | S | S | S | N | S |
| AM028 | Kharkof                     | 0.0   | S | S | S | N | S |
| AM031 | NE05548                     | 91.7  | S | S | S | N | S |
| AM036 | CO03W054                    | 0.0   | S | S | S | N | S |
| AM037 | TX03A0148                   | 0.0   | S | S | S | N | S |
| AM042 | OK Bullet                   | 0.0   | S | S | S | N | S |
| AM044 | OK00514-05806               | 0.0   | S | S | S | N | S |
| AM045 | AP06T3832                   | 3.5   | S | S | S | N | S |
| AM046 | HV9W02-942R                 | 6.7   | S | S | S | N | S |
| AM049 | TX03A0563                   | 0.0   | S | S | S | N | S |
| AM050 | Wesley                      | 0.0   | S | S | S | N | S |
| AM053 | Overley                     | 0.0   | S | S | S | N | S |
| AM062 | SD05118                     | 88.2  | S | S | S | N | S |
| AM064 | MT0495                      | 1.7   | S | S | S | N | S |
| AM065 | Fuller                      | 3.3   | S | S | S | N | S |
| AM066 | OK03522                     | 0.0   | S | S | S | N | S |
| AM078 | OK03305                     | 0.0   | S | S | S | N | S |
| AM083 | SD06069                     | 0.0   | S | S | S | N | S |
| AM087 | OK05830                     | 0.0   | S | S | S | N | S |
| AM092 | TX06A001239                 | 0.0   | S | S | S | N | S |

|       |                          |       |   |   |   |   |   |
|-------|--------------------------|-------|---|---|---|---|---|
| AM093 | TXHT006F8-CS06/472-STA34 | 21.7  | S | S | S | N | S |
| AM094 | MO011126                 | 8.3   | S | S | S | N | S |
| AM106 | NE06436                  | 86.7  | S | S | S | N | S |
| AM115 | SD07204                  | 100.0 | S | S | S | N | S |
| AM118 | CO04W210                 | 16.7  | S | S | S | N | S |
| AM122 | KS07HW25                 | 0.0   | S | S | S | N | S |
| AM138 | Branson                  | 100.0 | S | S | S | N | S |
| AM145 | IL02-19463               | 100.0 | S | S | S | N | S |
| AM158 | India exp.               | 0.0   | S | S | S | N | S |
| AM172 | OK01420W                 | 0.0   | S | S | S | N | S |
| AM176 | M04-4802                 | 0.0   | S | S | S | N | S |
| AM182 | KY97C-0519-04-07         | 0.0   | S | S | S | N | S |
| AM190 | TAM110                   | 0.0   | S | S | S | N | S |
| AM001 | Atlas66                  | 0.0   | S | N | S | S | S |
| AM086 | Guymon                   | 1.7   | S | N | S | S | S |
| AM097 | OK04507                  | 0.0   | S | N | S | S | S |
| AM100 | TX05A001334              | 0.0   | S | N | S | S | S |
| AM144 | Bess                     | 0.0   | S | N | S | N | S |
| AM002 | OK04505                  | 0.0   | N | N | S | S | S |
| AM014 | HV9W96-1271R-1           | 0.0   | N | N | S | S | S |
| AM055 | Century                  | 3.3   | N | N | S | S | S |
| AM077 | OK04525                  | 0.0   | N | N | S | N | S |
| AM085 | Chisholm                 | 80.0  | N | N | S | S | S |
| AM124 | KS010379M-2              | 0.0   | N | N | S | N | S |
| AM127 | OH02-12678               | 0.0   | N | N | S | N | S |
| AM004 | T158                     | 1.3   | N | S | S | N | S |
| AM017 | OK03825-5403-6           | 20.8  | N | S | S | N | S |
| AM023 | TAM-107                  | 0.0   | N | S | S | N | S |
| AM035 | NE05426                  | 3.5   | N | S | S | N | S |
| AM040 | NW04Y2188                | 96.7  | N | S | S | N | S |
| AM056 | KS05HW15-2               | 5.0   | N | S | S | N | S |
| AM057 | T151                     | 0.0   | N | S | S | N | S |
| AM068 | T153                     | 6.7   | N | S | S | N | S |
| AM071 | TX01V5134RC-3            | 0.0   | N | S | S | N | S |
| AM076 | Centerfield              | 0.0   | N | S | S | N | S |
| AM080 | T154                     | 0.0   | N | S | S | N | S |
| AM104 | OK05312                  | 0.0   | N | S | S | N | S |
| AM111 | P02444A1-23-9            | 0.0   | N | S | S | N | S |
| AM112 | VA05W-414                | 0.0   | N | S | S | N | S |

|       |                        |      |   |   |   |   |   |
|-------|------------------------|------|---|---|---|---|---|
| AM121 | LA01*425               | 0.0  | N | S | S | N | S |
| AM123 | SD07220                | 1.7  | N | S | S | N | S |
| AM131 | N02Y5117               | 58.3 | N | S | S | N | S |
| AM134 | NYCaIR-L               | 30.0 | N | S | S | N | S |
| AM136 | U07-698-9              | 0.0  | N | S | S | N | S |
| AM143 | TX06A001084            | 0.0  | N | S | S | N | S |
| AM147 | Pioneer Brand<br>26R61 | 55.9 | N | S | S | N | S |
| AM185 | MD01W233-06-1          | 0.0  | N | S | S | N | S |
| AM203 | OK05128                | 5.0  | N | S | S | N | S |
| AM103 | OH03-41-45             | 93.2 | N | N | N | N | S |
| AM128 | LA02-923               | 21.7 | N | N | N | N | S |
| AM162 | VA05W-78               | 0.0  | N | N | N | N | S |
| AM024 | AP05T2413              | 2.1  | N | S | N | N | S |
| AM060 | TX04A001246            | 0.0  | N | S | N | S | S |
| AM129 | SD05W148-1             | 71.7 | N | S | N | N | S |

\*R, accession has the resistant allele; S, accession has the susceptible allele; H, accession has the heterozygous allele; N, no genotype information.

## Appendix C - Supplementary data of Chapter 3

**Table C.1 List of primers used in chapter 3**

| Marker             | Sequence 5' – 3'            | Notes       |
|--------------------|-----------------------------|-------------|
| KASP-3B2479314F/H  | AGTGGGTCGTTACCTCT/G         | KASP marker |
| KASP-3B2479314C    | CATATTGCCAAAAGAAAAA         | KASP marker |
| KASP-3B4799538F/H  | CACAACCAATCACTACCGC/T       | KASP marker |
| KASP-3B4799538C    | TCGGTACAGTCAGAGCAAGA        | KASP marker |
| KASP-7A47772047F/H | CTTCTGCAGAGTGTCTTCATGTTT/T  | KASP marker |
| KASP-7A47772047C   | TAATGGAATTGGCCGGTAAG        | KASP marker |
| KASP-7A69907161F/H | CTTCGATTCTTGCAGTC/T         | KASP marker |
| KASP-7A69907161C   | TCTAGTAGCCAGGCCAGGTG        | KASP marker |
| K3289480F/H        | AAGGATAATGGCCGGAACCG/T      | KASP marker |
| K3289480C          | GCGGCATTTTCAGCACTTTTG       | KASP marker |
| K3351845F/H        | CCAGAGAAGCAGCGCCTG/A        | KASP marker |
| K3351845C          | CCATGGAGGAAGGTGGAGA         | KASP marker |
| K3492790F/H        | CCGCTATGTTCCGATCCCTTC/A     | KASP marker |
| K3492790C          | ATACAATGTCTGCAAGGCGC        | KASP marker |
| K3522143F/H        | CGCATAGATAGATTTGATCCGGC/T   | KASP marker |
| K3522143C          | GCGACACACATCTGCTCAAT        | KASP marker |
| K3528412F/H        | CACCTTGACAATAAAACAAAATGT/C  | KASP marker |
| K3528412C          | ACATGGTGATTCTCTTTTCCAATA    | KASP marker |
| K3531272F/H        | GGATAAGGGGAAGCAATGGAATG/C   | KASP marker |
| K3531272C          | AACCTCCACGGCAACATCTC        | KASP marker |
| K3787927F/H        | CCAGAAAGAGATGTAGCCTATGA/G   | KASP marker |
| K3787927C          | GGAGAGCGCACATACACTT         | KASP marker |
| K4088867F/H        | TGGGTATTGCATCTCGGGC         | KASP marker |
| K4088867C          | ACATACCATGTAGATGATAGTGCAA   | KASP marker |
| K4898344F/H        | GAGATCGTCCTTGATGGCAAT/C     | KASP marker |
| K4898344C          | GATCAAGGCAGCAGAGCGA         | KASP marker |
| K4997741F/H        | GGTCGATCCATCTCCGTGTTAATAA/T | KASP marker |
| K4997741C          | CGACCCGATGTACGACATGA        | KASP marker |
| K5558142F/H        | TCTCTATCTTTGTCTTCCTTAGGTC/A | KASP marker |
| K5558142C          | GCCTTGCCGCAAGTTTCT          | KASP marker |
| K6806633F/H        | TCTGCTTGGCGTATCCGA/G        | KASP marker |
| K6806633C          | CCCTTAACAACAATGTACAACACTT   | KASP marker |
| KR300-4F/H         | AGCGGGAAGATGGGCAAAAC/A      | KASP marker |
| KR300-4C           | CATGTTCTGGCATGTCTGTA        | KASP marker |
| KR300-26F/H        | CACGTACAGCTGGAGCAT/C        | KASP marker |
| KR300-26C          | CCCTCTGCCAGCTGGTCA          | KASP marker |

|             |                              |                       |
|-------------|------------------------------|-----------------------|
| KR300-32F/H | GTAGAACTCCTTTTCAAAATGCCC/A   | KASP marker           |
| KR300-32C   | AGCGGGAAGATGGGCAAA           | KASP marker           |
| KR300-40F/H | GACAAGGTGTTTCGCCTACG/A       | KASP marker           |
| KR300-40C   | GGCAGAGGGAGTTGACGTA          | KASP marker           |
| KR300-41F/H | TCAGAAGGCGTACCTGCA/CTCAGAAGG | KASP marker           |
| KR300-41C   | CGTACATGCT                   | KASP marker           |
| KB120-1F/H  | GTTCTCCCTCTGCTCGAAC          | KASP marker           |
| KB120-1C    | CAATGACAAGGGAGAGGAAATTGG/A   | KASP marker           |
| KB120-2F/H  | GAAATGTCTTAGGAACAGTGCATC     | KASP marker           |
| KB120-2C    | ACCGAGCATAGTTTCCCTCC/T       | KASP marker           |
| KB120-5F/H  | TCTCCAGATTGATCGTATTGCAC      | KASP marker           |
| KB120-5C    | TGTAACCGCAGTTTCAATTCTCT/C    | KASP marker           |
| KB120-6F/H  | ATGGCCTCGGTGTTGAAACT         | KASP marker           |
| KB120-6C    | CAAGCTAAGCCGAACCCTA/G        | KASP marker           |
| R300F1      | TCGCCCCGTCTAGCTACCACT        | KASP marker           |
| R300R1      | GCAATAACATGGCTGATGGGC        | Amplification of R300 |
| B120ATG     | ACTAGTCGGAGCCAAACATAGTC      | Amplification of R300 |
| B120TAA     | ATGACTCAGCTCCAAGCTGC         | Amplification of B120 |
|             | TTACTTCTGCGACTCCAGAGC        | Amplification of B120 |

F: FAM (GAAGGTGACCAAGTTCATGCT), H: HEX (GAAGGTGCGAGTCAACGGATT)

**Table C.2 Details for each of the 12 PacBio SMRT cells**

| Cell    | total bases | number of reads | coverage<br>(16 Gb) | average read length<br>(bp) | N50 (bp) |
|---------|-------------|-----------------|---------------------|-----------------------------|----------|
| M220511 | 20860370981 | 1445427         | 1.30                | 14431                       | 16822    |
| M220519 | 24152556755 | 1616795         | 1.51                | 14938                       | 17128    |
| M220521 | 24417930857 | 1690085         | 1.53                | 14447                       | 16407    |
| M220529 | 21134452496 | 1482737         | 1.32                | 14253                       | 14951    |
| M220607 | 28133699034 | 1919956         | 1.76                | 14653                       | 15325    |
| M220608 | 25650073943 | 2106306         | 1.60                | 12177                       | 12922    |
| M220623 | 14353739324 | 1116102         | 0.90                | 12860                       | 13694    |
| M220625 | 14953446754 | 1174423         | 0.93                | 12732                       | 13541    |
| M220705 | 21048855230 | 1556260         | 1.32                | 13525                       | 14232    |
| M220706 | 22663797582 | 1688522         | 1.42                | 13422                       | 14127    |
| M220709 | 17355261903 | 1310605         | 1.08                | 13242                       | 14127    |
| M220707 | 19356517610 | 1547686         | 1.21                | 12506                       | 13163    |