

Map-based cloning of the Hessian fly resistance gene *H13* in Wheat

by

Anupama Joshi

B.S., Panjab University, 2006

M.S., Panjab University, 2008

AN ABSTRACT OF A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Interdepartmental Genetics
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Abstract

H13, a dominant resistance gene transferred from *Aegilops tauschii* into wheat (*Triticum aestivum*), confers a high level of antibiosis against a wide range of Hessian fly (HF, *Mayetiola destructor*) biotypes. Previously, *H13* was mapped to the distal arm of chromosome 6DS, where it is flanked by markers *Xcfd132* and *Xgdm36*. A mapping population of 1,368 F₂ individuals derived from the cross: PI372129 (*h13h13*) / PI562619 (Molly, *H13H13*) was genotyped and *H13* was flanked by *Xcfd132* at 0.4cM and by *Xgdm36* at 1.8cM. Screening of BAC-based physical maps of chromosome 6D of Chinese Spring wheat and *Ae. tauschii* coupled with high resolution genetic and Radiation Hybrid mapping identified nine candidate genes co-segregating with *H13*. Candidate gene validation was done on an EMS-mutagenized TILLING population of 2,296 M₃ lines in Molly. Twenty seeds per line were screened for susceptibility to the *H13*-virulent HF GP biotype. Sequencing of candidate genes from twenty-eight independent susceptible mutants identified three nonsense, and 24 missense mutants for *CNL-1* whereas only silent and intronic mutations were found in other candidate genes. 5' and 3' RACE was performed to identify gene structure and CDS of *CNL-1* from Molly (*H13H13*) and Newton (*h13h13*). Increased transcript levels were observed for *H13* gene during incompatible interactions at larval feeding stages of GP biotype. The predicted coding sequence of *H13* gene is 3,192 bp consisting of two exons with 618 bp 5'UTR and 2,260 bp 3'UTR. It translates into a protein of 1063 amino acids with an N-terminal Coiled-Coil (CC), a central Nucleotide-Binding adapter shared by APAF-1, plant R and CED-4 (NB-ARC) and a C-terminal Leucine-Rich Repet (LRR) domain. Conserved domain analysis revealed shared domains in Molly and Newton, except for differences in sequence, organization and number of LRR repeat in Newton. Also, the presence of a transposable element towards the C terminal of *h13* was indicative of

interallelic recombination, recent tandem duplications and gene conversions in the *CNL* rich region near *H13* locus. Comparative analysis of candidate genes in the *H13* region indicated that gene duplications in *CNL* encoding genes during divergence of wheat and barley led to clustering and diversity. This diversity among *CNL* genes may have a role in defining differences in the recognition specificities of NB-LRR encoding genes. Allele mining for the *H13* gene in the core collection of *Ae. tauschii* and hexaploid wheat cultivars identified different functional haplotypes. Screening of these haplotypes using different HF biotypes would help in the identification of the new sources of resistance to control evolving biotypes of HF. Cloning of *H13* will provide perfect markers to breeders for HF resistance breeding programs. It will also provide an opportunity to study R-*Avr* interactions in the hitherto unexplored field of insect-host interaction.

Map-based cloning of the Hessian fly resistance gene *H13* in Wheat

by

Anupama Joshi

B.S., Panjab University, 2006

M.S., Panjab University, 2008

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Interdepartmental Genetics
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Approved by:

Major Professor
Dr. Bikram S. Gill

Copyright

© Anupama Joshi 2018.

Abstract

H13, a dominant resistance gene transferred from *Aegilops tauschii* into wheat (*Triticum aestivum*), confers a high level of antibiosis against a wide range of Hessian fly (HF, *Mayetiola destructor*) biotypes. Previously, *H13* was mapped to the distal arm of chromosome 6DS, where it is flanked by markers *Xcfd132* and *Xgdm36*. A mapping population of 1,368 F₂ individuals derived from the cross: PI372129 (*h13h13*) / PI562619 (Molly, *H13H13*) was genotyped and *H13* was flanked by *Xcfd132* at 0.4cM and by *Xgdm36* at 1.8cM. Screening of BAC-based physical maps of chromosome 6D of Chinese Spring wheat and *Ae. tauschii* coupled with high resolution genetic and Radiation Hybrid mapping identified nine candidate genes co-segregating with *H13*. Candidate gene validation was done on an EMS-mutagenized TILLING population of 2,296 M₃ lines in Molly. Twenty seeds per line were screened for susceptibility to the *H13*-virulent HF GP biotype. Sequencing of candidate genes from twenty-eight independent susceptible mutants identified three nonsense, and 24 missense mutants for *CNL-1* whereas only silent and intronic mutations were found in other candidate genes. 5' and 3' RACE was performed to identify gene structure and CDS of *CNL-1* from Molly (*H13H13*) and Newton (*h13h13*). Increased transcript levels were observed for *H13* gene during incompatible interactions at larval feeding stages of GP biotype. The predicted coding sequence of *H13* gene is 3,192 bp consisting of two exons with 618 bp 5'UTR and 2,260 bp 3'UTR. It translates into a protein of 1063 amino acids with an N-terminal Coiled-Coil (CC), a central Nucleotide-Binding adapter shared by APAF-1, plant R and CED-4 (NB-ARC) and a C-terminal Leucine-Rich Repet (LRR) domain. Conserved domain analysis revealed shared domains in Molly and Newton, except for differences in sequence, organization and number of LRR repeat in Newton. Also, the presence of a transposable element towards the C terminal of *h13* was indicative of

interallelic recombination, recent tandem duplications and gene conversions in the *CNL* rich region near *H13* locus. Comparative analysis of candidate genes in the *H13* region indicates that gene duplications in *CNL* encoding genes during divergence of wheat and barley led to clustering and diversity. This diversity among *CNL* genes may have a role in defining differences in the recognition specificities of NB-LRR encoding genes. Allele mining for the *H13* gene in the core collection of *Ae. tauschii* and hexaploid wheat cultivars identified different functional haplotypes. Screening of these haplotypes using different HF biotypes would help in the identification of the new sources of resistance to control evolving biotypes of HF. Cloning of *H13* will provide perfect markers to breeders for HF resistance breeding programs. It will also provide an opportunity to study R-*Avr* interactions in the hitherto unexplored field of insect-host interaction.

Table of Contents

List of Figures	xi
List of Tables	xii
Acknowledgments.....	xiii
Chapter 1 - Review of Literature	1
1.1 Importance of Wheat	1
1.2 Origin and evolution of Wheat	2
1.3 Need for Wheat improvement.....	3
1.4 Introgression of traits from <i>Ae. tauschii</i> for wheat improvement.....	4
1.4.1 Generation of Synthetic hexaploid.....	5
1.4.2 Direct hybridization	6
1.5 Mapping of genes in wheat.....	7
1.5.1 Genetic mapping	8
1.5.2 Physical Mapping.....	9
1.6 Gene Cloning	10
1.7 Plant pest interactions	12
1.7.1 Host plant resistance (HPR).....	13
1.7.2 Role of HPR in pest management.....	13
1.7.3 Influence of biotic and abiotic factors on induced resistance	14
1.8 Mechanism of plant immunity against pests.....	15
1.8.1 The gene-for-gene model	15
1.9 Pathways involved in plant-pathogen interactions	17
1.10 Structure of R genes.....	17
1.10.1 R genes with extracellular domains	18
1.10.2 R genes with NB-LRR domains.....	18
1.10.3 Functions of different domains in CC-NB-LRR proteins.....	20
1.10.4 Diversification of R genes.....	21
1.11 Hessian fly: A destructive pest of wheat	22
1.11.1 History and economic impact	22
1.11.2 Biology and Host Range	23

1.11.3 Biotypes of Hessian fly	23
1.11.4 Virulence genes encoding putative HF effectors	24
1.11.5 The phenotypic response in wheat upon HF attack:	25
1.11.5.1 Wheat response to an attack by virulent larvae: The compatible interaction	25
1.11.5.2 Wheat response to an attack by avirulent larvae: The incompatible interaction	26
1.11.6 HF management	27
1.11.6.1 Cultural control	27
1.11.6.2 Biological control.....	27
1.11.6.3 Chemical control.....	28
1.11.6.4 R gene control	28
1.12 Introduction to rice and rice pests the brown plant hopper (BPH) and gall midges	29
1.12.1 Brown planthopper (BPH)	29
1.12.1.1 Plant-insect interactions in BPH	30
1.12.1.2 Identification and mapping of Bph resistant genes.....	31
1.12.1.3 Map-based cloning and characterization of Bph resistant genes	31
1.12.2 Asian rice gall midge:	32
1.13 References.....	34
Chapter 2 - Map-based cloning of the Hessian fly resistance gene H13 in Wheat.....	62
2.1 Abstract.....	62
2.2 Introduction.....	64
2.2.1 Objective of study	69
2.3 Material and Methods	69
2.3.1 Plant materials and mapping populations	69
2.3.2 Linked SSR marker analysis	69
2.3.3 Hessian fly infestation screening	70
2.3.4 PCR-based screening of 3-D BAC pools of chromosomes 1D, 4D, 6D.....	71
2.3.5 Anchoring of Xgdm36 on the genetic and physical map of Ae. tauschii	72
2.3.6 Designing D Genome-specific primers for markers	72
2.3.7 Polymorphism identification and Mapping of the hexaploid mapping population	73
2.3.8 Mapping on radiation hybrid (RH) panels	74
2.3.9 Fine mapping of the H13 interval using markers designed from scaffold sequence ...	74

2.3.10 Ethyl methanesulphonate (EMS) mutagenesis and population development	75
2.3.11 Tissue collection and 4X DNA pools	76
2.3.12 Hessian fly screening of EMS population	76
2.3.13 Characterization of EMS mutants	77
2.3.14 Candidate gene validation	77
2.3.15 5' and 3' Rapid amplification of cDNA ends (RACE)	78
2.3.16 Assembly and cloning of the full-length gene	79
2.3.17 Haplotype analysis and allele mining	80
2.3.18 Expression analysis of H13 gene	80
2.3.19 Comparative Mapping of candidate genes	82
2.3.20 Phylogenetic analyses	82
2.4 Results	83
2.4.1 Development of a high-resolution mapping population:	83
2.4.2 Development of a high-resolution map and identification of co-segregating markers:	83
2.4.2.1 Anchoring of SSR markers on BAC contigs of chromosome 6D	83
2.4.2.2 Development of D-genome specific markers:	84
2.4.2.3 Genetic mapping of <i>Ae. tauschii</i> markers on the high-resolution mapping population	84
2.4.3 Radiation hybrid (RH) mapping	84
2.4.4 Marker enrichment for H13 region	85
2.4.5 Fine mapping of H13 region	85
2.4.6 Functional validation of candidate genes by TILLING	86
2.4.7 Identification of gene structure for H13	87
2.4.8 Phylogenetic and comparative analysis	88
2.4.9 Expression analysis	89
2.4.10 Haplotype analysis and allele mining	89
2.5 Discussion	90
2.6 Figures	98
2.7 Tables	127
2.8 References	178

List of Figures

Figure 2.1 Chromosomal location of different HFR genes on the D genome of wheat.	98
Figure 2.2 Linked SSR marker analysis and HF infestation.....	99
Figure 2.3 Anchoring of SSR markers on Chinese Spring (CS) BAC pools.....	100
Figure 2.4 Anchoring of SSR markers on Ae. tauschii map.....	101
Figure 2.5 Development and validation of D genome specific markers.....	102
Figure 2.6 Identification of polymorphic markers for genetic mapping.....	103
Figure 2.7 Genetic mapping of Ae. tauschii markers on high resolution mapping population ..	104
Figure 2.8 Radiation hybrid (RH) mapping.....	105
Figure 2.9 Comparative RH and genetic linkage map of H13 region	106
Figure 2.10 Fine mapping and identification of candidate genes co-segregating with H13.....	107
Figure 2.11 Developing TILLING population and HF screening of M ₃ population	108
Figure 2.12 Characterization of susceptible lines from M ₃ population.....	109
Figure 2.13 Validation of contaminant lines by sequencing	110
Figure 2.14 Summary of mutations identified in susceptible lines.....	112
Figure 2.15 Primer locations and vector used for 5'-RACE and 3'-RACE.....	113
Figure 2.16 Colony screening for positive transformants from 5'-RACE and 3-RACE for CNL-1	114
Figure 2.17 Colony screening for positive transformants for full length gDNA and cDNA for CNL-1	115
Figure 2.18 Genetic and physical map of H13 region	116
Figure 2.19 Comparative analysis of H13 region	117
Figure 2.20 Gene structure and protein-domain prediction of CNL-1	118
Figure 2.21 Sequence of H13 gene and its deduced amino acid sequence.....	120
Figure 2.22 Alignment of amino acid sequences from Molly and Newton.....	122
Figure 2.23 Expression profile of H13 gene.....	123
Figure 2.24 Neighbor-joining tree between CNL-1 and other candidate CNLs.....	124
Figure 2.25 Neighbor-joining tree showing relationships between H13 and other identified CC- NB-LRR proteins	125
Figure 2.26 Neighbor-joining tree showing relationships between Hap1 and other haplotypes	126

List of Tables

Table 2.1 HFR genes with their source and chromosome locations.....	127
Table 2.2 Plant materials used in this study.....	129
Table 2.3. Haplotype of H13 region in high resolution mapping population	136
Table 2.4 Comparative analysis of candidate genes	137
Table 2.5 Phenotypic data of HS infestation on EMS M ₃ lines.....	137
Table 2.6 Details of mutations found in candidate genes	141
Table 2.7 Full length haplotypes of H13 gene.....	143
Table 2.8 Sites of variation in the haplotypes of H13 gene	143
Table 2.9 List of accessions belonging to each haplotype.....	149
Table 2.10 List of primers used in this study.....	151

Acknowledgments

I would like to express my sincere gratitude to my advisor Dr. Bikram Gill for his insight, guidance, encouragement, patience, and support throughout my Ph.D. program. I am thankful to him for giving me the opportunity to work in his lab, and it was an enriching experience to discuss research with him, his passion for science motivates me a lot.

I would also like to thank my committee Dr. Ming-Shun Chen and Dr. Jeffrey Stuart for their continuous support and suggestions in addition to troubleshooting for my experiments. I am highly grateful to Dr. Ming-Shun Chen for providing material for Hessian fly infestation.

I am grateful to Dr. Akhunov and Dr. White for their valuable feedback, advice, and encouragement. I thank Dr. Susan Brown for serving as an outside chairperson.

I want to express my appreciation to Duane Wilson for helping me with all greenhouse related work and for his constant words of encouragement, John Raupp for providing seeds and sharing resources and Shauna Dendy for helping me with greenhouse HF infestation experiments. I am thankful to Dr. Bernd Friebe and Dr. Sanzhen Liu for their support and help with answering questions related to my research.

I want to thank Dr. Alina Akhunova for sharing her beautiful success and struggle experience and motivating me now and then. Being away from home and in tough times such as when I lost my grandmom and dog, both the times were difficult being away from but talking to her has always been very uplifting. She inspires to work harder and stay positive.

I want to thank all the teachers who taught different coursework and inspired and instilled my interest in research. I want to thank Dr. Draper for his support and encouragement and to all members of Plant Pathology department for providing friendly atmosphere at work.

Thank you to all the current and previous lab members for providing a supportive environment. I want to thank all of the undergraduate students who have helped with planting and harvesting.

I would like to thank my friends Likitha, Pavan, Nidhi, Vijay, Hannah, Bhanu and all my friends for their unconditional support and friendship.

I would like to express my deepest thanks and gratitude towards my family for their never-ending support and love.

Most of all, I am thankful to my husband, Karan and my mother in law Mrs. Balwinder S Aulakh for their emotional support, love, and blessings throughout my degree. I dedicate my dissertation to both.

I want to thank God for giving me inner peace, strength and for being my ray of sunshine forever.

Chapter 1 - Review of Literature

1.1 Importance of Wheat

Hexaploid or bread or common wheat (*Triticum aestivum* L. $2n = 6x = 42$) belongs to the tribe Triticeae and is one of the three major domesticated crops cultivated around the world (Faostat 2009). Due to its wide adaptability, wheat is cultivated almost everywhere on the globe including regions from 67° N in Scandinavia and Russia to 45° S in Argentina, the tropics, and the subtropics (Feldman, 1995). The major wheat cultivation regions of the world include the central plains of the U.S., southern Canada, southern Russia, the Mediterranean basin, India, Argentina, northern China and Australia.

As a crop, wheat has embedded itself into diverse cultures and societies and is a staple diet for 35% of the world population (Faostat; Reynolds et al. 2012). It is an important source of carbohydrates and protein in our diet and accounts for 20% of calories consumed (Porter et al., 2007). Bread wheat possesses a unique ability to make an elastic dough, which is used to make bread. Bread wheat is consumed around the world in the form of bread, cereals, cakes, cookies, and bagels, whereas tetraploid durum wheat (*T. turgidum* L. $2n = 4x = 28$) is mainly used for the production of macaroni and semolina products. It has an average annual harvest close to 750 million tons, and its annual yield is increasing at a rate of 1% (www.faostat.fao.org).

Wheat cultivation has played an important and significant role in the transition of human civilization from hunter and gatherer lifestyles to sedentary producer and agriculturist lifestyles. A major fraction of wheat produced is used for human consumption, and a small fraction is used for animal feed.

1.2 Origin and evolution of Wheat

The “Fertile Crescent” is believed to be the origin for the founder crops including diploid einkorn wheat *Triticum monococcum* (*T. monococcum ssp.monococcum* L.), tetraploid emmer wheat *T. dicoccum* (*T. turgidum ssp. dicoccum* L.) and barley (Heun et al., 1997; Nesbitt, 2001; Dubcovsky and Dvorak, 2007; Zohary and Hopf, 2000). This region stretches from Israel, Jordan, Lebanon, and Western Syria through Southeast Turkey, Northern Iraq and Western Iran (Zohary and Hopf, 1994; Gopher et al. 2002).

The cultivated wheat species form a polyploid series ranging in chromosome number of $2n=14$, 28 and 42. The cultivated form of the diploid A genome species named *T. monococcum* L. ($2n=2x=14$, A_mA_m), was domesticated from the wild species *T. boeoticum* ca 12,000 years ago (Heun et al., 1997). The cultivated tetraploid wheat *T. turgidum ssp dicoccum* L. ($2n=4x=28$, AABB), was domesticated from the wild species *T. dicoccoides* 9, 500 years ago (Ozkan et al., 2002). The allohexaploid bread wheat, *T. aestivum* L. ($2n = 6x = 42$, AABBDD) arose from hybridization of domesticated tetraploid wheat and a diploid goat grass (*Ae. tauschii*; $2n = 2x = 14$; DD) in a farmer’s field in Caspian Iran ca 8,000 years ago (Kihara 1944; McFadden & Sears, 1946; Nesbitt 2001)

The A, B and D genome diploid ancestors of polyploid wheat diverged from a common ancestor ca3 million years ago (MYA) (Huang et al.2002, Marcussen et al. 2014). The tetraploid wheat species *T. turgidum* (AABB) arose more recently, around 0.5 MYA, from hybridization between an A genome species, named *T.urartu* L. ($2n=2x=14$, AA) as a male and an unknown B genome species, which is closely related to the S genome of *Ae. speltoides* ($2n=2x=14$, SS), as a female

(Dvorák and Zhang, 1990; Lev-Yadun et al., 2000; Feldman et al.1995; Dvořák et al. 1993, Kilian et al. 2007; Maestra and Naranjo, 1998).

Spike or spikelet shattering at maturity is one of the important traits in the wild species which ensured seed dispersal, whereas a mutation in the brittle rachis gene (Br) resulted in the non-shattering trait and was selected to avoid seed loss during harvesting (Nalam et al. 2006). Another trait of free threshing, called naked seed, was selected in the modern cultivated diploid, tetraploid and hexaploid species over the hulled trait present in the wild species. The free threshing trait arose by a mutation in the *q* gene which further influenced recessive mutations in the tenacious glume (*Tg*) locus (Jantasuriyarat et al. 2004; Simons et al. 2006; Dubkovsky and Dvorak, 2007).

1.3 Need for Wheat improvement

With the increase in human population, modernization and industrialization there is a substantial increase in demand for food (United Nations et al. 2015; Brown, 2004). The world's population has already surpassed 7.4 billion (<http://www.census.gov>) and is estimated to reach 9.15 billion by 2050 (Alexandratos, N. and J. Bruinsma. 2012). To compensate for the increasing demand, global wheat production needs to increase by at least 2% annually.

Wheat domestication and selection has led to reduction in natural genetic variation by paving the way to standardized and high yielding cultivars (Jones et al. 1996; Gregova et al. 1997; Kam-Morgan et al. 1989; Lubbers et al. 1991; Akhunov et al. 2010), thus making them more vulnerable to different stresses and susceptible to a variety of pests (Dyke 1993).

Given the limited arable land, one way to improve wheat production includes genetic introgression of resistance to biotic and abiotic stress from wild relatives, which is regarded as the most economical, effective and environmentally friendly method. This, and other management practices would help in reducing the 25% loss in production that occurs due to both biotic (pests) and abiotic (cold, heat, drought, salinity) stress.

1.4 Introgression of traits from *Ae. tauschii* for wheat improvement

The wild relatives of modern and domesticated crop species have higher diversity for combating biotic as well as abiotic stresses. To increase the diversity and adaptability of modern wheat, different members of the Triticeae family have been utilized for genetic introgression into wheat by geneticists and breeders (Feldman and Sears, 1981; Gill et al. 2006). The ability to introgress traits from wild species into wheat depends upon the genetic makeup and phylogenetic relationship of the wild species with wheat.

As mentioned earlier, when bread wheat arose 8,000 years ago, only a few gametes of *Ae. tauschii* were involved in the origin of bread wheat (Kam-Morgan et al. 1989). Bread wheat is reproductively isolated from *Ae. tauschii*, and no natural genetic introgression has occurred (Gill and Raupp 1987). Thus, to broaden the genetic base of wheat's D genome, *Ae. tauschii* has been widely used for the introduction of useful traits into modern hexaploid wheat, and *Ae. tauschii* serves as an important source of genetic variation for wheat improvement (Gill and Raupp 1987, Gill et al. 2006). Wang et al. (2013) reported SNP analysis of *Ae. tauschii* gene pool and found that out of 7000 total SNPs, only a very small fraction is present in the D genome of hexaploid wheat. In addition to providing biotic and abiotic stress tolerance, *Ae. tauschii* also provides beneficial QTLs for grain protein and kernel weight (Fritz et al. 1995; Cox et al. 1990).

There are two different approaches used for the introgression of agronomically important genes from *Ae. tauschii* into hexaploid wheat, namely either by the generation of synthetic hexaploid wheat or by direct hybridization. Both the approaches involve back crossing into *T. aestivum* to avoid unwanted genes and linkage drag (Trethowan and van Ginkel, 2009; Trethowan and Mujeeb-Kazi, 2008).

1.4.1 Generation of Synthetic hexaploid

Synthetic hexaploid wheat is generated by crossing *T. turgidum* (AABB) with *Ae. tauschii* (DD) followed by spontaneous or induced chromosome doubling (Kazi et al., 2008, Yang et al., 2009, Blakeslee and Avery, 1937). McFadden and Sears (1946) reported the production of the first synthetic hexaploid wheat by crossing between primitive cultivated tetraploid species *T. turgidum* ssp emmer and diploid species *Ae. tauschii*. One of the advantages of synthetic hexaploid production is that it introduces variability into the gene pool of all three genomes A, B and D. This approach also has some disadvantages, including the introgression of deleterious genes from *T. turgidum* (Inagaki et al., 2014) and epistatic interactions between genes from *T. turgidum* and *Ae. tauschii*. (Ogbonnaya et al., 2013). These effects lead to poor quality traits and lower yield of the primary synthetics (Trethowan and van Ginkel 2009). Despite the limitations, several pathogen and pest resistant lines have been generated. Examples include Hessian fly (*Mayetiola destructor*) resistance, by Xu et al. (2006); greenbug (*Schizaphis graminum*) resistance, by Joppa and Williams (1982), Porter et al. (1989) and, Lazar et al. (1996); cereal leaf beetle (*Oulema melanopus*) resistance, by Joukhadar et al. (2013); cereal cyst nematode (*Heterodera* spp.) resistance, by Mulki et al. (2012); stem rust (*Puccinia graminis* f. sp. *tritici*) resistance, by Kerber and Dyck (1978) and, Marais et al. (1994); yellow rust (*Puccinia striiformis* f.sp. *tritici*) resistance, by Singh et al. (2000); *Septoria* leaf blotch (*Mycosphaerella*

graminicola) resistance, by Mujeeb-Kazi et al. (2000) and, Ghaffary et al. (2012); powdery mildew (*Blumeria graminis* f. sp. *tritici*) resistance, by Lutz et al. (1995); and karnal bunt (*Tilletia indica* Mitra) resistance by Mujeeb-Kazi et al. (2001).

Additionally, improvements in traits associated with grain productivity and end-use quality have been introgressed into the D genome using synthetic hexaploid wheat. Examples include grain yield, by Mujeeb-Kazi et al. (2008), Yang et al. (2009) and Li et al. (2014); kernel size, by Okamoto et al. (2013), Williams and Sorrells (2014), and Rasheed et al. (2014); preharvest sprouting, by Lan and Yen (1992) and, Imtiaz et al. (2008); new HMW glutenin subunits, by Lagudah et al. (1987), Mackie et al. 1996, Hsam et al. 2001 and, Tang et al. 2008; low polyphenol oxidase and lipoxygenase, by Yang et al. (2006), Mares and Mrva (2008); milling and baking quality, by Kunert et al. (2007) and, Tang et al. (2008); salinity tolerance, by Ogonnaya et al. (2013); and aluminum tolerance by Ryan et al. (2010).

1.4.2 Direct hybridization

This approach involves crossing *Ae. tauschii* directly with *T. aestivum* followed by embryo rescue and back crossing of F₁ plants with *T. aestivum* (Gill and Raupp 1987). Sehgal et al. (2010) reported that back crossed seeds of F₁ plants after colchicine treatment resulted in an observed increase in seed set when *T. aestivum* is used as the male parent and *Ae. tauschii* is used as the female. This approach has advantages over synthetic hexaploid approach in that it conserves the A and B genomes of the hexaploid while introducing diversity into the D genome (Gill and Raupp 1987; Cox et al., 2017). Though this approach results in low seed set, there is efficient and rapid introgression of traits from *Ae. tauschii* into hexaploid wheat by

recombination and segregation of D-genome chromosomes of wheat and *Ae tauschii* during back crossing (Gill and Raupp 1987; Fritz et al., 1995).

Introgression of pathogen and pest resistance genes from *Ae. tauschii* via direct hybridization have been reported. Examples include Hessian fly resistance by Gill et al. (1986), Raupp et al. (1993) and, Cox and Hatchett (1994); greenbug resistance by Gill et al. (1986), and (1991b); soilborne mosaic virus resistance by Gill et al. (1986) and, Cox et al. (1994); wheat spindle streak mosaic virus resistance by Cox et al. (1994); leaf rust resistance, by Raupp et al. (2001), Cox et al., (1994; 1997), Gill et al. (1986) and, Gill et al. (1991b); powdery mildew resistance by Gill et al. (1986), Cox et al. (1994), Murphy et al. (1998; 1999) and Yang et al. (2000); septoria leaf blotch resistance by Gill et al. (2006); stagonospora blotch resistance by Cox et al. (1999); and tan spot resistance by Brown-Guedira et al. (1999). Cox et al. (1995a; 1995b; 1997) and Brown-Guedira et al. (2005) have reported better grain yield, kernel weight and milling properties for back crossed lines derived by direct hybridization.

1.5 Mapping of genes in wheat

Physical and genetic mapping techniques are used to determine location, linear order, and the distance between genes and molecular DNA markers on wheat chromosomes. A physical map provides the distinct position of a locus on a particular chromosome, chromosome arm or deletion bin and the distance between loci in number of nucleotide base pairs. A genetic map, also known as a linkage map, provides a relative distance between loci based on recombination frequency.

1.5.1 Genetic mapping

Genetic maps are useful tools for use in breeding as they provide information on the location of both qualitative and quantitative trait loci (Mohan et al., 1997; Doerge, 2002; Yim et al., 2002). Genetic maps are a graphical representation of position and order of loci along chromosomes. (Collard et al. 2005). Genetic maps facilitate faster introgression of desired traits by monitoring the information of associated DNA markers via marker-assisted (MAS) selection. The dense genetic maps form the framework for comparative mapping as different members of tribe Triticeae (Ahn and Tanksley, 1993; Paterson et al., 2000). They act as an anchor for correlating physical map data with sequence, location, traits and translocations (Yim et al., 2002). Genetic maps help in fine mapping and cloning of genes (Mohan et al., 1997; Vuysteke et al., 1999).

Genetic maps are developed based on the principle of recombination and segregation of loci during meiosis and can be studied in different populations (Paterson, 1996). The primary requirement for genetic mapping of loci is the polymorphism existing between parents of the mapping population. The low level of recombination at centromeric regions leads to clustering of markers (Erayman et al., 2004).

Botstein et al. (1980) developed the first genetic maps in humans using restriction fragment length polymorphisms (RFLPs). Afterwards in addition to RFLPs, amplified fragment length polymorphisms (AFLPs), simple sequence repeat (SSR), single nucleotide polymorphism (SNP), expressed sequence tags (ESTs), and diversity array technology (DArT) markers were used to develop genetic maps in hexaploid and tetraploid wheat (Hart et al., 1993; Xie et al., 1993; Jia et al., 1996; Blanco et al., 1998; Roder et al 1998; Messmer et al., 1999; Peng et al., 2000; Somers et al. 2003; Akbari et al., 2006, Chao et al. 1989; Gill et al. 1991a; Liu and Tsunewaki 1991;

Devos et al. 1992; Anderson et al. 1992; Gale et al. 1995; Gill et al. 1996b; Gill et al. 1996a; Mingeot and Jacquemin 1999; Celso et al., 1996; Nelson et al. 1995b; Nelson et al. 1995c; Nelson et al. 1995a; Van Deynze et al. 1995; Somers et al. 2004; Chalmers et al. 2001; Semagn et al. 2006). Although RFLP maps are most informative, the limitations of using RFLPs was the requirement of a large quantity of DNA, low level of polymorphism and is laborious and time-consuming. Both the polymerase chain reaction (PCR) - based markers and hybridization-based DArT markers are high-throughput, highly polymorphic and faster, hence leading to dense genetic maps. The high reproducibility, ease for high-throughput screening and co-dominant inheritance of SSRs and the coding region representation by ESTs (Somers et al. 2003; Akbari et al., 2006; Somers et al. 2004; Roder et al. 1998; Stephenson et al., 1998). With the help of next-generation sequencing technology, genotype-by-sequencing and many SNP based microarrays have been utilized to capture the variation present in wheat exome to assist in breeding (Wang et al., 2014, Poland et al., 2012).

1.5.2 Physical Mapping

While genetic mapping is affected by wheat polyploidy, as most of the genes are present in all three genomes, the buffering capacity of the polyploid genome made it possible to isolate a large number of aneuploid stocks for physical mapping (Sears 1954, 1966; Endo and Gill 1996). The cytogenetic stocks obtained by deletion of whole chromosomes, chromosome arms or different fractions of sub-arms led to the development of nullisomic-tetrasomic (NT), ditelosomic (DT), and single-break deletion lines respectively for the physical localization of markers (Endo 1988; Endo and Gill 1996; Sears 1954; Sears and Sears 1978). In NT stocks, a pair of chromosomes is missing and compensated by an extra pair of one of the two homoeologous chromosomes. In DT stocks, one pair of chromosome arms is missing. The deletion lines with different fraction

lengths (FL) of deleted chromosome segments for each of the 21 chromosomes of wheat cultivar Chinese Spring were produced by Endo and Gill (1996). Out of a total of 436 deletion lines, 350 lines had a deletion in the homozygous state and have been extensively used for allocating different markers/loci to deletion bins. These stocks have been used to develop physical maps with ESTs mapped into 189 chromosome bins (Qi et al. 2003, 2004, Conley et al. 2004; Hossain et al. 2004; Linkiewicz et al. 2004; Miftahudin et al. 2004; Munkvold et al. 2004; Peng et al. 2004; Randhawa et al. 2004)

In addition to cytogenetic stocks, fluorescence in situ hybridization (FISH) and Radiation Hybrid Mapping (RH) are other physical mapping methods used in wheat. The fluorescent labeled DNA probes are directly visualized on chromosomal preparations using FISH (Jiang and Gill 1994, 2006). For RH mapping, radiation-induced chromosome breaks in irradiated pollen from hexaploid wheat when crossed with tetraploid wheat gives the D genome RH panels where markers are scored as either present or absent in panel lines depending upon the presence of retained or lost chromosome arm segments (Tiwari et al., 2012, Tiwari et al., 2014).

1.6 Gene Cloning

To understand the molecular basis of any specific trait, it is important to identify, fine map and clone the associated genes so that they can be used for crop improvement. Gene-derived markers help in marker-assisted selection. Gene cloning helps in understanding its mode of action and interaction with other genes. Many factors affect the development of markers for fine mapping and gene cloning in wheat. These include large genome size, the presence of a large amount of repetitive DNA, non-uniform recombination rates, the presence of genes in either gene-rich or gene-poor regions, and, sometimes, the lack of availability of sequence information of

homeologs to develop genome specific primers (Erayman et al. 2004; Qi et al. 2004, Gill et al. 1993; Hohmann et al. 1994; Delaney et al. 1995a,b; Mickelson-Young et al. 1995; Gill et al. 1996a, b; Sandhu and Gill 2002, Dvorak and Chen 1984; Kota et al. 1993; Lukaszewski and Curtis 1993; Dvorak et al. 1998). Recent developments in sequencing and assembly of whole genomes of wheat, *Ae tauschii* and gene editing technologies will aid in cloning and functional validation of genes in wheat (IWGSC: https://urgi.versailles.inra.fr/blast_iwgsc; Wang et al., 2018).

Nevertheless, many genes for disease resistance including *Lr10* (Feuillet et al., 2003), *Lr21* (Huang et al., 2003), *Lr22a* (Thind et al., 2017) and *Lr1* (Cloutier et al., 2007) against leaf rust; *Sr22* and *Sr45* (Steuernagel et al., 2016), *Sr35* (Saintenac et al., 2013), and *Sr33* (Periyannan et al., 2013) against stem rust; *Pm3b* (Yahiaoui et al., 2004) and *Pm8* (Hurni et al., 2013) against powdery mildew; *Yr36* (Fu et al., 2009), *Yr10* against stripe rust (Liu et al., 2014), *Fhb1* against fusarium head blight (Rawat et al., 2016), *Stb6* against Septoria tritici blotch (Saintenac et al., 2018), agronomic traits including *TaTGW-7A* (Hu et al., 2016), *TaSnRK2.10* (Zhang et al., 2017), *TaGS5-3A* (Ma et al., 2016), abiotic stress namely *CRT* (Jia et al., 2008), *TaBAG* (Ge et al., 2016), *TaSTK* (Ge et al., 2007), domestication related traits such as *Q* gene (Faris et al., 2003, Simons et al. 2006), grain including *Gpc-B1* (Uauy et al., 2006) , vernalization genes *VRN1* (Yan et al., 2003), *VRN2* (Yan et al., 2004), *VRN3* (Yan et al., 2006) and flowering related genes such as *TaEG-4A,B,D* (Yang et al., 2013a), *TaAGO1b* and *TaAGO4* (Meng et al., 2013) have been cloned in wheat.

1.7 Plant pest interactions

Plants and arthropod insects are believed to have coevolved to ever-changing and different kinds of abiotic and biotic stresses (Futuyma and Agrawal 2009, Holden et al., 1993). About 400 million years ago, arthropods speciation started after the land plants evolved from their aquatic ancestors (Bateman et al., 1998, Ehrlich and Raven 1964, Gatehouse 2002). Herbivory of the arthropods led to the natural selection of plants. The interaction between plants and herbivorous arthropods in agriculture are complex and involve many factors (Rasmann and Agrawal, 2009). For the herbivore, the search for a suitable host typically starts with visual or odor cues followed by a more complex physical and chemical interaction with the host plant, which depends on various factors including toxins and nutrients (Duffey and Stout, 1996; Schoonhoven et al., 1998).

Plants are faced with different kinds of parasites some of which include bacteria, viruses, fungi and insects (Sacco and Moffett 2009). Plants employ a wide array of direct and indirect defense mechanisms like antibiosis, antixenosis, tolerance and physical barriers like thorns and trichomes to evade the invading pest species (Fritz and Simms 1992, Panda and Khush 1995; Schoonhoven et al., 2005; Smith and Clement, 2012; Kogan and Ortman in 1978, Horber, 1980; Smith, 2005). The defense response employed by the plant is constantly evolving and is determined by the selection pressure of the herbivore (Thompson 2005). To adapt to the evolving defense mechanism of plants, arthropods have also evolved (Fritz and Simms 1992, Futuyma and Agrawal 2009).

1.7.1 Host plant resistance (HPR)

General mechanisms of plant defense against pests and pathogens is defined as the cumulative effect of genetically inherited resistance related traits that determine the extent of harm caused to the plant by the pathogen (Painter 1951; Smith and Clement, 2012). Variations in host plant resistance (HPR) against a particular pathogen is determined to a major extent by the genetic background or genotype of the host. Genetic variation has a beneficial effect in pest management, but repeated exposure to same pest biotype results in artificial selection for a particular R gene leading to a reduction in genetic diversity (Schoonhoven et al., 2005). This suggests an important role of differences in natural or artificial selection affecting the generation of populations with varying degree of resistance (Ranger et al., 2007; Murugan et al., 2010; Diamond 2002). Genetic diversity also plays a very important role in modern agriculture and is sought after by entomologists and plant breeders to develop high yielding arthropod-resistant crops (Clement and Quisenberry 1999, Smith 2005, Stout 2007, Yencho et al., 2000).

1.7.2 Role of HPR in pest management

HPR is an efficient method for pest management and an important part of integrated pest management programs (Panda and Khush 1995, Smith 2005, Stout 2007; Adkisson and Dyck, 1980). Some of the advantages of using HPR include the elimination of insecticide use, and lakes and streams free of residues from insecticides, reduced mortality of useful insects. HPR is compatible with biological control agents & insecticides and is simple and cost-effective for the farmers (Smith 2005; Wilde, 2002; Wiseman, 1994). HPR also presents some disadvantages by not showing compatibility with all useful agents employed for biological control (Smith 2005). Also, single dominant gene resistance is transitory which may lead to the development of

virulent pests which are resistant to HPR (Fritz and Simms 1992, Smith 2005, Stout 2007, Van Emden 2007). For the effectiveness of the management program, in addition to the presence of a resistant gene (R) in the host other management practices like biological control, insecticide application and cultural practices must be integrated (Quisenberry and Schotzko 1994; Wiseman 1994)

1.7.3 Influence of biotic and abiotic factors on induced resistance

Studies have shown that after the initial attack of pests, the plant responds using immune mechanisms, which impart direct or indirect resistance against subsequent herbivores (Stout 2014). Studies on factors affecting pest induced resistance in plants by Painter et al.(1951) and others have suggested the important role of abiotic factors like soil fertility and temperature in imparting resistance (Panda and Khush 1995, Smith 2005) Plants respond to pest infestation by acquiring a resistant phenotype via multiple pathways including those involving reactive oxygen species and others that are mediated by hormones like jasmonic acid (JA), ethylene, salicylic acid (SA), abscisic acid (ABA), and gibberellic acid (GA) (Couldridge et al., 2007; Kielkiewicz 2002, Li et al., 2008, Liu et al., 2007). Particularly JA and SA are mutually inhibitory and act as mediators of plant response to insects and pathogen infestations, respectively (Howe and Jander, 2008). Increases in JA and SA levels precedes the elevated expression of resistance related genes (Wu and Baldwin 2010). Additionally, the application of exogenous JA or SA results in increased expression of resistance genes. Lines mutant for the biosynthesis of these hormones or their receptors are unable to confer resistance (Wu and Baldwin, 2010; Zhou et al., 2009).

1.8 Mechanism of plant immunity against pests

Plants lack specialized circulating immune cells and, thus, exhibit resistance mechanisms that rely on the perception of signals in the form of molecules derived directly or indirectly from the pathogen. Two main types of defense mechanisms, depending on whether they are constitutive or induced exist in plants and are called: Pathogen Associated Molecular Patterns (PAMPs) Triggered Immunity (PTI) and Effector-Triggered Immunity (ETI). PTI, also known as basal immunity, is a low-intensity response and applies to detection and defense against the majority of pathogens and works by perceiving PAMPs (Sacco and Moffett 2009; Chisholm et al., 2006). PAMPs include molecules such as bacterial flagellin, bacterial and fungal cell wall components, lipopolysaccharides or dsRNA (Nurnberger et al., 2004). On the other hand, pathogens that are adapted to the host defense act by suppressing PTI via the release of effector proteins. To confer resistance against these pests, plants have evolved ETI, which uses the products of R genes that act as receptors to detect effector proteins and elicit an even stronger immune response (Chisholm et al., 2006).

1.8.1 The gene-for-gene model

Flor (1971) first reported the study of the gene-for-gene interaction between different varieties of flax (*Linum usitatissimum* L.) and corresponding races of its fungal pathogen (*Melampsora lini*). For every *Avr* gene present in the pathogen, there is a corresponding R gene present in the host plant (Kerr 1987). The gene-for-gene model states that the genotype of both the plant host and the pathogen determines whether the host-pathogen interaction would result in susceptibility or resistance in the host. When the R gene in the host interacts with the *Avr* gene of the pathogen in incompatible interactions, it leads to resistance in the host.

On the other hand, compatible interactions lead to susceptibility in the host, which is caused by the lack of interaction between R gene of the host and *Avr* gene of the pathogen occurring either due to absence of the R or *Avr* gene or due to the evolution of the *Avr* gene in pathogen to evade recognition by the R gene (Sacco and Moffett 2009, Dodds and Rathjen, 2010). One of many plant defense responses initiated upon pathogen infection is called the hypersensitive response (HR) (Heath, 2000). HR deprives the pathogen of the nutrients it requires from the host to survive and propagate to a new site of infection. Most R genes belong to a small group of protein classes and, thus, imply that they function through similar pathways (Sacco and Moffett 2009; Cheng et al., 2013).

One of the limitations of R genes, which follow the gene-for-gene model is that during the course of time, the *Avr* gene from avirulent pathogens undergoes evolutionary changes, which renders it compatible with the host R gene and, thus, leads to disease and susceptibility (Kaloshian 2004). The newly evolved virulent *Avr* gene imparts a selection pressure on the host, which results in the evolution of resistant gene (R) in the host. Different biotypes of the pest are classified as virulent or avirulent based on their compatible and incompatible interactions respectively with the plant cultivars having different resistant genes (Smith 2005; Smith et al., 1994).

Plant host resistance may be qualitative or quantitative based on the number of loci encoding for the resistance genes (Yencho et al., 2000; Keurentjes et al., 2008). Besides gene-for-gene resistance which may be qualitative or quantitative, resistance may be race-specific or non-race-specific. Example of latter type resistance is *Lr34* which encodes for an ABC transporter and provides an important source of resistance against multiple fungal pathogens in wheat (Krattinger et al. 2009). The non-race-specific resistance gene alone may not provide a high level

of resistance, but it is durable. Many R genes have been identified that confer resistance against insects in wheat and rice. Some of the genes include Hessian fly resistance (*H*) genes, greenbug resistant genes (*Gb*), Russian wheat aphid (*Dn*) resistance in wheat and Brown plant hopper genes (*Bph*) and gall-midge resistance (*Gm*) genes in rice.

1.9 Pathways involved in plant-pathogen interactions

The mechanism of action for R genes in different plant species have been studied. In wheat-hessian fly interaction, one of the important components of resistance pathway is to maintain the cuticle integrity (Kosma et al., 2010). For the resistance against brown planthopper in rice, defense mechanism employs callose deposition in the phloem as means to avoid feeding of the phloem by the pest (Chen et al., 2008; Du et al., 2009). In melon, wound healing via enhancement of sieve elements acts as defense mechanism through a resistance gene named *Vat* (Martin et al., 2003).

1.10 Structure of R genes

Some effectors have been identified, that are secreted by different insects (Ramsey et al., 2007; Harmel et al., 2008; Carolan et al., 2009; Bos et al., 2010; Rodriguez and Bos, 2013). Many R genes, which confer resistance against pests and pathogens, belong to a member of the nucleotide-binding, leucine-rich repeat (NBS-LRR) family (Kaloshian, 2004). NBS-LRR proteins work by direct and indirect binding to the effectors released by insects either during probing or prolonged feeding (Will et al., 2007; Moreno et al., 2011). In different plant species, R genes exhibit variation in both number and function of different domains which explains the diversification and adaptability of R genes to evolving effector proteins.

1.10.1 R genes with extracellular domains

One class of membrane-bound R genes are comprised of both extracellular and intracellular domains (Hammond-Kosack et al., 1994; Dixon et al., 1996, 1998). Intracellular domain is made up of a short peptide, whereas Leucine-Rich Repeat (LRR) domain is extracellular and serves as a ligand binding domain and are, thus, also called receptor-like proteins (RLPs). LRR domains consist of a consensus sequence LxxLxxLxxLxLxxNxLxGxIPxx with a number of repeats ranging from 25 to 38 (Jones and Jones, 1997). Given the variability in the number of repeats and amino acid sequence, LRR domain serves as an important determinant of recognition specificity (Wulff et al., 2001).

In rice, some RLP encoded by genes like *Xa21* and *Xa26*, contain cytoplasmic kinase domain, in addition to LRR domains, and are thus called receptor-like kinases (RLKs) (Gomez-Gomez and Boller, 2000; Xiang et al., 2006). Among different accessions of *Arabidopsis*, genes encoding for RLKs are one of the most diverse gene families (Fritz-Laylin et al., 2005; Clark et al., 2007). Kinase domains in the RLK proteins are proposed to be involved in different signaling pathways with examples of RLK proteins like EFR1 serving as PAMP receptors and are involved in PTI response to bacterial flagellin (Gomez-Gomez and Boller, 2000; Zipfel et al., 2006).

1.10.2 R genes with NB-LRR domains

Plant NBS-LRR proteins consist of nucleotide binding domain (NB) and a carboxy-terminal LRR domain (Sacco and Moffett 2009) and have common structural and functional properties with an important component of innate animal immunity called NOD-like receptor (NLR) family (Rairdan and Moffett, 2007). LRR domain present in the NB-LRR proteins is different from LRR domain found in the extracellular domains of certain R proteins described previously. LRR

domain of NB-LRR proteins consists of a consensus sequence LxxLxxLxxLxLxx (N/C/T) x(x) LxxIPxx (Jones and Jones, 1997; Kajava, 1998).

The classification of different NB-LRR proteins is based on the type of domain present in the amino terminal of this protein. Between the two major classes of NB-LRR proteins namely: TIR-NB-LRR and CC-NB-LRR, the main difference lies in the presence or absence of a domain at the amino terminus, which shows homology to animal Toll and interleukin-1 receptors (TIR) (Whitham et al., 1994). Additionally, the two classes also differ by having a distinct set of consensus sequences in their NB and ARC domains (Meyers et al., 1999; Cannon et al., 2002). CC (non-TIR)-NB-LRR proteins have a coiled-coil domain at the amino-terminal instead of the Toll and interleukin-1 receptors (TIR) domain.

In different plant species, there are variations in the number of genes encoding for NB-LRR proteins with further differences in a number of NB-LRR proteins belonging to each class within the plant species. A total of 333 and 399 NB-LRR encoding genes are present in *Medicago truncatula* L. and *Populus trichocarpa* L., respectively (Tuskan et al. 2006; Ameline-Torregrosa et al. 2008; Kohler et al. 2008). In *Arabidopsis*, out of the total 149 genes encoding for NB-LRR proteins, 94 genes encode for the TIR-NB-LRR class of proteins (Meyers et al., 2003), whereas none of 400 NB-LRR encoding genes in rice belong to TIR-NB-LRR class (Monosi et al., 2004; Zhou et al., 2004; Bai et al., 2002). The papaya genome has been shown to contain 58 NB-LRR encoding genes with the majority of them belonging to the CC-NB-LRR class of proteins (Ming et al. 2008).

1.10.3 Functions of different domains in CC-NB-LRR proteins

The domain localized to amino terminus of CC-NB-LRR proteins is called a CC domain, which is not well conserved but consists of a conserved motif EDVID (Ellis et al., 2000; Shen et al., 2003; Dodds et al., 2006; Qu et al., 2006; Rairdan and Moffett, 2006). Primarily CC domain is involved in protein-protein interactions, which is supported by mutational analysis of conserved motif in CC domain of Rx protein in pepper (Rairdan et al., 2008). The important role of CC domain in intramolecular interactions is demonstrated by activation of Rx protein resulting from the loss of interaction between CC domain and NBS-LRR domain (Moffett et al. 2002). The expression of CC domain by itself from N requirement gene 1 (NRG1) and activated disease resistance 1 (ADR1) resulted in hyper sensitive response (HR) indicating an active role of CC domain in signaling (Collier et al., 2011).

The NB domain is considered to have a conserved function and sequence among the classes of NB-LRR proteins (Tameling et al., 2002). The ARC domain is named after the proteins APAF-1, plant R and CED-4 proteins which share this domain (Van der Biezen and Jones, 1998a). The conserved arrangement of NB-ARC domains is also named as NB-ARC regions (NBS), and the motifs present in NB and ARC domains are distinct between TIR and CC classes of NB-LRR proteins (Meyers et al., 1999). NBS domains play an important role in signaling and is demonstrated by the expression and mutational studies described below. An *Avr*-independent response similar to HR was observed when NB domain of Rx protein was transiently overexpressed in tobacco leaves (Rairdan et al., 2008). Similarly, *Avr*-independent HR response was observed when different NB-LRR proteins with either mutated or deleted domains like LRR and/or ARC were expressed (Hwang et al., 2000; Bendahmane et al., 2002; Shirano et al., 2002; Zhang et al., 2003; Howles et al., 2005; Tameling et al., 2006).

The functional diversity in R gene is attributed to variations in the LRR domain and is demonstrated by differences in recognition specificities resulting from swapping of LRR domains between closely related R genes (Ellis et al., 2000; Shen et al., 2003; Dodds et al., 2006; Qu et al., 2006; Rairdan and Moffett, 2006). In addition to being an important component of NB-LRR proteins, LRR domains in addition to ATP hydrolysis have a very important role in maintaining the inactive state of NBS domain (Moffett et al. 2002).

1.10.4 Diversification of R genes

In several plant species, alternate configurations for domains similar to NB-LRR consisting of TIR-NB, TIR and CC-NB have been reported for proteins (Bai et al., 2002; Meyers et al., 2002, 2003). In *Arabidopsis*, R genes like RRS1-R involved in the signaling pathway of disease resistance against *Ralstonia solanacearum* have been reported to contain additional domains like nuclear localization signal (NLS) (Deslandes et al., 2002).

During evolution, events of gene duplication and divergence lead to clustering and diversity in NB-LRR encoding genes (Michelmore and Meyers, 1998). Grube et al. (2000) reported differences in pathogen recognition specificities for R genes, which either arose by duplication or are related by descent. One such example exists in potato, where resistance against a nematode and pest arises from two R genes namely *Gpa2* and *Rx* which belong to the same gene cluster (Bakker et al., 2003). Also, the alleles of the same R gene can confer resistance against different variants of an *Avr* gene from the same pathogen, which is demonstrated by alleles of *Mla* locus against *Avr* variants of powdery mildew-causing pathogen *Blumeria graminis f. sp. hordei* (Haltermann and Wise, 2004). Alleles of same R gene can provide resistance against different pathogens as stated by different alleles namely *HRT*, *RCY1*, and *RPP8* from the same locus

conferring resistance towards different genera of viruses TCV & CMV and oomycete *Hyaloperonospora parasitica*, respectively (Cooley et al., 2000; Takahashi et al., 2002). Also, allelic variants of R gene can also impart resistance towards different effectors encoded by the same *Avr* gene as reported by recognition of AvrRPM1 and AvrB effectors from *Pseudomonas syringae* by RPM1 of *Arabidopsis* (Grant et al., 1995).

1.11 Hessian fly: A destructive pest of wheat

1.11.1 History and economic impact

The Hessian fly (HF) (*Mayetiola destructor* L.) is a gall-forming insect. It belongs to the family Cecidomyiidae, order: Diptera. The HF is thought to have originated in the Fertile Crescent (south-western Asia) where wheat also originated and spread into Europe and Central Asia. (Barnes 1956). The HF was introduced in North America by soldiers of American revolutionary war via straw bedding for horses in 1777. The British hired these soldiers from Hesse, Germany to fight against the American colonists and hence the name Hessian fly for the insect. (Pauly 2002, Hunter B 2001)

HF is a serious pest and causes widespread economic losses to the wheat crop. HF is present in North America, Asia, Europe, New Zealand and North Africa. (Chen et al., 2009; Lidell and Schuster 1990; Ratcliffe et al., 2000; Smiley et al., 2004; Watson 2005; Hatchett et al. 1987; Pauly 2002; Harris et al. 2003). HF infestation causes millions of dollars in the loss by affecting both the yield and quality of wheat grain (Smiley et al., 2004; Buntin 1999). Multiple generations of HF can infest within a year affecting both winter and spring wheat (Buntin 1999; Cambron et al., 2010; Flanders and Buntin 2008; Pedigo 2002). Crops infested at one leaf stage are killed

whereas if infested at a later stage result in reduced seed count and seed weight. (Byers & Gallun 1971; Buntin & Chapin 1990)

1.11.2 Biology and Host Range

The life span of adult HF is 1-4 days during which females mate only once and deposit eggs in the grooves on the adaxial leaf surface. Eggs hatch in 3-4 days and neonate larvae move to the base of the nearest node to feed (Haseman 1930). First and second instar larvae feed, and third instar larvae grow within a puparium. If environmental conditions are unfavorable, third-instar larvae go into diapause, else flies emerge from pupae. The non-feeding third instar larvae can be maintained at 4°C for more than a year. The life cycle from oviposition to adult flies takes around 28 days. (Harris et al., 2003; Gagne & Hatchett 1989; Stuart et al. 2008). In addition to hexaploid wheat, HF can infest barley, rye, triticale and other genera of tribe Triticeae such as *Aegilops* and *Agropyron*. (Harris et al., 2006; Jones 1936; Jones 1938).

1.11.3 Biotypes of Hessian fly

Biotypes have been defined as populations within an arthropod species that differ in their ability to utilize a particular trait such as survival and development on a particular plant genotype (Smith 2005; Gallun and Khush 1980; Wilhoit 1992; Pedigo 1999). In addition to the 16 known biotypes of HF that are designated as Great Plains (GP), and A through O, newly emerging biotypes such as *vH9* and *vH13* have been identified. Each biotype is classified by its specific pattern of resistance and susceptibility on different wheat cultivars. As there are more than 35 different R genes against HF in wheat, so potentially 235 (34359738368) genotypes of HF are possible. As it is impractical to designate these many biotypes, so only four resistant wheat cultivars are used for identification of biotypes, a total of 24 or 16 possible combinations are

there, hence 16 biotypes (Gallun 1977; Diehl & Bush 1994). Virulence in HF is controlled by recessive, non-allelic genes. The virulence to *H9* and *H13* is controlled by virulence genes *vH9*, *vH13*, respectively.

1.11.4 Virulence genes encoding putative HF effectors

Differences in HF biotypes is attributed to the diversity of proteins encoded by genes present in the salivary glands of Hessian fly. There are ~ 2000 unique genes, which have been identified from salivary glands of Hessian fly from different geographical populations and biotypes; these encode putative effector proteins, the Secretory Salivary Gland Proteins (SSGPs) (Chen et al., 2004; Chen et al., 2006). This superfamily has a non-conventional conservation pattern, with non-protein-coding regions having more sequence similarity than the coding region of these SSGP's (Chen et al., 2010). The 5'-UTR, signal peptide coding, introns and 3'-UTR are conserved whereas the mature protein coding region has diversification. The diversification of the mature coding region hints that these genes are under high selection pressure for functional adaptation. These SSGPs could be determinants of HF virulence (Chen et al., 2008). SSGPs are categorized into 78 families; each family has a unique structure, e.g., one family has a higher content of cysteine residues. Cysteine can form disulfide linkages and conjugate with host proteins.

Avirulence to *H13* is due to the recessive *vH13* gene, which is the first *Avr* gene to be cloned in insects (Aggarwal et al., 2014). It is transcribed in salivary glands of *H13*-avirulent larvae. In *H13*-virulent larvae from field populations, two independent insertions of lengths 4.7kb and 254bp were found at the exon-intron junctions, whereas another insertion of 461 bp was found in the coding region of the second exon. (Aggarwal et al., 2014). The *vH13* sequence is unique,

and neither homologs nor conserved domains were identified with any protein in the NCBI database. (Aggarwal et al., 2014; Rider et al., 2002; Lobo et al., 2006). Another *Avr* gene *vH24* cognate to *H24* encodes a type-2 serine/threonine protein phosphatase (PP2C) domain along with a secretion signal peptide. In *H24*-virulent flies, insertion in the promoter of *vH24* gene prevents transcription (Zhao et al., 2016b). Two more *Avr* genes *vH9* and *vH6* have been identified. A premature stop codon in *Avr* gene *vH9* and insertion in *Avr* gene *vH6* leads to silencing of gene and makes flies virulent to respective R genes (Stuart J. 2015).

1.11.5 The phenotypic response in wheat upon HF attack:

1.11.5.1 Wheat response to an attack by virulent larvae: The compatible interaction

The virulent HF larvae induce a nutritive tissue for feeding. When larvae infest and feed on wheat seedling, it causes irreversible morphological changes in wheat. The epidermal and mesophyll sheath cells are altered to become nutritive cells, which have exceptionally large nuclei, increased numbers of mitochondria and Golgi apparatus, endoplasmic reticulum and thin cell wall that breaks down to provide nutrients to larvae. (Harris et al., 2006). Nutritive tissue acts as a sink tissue within the wheat seedling. The susceptible wheat seedlings become dark green with stunted growth (Miller et al., 1958; Cartwright et al., 1959; Robinson et al., 1960). There are fewer heads and reduced seed set. HF attack induces the genes, which have a role in nutrient metabolism and transport. Heat shock proteins, transcription factors, genes from primary metabolic pathways are up-regulated in the plant post-HF attack. Carbon/Nitrogen ratio is decreased at the feeding sites. Membrane permeability increases and nutrients become more accessible to the feeding larvae. Genes encoding defense proteins and histones are down-

regulated as plant growth is retarded, so plant needs less structural proteins. Genes involved in cell wall metabolism and lipid transfer proteins are down-regulated (Liu et al., 2007)

1.11.5.2 Wheat response to an attack by avirulent larvae: The incompatible interaction

In incompatible interactions, HF larvae die within few days of infestation, and there is no larval growth. There are localized, induced responses in the plant upon attack by HF larvae. The resistance to HF has common features with penetration resistance against fungi (Harris et al., 2010).

The initial attack by larvae using its mandibles to attack the cell wall remains the same in both compatible and incompatible interactions. However, larvae are unable to induce the susceptibility features, such as the formation of nutritive tissue and permeable cell wall in incompatible interactions. Epidermal cells at the site of infection show plasmolysis & death in addition to wall reinforcement in adjacent cells. The reactive oxygen species (ROS) are formed, adjacent plant cells show swollen mitochondria, the proliferation of the endoplasmic reticulum and Golgi apparatus.

Reinforcement of cell wall involves repair of the tangential wall from the physical damage done by HF's mandibles and acts as a barrier to effectors produced by salivary glands of HF larvae. Granular material is formed in the paramural space between the wall and plasma membrane and within pockets localized in the wall matrix. These granular materials may contain toxins like lectins, wheat protease inhibitors, and enzymes that synthesize phenyl propanoids. The abundance of ER-Golgi bodies suggests active synthesis and secretions with transport to the wall via numerous vesicles (Harris et al., 2010)

The genes that are up-regulated in wheat during incompatible interactions are involved in cell wall metabolism. These include gluconases and cellulose synthases. Peroxidases are also up-regulated and have a role in ROS synthesis, signaling and cell wall cross-linking. Genes including lipid transfer proteins which are associated with rapid mobilization of membrane lipids at the infested site are also up-regulated. (Liu et al., 2007)

1.11.6 HF management

In addition to the use of resistant wheat cultivars, there are cultural, biological and chemical methods for controlling HF.

1.11.6.1 Cultural control

Cultural practices include planting after the “fly-free date” so that the HF adults that emerge during that Fall season die before wheat is available for infestation (Dean and McCulloch 1915; Horton et al. 1945; Davis et al. 2009). Crop rotation helps in killing the population of HF from previous year’s infested stubble. Growing wheat every year on the same field should be avoided. This helps in the killing of HF from the previous season as HF does not infest plants outside of Triticeae tribe. Volunteer wheat should be destroyed, as they act as green-bridge for HF population from one season to another as they act to increase HF populations and reduce the effectiveness of fly-free date as control (Whitworth et al. 2009).

1.11.6.2 Biological control

Natural enemies of HF are known, but these are not always effective at maintaining HF populations below economic thresholds (Gahan 1933, Hill et al. 1939, Hill 1953, Schuster and Lidell 1990, Rojas et al. 2000, Knutson et al. 2002). Parasitoids such as *Platygaster hiemalis* L.,

Homoporus destructor L. (Say) and *Eupelmus allynii* L. are lethal to HF in laboratory experiments but not have been documented to be effective in the field (Schuster and Lidell 1990; Rojas et al. 2000; Knutson et al. 2002)

1.11.6.3 Chemical control

Use of insecticide-treated seeds may provide control for a short period and reduce the effect of infestation depending on the timing of planting and emergence of HF but is not effective for a complete season (Wilde et al., 2001). Application of foliar insecticides on crops also may provide some control, but inconsistent results and cost ineffectiveness lead to less usage of insecticides. Eggs are present on the leaf surface for a very short time, and larvae are present inside leaf sheath which makes it difficult for timely and effective application of insecticides (Buntin & Hudson 1991)

1.11.6.4 R gene control

Use of HF resistant wheat cultivars is the most effective and cost-efficient approach to HF management (Ratcliff & Hatchet 1997). There are more than 35 genes known that provide resistance against HF. Currently, *H13* is most effective in Alabama, Georgia, North Carolina, South Carolina and Louisiana. Among the other HFR genes *H12*, *H18*, *H24*, *H25*, *H26* are most effective, and *H33*, *H9*, *H10*, *H22*, *H31* are moderately effective. Single dominant resistance genes *H3*, *H5*, and *H6*, were deployed in Indiana in 1960s, but they lost their effectiveness within ten years (Foster et al., 1991; Gould 1998). Therefore, it is important to clone and find novel resistance genes and understand the mechanism by studying interactions between R-*Avr* by cloning of a cognate pair.

1.12 Introduction to rice and rice pests the brown plant hopper (BPH) and gall midges

Rice, similar to wheat, is grass and rice pests BPH and gall midge are similar to Hessian fly forming galls on the host plant and sharing similar co-evolutionary histories. Rice (*Oryza sativa* L.) is considered as one of the top three important cereal crops in the world. Rice serves as a staple food for more than 50% of the population worldwide (Khush 2005) and contributes 20% of calories consumed by humans annually (Peng et al., 2016). With the increasing population and food demand, it is very important to increase the production of rice in a sustainable manner (Normile 2008). There exist two main rice varieties, which are cultivated worldwide named *Oryza glaberrima*, which is primarily grown in Africa, and *Oryza sativa* which is cultivated across all the continents in temperate and tropical regions (Khush, 1997; Heinrichs, 2009). Majority of rice is production and consumption (90%) is done in temperate and tropical Asia (Khush 1997).

Rice is attacked by a variety of arthropod species, and this affects the yield by 1-1.5 billion kg/year resulting in a big economic loss. The effective management of different pests is one of the ways to increase the global production of rice (Zhang 2007). Many insects attack the rice plant and are classified based on the site and mode of feeding: root feeders, stem feeders, planthoppers, leafhoppers, defoliators, grain sucking and gall midges (Chen et al., 2012).

1.12.1 Brown planthopper (BPH)

Among the rice pests, the brown planthopper (BPH), *Nilaparvata lugens* Stål L. (Hemiptera: Delphacidae) is one of the major monophagous pests and has affected the cultivated rice (*Oryza sativa*) production in Asia for a long period (Dyck and Thomas 1979, Heong and Hardy 2009).

BPH uses a stylet to feed on phloem of the rice plant and results in the drying of the susceptible plant. Different management practices including the application of insecticides are used, which eventually lead to pesticide resistance and additionally has a hazardous effect on the environment. The deployment of rice varieties resistant to BPH is considered the environmentally and economically friendly approach (Brar et al., 2009). The first report of resistant rice cultivar against the BPH was in 1969 (Pathak et al., 1969). Based on the phenotype of plant-insect interactions, BPH biotypes were identified in different regions around the world (Khush et al., 1985). Biotype 1 exists in East and Southeast Asia; Biotype 2 exists as a dominant biotype in Indonesia, Vietnam, and the Philippines; Biotype 3 was isolated by the International Rice Research Institute; Biotype 4 is present in the Indian subcontinent. Over the course of evolution, selection pressure exerted by resistant cultivars may lead to the emergence of virulent biotypes of BPH (Claridge and Hollander, 1980). With the genomes of rice sequences (Goff et al., 2002, Yu et al., 2002) and BPH (Xue et al., 2014), the application of modern molecular approaches has generated an excellent rice-BPH model for studying co-evolution and insect-crop interaction.

1.12.1.1 Plant-insect interactions in BPH

Studies have reported that the BPH-resistant genes act on the physiology of the insect and affects the survival rate, nymphal periods, oviposition and weight gain (Horgan 2009; Sogawa and Pathak 1970). Comparative analysis was done between the resistant and susceptible lines during BPH feeding, and the results indicated distinct changes in both transcriptome (Lv et al. 2014) and metabolites (Wei et al. 2009; Du et al. 2015) in the host plant. Host defense response against the insect is primarily determined by the secondary metabolites, including glucosinolate and phenolics (Liu et al. 2010; Uawisetwathana et al. 2015; Leiss et al. 2009; Riach et al. 2015;

Jansen et al. 2009) produced from primary metabolism and affecting the growth and development of the plant (Schwachtje and Baldwin 2008).

1.12.1.2 Identification and mapping of Bph resistant genes

Using different rice varieties, a total of 31 BPH-resistant genes have been identified (Cheng et al., 2013, Fujita et al., 2013, Hu et al., 2016, Kobayashi 2016, Yang and Zhang 2016, Wang et al., 2015). Traditional *indica* rice varieties were used to identify 15 resistant genes namely *Bph1-9*, *Bph17*, *Bph19*, *Bph25-26*, *Bph28* and *Bph32*. Seven wild varieties of rice were used to identify other 16 resistant genes (Cheng et al., 2013, Fujita et al., 2013). Using the sequence information and molecular markers, except *Bph5* and *Bph8*, 29 out of 31 identified resistant genes have been mapped in the form of clusters having closely linked loci on different rice chromosomes. Twenty-five genes have been mapped on chromosomes 3, 4, 6, 12; 8 genes including *Bph1*, *bph2*, *bph7*, *Bph9*, *Bph10*, *Bph18*, *Bph21* and *Bph26* have been mapped to chromosome 12L (Fujita et al., 2013, Hu et al., 2016); Five genes including *Bph12*, *Bph15*, *Bph17*, *Bph20*, and *bph22* are mapped to chromosome 4S whereas 3 genes *Bph6*, *bph16* and *Bph27* are mapped to chromosome 4L (Fujita et al., 2013, Hu et al., 2016, Huang et al., 2013); five genes including *Bph3*, *bph4*, *Bph25*, *bph29*, and *Bph32* have been mapped to chromosome 6S (Wang et al., 2015, Myint et al., 2012, Kawaguchi et al., 2001, Jairin et al., 2007).

1.12.1.3 Map-based cloning and characterization of Bph resistant genes

Bph14 is the first BPH-resistant gene to be characterized and encodes for CC-NB-LRR protein (Du et al., 2009) Using map-based cloning approach, 13 BPH-resistant genes have been cloned including *Bph14* (Du et al., 2009), *Bph3* (Du et al., 2009, Liu et al., 2015), *Bph15* (Cheng et al., 2013), *Bph26/2* (Tamura et al., 2014), *bph29* (Wang et al., 2015), *Bph18*

(Ji et al., 2016), *Bph9/1/7/10/21* (Zhao et al., 2016a) and *Bph32* (Ren et al., 2016). Based on the results of gene characterization, most of the BPH-resistant genes encode for receptors which are involved in signaling pathways for plant defense (Jing et al., 2017). *Bph3* and *Bph15* encode for a membrane-localized lectin receptor kinase (LPK) (Liu et al., 2015, Cheng et al., 2013). Upon insect attack, BPH-resistant genes induce resistance through two major phases of a defense response. The first response involves the perception of insect-related herbivore-associated molecular patterns (HAMPs) followed by effector recognition and induction of effector-triggered immunity (ETI) (Jones and Dangl 2006). Studies have reported that deployment of BPH-resistant genes named *Bph14* (CC-NB-LRR) and *Bph15* (LecRK) provides durable resistance against BPH (Li et al., 2006, Hu et al., 2012).

1.12.2 Asian rice gall midge:

The Asian rice gall midge (RGM, *Orseolia oryzae* Wood-Mason) is a gall-forming insect which infests rice. It belongs to the family Cecidomyiidae, order: Diptera. The first instar larvae of RGM feed at the apical meristematic tissue and cause enlargement of host cells (Sain, 1988). It leads to the formation of nutritive tissue; gall tissue insulates larvae and elongates to form tubular gall tissue called as silver shoot. Further growth of the plant is affected leading to the formation of new tillers only which gets infested with more larvae (Sardesai et al., 2001).

Plant resistance plays an important role in controlling damage caused by RGM in rice. Genes conferring resistance to RGM are mostly dominant R genes, except gm3 which is a recessive resistance gene (Kumar et al., 1998). Eleven R genes, designated as Gm1 to Gm11 have been mapped, and seven biotypes of RGM have been identified (Shrivastava et al., 2003, Himabindu et al., 2010, Vijaya Lakshmi et al., 2006). A gene-for-gene relationship was observed; R genes

interact with Avr genes of insect and lead to either hypersensitive reaction (HR) or no HR as its defense response (Rawat et al., 2012, Bentur and Kalode, 1996). Deployment of R gene carrying rice varieties led to the evolution of new biotypes which survive on rice carrying a different set of R genes (Divya et al., 2014). Virulence is controlled by recessive genes in the insect, and gall-inducing effector proteins have been identified (Sinha et al., 2012a, Sinha et al., 2012b, Bentur et al., 2008).

The above brief review of wheat genetics and wheat-Hessian fly, rice-BPH, and rice-RGM forms the background and provides a rationale for the research described in the next chapter. In the wheat-Hessian fly pathosystem, several Avr genes have been cloned in Hessian fly, but corresponding wheat R gene has not been cloned. In rice, several R genes have been cloned from rice, but no Avr genes have been cloned from the BPH. For the rice-RGM system, no R or Avr genes have been isolated. Therefore, the focus of research for this dissertation was the cloning of the H13 R gene in wheat for which the Avr gene was cloned previously (Aggarwal et al., 2014). The plants and insects have been co-evolving with cycles of resistance and susceptibility in the plant and corresponding emerging new avirulence and virulence in the insect biotypes. Any single gene deployed over a long time in a specific geographical region loses its effectiveness. Hence it is important to study a cloned cognate pair of R-Avr in insects to study the mechanism of resistance. The host-pest interactions between wheat-HF can serve as a model system for understanding the molecular mechanisms that allow insects to exploit plants and will help in the discovery of new targets for both resistance and susceptibility in wheat for durable resistance against insects.

1.13 References

- Adkisson, P.L., Dyck, V.A., 1980. Resistant varieties in pest management systems. In: Maxwell, F.G., Jennings, P.R. (Eds.), *Breeding Plants Resistant to Insects*. John Wiley & Sons, New York: 233–252.
- Aggarwal et al., 2014. Avirulence effector discovery in a plant galling and plant parasitic arthropod, the Hessian fly (*Mayetiola destructor*). *PLoS One*, 9: e100958.
- Ahn S, Tanksley SD. 1993. Comparative linkage maps of the rice and maize genomes. *Proc. Natl. Acad. Sci.*, 90: 7980–7984.
- Akbari, et al., 2006. Diversity arrays technology (DART) for high throughput profiling of the hexaploid wheat genome. *Theoretical and Applied Genetics*, 113: 1409–1420.
- Akhunov et al., 2010. Nucleotide diversity maps reveal variation in diversity among wheat genomes and chromosomes. *BMC Genomics*, 11: 702.
- Alexandratos, N. and J. Bruinsma. 2012. World agriculture towards 2030/2050: the 2012 revision. ESA Working paper No. 12-03. Rome, FAO
- Ameline-Torregrosa et al., 2008. Identification and characterization of nucleotide binding site-leucine-rich repeat genes in the model plant *Medicago truncatula*. *Plant Physiology*, 146: 5–21.
- Anderson et al., 1992. Development of a chromosomal arm map for wheat based on RFLP markers. *Theor Appl Genet*, 83: 1035–1043.
- Bai et al., 2002. Diversity in nucleotide binding site-leucine-rich repeat genes in cereals. *Genome Research*, 12: 1871–1884.
- Bakker et al., 2003. Genetic and physical mapping of homologues of the virus resistance gene Rx1 and the cyst nematode resistance gene Gpa2 in potato. *Theoretical and Applied Genetics*, 106: 1524–1531.
- Barnes HF. 1956. Gall midges of economic importance. VII. Cereal crops. Crosby Lockwood & Son London, UK.
- Bateman et al., 1998. Early evolution of land plants: phylogeny, physiology, and ecology of the primary terrestrial radiation. *Annu. Rev. Ecol. Syst.*, 29: 263–292.
- Bendahmane et al., 2002. Constitutive gain-of-function mutants in a nucleotide binding site-leucine-rich repeat protein encoded at the Rx locus of potato. *The Plant Journal*, 32: 195–204.
- Bentur JS, Cheralu C, Rao PRM 2008. Monitoring virulence in Asian rice gall midge populations in India. *Entomol Exp Appl*, 129: 96–106.

- Bentur, J.S., Kalode, M.B., 1996. Hypersensitive reaction and induced resistance in rice against Asian rice gall midge *Orseolia oryzae*. Entomol. Exp. Appl., 78: 77–81.
- Blakeslee, A. F., and Avery, A. G. 1937. Methods of inducing doubling of chromosomes in plants. Journ. of Hered., 28: 393-411.
- Blanco et al., 1998. A genetic linkage map of durum wheat. Theor Appl Genet, 97: 721.
- Bos et al., 2010. A functional genomics approach identifies candidate effectors from the aphid species *Myzus persicae* (green peach aphid). PLoS Genet., 6: e1001216.
- Botstein et al., 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet., 32(3): 314-331.
- Brar et al., 2009. Breeding for resistance to planthoppers in rice. In K. L. Heong & B. Hardy (Eds.), Planthoppers: New threats to the sustainability of intensive rice production systems in Asia. Los Ban~os, Philippines: Int. Rice Res. Inst: 401–427.
- Brown, L. 2004. Outgrowing the earth: the food security challenge in the age of falling water tables and rising temperatures. Earth Policy Institute. W.W. Norton & Co. New York, NY.
- Brown-Guedira et al., 1999. Registration of KS96WGRC38 and KS96WGRC39 tan spot-resistant hard red winter wheat germplasms. Crop Sci. 39: 596.
- Brown-Guedira et al., 2005. Notice of release of KS04WGRC49 hard winter wheat germ plasm with unique glutenin and gliadin proteins, Annu. Wheat Newslett., 51: 190.
- Buntin G. D, Chapin JW. 1990. Biology of Hessian fly (Diptera: Cecidomyiidae) in the southeastern United States: geographic variation and temperature-dependent phenology. J Econ Entomol, 83: 1015-1024.
- Buntin G. D. 1999. Hessian fly (Diptera: Cecidomyiidae) injury and loss of winter wheat grain yield and quality. J Econ Entomol., 92: 1190–1197.
- Buntin, G. D., and R. D. Hudson. 1991. Spring control of the Hessian fly (Diptera: Cecidomyiidae) in winter wheat using insecticides. J. Econ. Entomol., 84: 1913-1919.
- Byers RA and Gallun RL. 1971. The ability of the Hessian fly to stunt winter wheat. I. Effect of larval feeding on the elongation of leaves. J Econ Entomol., 65: 955-958.
- Cambron et al., 2010. Virulence in Hessian fly (Diptera: Cecidomyiidae) field collections from the southeastern United States to 21 resistance genes in wheat. J. Econ. Entomol., 3: 2229–2235.
- Cannon et al., 2002. Diversity, distribution, and ancient taxonomic relationships within the TIR and nonTIR NBS-LRR resistance gene subfamilies. Journal of Molecular Evolution, 54: 548–562.

- Carolan et al., 2009. The secreted salivary proteome of the pea aphid *Acyrtosiphon pisum* characterized by mass spectrometry. *Proteomics*, 9: 2457–2467.
- Cartwright, W. B., R. M. Caldwell, and L. E. Compton. 1959. The response of resistant and susceptible wheat to Hessian fly attack. *Agron. J.*, 51: 529-531.
- Cassman KG. 1999. Ecological intensification of cereal production systems: Yield potential, soil quality, and precision agriculture. *Proc Natl Acad Sci*, 96: 5952-5959.
- Celso et al., 1996. Molecular genetic maps of the group 6 chromosomes of hexaploid wheat (*Triticum aestivum* L. em. *Thell.*). *Genome*, 39(2): 359-366.
- Chalmers et al., 2001. Construction of three linkage maps in bread wheat, *Triticum aestivum*. *Aust J Agric Res*, 52: 1089-1119.
- Chao et al., 1989. RFLP-based genetics maps of wheat homoeologous group 7 chromosomes. *Theor Appl Genet*, 78: 495- 504.
- Chen et al., 2004. A group of related cDNAs encoding secreted proteins from Hessian fly [*Mayetiola destructor* (Say)] salivary glands. *Insect Mol. Biol.*, 13: 101–108.
- Chen et al., 2006. A super-family of genes coding for secreted salivary gland proteins from the Hessian fly, *Mayetiola destructor*. *J. Insect Sci.*, 6: 12.
- Chen et al., 2008. Analysis of transcripts and proteins expressed in the salivary glands of Hessian fly (*Mayetiola destructor*) larvae. *J. Insect Physiol.*, 54: 1–16.
- Chen et al., 2008. Herbivore-induced callose deposition on the sieve plates of rice: an important mechanism for host resistance. *Plant Physiol.*, 146L: 1810–1820.
- Chen et al., 2009. Virulence analysis of Hessian fly populations from Texas, Oklahoma, and Kansas. *J. Econ. Entomol.*, 102: 774–780.
- Chen et al., 2010. Unusual conservation among genes encoding small secreted salivary gland proteins from a gall midge. *BMC Evol. Biol.*, 10: 296.
- Chen et al., 2012. Genetic, Molecular and Genomic Basis of Rice Defense against Insects, *Critical Reviews in Plant Sciences*, 31(1): 74-91.
- Cheng et al., 2013. A rice lectin receptor-like kinase that is involved in innate immune responses also contributes to seed germination. *Plant J*, 76: 687-698.
- Cheng X, Zhu L, He G. 2013. Towards an understanding of molecular interactions between rice and the brown planthopper. *Mol Plant*, 6: 621-634.
- Chisholm et al., 2006. Host-microbe interactions: shaping the evolution of the plant immune response. *Cell*, 124: 803–814.

- Claridge, M. F., & Hollander, J. D. 1980. The “biotypes” of the rice brown planthopper, *Nilaparvata lugens*. *Entomologia experimentalis et applicata*, 27(1): 23–30.
- Clark et al., 2007. Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. *Science*, 317: 338–342.
- Clement SL, Quisenberry SS, eds. 1999. Global Plant Genetic Resources for Insect-Resistant Crops. Boca Raton, FL: CRC Press.: 295.
- Cloutier et al., 2007. Leaf rust resistance gene *Lr1*, isolated from bread wheat (*Triticum aestivum* L.) is a member of the large *psr567* gene family. *Plant Mol. Biol.*, 65: 93–106.
- Collard et al., 2005.) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica*, 142, 169–196.
- Collier et al., 2011. Cell Death Mediated by the N Terminal Domains of a Unique and Highly Conserved Class of NB-LRR Protein. *Molecular Plant-Microbe Interactions: MPMI*, 24 (8): 918–931.
- Conley et al., 2004. A 2600-locus chromosome bin map of wheat homoeologous group 2 reveals interstitial gene-rich islands and colinearity with rice. *Genetics*, 168: 625–637.
- Cooley et al., 2000 Members of the *Arabidopsis* HRT/RPP8 family of resistance genes confer resistance to both viral and oomycete pathogens. *The Plant Cell*, 12: 663–676.
- Couldridge et al., 2007. Exploring plant responses to aphid feeding using a full *Arabidopsis* microarray reveals a small number of genes with significantly altered expression. *Bull. Entomol. Res.*, 97: 523–532.
- Cox and Hatchett 1994. Hessian fly-resistance gene H26 transferred from *Triticum tauschii* to common wheat. *Crop Sci.*, 34: 958–960.
- Cox et al., 1990. Agronomic performance of hexaploid wheat lines derived from direct crosses between wheat and *Aegilops squarrosa*. *Plant Breeding*, 105: 271–277.
- Cox et al., 1994. Leaf rust-resistance genes *Lr41*, *Lr42*, and *Lr43* transferred from *Triticum tauschii* to common wheat. *Crop Sci.* 34: 339–343.
- Cox et al., 1994. Registration of KS92WGRC21 and KS92WGRC22 hard red winter wheat germplasms resistant to wheat spindle-streak mosaic virus, wheat soil borne mosaic virus, and powdery mildew. *Crop Sci.*, 34: 546.
- Cox et al., 1995a. Use of winter wheat $\times$ *Triticum tauschii* backcross populations for germplasm evaluation, *Theor. Appl. Genet.*, 90, 571–577.
- Cox et al., 1995b. Germplasm enhancement in winter wheat $\times$ *Triticum tauschii* backcross populations, *Crop Sci.*, 35, 913–919.

- Cox et al., 1997. Grain yield and breadmaking quality of wheat lines with the leaf rust resistance gene *Lr41*. *Crop Sci.*, 37: 154–161. Not in text
- Cox et al., 1999. Registration of KS96WGRC40 hard red winter wheat germplasm resistant to wheat curl mite, stagonospora leaf blotch, and Septoria leaf blotch. *Crop Sci.*, 39: 597.
- Cox et al., 2017. Comparing two approaches for introgression of germplasm from *Aegilops tauschii* into common wheat. *The Crop Journal*, 5: 355-362.
- Davis et al., 2009. Hessian fly-free date in Kansas: Is it still valid after 70+ years? Poster Session WERA Meeting. Las Vegas, Nevada.
- Dean, G. O., and J. W. McCulloch. 1915. The Hessian fly situation in Kansas. Kansas State Agricultural College, Manhattan, Kansas. Circular No. 37.
- Delaney et al., 1995a. Cytologically-based physical maps of the group-2 chromosomes of wheat. *Theor. Appl. Genet.*, 91: 568–573.
- Delaney, et al., 1995b. Cytologically-based physical maps of the group-3 chromosomes of wheat. *Theor. Appl. Genet.*, 91: 780–782
- Deslandes et al., 2002. Resistance to *Ralstonia solanacearum* in *Arabidopsis thaliana* is conferred by the recessive RRS1-R gene, a member of a novel family of resistance genes. *Proceedings of the National Academy of Sciences*, 99: 2404–2409.
- Devos et al., 1992. RFLP-based genetic map of the homoeologous group 3 chromosomes of wheat and rye. *Theor Appl Genet*, 83: 931-939.
- Diamond, J. 2002. Evolution, consequences and future of plant and animal domestication. *Nature*, 418: 700–707.
- Diehl and Bush. 1984. An evolutionary and applied perspective of Insect biotypes. *Ann. Rev. Entomol.*, 29: 471-504.
- Divya et al., 2014. Cloning of a gene encoding LRR protein and its validation as candidate gall midge resistance gene, Gm4, in rice. *Euphytica*: 1-11.
- Dixon et al., 1996. The tomato Cf-2 disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. *Cell*, 84: 451–459.
- Dixon et al., 1998. The tomato Cf-5 disease resistance gene and six homologs show pronounced allelic variation in leucine-rich repeat copy number. *The Plant Cell*, 10: 1915–1925.
- Dodds et al., 2006. Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proceedings of the National Academy of Sciences*, 103: 8888–8893.

- Dodds, P.N. and Rathjen, J.P. 2010. Plant immunity: towards an integrated view of plant-pathogen interactions. *Nat. Rev. Genet.*, 11: 539–548.
- Doerge R. 2002. Mapping and analysis of quantitative trait loci in experimental populations. *Nature Rev.*, 3: 43-52.
- Du et al., 2009. Identification and characterization of *Bph14*, a gene conferring resistance to brown planthopper in rice. *Proc. Natl Acad. Sci.*, 106: 22163–22168.
- Du et al., 2015. Phloem-exudate proteome analysis of the response to insect brown plant-hopper in rice. *Journal of Plant Physiology*, 183: 13– 22.
- Dubcovsky and Dvorak. 2007. Genome plasticity a Key key factor in the Success of polyploid wheat Under domestication. *Science*, Vol. 316, Issue 5833: 1862-1866.
- Duffey, S.S., Stout, M.J., 1996. Antinutritive and toxic components of plant defense against insects. *Arch. Insect Biochem. Physiol.*, 32: 3–37.
- Dvorák and Zhang. 1990. Variation in repeated nucleotide sequences sheds light on the phylogeny of the wheat B and G genomes. *PNAS*, 87 (24): 9640-9644.
- Dvorák et al., 1993. The evolution of polyploid wheat: identification of the A genome donor species. *Genome*, 36(1): 21-31.
- Dvorak, J. and Chen, K.-C. 1984. Distribution of nonstructural variation between wheat cultivars along chromosome arm 6Bp: Evidence from the linkage map and physical map of the arm. *Genetics*, 106: 325–333.
- Dvorak, J., Luo, M.-C., and Yang, Z.-L. 1998. Restriction fragment length polymorphism and divergence in the genomic regions of high and low recombination in self-fertilizing and cross-fertilizing *Aegilops* species. *Genetics*, 148: 423–434.
- Dyck VA, Thomas B. 1979. The brown planthopper problem. *Brown Planthopper: Threat to Rice production in Asia*; I. R. R. Institute: Los Banos: International Rice Research Institute: 3-17.
- Dyke GV. 1993. John Lawes of Rothamsted. Pioneer of science farming and industry. Harpenden, UK: Hoos Press.
- Eastwood et al., 1991. *Triticum tauschii*: a novel source of resistance to cereal cyst nematode (*Heterodera avenae*). *Aust. J. Agric. Res.*, 42: 69–77.
- Ehrlich PR, Raven PH. 1964. Butterflies and plants: a study in coevolution. *Evolution*, 18: 586–608.
- Ellis, J., Dodds, P. and Pryor, T. 2000. Structure, function, and evolution of plant disease resistance genes. *Current Opinion in Plant Biology*, 3: 278–284.

- Endo TR, Gill BS. 1996. The deletion stocks of common wheat. *J Hered*, 87: 295-307.
- Endo TR. 1988. Induction of chromosomal structural changes by a chromosome of *Aegilops cylindrica* L. in common wheat. *J Hered*, 79: 366-370.
- Erayman M et al., 2004. Demarcating the gene-rich regions of the wheat genome. *Nucleic Acids Res*, 32: 3546–3565.
- F. Gould. 1998. Sustainability of transgenic insecticidal cultivars: Integrating pest genetics and ecology. *Annu. Rev. Entomol.*, 43: 701–726.
- Faris et al., 2003. A bacterial artificial chromosome contig spanning the major domestication locus Q in wheat and identification of a candidate gene. *Genetics* 164: 311–321.
- Feldman M. 1995. Wheat. In: Smartt J, Simmonds NW, eds. *Evolution of crop plants*. Harlow, UK: Longman Scientific and Technical, 185–192.
- Feldman M. 2001. Origin of cultivated wheat. In: Bonjean AP, Angus WJ, eds. *The world wheat book: a history of wheat breeding*. Paris, France: Lavoisier Publishing, 3–56.
- Feldman, M. and E.R. Sears. 1981. The wild gene resources of wheat. *Scientific Amer.*, 244: 102-112.
- Feuillet et al., 2003. Map-based isolation of the leaf rust disease resistance gene Lr10 from the hexaploid wheat (*Triticum aestivum* L.) genome. *Proc. Natl Acad. Sci.*, 100: 15253–15258.
- Flanders and Buntin 2008. Biology and management of Hessian fly in wheat. *Ala. Coop. Ext. Serv. Bull. ANR*, 1069: 4.
- Flor, H. H. 1971. The current status of the gene-for-gene concept. *Ann. Rev. Phytopathol.*, 9: 275–296.
- Food and Agricultural Organization of the United Nations (FAO). 2009. FAOSTAT-Agriculture. <http://faostat.fao.org/site/339/default.aspx> Accessed 2010 June 28.
- Foster et al., 1991. The effectiveness of deploying single gene resistances in wheat for controlling the damage by the Hessian fly (Diptera: Cecidomyiidae). *Environ. Entomol.*, 20: 964–969.
- Fritz RS, Simms EL, eds. 1992. *Plant Resistance to Herbivores and Pathogens: Ecology, Evolution, and Genetics*. Chicago: Univ. Chicago Press.: 590.
- Fritz, A.K., T.S. Cox, B.S. Gill, and R.G. Sears. 1995. Marker-based analysis of quantitative traits in winter wheat by *Triticum tauschii* populations. *Crop Sci.*, 35: 1695-1699.
- Fritz-Laylin et al., 2005. Phylogenomic analysis of the receptor-like proteins of rice and *Arabidopsis*. *Plant Physiology*, 138: 611–623.

- Fu et al., 2009. A novel kinase-START gene confers temperature-dependent resistance to wheat stripe rust. *Science*. 323(5919): 1357–1360.
- Fufa et al., 2005. Genetic improvement trends in agronomic performances and end-use quality characteristics among hard red winter wheat cultivars in Nebraska. *Euphytica*, 144: 187-198.
- Fujita D, Kohli A, Horgan FG. 2013. Rice resistance to planthoppers and leafhoppers. *Crit Rev Plant Sci*, 32: 162-191.
- Futuyma DJ, Agrawal AA. 2009. Macroevolution and the biology diversity of plants and herbivores. *Proc. Natl. Acad. Sci.*, 106: 18054–18061.
- Gagné RJ, Hatchett JH. 1989. Instars of the Hessian fly (Diptera: Cecidomyiidae). *Ann Entomol Soc Am*, 82: 73-79.
- Gahan, A. B. 1933. The serphoid and chalcidoid parasites of Hessian fly. *USDA Misc. Pub.* 174: 1-148: 222-229.
- Gale et al., 1995. Genetic maps of hexaploid wheat. *Proceedings 8th International Wheat Genetic Symposium*. In: Li ZS, Xin ZY (eds). China Agricultural Sciencetech Press, Beijing: 29-40.
- Gallun R. L. 1977. Genetic basis of Hessian fly epidemics. *Ann. N. Y. Acad. Sci.*, 287: 222–29.
- Gallun, R. L., and G. S. Khush. 1980. Genetic factors affecting expression and stability of resistance. In F. G. Maxwell and P. R. Jennings (Eds.), *Breeding Plants Resistant to Insects*. John Wiley: 64–85.
- Gatehouse JA. 2002. Plant resistance towards insect herbivores: a dynamic interaction. *New Phytol*, 156: 145–169.
- Ge et al., 2007. Cloning and functional characterization of a wheat serine/threonine kinase gene (TaSTK) related to salt-resistance. *Plant Science* 173: 55–60.
- Ge et al., 2016. Cloning and function analysis of BAG-family genes in wheat. *Functional Plant Biology* 43.5: 393-402.
- Ghaffary et al., 2012. New broad-spectrum resistance to *Septoria tritici* blotch derived from synthetic hexaploid wheat. *Theor. Appl. Genet.*, 124: 125–142.
- Gill BS, Raupp WJ. 1987. Direct Genetic Transfers from *Aegilops-Squarrosa* L to Hexaploid Wheat. *Crop Science*, 27(3): 445-450.
- Gill et al., 1986. Resistance in *Aegilops squanosa* to wheat leaf rust, wheat powdery mildew, green bug and Hessian fly. *Pl Dis*, 70: 553–556.
- Gill et al., 1991a. A genetic linkage map of *Triticum tauschii* (DD) and its relationship to the D genome of bread wheat (AABBDD). *Genome*, 34: 362-374.

- Gill et al., 1991b. Registration of KS89WGRC4 hard red winter wheat germplasm lines with resistance to Hessian fly, greenbug, and soil borne mosaic virus. *Crop Sci.*, 31: 246.
- Gill et al., 1993. A chromosome region-specific mapping strategy reveals gene-rich telomeric ends in wheat. *Chromosoma*, 102: 374-381.
- Gill et al., 1996a Identification and high-density mapping of gene-rich regions in chromosome group 1 of wheat. *Genetics*, 144: 1883-1891.
- Gill et al., 1996b. Identification and high-density mapping of gene-rich regions in chromosome group 5 of wheat. *Genetics*, 143: 1001-1012.
- Gill et al., 2006. Wheat Genetics Resource Center: the first 25 years. *Adv. Agron.*, 89: 73–136.
- Goff et al., 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science*, 296: 92-100.
- Gomez-Gomez, L. and Boller, T. 2000. FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. *Mol Cell*, 5: 1003–1011.
- Gopher, A., S. Abbo, and S. Lev-Yadun. 2002. The “when”, the “where” and the “why” of the Neolithic revolution in the Levant. *Documenta Prehistorica*, 28: 49-62.
- Grant et al., 1995. Structure of the *Arabidopsis* RPM1 gene enabling dual specificity disease resistance. *Science*, 269: 843–846.
- Gregova, E., Tisova, V., Kraic, J. 1997. “Genetic variability at the Glu-1 loci in old and modern wheat (*Triticum aestivum* L.) cultivated in Slovakia” in *Genetic Resources and Crop Evolution*, 44: 301-306.
- Grube, R.C., Radwanski, E.R. and Jahn, M. 2000. Comparative genetics of disease resistance within the Solanaceae. *Genetics*, 155: 873–887.
- Halterman, D.A. and Wise, R.P. 2004. A single-amino-acid substitution in the sixth leucine-rich repeat of barley MLA6 and MLA13 alleviates dependence on RAR1 for disease resistance signaling. *The Plant Journal*, 38: 215–226.
- Hammond-Kosack, K.E., Jones, D.A. and Jones, J. 1994. Identification of two genes required in tomato for full Cf-9-dependent resistance to *Cladosporium fulvum*. *The Plant Cell*, 6: 361–374.
- Harmel et al., 2008. Identification of aphid salivary proteins: a proteomic investigation of *Myzus persicae*. *Insect Mol. Biol.*, 17: 165–174.
- Harris et al., 2003. Grasses and gall midges: plant defense and insect adaptation. *Annu. Rev. Entomol.*, 48: 549–77.

- Harris et al., 2006. Virulent Hessian fly (Diptera: Cecidomyiidae) larvae induce a nutritive tissue during compatible interactions with wheat. *Ann. Entomol. Soc. Am.*, 99: 305–316.
- Harris et al., 2010. H gene-mediated resistance to Hessian fly exhibits features of penetration resistance to fungi. *Phytopathology*, 100: 279–289.
- Hartl et al., 1993. Use of RFLP markers for the identification of alleles of the Pm3 locus conferring powdery mildew resistance in wheat (*Triticum aestivum* L.). *Theoret. Appl. Genetics*, 86: 959.
- Haseman L. 1930. The Hessian fly larva and its method of taking food. *J Econ Ent*, 23: 316-319.
- Hatchett JH, Starks KJ, Webster JA. 1987. Insect and mite pests of wheat. In *Wheat and Wheat Improvement*. *Agron Mono*, 13: 625-675.
- Heath, M.C. 2000. Hypersensitive response-related death. *Plant Molecular Biology*, 44: 321–334.
- Heinrichs, E. A. 2009. Management of rice insect pests. In: Radcliffe's IPM World Textbook. Radcliffe, E. B., Hutchison, W. D., and Cancelado, R. E., Eds. University of Minnesota, St. Paul, MN.
- Heong and Hardy 2009. Planthoppers: New Threats to the Sustainability of Intensive Rice Production Systems in Asia. Los Ban~os: International Rice Research Institute.
- Heun et al., 1997. The site of Einkorn einkorn Wheat wheat Domestication domestication Identified identified by DNA Fingerprintingfingerprinting. *Science*, 278(5341):1312-1314.
- Hill, C. C. 1953. Parasites of the Hessian fly in the North Central states. *USDA Cir.* 923: 1-15.
- Hill, C. C., J. S. Pinckney, and E. J. Udine. 1939. Status and relative importance of the parasites of the Hessian fly. *J. Agric. Res.* 55: 199 – 213.
- Himabindu et al., 2010. A new rice gall midge resistance gene in the breeding line CR57-MR1523, mapping with flanking markers and development of NILs. *Euphytica*, 174: 179–187.
- Himabindu, K., 2009. Identification, tagging, and mapping of rice gall midge resistance genes using microsatellite markers (Ph.D. thesis). Acharya Nagarjuna University, Hyderabad, India
- Hohmann et al., 1994. Comparison of genetic and physical maps of group 7 chromosomes from *Triticum aestivum* L. *Mol. Gen. Genet.*, 245: 644–653.
- Holden J, Peacock J, Williams T. 1993. *Genes, Crops and the Environment*. Cambridge/New York: Cambridge Univ. Press: 162.

- Horber, E., 1980. Types and classification of resistance. In: Maxwell, F.G., Jennings, P.R. (Eds.), *Breeding Plants Resistant to Insects*. John Wiley & Sons, New York: 15–21.
- Horgan, F. 2009. Mechanisms of resistance: A major gap in understanding planthopper–rice interactions. In K. L. Heong & B. Hardy (Eds.), *Planthoppers: New threats to the sustainability of intensive rice production systems in Asia*. Los Ban~os, Philippines: Int. Rice Res. Inst: 281–302.
- Horton et al., 1945. Wheat-seeding dates and the 73 Hessian fly in Kansas. *Agricultural Experiment Station Bulletin* 59. Kansas State University. Manhattan, Kansas.
- Hossain et al., 2004. A chromosome bin map of 2148 expressed sequence tag loci of wheat homoeologous group 7. *Genetics*, 168: 687–699.
- Howe, G.A., Jander, G. 2008. Plant immunity to insect herbivores. *Annu. Rev. Plant Biol.*, 59: 46–66.
- Howles et al., 2005. Auto active alleles of the flax L6 rust resistance gene induce non-race-specific rust resistance associated with the hypersensitive response. *Molecular Plant-Microbe Interactions*, 18: 570–582.
- Hsam et al., 2001. Significance of *Aegilops tauschii* glutenin genes on breadmaking properties of wheat. *Cereal Chem.* 78: 521–525.
- Hu et al., 2012. Pyramiding and evaluation of the brown planthopper resistance genes *Bph14* and *Bph15* in hybrid rice. *Mol Breed*, 29: 61–69.
- Hu et al., 2016. Cloning and Characterization of TaTGW-7A Gene Associated with Grain Weight in Wheat via SLAF-seq-BSA. *Front Plant Sci.* 20. 7: 1902.
- Hu J, Xiao C, He Y. 2016. Recent progress on the genetics and molecular breeding of brown planthopper resistance in rice. *Rice*, 9: 30.
- Huang et al., 2002a. Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. *Theoretical and Applied Genetics*, 105(5): 699–707.
- Huang et al., 2002b. Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. *PNAS*, 99 (12): 8133–8138.
- Huang et al., 2003. Map-based cloning of leaf rust resistance geneLr21 from the large and polyploid genome of bread wheat. *Genetics*, 164: 655–664.
- Huang et al., 2013. Fine mapping and characterization of *BPH27*, a brown planthopper resistance gene from wild rice (*Oryza rufipogon* Griff.). *Theor Appl Genet*, 126: 219–229.
- Hunter B. 2001. Rage for Grain: Flour Milling in the Mid-Atlantic, 1750–1815. Ph.D. thesis. Univ. Delaware: 322.

- Hurni et al., 2013. Rye Pm8 and wheat Pm3 are orthologous genes and show evolutionary conservation of resistance function against powdery mildew. *Plant J.*, 76: 957–969.
- Hwang et al., 2000. Evidence for a role of the N terminus and leucine-rich repeat region of the Mi gene product in the regulation of localized cell death. *The Plant Cell*, 12: 1319–1329.
- Imtiaz et al., 2008. Characterization of QTLs controlling genetic variation for pre-harvest sprouting in synthetic backcross-derived wheat lines. *Genetics* 178: 1725–1736.
- Inagaki et al., 2014. Some constraints on the interspecific crossing of durum wheat with *Aegilops tauschii* accessions screened under water-deficit stress. *J. Plant Breed. Genet.*, 2: 7–14.
- IWGSC 2014. A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome *Science*, 345(6194): 1251788.
- IWGSC Ref Seq assembly, wheat genome assembly https://urgi.versailles.inra.fr/blast_iwgsc)
- Jairin et al., 2007. Physical mapping of Bph3, a brown planthopper resistance locus in rice. *Maejo Int J Sci Technol*, 1: 166-177.
- Jansen et al., 2009. Metabolomic analysis of the interaction between plants and herbivores. *Metabolomics*, 5(1): 150–161.
- Jantasuriyarat et al., 2004. Identification and mapping of genetic loci affecting the free-threshing habit and spike compactness in wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, 108: 261–273.
- Ji et al., 2016. Map-based cloning and characterization of the BPH18 gene from wild rice Conferring Resistance to Brown Planthopper (BPH) Insect Pest. *Sci Rep*, 6: 34376.
- Jia et al., 1996. RFLP-based maps of the homoeologous group-6 chromosomes of wheat and their application in the tagging of Pm12, a powdery mildew resistance gene transferred from *Aegilops speltoides* to wheat. *Theor Appl Genet.*, 92(5): 559-565.
- Jia et al., 2008. Molecular cloning and characterization of wheat calreticulin (CRT) gene involved in drought-stressed responses. *Journal of Experimental Botany*. 59. 4: 739–751.
- Jiang J, Gill B S. 1994. Nonisotopic in situ hybridization and plant genome mapping: The first 10 years. *Genome*, 37: 717–725.
- Jiang J, Gill B S. 2006. Current status and the future of fluorescence in situ hybridization (fish) in plant genome research. *Genome*, 49: 1057–1068.
- Jing et al., 2017. Genomics of interaction between the brown planthopper and rice. *Current Opinion in Insect Science*. 19: 82–87.
- Jones ET. 1936. Hordeum grasses as hosts of the Hessian fly. *J Econ Entomol*, 29: 704-710.

- Jones ET. 1938. Infestation of grasses of the genus *Aegilops* by the Hessian fly. J Econ Entomol., 31: 333-337.
- Jones JD, Dangl JL. 2006. The plant immune system. Nature, 444: 323-329.
- Jones, D.A. and Jones, J.D.G. 1997. The role of leucine-rich repeat proteins in plant defenses. Advances in Botanical Research, 24: 89–167.
- Jones, M., Brown, T., Allaby, R. 1996. “Tracking early crops and early farmers: the potential of biomolecular archaeology” in D. Harris (ed.), The origins and spread of agriculture and pastoralism in Eurasia, Smithsonian Institution Press, Washington D.C.: 93-100.
- Joppa and Williams. 1982. Registration of Largo, a greenbug resistant hexaploid wheat. Crop Sci., 22: 901–902.
- Joukhadar et al., 2013. Genome-wide association mapping for five major pest resistances in wheat. Mol. Breed., 32: 943–960.
- Kajava, A.V. 1998. Structural diversity of leucine-rich repeat proteins. Journal of Molecular Biology, 277: 519–527.
- Kaloshian I. 2004. Gene-for-gene disease resistance: bridging insect pest and pathogen defense. J Chem Ecol, 30: 2419-2438.
- Kam-Morgan L, Gill B, Muthukrishnan S. 1989. DNA restriction fragment length polymorphisms: a strategy for genetic mapping of D genome of wheat. Genome, 32(4): 724-732.
- Kawaguchi et al., 2001. Assignment of brown planthopper (*Nilaparvata lugens* Stål) resistance gene bph4 to the rice chromosome 6. Breed Sci, 51: 13-18.
- Kerber and Dyck. 1978. Resistance to stem and leaf rust of wheat in *Aegilops squarrosa* and transfer of a gene for stem rust resistance to hexaploid wheat, in: S. Ramanujam (Ed.), Proceedings of the 5th International Wheat Genetics Symposium, February 23–28, Indian Society of Genetics & Plant Breeding, New Delhi, India 1978: 358–364.
- Kerr A. 1987: The impact of molecular genetics on plant pathology. Annu Rev Phytopathol, 25: 87-110.
- Keurentjes, J.J.B., Koornneef, M. and Vreugdenhil, D. 2008. Quantitative genetics in the age of omics. Curr. Opin. Plant Biol., 11: 123–128.
- Khush, G. S. 2005. What it will take to feed 5.0 billion rice consumers in 2030. Plant Mol. Biol., 59: 1–6.
- Khush, G. S., Rezaul-Karim, A. N. M., and Angeles, E. R. 1985. Genetics of resistance of rice cultivar ARC10550 to Bangladesh brown planthopper biotype. J. Genet, 64: 121–125.

- Khush, G. V. 1997. Origin, dispersal, cultivation, and variation of rice. *Plant Mol. Biol.*, 35: 25–34.
- Kielkiewicz M. 2002. Influence of carmine spider mite *Tetranychus cinnabarinus* Boisd. (Acarida: Tetranychidae) feeding on ethylene production and the activity of oxidative enzymes in damaged tomato plants. In *Acarid Phylogeny and Evolution: Adaptation in Mites and Ticks—Proc. IV Symp. Eur. Assoc. Acarol.*, ed. F Bernini, R Nannelli, G Nuzzaci, E de Lillo.: 389–92. Dordrecht, The Netherlands: Kluwer: 452.
- Kihara H. 1944. Discovery of the DD-analyser, one of the ancestors of vulgare wheat. *Agric Hortic*, 19: 889-890.
- Kilian et al., 2007. Independent wheat B and G genome origins in outcrossing *Aegilops* progenitor haplotypes. *Molecular Biology and Evolution*, 24(1): 217-227.
- Knutson et al., 2002. Interaction of parasitoids and resistant cultivars of wheat on Hessian fly, *Mayetiola destructor* Say (Cecidomyiidae). *Southwestern Entomol.*, 27: 1-10.
- Knutson, A. E., E. A. Rojas, D. Marshal, F. E. Gilstrap. 2002. Interaction of parasitoids and resistant cultivars of wheat on Hessian fly, *Mayetiola destructor* Say (Cecidomyiidae). *Southwestern Entomol.* 27: 1-10.
- Kobayashi T. 2016. Evolving ideas about genetics underlying insect virulence to plant resistance in rice–brown planthopper interactions. *J Insect Physiol*, 84: 32-39.
- Kogan, M., Ortman, E., 1978. Antixenosis—a term proposed to define painter’s ‘Nonpreference’ modality of resistance. *Bull. Entomol. Soc. Am.*, 24: 175–176.
- Kohler et al., 2008. Genome-wide identification of NBS resistance genes in *Populus trichocarpa*. *Plant Molecular Biology*, 66: 619–636.
- Kosma et al., 2010. Changes in properties of wheat leaf cuticle during interactions with Hessian fly. *Plant J.*, 63: 31–43.
- Kota et al., 1993. A cytogenetically based physical map of chromosome-1B in common wheat. *Genome*, 36: 548–554.
- Krattinger et al., 2009. A Putative ABC Transporter Confers Durable Resistance to Multiple Fungal Pathogens in Wheat. *Science*, 323(6), 1360–1363.
- Kumar, A., Srivastava, M. N., and Shukla, B. C. 1998. Inheritance and the allelic relationship of gall midge biotype-1 resistant gene(s) in some new donors. *Oryza*, 35: 70–73.
- Kunert et al., 2007. The AB-QTL analysis in winter wheat: I. Synthetic hexaploid wheat (*T. turgidum* ssp. *dicoccoides* × *T. tauschii*) as a source of favorable alleles for milling and baking quality traits. *Theor. Appl. Genet.* 115: 683–695.

- Lagudah et al., 1987. The influence of high-molecular-weight subunits of glutenin from *Triticum tauschii* on flour quality of synthetic hexaploid wheat. *J. Cereal Sci.* 5: 129–138.
- Lan and Yen 1992. An amphidiploid derived from a Chinese landrace of tetraploid wheat, *Ailanmai*, crossed with *Aegilops tauschii* native to China with reference to its utilization in wheat breeding. *J. Sichuan Agric. Univ.* 10: 581–585.
- Lazar et al., 1996. Registration of eight closely related wheat germplasm lines differing in biotype E greenbug resistance. *Crop Sci.*, 36: 1419.
- Leiss et al., 2009. NMR metabolomics of thrips (*Frankliniella occidentalis*) resistance in Senecio hybrids. *Journal of Chemical Ecology*, 35(2): 219–229.
- Lev-Yadun S, Gopher A, Abbo S 2000. The cradle of agriculture. *Science*, 288: 1602–1603.
- Li et al. 2008. Soybean defense responses to the soybean aphid. *New Phytol.*, 179: 185–195.
- Li et al., 2006. Marker-assisted selection for brown planthopper (*Nilaparvata lugens* Stål) resistance genes Bph14 and Bph15 in rice. *Sci Agric Sin*, 39: 2132–2137.
- Li et al., 2014. Synthetic hexaploid wheat enhances variation and adaptive evolution of bread wheat in breeding processes. *J. Syst. Evol.* 52: 735–742.
- Lidell and Schuster 1990. The effectiveness of wheat genes for Hessian fly (Diptera: Cecidomyiidae) resistance in Texas. *J. Econ. Entomol.*, 83: 1135–1139.
- Linkiewicz et al., 2004. A 2500-locus bin map of wheat homoeologous group 5 provides insights on gene distribution and colinearity with rice. *Genetics*, 168: 665–676.
- Liu et al., 2007. Gene expression of different wheat genotypes during an attack by virulent and avirulent Hessian fly (*Mayetiola destructor*) larvae. *J. Chem. Ecol.*, 33: 2171–2194.
- Liu et al., 2010. Revealing different systems responses to brown planthopper infestation for pest susceptible and resistant rice plants with the combined metabonomic and gene-expression analysis. *Journal of Proteome Research*, 9(12): 6774–6785.
- Liu et al., 2014. The Stripe Rust Resistance Gene Yr10 Encodes an Evolutionary-Conserved and Unique CC–NBS–LRR Sequence in Wheat. *Molecular Plant*, 7: 1740–1755.
- Liu et al., 2015. A gene cluster encoding lectin receptor kinase confers broad-spectrum and durable insect resistance in rice. *Nat Biotechnol*, 33: 301–305.
- Liu YG, Tsunewaki K. 1991. Restriction fragment length polymorphism (RFLP) analysis in wheat. II. Linkage maps of the RFLP sites in common wheat. *Jpn J Genet.*, 66(5): 617–633.
- Lobo et al., 2006. Genomic analysis of a 1 Mb region near the telomere of Hessian fly chromosome X2 and avirulence gene *vH13*. *BMC Genomics*, 7: 7.

- Lubbers EL, Gill KS, Cox TS, Gill BS. 1991. Variation of Molecular Markers among Geographically Diverse Accessions of *Triticum-Tauschii*. *Genome*, 34(3): 354-361.
- Lukaszewski, A.J. and Curtis, C.A. 1993. Physical distribution of recombination in B-genome chromosomes of tetraploid wheat. *Theor. Appl. Genet.*, 84: 121–127.
- Lutz et al., 1995. Chromosomal location of powdery mildew resistance genes in *Triticum aestivum* L. (common wheat). 2. Genes *Pm2* and *Pm19* from *Aegilops squarrosa* L, *Heredity* 74: 152–156.
- Lv et al., 2014. BAC and RNA sequencing reveal the brown planthopper resistance gene BPH15 in a recombination cold spot that mediates a unique defense mechanism. *BMC Genomics*, 15(1): 674.
- Ma et al., 2016. TaGS5-3A, a grain size gene selected during wheat improvement for larger kernel and yield. *Plant Biotechnol J.* 14.5:1269-80.
- Mackie et al., 1996. The nucleotide and derived amino acid sequence of an HMW-glutenin gene from *Triticum tauschii* and comparison with those from the D-genome of bread wheat. *J. Cereal Sci.*, 24: 73–78.
- Maestra, B. and T. Naranjo. 1998. Homoeologous relationships of *Aegilops speltoides* chromosomes to bread wheat. *Theor. Appl. Genet.*, 97: 181-186.
- Marais et al., 1994. An assessment of the variation for stem rust resistance in the progeny of a cross involving the *Triticum* species *aestivum*, *turgidum*, and *tauschii*. *S. Afr. J. Plant Soil*, 11: 15–19.
- Marcussen et al., 2014. Ancient hybridizations among the ancestral genomes of bread wheat. *Science*, 345 (6194), 1250092.
- Mares and Mrva. 2008. Genetic variation for quality traits in synthetic wheat germplasm. *Aust. J. Agric. Res.* 59: 406–412.
- Martin, B., Rahbe', Y. and Fereres, A. 2003. Blockage of stylet tips as the mechanism of resistance to virus transmission by *Aphis gossypii* in melon lines bearing the Vat gene. *Ann. Appl. Biol.*, 142: 245–250.
- McFadden and Sears. 1946. The origin of *Triticum spelta* and its free-threshing hexaploid relatives. *J. Hered.*, 37 (1946): 81–89 (107–116).
- Meng et al., 2013. Cloning and characterization of two Argonaute genes in wheat (*Triticum aestivum* L.). *BMC Plant Biol.* 4. 13-18.
- Messmer et al., 1999. Genetic linkage map of a wheat×spelt cross. *Theor Appl Genet*, 98: 1163.

- Meyers et al., 1999. Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *The Plant Journal*, 20: 317–332.
- Meyers et al., 2003. Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. *The Plant Cell*, 15: 809–834.
- Meyers, B.C., Morgante, M., and Michelmore, R.W. 2002. TIR-X and TIR-NBS proteins: two new families related to disease resistance TIR-NBS-LRR proteins encoded in *Arabidopsis* and other plant genomes. *The Plant Journal*, 32: 77–92.
- Michelmore, R.W. and Meyers, B.C. 1998. Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Research*, 8: 1113–1130.
- Mickelson-Young, L., Endo, T.R., and Gill, B.S. 1995. A cytogenetic ladder-map of the wheat homoeologous group-4 chromosomes. *Theor. Appl. Genet.*, 90: 1007–1011.
- Miftahudin et al., 2004. Analysis of expressed sequence tag loci on wheat chromosome group 4. *Genetics*, 168: 651–663.
- Miller et al., 1958. Lipid-soluble pigments of wheat plants as related to Hessian fly infestation. *Plant Physiol.*, 33: 413–416.
- Ming et al., 2008. The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature*, 452: 991–996.
- Mingeot D, Jacquemin JM. 1999. Mapping of RFLP probes characterized for their polymorphism on wheat. *Theor Appl Genet*, 98: 1132–1137.
- Moffett et al., 2002. The interaction between domains of a plant NBS-LRR protein in disease resistance-related cell death. *EMBO Journal*, 21: 4511–4519.
- Mohan et al., 1997. Genome mapping, molecular markers and marker-assisted selection in crop plants. *Mol. Breed.* 3: 87–103.
- Monosi et al., 2004. Full-genome analysis of resistance gene homologues in rice. *Theoretical and Applied Genetics*, 109: 1434–1447.
- Moreno et al., 2011. Aphids secrete watery saliva into plant tissues from the onset of stylet penetration. *Entomol. Exp. Appl.*, 139: 145–153.
- Mujeeb-Kazi et al., 2000. Registration of 10 wheat germplasms resistant to *Septoria tritici* leaf blotch. *Crop Sci.*, 40: 590–591.
- Mujeeb-Kazi et al., 2001. Registration of 10 synthetics hexaploid wheat and six bread wheat germplasms resistant to Karnal bunt. *Crop Sci.*, 41: 1652–1653.

- Mujeeb-Kazi et al., 2008. Rebirth of synthetic hexaploids with global implications for wheat improvement. *Crop Pasture Sci.*, 59: 391–398.
- Mulki et al., 2012. Association mapping for soil-borne pathogen resistance in synthetic hexaploid wheat. *Mol. Breed.*, 31: 299–311.
- Munkvold et al., 2004. Group 3 chromosome bin maps of wheat and their relationship to rice chromosome 1. *Genetics*, 168: 639–650.
- Murphy et al., 1998. Registration of NC96BGTD1, NC96BGTD2 and NC96BGTD3 wheat germplasm resistant to powdery mildew. *Crop Sci.*, 38: 570–571.
- Murphy et al., 1999. Registration of NC97BGTD7 and NC97BGTD8 wheat germplasms resistant to powdery mildew. *Crop Sci.*, 39: 884–885.
- Murugan et al., 2010. Variation of resistance in barley against biotypes 1 and 2 of the Russian wheat aphid (Hemiptera: Aphididae). *J. Econ. Entomol.*, 103: 938–948.
- Myint et al., 2012. Mapping and pyramiding of two major genes for resistance to the brown planthopper (*Nilaparvata lugens* [Stål]) in the rice cultivar ADR52. *Theor Appl Genet*, 124: 495–504.
- Nalam et al., 2006. Map-based analysis of genes affecting the brittle rachis character in tetraploid wheat (*Triticum turgidum* L.). *Theoretical and Applied Genetics*, 112: 373–381.
- Nelson et al., 1995a. Molecular mapping of wheat: major genes and rearrangements in Homoeologous groups 4, 5, and 7. *Genetics*, 141: 721–731.
- Nelson et al., 1995b. Molecular mapping of wheat. Homoeologous group 2. *Genome*, 38: 516–524.
- Nelson et al., 1995c. Molecular mapping of wheat. Homoeologous group 3. *Genome*, 38: 525–533.
- Nesbitt, M. 2001. “Wheat evolution: integrating archaeological and biological evidence” in *Wheat Taxonomy: The Legacy of John Percival*. The Linnean Special Issue 3, The Linnean Society of London, Academic Press, London, GB: 37–59.
- Normile, D. 2008. Reinventing rice to feed the world. *Science*, 321: 330–333.
- Nurnberger et al., 2004. Innate immunity in plants and animals: striking similarities and obvious differences. *Immunol Review*, 198: 249–266.
- Ogbonnaya et al., 2013. Synthetic hexaploids: harnessing species of the primary gene pool for wheat improvement. *Plant Breeding Reviews*, 37: 35–122.
- Okamoto et al., 2013. Identification of quantitative trait loci controlling grain size and shape in the D genome of synthetic hexaploid wheat lines. *Breed. Sci.* 63: 423–429.

- Ozkan, H., A. Brandolini, R. Schafer-Pregl, and F. Salamini. 2002. AFLP analysis of a collection of tetraploid wheat indicates the origin of emmer and hard wheat domestication in southeast Turkey. *Molecular Biology and Evolution*, 19(10): 1797-1801.
- Painter, R.H., 1951. *Insect Resistance in Crop Plants*. The University Press of Kansas, Lawrence.
- Panda N, Khush GS. 1995. *Host Plant Resistance to Insects*. Wallingford, UK: CABI/IRRI.: 431.
- Paterson et al., 2000. Comparative genomics of plant chromosomes. *Plant Cell* ,12: 1523-1539.
- Paterson. 1996. Making genetic maps. In: Paterson AH (ed.) *Genome mapping in plants*, San Diego, California: Academic Press, Austin, Texas: 23–39.
- Pathak MD, Cheng CH, Fortuno ME. 1969. Resistance to *Nephotettix impicticeps* and *Nilaparvata lugens* in varieties of rice. *Nature*, 223: 502-504.
- Pauly PJ. 2002. Fighting the Hessian fly. *Environmental History*, 7: 385-507.
- Pedigo, L. 1999. *Entomology and Pest Management*. Prentice Hall, Upper Saddle River, NJ: 742.
- Pedigo, L.P. 2002. *Entomology and pest management*. 3rd ed. Prentice-Hall, Englewood Cliffs, NJ.
- Peng et al., 2000. Molecular Genetic Maps in Wild Emmer Wheat, *Triticum dicoccoides*: Genome-Wide Coverage, Massive Negative Interference, and Putative Quasi-Linkage. *Genome Research*, 10(10): 1509-1531.
- Peng et al., 2004. Chromosome bin map of expressed sequence tags in homoeologous group 1 of hexaploid wheat and homoeology with rice and *Arabidopsis*. *Genetics*, 168: 609-623.
- Peng et al., 2016. Comparative metabolomics of the interaction between rice and the brown planthopper. *Metabolomics*, 12: 132.
- Periyannan et al., 2013. The gene Sr33, an ortholog of barley Mla genes, encodes resistance to wheat stem rust race Ug99. *Science*, 341(6147): 786-788.
- Poland et al., 2012. Genomic selection in wheat breeding using genotyping-by-sequencing. *The Plant Genome Journal* 5: 103.
- Porter et al., 1989. Registration of TXGH10563B, TXGH10989, and TXGH13622 greenbug-resistant wheat germplasm lines. *Crop Sci.*, 29: 1585–1586.
- Porter, J., P. Jameison, and P. Grace. 2007. *Wheat Production and Global Climate Change*. In *Terrestrial Ecosystems in a Changing World*. Springer Berlin, Heidelberg.
- Qi et al., 2004. A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. *Genetics*, 168: 701-712.

- Qi L, Echaliier B, Friebe B, et al. 2003. Molecular characterization of a set of wheat deletion stocks for use in chromosome bin mapping of ESTs. *Funct. Integr. Genomics*, 3: 39–55.
- Qu et al., 2006. The broad-spectrum blast resistance gene Pi9 encodes a nucleotide-binding site-leucine-rich repeat protein and is a member of a multigene family in rice. *Genetics*, 172: 1901–1914.
- Quisenberry, S.S., Schotzko, D.J., 1994. Integration of plant resistance with pest management methods in crop production systems. *J. Agric. Entomol.*, 11: 279–290.
- Rairdan et al., 2008. The coiled-coil and nucleotide binding domains of the potato Rx disease resistance protein function in pathogen recognition and signaling. *The Plant Cell*, 20: 739–751.
- Rairdan, G. and Moffett, P. 2007. Brothers in arms? Common and contrasting themes in pathogen perception by plant NB-LRR and animal NACHT-LRR proteins. *Microbes and Infection*, 9: 677–686.
- Rairdan, G.J., and Moffett, P. 2006. Distinct domains in the ARC region of the potato resistance protein Rx mediate LRR binding and inhibition of activation. *The Plant Cell*, 18: 2082–2093.
- Ramsey et al., 2007. Genomic resources for *Myzus persicae*: EST sequencing, SNP identification, and microarray design. *BMC Genomics*, 8: 423.
- Randhawa et al., 2004. Deletion mapping of homoeologous group 6-specific wheat expressed sequence tags. *Genetics*, 168: 677–686.
- Ranger et al., 2007. Intraspecific variation in aphid resistance and constitutive phenolics exhibited by the wild blueberry *Vaccinium darrowi*. *J. Chem. Ecol.*, 33: 711–729.
- Rasheed et al., 2014. Genome-wide association for grain morphology in synthetic hexaploid wheat using digital imaging analysis, *BMC Plant Biol.*, 14: 128.
- Rasmann, S., Agrawal, A.A., 2009. Plant defense against herbivory: progress in identifying synergism, redundancy, and antagonism between resistance. *Curr. Opin. Plant Biol.*, 12: 1–6.
- Ratcliffe et al., 2000. Biotype composition of Hessian fly (Diptera: Cecidomyiidae) populations from the southeastern, midwestern, and northwestern United States and virulence to resistance genes in wheat. *J. Econ. Entomol.*, 93: 1319–1328.
- Ratcliffe, R. H., and J. H. Hatchett. 1997. Biology and genetics of the Hessian fly and resistance in wheat. In K. Bondari (ed.). *New Developments in Entomology*: 47–56. Research Signpost, Scientific Information Guild, Trivandrum, India.
- Raupp et al., 1993. Chromosomal location of Hessian fly-resistance genes *H22*, *H23*, and *H24* derived from *Triticum tauschii* in the D genome of wheat. *J. Hered.*, 84: 142–145.

- Raupp et al., 2001. Cytogenetic and molecular mapping of the leaf rust resistance gene *Lr39* in wheat. *Theor. Appl. Genet.*, 102: 347–352.
- Rawat et al., 2016. Wheat *Fhb1* encodes a chimeric lectin with agglutinin domains and a pore-forming toxin-like domain conferring resistance to *Fusarium* head blight. *Nat Genet.* 48. (12): 1576-1580.
- Rawat N et al., 2012. A novel mechanism of gall midge resistance in the rice variety Kavya revealed by microarray analysis. *Funct Integr Genomics*, 12: 249-264.
- Ren et al., 2016. *Bph32*, a novel gene encoding an unknown SCR domain-containing protein, confers resistance against the brown planthopper in rice. *Sci Rep*, 6: 37645.
- Reynolds et al., 2012. Achieving yield gains in wheat. *Plant Cell Environ*, 35(10): 1799-1823.
- Riach et al., 2015. Analysis of plant leaf metabolites reveals no common response to insect herbivory by *Pieris rapae* in three related host–plant species. *Journal of Experimental Botany*, 66(9): 2547–2556.
- Rider et al., 2002. Chromosome landing near avirulence gene *vH13* in the Hessian fly. *Genome*, 45: 812–822.
- Robinson et al., 1960. Chloroplast number in leaves of normal wheat plants and those infested with Hessian fly, or treated with maleic hydrazide. *J. Econ. Entomol.*, 53: 560-567.
- Roder et al., 1998. A microsatellite map of wheat. *Genetics*, 149(4): 2007-2023.
- Rodriguez, P. A., and Bos, J. I. 2013. Toward understanding the role of aphid effectors in plant infestation. *Mol. Plant-Microbe Interact.*, 26: 25–30.
- Rojas, E. A., A. E. Knutson, and F. E. Gilstrap. 2000. The release of *Platygaster hiemalis* Forbes (Platygasteridae) in Texas for biological control of the Hessian fly, *Mayetiola destructor* Say (Cecidomyiidae). *Southwestern Entomol.* 25: 175-184.
- Ryan et al., 2010. The multiple origins of aluminum resistance in hexaploid wheat include *Aegilops tauschii* and more recent cis mutations to TaALMT1. *Plant J.*, 64: 446–455.
- Sacco and Moffett. 2009. Disease resistance genes: Form and function. Chp5 in *Molecular Plant-Microbe Interactions*. 10.1079/9781845935740.0094
- Sain, M., 1988. Studies on varietal resistance and bionomics of rice gall midge, *Orseolia oryzae* (Wood-Mason) (Ph.D. thesis). Osmania University, Hyderabad India: 239.
- Saintenac et al., 2013. Identification of wheat gene *Sr35* that confers resistance to Ug99 stem rust race group. *Science*, 341: 783–786.
- Saintenac et al., 2018. Wheat receptor-kinase-like protein *Stb6* controls gene-for-gene resistance to fungal pathogen *Zymoseptoria tritici*. *Nature Genetics*. 18: 368–374.

- Sandhu, D. and Gill, K.S. 2002. Gene-containing regions of wheat and the other grass genomes. *Plant Physiol.*, 128: 803–811.
- Sardesai et al., 2001. Genetic, physiological and molecular interactions of rice and its major dipteran pest, gall midge. *Plant Cell Tissue Organ Cult*, 64: 115-131.
- Schoonhoven, L.M., Jermy, T., van Loon, J.J.A., 1998. *Insect-Plant Biology: From Physiology to Evolution*. Chapman & Hall, London.
- Schoonhoven, L.M., van Loon, J.J.A. and Dicke, M. 2005. *Insect-Plant Biology*. New York: Oxford University Press.
- Schuster, M. F., and M. C. Lidell. 1990. Distribution and season abundance of the Hessian fly (Diptera: Cecidomyiidae) parasitoids in Texas. *J. Econ. Entomol.* 83: 2269-2273.
- Schwachtje, J., & Baldwin, I. T. 2008. Why does herbivore attack reconfigure primary metabolism? *Plant Physiology*, 146(3): 845–851.
- Sears ER, Sears LMS. 1978. The telocentric chromosomes of common wheat. In: Ramanujam S (ed) *Proceedings of the 5th International Wheat Genetics Symposium*: 389-407.
- Sears ER. 1954. The aneuploids of common wheat. University of Missouri Agricultural Experiment Station, Research Bulletin, 572: 1-58.
- Sears ER. 1966. Chromosome mapping with the aid of telocentrics. *Proceedings of the 2nd International Wheat Genetics Symposium*: 370-381.
- Sehgal et al., 2010. A direct hybridization approach to gene transfer from *Aegilops tauschii* Coss. To *Triticum aestivum* L. *Plant Breeding*, 130(1): 98-100.
- Semagn et al., 2006. Distribution of DArT, AFLP, and SSR markers in a genetic linkage map of a doubled-haploid hexaploid wheat population. *Genome*, 49: 545-555.
- Shen et al., 2003. Recognition specificity and RAR1/SGT1 dependence in barley *Mla* disease resistance genes to the powdery mildew fungus. *The Plant Cell*, 15: 732–744.
- Shirano et al., 2002. A gain-of-function mutation in an Arabidopsis Toll Interleukin1 receptor-nucleotide binding site-leucine-rich repeat type R gene triggers defense responses and results in enhanced disease resistance. *The Plant Cell*, 14: 3149–3162.
- Shrivastava et al., 2003. A new gene for resistance in rice to Asian rice gall midge (*Orseolia oryzae* Wood-Mason) biotype one population at Raipur, India. *Euphytica*, 130: 143–145.
- Simons et al., 2006. Molecular characterization of the major wheat domestication gene Q. *Genetics*, 172: 547–555.
- Singh et al., 2000. Mapping *Yr28* and other genes for resistance to stripe rust in wheat. *Crop Sci.*, 40: 1148–1155.

- Sinha et al., 2012a. Pyrosequencing-based transcriptome analysis of the Asian rice gall midge reveals differential response during the compatible and incompatible interaction. *Int J Mol Sci*, 13: 13079-13103.
- Sinha et al., 2012b. Expression of *Orseolia oryzae* nucleoside diphosphate kinase (OoNDPK) is enhanced in rice gall midge feeding on susceptible rice hosts, and its over-expression leads to salt tolerance in *Escherichia coli*. *Insect Mol Biol*, 21: 593-603.
- Smiley et al., 2004. The economic impact of Hessian fly (Diptera: Cecidomyiidae) on spring wheat in Oregon and additive yield losses with *Fusarium* crown rot and lesion nematode. *J Econ Entomol.* 97:397–408.
- Smith CM. 2005. Plant Resistance to Arthropods: Molecular and Conventional Approaches. Dordrecht, The Netherlands: Springer: 423.
- Smith et al., 1994. Techniques for Evaluating Insect Resistance in Crop Plants. CRC Press, Boca Raton, FL.: 320.
- Smith, C. M., and Clement, S. L. 2012. Molecular bases of plant resistance to arthropods. *Annu. Rev. Entomol.*, 57: 309–328.
- Sogawa, K., & Pathak, M. D. 1970. Mechanisms of brown planthopper resistance in Mudgo variety of rice (Hemiptera: Delphacidae). *Applied Entomology and Zoology*, 5(3): 145–158.
- Somers et al., 2003. Mining single-nucleotide polymorphisms from hexaploid wheat ESTs *Genome*, 46: 431-437.
- Somers, D.J., Isaac, P. and Edwards, K. 2004. A high-density microsatellite consensus map For bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, 109: 1105–1114.
- Stephenson et al., 1998. Fifty new microsatellite loci for the wheat genetic map. *Theor Appl Genet*, 97: 946.
- Steuernagel et al., 2016. Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nature Biotechnology.*, 34: 652-655.
- Stout MJ. 2007. Types and mechanisms of rapidly induced plant resistance to herbivorous arthropods. In *Induced Resistance for Plant Defence*, ed. D Walters, A Newton, G Lyon, pp. 89–107. Oxford: Blackwell.: 271.
- Stout, Michael. 2014. Host-Plant Resistance in Pest Management. *Integrated Pest Management: Current Concepts and Ecological Perspective*. 1-21.
- Stuart JJ, Chen MS, Harris M. 2008. In *Genome mapping and genomics in arthropods* Ed. Hunter W, Kole C Springer Berlin Heidelberg, Germany. Hessian fly: 93-102

- Stuart, J., 2015. Insect effectors and gene-for-gene interactions with host plants. *Curr. Opin. Insect Sci.*, 9: 1–6.
- Takahashi et al., 2002. RCY1, an *Arabidopsis thaliana* RPP8/HRT family resistance gene, conferring resistance to cucumber mosaic virus requires salicylic acid, ethylene, and a novel signal transduction mechanism. *The Plant Journal*, 32: 655–667.
- Tameling et al., 2002. The tomato R gene products I-2 and MI-1 are functional ATP binding proteins with ATPase activity. *The Plant Cell*, 14: 2929–2939.
- Tameling et al., 2006. Mutations in the NB-ARC domain of I-2 that impair ATP hydrolysis cause autoactivation. *Plant Physiology*, 140: 1233–1245.
- Tamura et al., 2014. Map-based cloning and characterization of a brown planthopper resistance gene BPH26 from *Oryza sativa* L. *ssp. Indica* cultivar ADR52. *Sci Rep*, 4: 5872.
- Tang et al., 2008. Effect of HMW-GS 6 + 8 and 1.5 + 10 from synthetic hexaploid wheat on wheat quality traits. *Agric. Sci. China*, 7: 1161–1171.
- Thind et al., 2017. Rapid cloning of genes in hexaploid wheat using cultivar-specific long-range chromosome assembly. *Nature Biotechnology*, 35: 793–796.
- Thompson, J.N. 2005. Coevolution: the geographic mosaic of coevolutionary arms races. *Curr. Biol.*, 15: R992–R994.
- Tiwari et al., 2012. Endosperm tolerance of paternal aneuploidy allows radiation hybrid mapping of the wheat D-genome and a measure of c ray-induced chromosome breaks. *PLoS ONE*, 7: e48815.
- Tiwari et al., 2014. A whole-genome, radiation hybrid mapping resource of hexaploid wheat. *The Plant Journal*, 86, 195–207.
- Trethowan R.M. and Mujeeb-Kazi A. 2008. Novel germplasm resources for improving environmental stress tolerance of hexaploid wheat. *Crop Science*, 48: 1255–1265.
- Trethowan R.M. and van Ginkel M. 2009. Synthetic wheat-an emerging genetic resource. In: Carver B. (ed.). *Wheat - Science and Trade*: 369–386. Wiley-Blackwell, Ames, IO, USA.
- Tuskan et al., 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*, 313: 1596–1604.
- Uauy et al., 2006. A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. *Science* 314, 1298–1301.
- Uawisetwathana et al., 2015. Quantitative ¹H NMR metabolome profiling of Thai Jasmine rice (*Oryza sativa*) reveals primary metabolic response during brown planthopper infestation. *Metabolomics*, 11(6): 1640–1655.

- United Nations, Department of Economic and Social Affairs, Population Division. 2015. World Population Prospects: The 2015 Revision, Key Findings and Advance Tables.
- Van der Biezen, E.A., and Jones, J.D. 1998a. The NB-ARC domain: a novel signaling motif shared by plant resistance gene products and regulators of cell death in animals. *Current Biology*, 8: R226–R227.
- Van Deynze et al., 1995. Molecular-genetic maps for group 1 chromosomes of Triticeae species and their relation to chromosomes in rice and oat. *Genome*, 38: 45-59.
- Van Emden HF. 2007. Host-plant resistance. In *Aphids as Crop Pests*, ed. HF van Emden, R Harrington: 447–68. Wallingford, UK: CABI.: 717.
- Vijaya Lakshmi et al., 2006. A new biotype of the Asian rice gall midge *Orseolia oryzae* (Diptera: Cecidomyiidae) characterized from the Warangal population in Andhra Pradesh, India. *Int. J. Trop. Insect Sci.*, 26: 207–211.
- Vuysteke et al., 1999. Two high-density AFLP linkage maps of *Zea mays* L.: analysis of the distribution of AFLP markers. *Theor. Appl. Genet.*, 99: 921–935.
- Wang et al., 2013. *Aegilops tauschii* single nucleotide polymorphisms shed light on the origins of wheat D-genome genetic diversity and pinpoint the geographic origin of hexaploid wheat. *New Phytol.*, 198: 925–937.
- Wang et al., 2014. Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12:787–796.
- Wang et al., 2015. Map-based cloning and characterization of BPH29, a B3 domain-containing recessive gene conferring brown planthopper resistance in rice. *J Exp Bot*, 66: 6035-6045.
- Wang et al., 2018. Transgenerational CRISPR-Cas9 Activity Facilitates Multiplex Gene Editing in Allopolyploid Wheat. *The CRISPR Journal*, 1(1): 65-74.
- Watson 2005. Hessian fly problems have been increasing in recent years in the Central Plains. *Wheat Farmer Row Crop Farmer* 9: 4–5.
- Wei et al., 2009. Understanding rice plant resistance to the brown planthopper (*Nilaparvata lugens*): A proteomic approach. *Proteomics*, 9(10): 2798–2808.
- Whitham et al., 1994. The product of the tobacco mosaic virus resistance gene N: similarity to toll and the interleukin-1 receptor. *Cell*, 78: 1101–1115.
- Whitworth et al., 2009. Hessian fly. Kansas State University, MF-2866. Manhattan, Kansas.
- Wilde et al., 2001. Seed treatment for control of wheat insects and its effect on yield. *J. Agric. Urban Entomol.*, 18: 1-11.

- Wilde, G., 2002. Arthropod host plant resistant crops. In: Pimentel, D. (Ed.), *Encyclopedia of Pest Management*. CRC Press, Boca Raton, FL: 33–35.
- Wilhoit, L. R. 1992. Evolution of virulence to plant resistance: Influence of variety mixtures. In: R. S. Fritz and E. L. Simms (Eds.), *Plant Resistance to Herbivores and Pathogens: Ecology, Evolution, and Genetics*, University of Chicago Press, Chicago: 91–119.
- Will et al., 2007. Molecular sabotage of plant defense by aphid saliva. *Proc. Natl. Acad. Sci.*, 104: 10536–10541.
- Williams and Sorrells 2014. Three-dimensional seed size and shape QTL in hexaploid wheat (*Triticum aestivum* L.) populations. *Crop Sci.*, 54: 98–110.
- Wiseman, B.R., 1994. Plant resistance to insects in integrated pest management. *Plant Dis.*, 78: 927–932.
- Wu, J., Baldwin, I.T. 2010. New insights into plant responses to the attack from insect herbivores. *Annu. Rev. Genet.*, 44: 1–24.
- Wulff et al., 2001. Domain swapping and gene shuffling identify sequences required for induction of an Avr-dependent hypersensitive response by the tomato Cf-4 and Cf-9 proteins. *The Plant Cell*, 13: 255–272.
- Xiang et al., 2006. Xa3, conferring resistance for rice bacterial blight and encoding a receptor kinase-like protein, is the same as Xa26. *Theoretical and Applied Genetics*, 113: 1347–1355.
- Xie et al., 1993. RFLP-based genetic maps of the homoeologous group 5 chromosomes of bread wheat (*Triticum aestivum* L.). *Theor Appl Genet.*, 87(1-2): 70–74.
- Xu et al., 2006. Registration of two synthetic hexaploid wheat germplasms resistant to Hessian fly. *Crop Sci.*, 46: 1401–1402.
- Xue et al., 2014. Genomes of the rice pest brown planthopper and its endosymbionts reveal complex complementary contributions for host adaptation. *Genome Biol*, 15: 1.
- Yahiaoui et al., 2004. Genome analysis at different ploidy levels allows cloning of the powdery mildew resistance gene Pm3b from hexaploid wheat. *Plant J.*, 37: 528–538.
- Yan et al., 2003. Positional cloning of the wheat vernalization gene Vrn1. *Proc. Natl. Acad. Sci.* 100:6263–6268.
- Yan et al., 2004. The wheat Vrn2 gene is a flowering repressor down-regulated by vernalization. *Science* 303: 1640–1644.
- Yan et al., 2006. The wheat and barley vernalization gene Vrn3 is an orthologue of FT. *Proc. Natl. Acad. Sci.* 103: 19581–19586.

- Yang et al., 2000. Direct transfer of powdery mildew resistant gene from *Aegilops tauschii* into bread wheat and its genetic expression, Southwest China. J. Agric. Sci., 13: 15–20.
- Yang et al., 2006. Genetic detection of PPO gene in synthetic hexaploid wheat-derived populations. Mol. Plant Breed., 4: 772–778.
- Yang et al., 2009. Synthetic hexaploid wheat and its utilization for wheat genetic improvement in China, J. Genet. Genomics, 36: 539–546.
- Yang et al., 2013a. Cloning and characterization of endo- β -1,4-glucanase genes in the common wheat line three pistils. Genet Mol Biol. 36 3:400-407.
- Yang et al., 2013b. Rapidly evolving R genes in diverse grass species confer resistance to rice blast disease. Proc Natl Acad Sci, 110: 18572-18577.
- Yang L, Zhang W. 2016. Genetic and biochemical mechanisms of rice resistance to planthopper. Plant Cell Rep, 35(8): 1559- 1572.
- Yencho GC, Cohen MB, Byrne PF. 2000. Applications of tagging and mapping insect resistance loci in plants. Annu. Rev. Entomol., 45: 393–422.
- Yim et al., 2002. Characterization of three maize bacterial artificial chromosome libraries toward anchoring of the physical map to the genetic map using high-density bacterial artificial chromosome filter hybridization. Plant Physiol., 130: 1686–1696.
- Yu et al., 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science, 296: 79-92.
- Zhang 2007. Strategies for developing green super rice. Proc. Natl. Acad. Sci., 104: 16402–16409.
- Zhang et al., 2003. A gain-of-function mutation in a plant disease resistance gene leads to constitutive activation of downstream signal transduction pathways in suppressor of npr1-1, constitutive 1. The Plant Cell, 15: 2636– 2646.
- Zhang et al., 2004. BAC-FISH in wheat identifies chromosome landmarks consisting of different types of transposable elements. Chromosoma, 112(6): 288-299.
- Zhang et al., 2017. Isolation and characterization of the TaSnRK2.10 gene and its association with agronomic traits in wheat (*Triticum aestivum* L.). PLoS One.12. 3:e0174425.
- Zhao et al., 2016a. Allelic diversity in an NLR gene BPH9 enables rice to combat planthopper variation. Proc Natl Acad Sci, 113: 12850-12855.
- Zhao et al., 2016b. Avirulence gene mapping in the Hessian fly (*Mayetiola destructor*) reveals a protein phosphatase 2C effector gene family. J Insect Physiol., 84: 22-31.

- Zhou et al., 2004. Genome-wide identification of NBS genes in japonica rice reveals the significant expansion of divergent non-TIR NBS-LRR genes. *Molecular Genetics and Genomics*, 271: 402–415.
- Zhou et al., 2009. Silencing OsHI-LOX makes rice more susceptible to chewing herbivores, but enhances resistance to a phloem feeder. *Plant J.*, 60: 638–648.
- Zipfel et al., 2006. Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts *Agrobacterium*-mediated transformation. *Cell*, 125: 749–760.
- Zohary, D. and M. Hopf. 1994. “The origin and spread of cultivated plants in West Asia, Europe, and the Nile valley.” *Domestication of plants in the Old World* 2nd ed. Oxford Univ. Press, New York.
- Zohary, D., and M. Hopf. 2000. *Domestication of plants in the Old World*. 3rd ed. Oxford University Press, Oxford.

Chapter 2 - Map-based cloning of the Hessian fly resistance gene

H13 in Wheat

2.1 Abstract

H13, a dominant resistance gene transferred from *Aegilops tauschii* into wheat (*Triticum aestivum*), confers a high level of antibiosis against a wide range of Hessian fly (HF, *Mayetiola destructor*) biotypes. Previously, *H13* was mapped to the distal arm of chromosome 6DS, where it is flanked by markers *Xcfd132* and *Xgdm36*. A mapping population of 1,368 F₂ individuals derived from the cross: PI372129 (*h13h13*) / PI562619 (Molly, *H13H13*) was genotyped and *H13* was flanked by *Xcfd132* at 0.4cM and by *Xgdm36* at 1.8cM. Screening of BAC-based physical maps of chromosome 6D of Chinese Spring wheat and *Ae. tauschii* coupled with high resolution genetic and Radiation Hybrid mapping identified nine candidate genes co-segregating with *H13*. Candidate gene validation was done on an EMS-mutagenized TILLING population of 2,296 M₃ lines in Molly. Twenty seeds per line were screened for susceptibility to the *H13*-virulent HF GP biotype. Sequencing of candidate genes from twenty-eight independent susceptible mutants identified three nonsense, and 24 missense mutants for *CNL-1* whereas only silent and intronic mutations were found in other candidate genes. 5' and 3' RACE was performed to identify gene structure and CDS of *CNL-1* from Molly (*H13H13*) and Newton (*h13h13*). Increased transcript levels were observed for the *H13* gene during incompatible interactions at larval feeding stages of GP biotype. The predicted coding sequence of the *H13* gene is 3,192 bp consisting of two exons with 618 bp 5'UTR and 2,260 bp 3'UTR. It translates into a protein of 1063 amino acids with an N-terminal Coiled-Coil (CC), a central Nucleotide-Binding adapter shared by APAF-1, plant R and CED-4 (NB-ARC) and a C-terminal Leucine-

Rich Repeat (LRR) domain. Conserved domain analysis revealed shared domains in Molly and Newton, except for differences in sequence, organization and number of LRR repeat in Newton. Also, the presence of a transposable element towards the C terminal of *h13* is indicative of interallelic recombination, recent tandem duplications and gene conversions in the *CNL* rich region near *H13* locus. Comparative analysis of candidate genes in the *H13* region indicates that gene duplications in *CNL* encoding genes during divergence of wheat and barley led to clustering and diversity. This diversity among *CNL* genes may have a role in defining differences in the recognition specificities of NB-LRR encoding genes. Allele mining for the *H13* gene in the core collection of *Ae. tauschii* and hexaploid wheat cultivars identified different functional haplotypes. Screening of these haplotypes using different HF biotypes would help in the identification of the new sources of resistance to control evolving biotypes of HF. Cloning of *H13* will provide perfect markers to breeders for HF resistance breeding programs. It will also provide an opportunity to study R-*Avr* interactions in the hitherto unexplored field of insect-host interaction.

2.2 Introduction

Wheat provides about 20% of the calories consumed worldwide and is grown on more land than any other crop plant (<http://www.fao.org/resources/infographics/infographics-details/en/c/240943/>). The world's population has already surpassed 7.4 billion (<http://www.census.gov>) and is estimated to reach 9.15 billion by 2050 (Alexandratos, N. and J. Bruinsma. 2012). To keep pace with this population increase, the wheat production needs to be increased by at least 2% every year to ensure sustainable agriculture (Alexandratos, N. and J. Bruinsma 2012). It will be challenging to meet yield projections of wheat with the scarcity of water, nutrients, and arable crop land along with climatic changes and biotic factors. Biotic factors include different insects and disease-causing pathogens which affect the yield and end-use-quality of wheat. There has been tremendous progress in the identification of new genes and alleles in wheat for various biotic and abiotic factors, and these advances are allowing the breeders to develop resistant cultivars to address key challenging issues.

Along with plant diseases and abiotic stresses, wheat production is adversely affected by insect pests. Hessian fly (HF) (*Mayetiola destructor*), a gall-forming insect, is one of the most important pests of wheat. Both wheat and HF are believed to have originated in the Fertile Crescent and have coevolved since ancient times. HF causes widespread economic losses to both bread wheat (*Triticum aestivum* L., AABBDD, $2n = 6x = 42$) and durum wheat (*T. turgidum* L. *subsp. durum*, AABB, $2n = 4x = 28$) crops (Barnes 1956). Multiple generations of HF can affect wheat during the same growing season (Buntin 1999; Cambron et al., 2010; Flanders et al., 2008; Pedigo 2002). The HF infestation can affect wheat yield by resulting in stunted growth, broken culms, shriveled seeds and reduced grain filling (Byers & Gallun 1971; Buntin & Chapin 1990). For HF management, cultural practices such as planting after the fly-free date, crop rotation, and

destruction of volunteer wheat are followed to reduce HF infestations. However, the use of HF resistant wheat cultivars is the most effective, environmentally favorable and cost-efficient approach to HF management (Ratcliff & Hatchet 1997).

There are more than 35 genes known which provide resistance against HF (McIntosh et al., 2012). Among the HF resistance (HFR) genes, eight were identified in hexaploid wheat; seventeen were derived from tetraploid wheat, eight from diploid *Ae. tauschii* (DD, $2n = 2x = 14$), two from rye (RR, $2n = 2x = 14$), one each from *Ae. ventricosa* ($D^vD^vM^vM^v$, $2n = 4x = 28$), and *Ae. triuncialis* (CCUU, $2n = 4x = 28$) (Noble et al., 1943; Caldwell et al., 1946; Suneson et al., 1950; Shands and Cartwright 1953; Gallun and Patterson 1977; Oellermann et al., 1983; Allan et al., 1959; Stebbins et al., 1982, 1983; Mass et al., 1989; Patterson et al., 1988; Obanni et al., 1988, 1989; Amri et al., 1990; Cebert et al., 1996; Williams et al., 2003; Liu et al., 2005a, 2005b; Gill et al., 1987; Zhao et al., 2006; Gill et al., 1991; Wang et al., 2006; Sardesai et al., 2005; Cainong et al., 2010; Friebe et al., 1991; Delibes et al., 1997; Martin et al., 2003). Using monosomic analysis and simple sequence repeat (SSR) markers, these HFR genes were mapped on different chromosomes across the A, B, D genomes of wheat (Table 2.1). Among the HFR genes mapped on D genome of wheat (Figure 2.1), *H22* is mapped on chromosome 1D; *H24*, *H26*, and *H32* are mapped on chromosome 3D; *H27* is mapped on chromosome 4D; *H7* and *H8* are mapped on chromosome 5D; *H13*, *H23*, *HWGRC4* and *HNC09MDD14* are mapped to chromosome 6D (Zhao et al., 2006; Gill et al., 1991; Wang et al., 2006; Sardesai et al., 2005; Delibes et al., 1997; Amri et al., 1990; Liu et al., 2005; Miranda et al., 2010). Different HFR genes provide resistance against different biotypes of HF. Interestingly, many HFR genes are present in clusters on chromosome 1AS and 6DS. No HFR gene has been cloned yet.

As planting resistant varieties provide the best agronomic strategy for the sustainable agricultural production, many breeders have deployed different HFR genes in elite cultivars. In 1955, the first HFR gene deployed in hexaploid wheat was *H3* in wheat cultivar Dual (Caldwell et al., 1957). Additionally, different HFR genes such as *H5*, *H6*, *H9*, *H13*, *H21*, *H25*, *H26* and others have been deployed for successful HF management (Cainong et al. 2010; Cambron et al. 2010; Chen et al. 2009a; Johnson et al. 2009; Ratcliffe 2012).

Cloning of HFR genes will be valuable for efficient introgression into different cultivars using marker-assisted selection (MAS) and allele mining which would help in identification of other known functional resistant alleles of the gene. In response to selection pressure exerted by the HFR genes, *Avr* genes from HF constantly evolve to evade host plant resistance resulting in an increase of virulent biotypes and the breakdown of HFR gene (Ratcliffe and Hatchett 1997). Over the time, the effectiveness of different HFR genes *H3*, *H5* and *H6* was lost due to the adaptability of flies which rendered them virulent (Gould 1998; Foster et al., 1991). Due to differences in survival of different biotypes on resistant wheat, there exists diversity in HF population in a given geographical area (Ratcliffe et al. 2000; Naber et al. 2003; Chen et al. 2009b). Pyramiding of different HFR genes may provide effective and durable resistance (Porter et al., 2009). R-Avr interaction studies between wheat and Hessian fly could serve as a good model system for understanding the underlying mechanisms of resistance and pathogenicity.

H13 is a dominant resistant gene which confers resistance against a variety of HF biotypes (Hatchett et al. 1981; Gill et al. 1987; Ratcliffe and Hatchett 1997; El Bouhssini et al. 1999). *H13* was introgressed into hexaploid wheat from *Ae. tauschii*. Synthetic amphiploids TA3386 (KU221-14) and TA3387 (KU221-19) were developed from an interspecific cross between

diploid *Ae. tauschii* TA2452 (KU2076) and tetraploid *T. durum* Desf. 'Gulab' (KU134) and *T. persicum* L. var. *stramineum* Zhuk (KU138) respectively at Kyoto University, Japan (Tanaka, M. 1961, Hatchett, J. H 1981). Single dominant gene resistance of *H13* was determined by making a cross between synthetic hexaploids TA 3386 and TA3387 with susceptible wheat cultivars Amigo and Eagle. As expected for a single gene resistance, segregation of resistant and susceptible plants in F₂ population was in the ratio of 3 R:1 S when infested with HF biotype D (Hatchett, J. H 1981). When the resistant F₂ plants were progressed to F₃ generation, the ratio of homozygous resistant to segregating lines was 1:2, as expected for a resistance controlled by a single dominant gene (Hatchett, J. H 1981). Hessian fly resistant hard red winter wheat germplasm KS81H164OHF (CI 17960) which was bulk of seeds from homozygous HF resistant F₂ plants obtained from the cross KU221-14/ 'Eagle' // NE73640 ('Bennett' sib) / three/ 'Cheney' was released (Martin T. J. et al. 1982).

Homozygous resistant progeny obtained from the cross between amphiploid and cultivar Eagle were backcrossed six times with susceptible hard red winter wheat cultivar Newton TA 3002, followed by a few generations of selecting the homozygous resistant seeds by screening with HF biotype D and germplasm Molly (PI 562619) homozygous for *H13* gene was released (Patterson, F. L. 1994). For Molly, the name was selected beginning with the letter M (13th letter) which corresponds to the 13 in *H13* (Patterson, F. L. 1994). When screened for HF resistance with four HF biotypes B, C, D and L, Molly was resistant to all the biotypes (Patterson, F. L. 1994). For chromosomal localization of *H13*, crosses were made between KS81H164OHF and seven (1D to 7D) D-genome monosomics (2n=41). The F₁ plants carrying 41 chromosomes identified cytologically, were selfed and F₂ plants from each of 1D to 7D monosomic hybrids were screened for resistance to Hessian fly biotype D (Gill et al., 1987). Segregation ratio of resistant

and susceptible plants in F₂ monosomic families was in the ratio of 3:1 for six non-critical families 1D to 5D, and 7D whereas the critical 6D monosomic family had more than 90% resistant plants, and a few susceptible plants that were nullisomics for 6D (Gill et al., 1987). Therefore, *H13* was mapped to chromosome 6D. For arm localization of *H13*, the telocentric analysis was done. Crosses were made between KS81H164OHF (*H13H13*) and telosomic stocks 6DS (TA3128) and 6DL (TA3129) (Gill et al., 1987) and *H13* was mapped to chromosome arm 6DL.

Liu et al. 2005b undertook molecular mapping of *H13* using SSR markers. Polymorphism survey using 6D-specific SSR markers showed that none of the SSR markers from 6DL were polymorphic between Molly and near-isogenic line Newton. Instead, SSR markers from 6DS were polymorphic and were used to map the *H13* gene. Five codominant SSR markers were linked to *H13*. *Xcfd132* was co-segregating with *H13*, and *Xgdm36* was mapped 2.7cM distal to *H13* in F₂:3 families (Liu et al., 2005b). For physical mapping of *H13* gene, these polymorphic markers from 6DS were tested on deletion lines of the short arm of chromosome 6D including TA4544 del6DS-2-0.45 (breakpoint was at fraction length of 45% from the centromere), del6DS-4-0.79, and del6DS-6-0.99. The SSR markers *Xcfd42* and *Xgdm141* which are distal to *Xgdm36* did not amplify in del6DS-6-0.99 whereas *Xgdm36* and *Xcfd132* amplified in del6DS-6-0.99, so *H13* was mapped proximal to the breakpoint of del6DS-6-0.99 (Liu et al., 2005b).

Avirulence to *H13* is due to the recessive *vH13* gene which is the first *Avr* gene to be cloned in insects (Aggarwal et al., 2014). It is transcribed in salivary glands of *H13*-avirulent larvae. In *H13*-virulent larvae from field populations, two independent insertions of lengths 4.7kb and 254bp were found at the exon-intron junctions, whereas another insertion of 461 bp was found in the coding region of the second exon. (Aggarwal et al., 2014). the *vH13* sequence is unique, and

neither homologs nor conserved domains were identified with any protein in the NCBI database. (Aggarwal et al., 2014). Cloning of *H13* would further aid in the understanding of direct or indirect interactions between effector proteins and R gene.

2.2.1 Objective of study

The main objectives of this study were: i) Fine mapping of *H13* gene ii) Identification of candidate genes iii) Generation of TILLING population and functional validation of candidate genes iv) Generation of full length genomic and cDNA of *H13* gene v) Expression analysis of *H13* gene at different time points from both feeding site and leaf tissue vi) Allele mining and haplotype analysis to identify allelic variants present in *Ae. tauschii* core collection, resistant *Ae. tauschii* accessions and different cultivars.

2.3 Material and Methods

2.3.1 Plant materials and mapping populations

Information about all the plant materials used in this study is provided in Table 2.2. A mapping population of 1368 F₂ lines was developed by selfing of F₁ individuals from a cross between PI372129 (Turcikum57; *h13h13*; susceptible) and PI562619 (Molly; *H13H13*; resistant). The F₂ population was grown in the greenhouse and advanced to F₃ generation. All the primers used in this study are listed in Table 2.10.

2.3.2 Linked SSR marker analysis

Leaf tissue from plant materials was collected in microtubes in the 96-well box. DNA was isolated from lyophilized leaf tissue using QiagenBioSprint 96 workstation with Qiagen

BioSprint 96 DNA Plant Kit (Qiagen, Valencia, CA) as per the manufacturer's instructions. DNA quantification was done on NanoDrop spectrophotometer, and all samples were normalized to the concentration of 25 ng/ μ L. The F₂ population was genotyped with previously identified linked co-dominant microsatellite markers *Xgdm36* (2.7cM distal to *H13*) and co-segregating *Xcfd132* (Liu et al., 2005). Primers for polymorphic marker *Xcfd132* amplified a 160-bp size product from the resistant parent (Molly) and a 120-bp segment from the susceptible parent (Turcikum57). Primers for polymorphic marker *Xgdm36* amplified a 170-bp product from Molly and 130-bp from Turcikum57.

PCR cocktail per reaction was as follows: 10X reaction buffer - 1.8 μ L, 50mM MgCl₂ - 0.6 μ L, 2.5mM dNTPs - 1.8 μ L, 4 μ M primer - 1 μ L each of forward and reverse, ddH₂O - 8.75 μ L, Taq polymerase (Bioline) - 0.05 μ L (0.25 U) and DNA template - 3 μ L. Thermocycler conditions steps included initial denaturation at 96°C for 5 minutes, followed by 35 cycles of 3 steps- denaturation at 95°C for 45 seconds, annealing at primer-specific T_m for 1 minute and extension at 72°C for 1 minute, followed by final extension of 10 minutes at 72°C, performed on a BioRad thermocycler (BioRad, Hercules, CA, USA). PCR products were visualized on 3% high-resolution agarose gels in 1X TBE containing ~0.1 μ g/mL of ethidium bromide in the final volume. Genetic linkage map analysis was performed using Inclusive composite interval mapping software (IciMapping 3.3). The high-resolution (HR), F₂ mapping population, was identified for recombinants between markers *Xcfd132* and *Xgdm36*

2.3.3 Hessian fly infestation screening

Thirty F₃ seeds of each recombinant F₂ plant along with Molly, Caldwell, and Carol as resistant checks whereas Karl92 as susceptible check were evaluated for reaction to Hessian fly

infestation. Laboratory stocks of biotype Great Plains (GP) of Hessian fly were used for infestation. Thirty seeds of each line along with the controls were sown in equally spaced furrows in flats of 1:1 mixture of soil and Metromix in the greenhouse. Seedlings were fertilized on the fifth day. At 1.5 leaf stage, seedlings were infested with GP flies and 17 days post-infestation, the observed phenotype was recorded as susceptible, and resistant plants. Plants with stunted growth, dark green leaves, missing internodes between leaves and presence of 3rd instar live HF larvae were scored as susceptible plants. Plants with normal growth, light green leaves, the presence of internodes between leaves and dead 1st instar larvae were scored as HF resistant plants.

2.3.4 PCR-based screening of 3-D BAC pools of chromosomes 1D, 4D, 6D

Wheat chromosomes were flow sorted, and fraction one containing chromosomes 1D, 4D and 6D were used to develop BAC libraries for fingerprinting. One hundred thirty 3-dimensional BAC pools were developed from minimal tiling path (MTP) for chromosomes 1D, 4D, and 6D fingerprinted contig (FPC) assembly of Chinese spring (Sehgal et al., 2012).

BAC pools were screened for flanking simple sequence repeat markers (*Xcfd132*, *Xcfd213*, *Xgdm36*). *Xcfd213* is 1.1cM proximal to *Xcfd132*. PCR reactions were performed using the same conditions as mentioned before. PCR products were visualized on 3% high-resolution agarose gels in 1X TBE containing ~ 0.1µg/mL of ethidium bromide in the final volume. Positive BAC pool data of row pools, column pools, and plate pools were deconvoluted for *Xcfd132*, *Xcfd213*, *Xgdm36* markers to identify candidate BAC clones and their corresponding contigs from minimal tiling path (MTP) for chromosomes 1D, 4D, and 6D fingerprinted contig (FPC) assembly of Chinese spring.

2.3.5 Anchoring of *Xgdm36* on the genetic and physical map of *Ae. tauschii*

Markers from physical and genetic maps of *Ae. tauschii* (Luo et al., 2013) were also mapped on minimal tiling path (MTP) for chromosomes 1D, 4D, and 6D fingerprinted contig (FPC) assembly of Chinese spring (Sehgal et al., 2012). The contigs which were identified for *Xcfd132*, *Xcfd213*, *Xgdm36* SSR markers by PCR based screening of 3-D BAC pools of chromosomes 1D, 4D, 6D (CS) were screened for *Ae. tauschii* markers. *Ae. tauschii* marker AT6D5315 and SSR marker *Xgdm36* were mapped on different BAC clones of the same contig. Locus markers which were 15cM distal as well as proximal to AT6D5315 in the genetic and physical map of *Ae. tauschii* (Luo et al., 2013) were selected to design D-genome specific primers.

2.3.6 Designing D Genome-specific primers for markers

Ae. tauschii marker sequences were used for BLASTn tool on International Wheat Genome Sequencing Consortium (IWGSC; https://urgi.versailles.inra.fr/blast_iwgsc) site to identify homeologs for each marker in A, B, D genomes of wheat. To design D-genome specific markers, primers were designed using polymorphic sites among A, B and D genome homeologs. Primers were selected manually with polymorphism at 3' end, and their parameters were checked with OligoAnalyzer tool on Integrated DNA Technologies (IDT; <https://www.idtdna.com/calc/analyzer>) site for T_m mismatch, the presence of hairpin and the formation of self-dimers to ensure specific binding of primers to the target template. Primers were picked from 18 to 28 bases in length and melting temperature between 54°C to 66°C. The preferred amplicon length was in the range of 800 bp to 1100 bp.

The genome specificity and localization of markers on the chromosomal arm and deletion bin was confirmed by testing on the genetic stocks Nullisomic-6D, dt6DS, dt6DL, and deletion lines 6DS-2, 6DS-4, 6DS-6 (Sears 1954; Sears 1966; Sears and Sears 1978; Endo 1988; Endo and Gill 1996). Additionally, these markers were amplified from Molly, Newton, and Turcikum57 and tested for polymorphism. PCR was performed with a touchdown profile from either of TD- 65-60°C, TD- 63-58°C or TD- 60-55°C depending upon T_m of primer pairs. Thermocycler conditions steps included initial denaturation at 96°C for 5 minutes followed by 5 cycles of 3 steps- denaturation at 95°C for 1 minute, annealing at primer-specific T_m for 1 minute with a decrease of 1°C per cycle and extension at 72°C for 1 minute 15 seconds, followed by 30 cycles of 3 steps- denaturation at 95°C for 45 seconds, annealing at primer-specific T_m for 1 minute and extension at 72°C for 1 minute 15 seconds and a final extension of 10 minutes at 72°C, performed on a BioRad thermocycler (BioRad, Hercules, CA, USA). PCR cocktail per reaction was as follows: 10X reaction buffer – 2.5 μ L, 50mM $MgCl_2$ – 0.75 μ L, 2.5mM dNTPs – 2.5 μ L, 4 μ M primer – 2 μ L each of forward and reverse, ddH₂O – 10.15 μ L, Taq polymerase (Bioline) – 0.1 μ L (0.5 U) and DNA template – 5 μ L. PCR products were visualized on 2% agarose gels in 1X TBE containing ~ 0.1 μ g/mL of ethidium bromide in the final volume. Genome-specific primers were selected for screening for the presence of polymorphism between Molly, Newton, and Turcikum57.

2.3.7 Polymorphism identification and Mapping of the hexaploid mapping population

Primer pairs which were genome-specific for the markers were selected for Sanger sequencing. PCR products amplified on Molly, Newton, and PI372129 were purified using 2 μ L of Exo-SAP-

IT (Affymetrix) to 10 μ L of PCR product and incubating at 37°C for 30 minutes and inactivating at 80°C for 15 minutes. After enzymatic PCR clean-up, samples were diluted to twice its volume. Samples were submitted to Genewiz for Sanger sequencing. Sequencing results were aligned using Clustal Omega software to identify types of polymorphism including SNPs and InDels. The sequence trace files were analyzed using BioEdit software to confirm the polymorphism observed in alignment. Polymorphic markers were mapped on the HR mapping population. PCR and sequencing were done as described earlier. Genetic linkage map analysis was performed using Inclusive composite interval mapping software (IciMapping 3.3).

2.3.8 Mapping on radiation hybrid (RH) panels

Radiation hybrid mapping is a very useful approach in fine-mapping regions with a low level of polymorphism (Tiwari et al., 2012). To further fine map *H13* region, gamma-ray (γ) induced endosperm and pollen RH panels were used. Irradiated pollens from the hexaploid Chinese Spring were used to pollinate tetraploid cultivar Altar84. DNA was extracted from endosperm dissected from F₁ hybrid seeds and leaf tissue of F₁ hybrid plants for endosperm and pollen panel respectively. PCR cocktail of 10 μ L was used, and PCR products were visualized on 2% agarose gels. As the markers are genotyped as either presence or absence of bands, so an extra primer pair was added in each reaction to act as an internal PCR reaction control. The ordering of markers on the RH panel was done using Carthagene software (Givry et al., 2005).

2.3.9 Fine mapping of the *H13* interval using markers designed from scaffold sequence

Scaffold sequences from *Ae. tauschii* were used to design markers (Luo et al., 2017). Interspersed repeats were masked and replaced by Ns using RepeatMasker software (Smit et al.,

2013-2015). Annotation of the sequence was done using TriAnnot (Leroy P. et al., 2012) and FGENESH2 software (Solovyev et al., 2006). D-genome specific primers were designed (as described before) for genes annotated on the scaffold sequences. The genome specificity and localization of markers on the chromosomal arm and deletion bin were confirmed by testing on the genetic stocks nullisomic-6D, dt6DS, dt6DL, and deletion lines 6DS-2, 6DS-4, 6DS-6. The D-genome specific markers were screened for the presence of SNPs and InDels between Molly, Newton, and Turcikum57 by Sanger sequencing. The polymorphic markers between Molly and Turcikum57 were amplified, sequenced and mapped onto HR mapping population.

2.3.10 Ethyl methanesulphonate (EMS) mutagenesis and population development

For the functional validation of *H13* gene candidates, a TILLING population was developed in Molly. To determine the optimum dosage of EMS for developing mutagenized population for Molly, different concentrations of EMS were used to obtain around 50% kill rate in M₁ plants (Figure 2.11). A Hundred seeds in each set were first soaked in 75 ml water in 250 ml flasks for 8 hours at room temperature on a shaker at 75rpm; water was decanted, and different concentrations of EMS were added in each flask and were kept on a shaker at 75 rpm for 16 hours. In the first test, five sets of 100 seeds each of Molly were treated with 0, 0.7, 0.8, 0.9, 1.0 % EMS; EMS solution was decanted, and seeds were extensively rinsed under running tap water in cheese cloth for 8 hours before sowing individually into root trainers in the greenhouse. Kill rate was recorded after 2 to 3 weeks, but none of the treated M₁ seeds survived. In the second testing, six sets of 100 seeds of Molly were treated with 0, 0.25, 0.3, 0.35, 0.4, 0.45 % EMS. The desirable kill rate of 46% was observed with 0.45% EMS treatment (Figure 2.11). A total of

7158 M₀ seeds of Molly were treated with 0.5% EMS to obtain desirable kill rate. M₁ plants were selfed, and 6168 M₁ spikes were harvested. Around one-third of the population was progressed into the next generation. A single M₂ seed from 2448 lines was grown, out of which 115 lines were sterile, and 37 lines had very few (less than 6) seeds. 2296 M₃ lines were used for HF screening.

2.3.11 Tissue collection and 4X DNA pools

Tissue was collected from each M₂ individual in microtubes in the 96-well box. DNA was isolated from lyophilized leaf tissue using QiagenBioSprint 96 workstation with Qiagen BioSprint 96 DNA Plant Kit (Qiagen, Valencia, CA) as per the manufacturer's instructions. DNA quantification was done on NanoDrop spectrophotometer, and all samples were normalized to the concentration of 25 ng/ul in the 96-well box before 4X pooling. The 4X pooling of DNA was done by combining the DNA from each row of four boxes. The mutation frequency of the population was calculated using A, B, D-genome specific primers for starch synthase IIa gene. The mutation frequency of 1/36kb was observed for this gene.

2.3.12 Hessian fly screening of EMS population

Laboratory stocks of biotype Great Plains (GP) of Hessian fly were used to screen 20 seeds of each of the 2296 M₃ lines in the greenhouse. Biotype GP is avirulent on Molly and virulent on susceptible Newton. Twenty lines were grown in each flat with equally spaced furrows as separator along with four checks- Molly, Caldwell, Carol, and Karl92. Carol has *H3* gene, and Caldwell has *H6* gene. Seeds were sown in an equal mixture of soil and Metromix. Seedlings were fertilized on the fifth day. At 1.5 leaf stage, plants were infested with Hessian flies. Screening of plants was done 17 days after infestation. Tissue was collected from all susceptible

EMS mutagenized Molly lines along with six resistant EMS mutagenized Molly lines in microtubes in the 96-well box. DNA was isolated from lyophilized leaf tissue using Qiagen 96 DNA plant kit as described before.

2.3.13 Characterization of EMS mutants

As the EMS population was developed with field grown Molly seeds, so all susceptible EMS mutagenized Molly lines were tested for flanking SSR and genic markers to confirm that the susceptible phenotype is not due to contaminating seeds of different germplasm lines. The susceptible lines along with WT-Molly and six resistant EMS mutagenized lines were genotyped with flanking co-dominant microsatellite markers *Xgdm36* and *Xcfd132*. PCR reactions were performed using the same conditions as mentioned before. PCR products were visualized on 3% high-resolution agarose gels in 1X TBE containing ~ 0.1µg/mL of ethidium bromide in the final volume. Also, these lines were genotyped with flanking genic markers 19-34 and co-segregating genic marker for *CNL-1*. The PCR products for *CNL-1* were purified using 2µL of Exo-SAP-IT (Affymetrix) to 10µL of PCR product and incubated at 37°C for 30 minutes and inactivated at 80°C for 15 minutes. After enzymatic PCR clean-up, samples were diluted to twice its volume. Samples were submitted to Genewiz for Sanger sequencing. The sequencing results from all susceptible EMS mutagenized lines along with WT-Molly and six resistant EMS mutagenized lines were aligned using Clustal Omega software to identify any polymorphism.

2.3.14 Candidate gene validation

D-genome specific primers were designed for nine candidate genes as described earlier. Different primer combinations were designed to cover the complete sequence of the gene. PCR amplicons were submitted to Genewiz for Sanger sequencing. The results from sequencing were

aligned with wild-type (WT) Molly using Clustal Omega software to identify SNPs. The sequence trace files were analyzed using BioEdit software (Hall et al., 1999) to confirm the SNPs. Confirmed mutant sequences were translated into amino acid sequences using ExPASy translate tool (<http://web.expasy.org/translate>). Amino acid sequence from mutants was compared with WT Molly to identify possible missense, nonsense and silent mutations. The effect of missense mutations was analyzed using PROVEAN tool (Choi Y. et al., 2015).

2.3.15 5' and 3' Rapid amplification of cDNA ends (RACE)

RACE was performed to identify transcription initiation site at 5' and termination site at 3' using Switching Mechanism at 5' End of RNA Template (SMART)er RACE 5'/3' kit (Clontech). Total RNA was extracted from the feeding site tissue of Molly, and Newton infested with GP biotype flies at 60 hours post egg hatching using RNeasy Plant Mini kit (Qiagen) as per the manufacturer's instructions. RNA quantification and integrity assessment were done as described earlier. 1 µg of RNA was used as the starting material. Both 5' and 3' RACE-ready first-strand cDNA were synthesized individually using reagents and primers (5'CDS primer A and 3' CDS primer A) provided with the kit as per manufacturer's instructions. Two anti-sense gene-specific primers (GSPs) were designed from 1st and 2nd exon (GSP2 and GSP3) for the 5'RACE PCR and one sense GSP was designed from 1st exon (GSP1) for the 3'RACE PCR and 15bp sequence GATTACGCCAAGCTT was added to the 5' ends of all GSPs for overlap with the linearized vector for cloning step. Location of GSPs is shown in Figure 2.15. 5' and 3'RACE PCR was done using Clontech's SeqAmp DNA polymerase with UPM primer mix and GSPs. The PCR cocktail per reaction was as follows: 5'- or 3'-RACE-ready cDNA 2.5 µL, 10X UPM primer 5 µL, 5' or 3' GSP (10 µM) 1 µL, 2X SeqAmp buffer 25 µL, SeqAmp DNA polymerase one µL, Water- 15.5 µL. Thermocycler conditions steps included initial denaturation at 96°C for 2 minutes

followed by 35 cycles of 3 steps- denaturation at 94°C for 30 seconds, annealing at 68°C for 30 seconds and extension at 72°C for 5 minutes and a final extension of 10 minutes at 72°C was performed on a BioRad thermocycler (BioRad, Hercules, CA, USA). PCR products from the 5'- and 3'-RACE reactions were purified with Nucleospin gel extraction kit (Clontech) as per the kit's protocol. Purified PCR fragments along with linearized pRACE vector as a negative control and non-digested pRACE vector as positive control were sub-cloned into linearized pRACE vector (Clontech) and transformed into chemically- competent stellar *E.coli* cells (Clontech). The cloning sites on pRACE vector map are shown in Figure 2.15. Transformation mixture was incubated by shaking for 1 hour at 37°C. Different dilutions of the mixture were plated on pre-warmed Luria-Bertani (LB) plates containing 100 µg/mL of carbenicillin and incubated at 37°C. A total of 45 colonies were picked for each RACE PCR transformations. Screening of colonies was done by PCR using M13 F-R primers and M13F-corresponding GSP. Out of all positive colonies from each transformation, five colonies were randomly selected, re-streaked and inoculated into five mL of LB broth containing 100 µg/mL of carbenicillin. Plasmid DNA was isolated, and samples were submitted to Genewiz for Sanger sequencing using M13 primers, GSP, and internal gene primers.

2.3.16 Assembly and cloning of the full-length gene

Primers were designed for 5' end of the gene using sequence information from 5'RACE products and similarly for 3' end of the gene using sequence information from 3'RACE products. Using these primers, the gene was amplified from genomic DNA (gDNA) and cDNA. Sanger sequencing was done for PCR products. Sequences from gDNA and cDNA were aligned to identify gene structure.

2.3.17 Haplotype analysis and allele mining

For allele mining, *Ae. tauschii* accessions from WGRC stock were selected including accessions which are resistant to HF, Minicore collection representing 574 *Ae. tauschii* accessions and commonly deployed wheat cultivars available at WGRC (Singh 2017). Also, hexaploid germplasm lines carrying different HFR genes were selected and were provided by Dr. Ming Shun Chen, USDA ARS, Entomology, KSU. *H13* carrying lines 9346A1-2-5-5-2-1 and INW9811 were provided by Dr. Christie Williams, USDA, Purdue University. Details for lines mentioned above are listed in Table 2.2.

Tissue was collected, and DNA was isolated as described earlier. The complete *H13* gene was amplified, and Sanger sequencing was done for PCR products. The results from sequencing of all the accessions were compared by alignment with Molly using Clustal Omega software to identify SNPs. The sequence trace files were analyzed using BioEdit software to confirm the SNPs. Sequence reads were translated into amino acid sequences using ExPASy translate tool (<http://web.expasy.org/translate>). The different domains of *H13* protein were determined by searching for conserved domains using the NCBI database (Marchler-Bauer et al. 2015). The amino acid sequence alignments were performed using MS Excel.

2.3.18 Expression analysis of *H13* gene

Seeds of near-isogenic lines, Molly (*H13*) and Newton (*h13*) were used for expression profiling of *H13* gene. Laboratory stocks of biotype Great Plains (GP) and *vH13* of Hessian fly were used for infestation. Ten seeds were grown in each pot with Metromix in a growth chamber set at 20±1°C for 14 hours light and 18±1°C for 10 hours dark and around 70% relative humidity. Pots were covered with plastic cages with vent. At 1.5 leaf stage, ten mated female flies were added

manually in each pot along with two male flies. Additional pots were planted and infested at the same time to monitor hatching time of HF larvae. After larvae hatch, they start migrating towards the base of leaf and time was recorded when they reach feeding site. Leaf and feeding site tissue were collected in three biological replicates from control un-infested plants and treated infested plants. Samples were collected at different time points (hours) post-hatching :12, 48, 60, 72, 108, 150, 192. Larvae were removed in chilled deionized water using chilled forceps and a pair of scissors before plant tissue was flash frozen in liquid nitrogen and stored at -80°C until RNA was extracted.

Before RNA extraction, benchtop, pipettes, glassware, spatula, pestle, and mortar were cleaned with RNase AWAY (Molecular Bioproducts) to eliminate any RNase and DNA. Total RNA was extracted using the RNeasy Plant Mini kit (Qiagen) as per the manufacturer's instructions. RNA quantification was done on NanoDrop spectrophotometer followed by RNA quality assessment to calculate RNA integrity number (RIN) using 2100 Bioanalyzer (Agilent Technologies). 1µg of total RNA was used for cDNA synthesis using the QuantiTect Reverse Transcription kit (QIAGEN). RT-PCR reaction was prepared using SsoAdvanced Universal Inhibitor-Tolerant SYBR Green Supermix (Bio-Rad), and the reaction was run on a CFX96 (Bio-Rad) instrument. Tubulin gene was used as a reference gene, and transcript levels of *H13* gene were quantified using $2^{-\Delta CT}$ method (Morse et al., 2005; Livak and Schmittgen, 2001). Both exonic and intron-spanning primers were designed for expression analysis. All samples had three biological and two technical replicates.

2.3.19 Comparative Mapping of candidate genes

The nucleotide sequence of annotated genes from 6DS near *H13* region was used for homology search in *Ae. tauschii*, Barley, Brachypodium, Rice, Maize and Sorghum to identify orthologous genomic regions. BLAST was performed on EnsemblPlants database, and resulting hits with the lowest E-value were selected, and their mapping coordinates were determined. Previous information regarding synteny of chromosomes was used to differentiate gene hits which belong to either collinear or non-collinear genes. The coordinates of gene hits were used to construct a comparative map of *H13* region. Additionally, a comparative analysis was done using annotated genes from 6DS near *H13* region to identify homeologs from chromosome 6A and 6B of CS available at International Wheat Genome Sequencing Consortium website (IWGSC; https://urgi.versailles.inra.fr/blast_iwgsc).

2.3.20 Phylogenetic analyses

Phylogenetic analysis was performed using the protein sequence of CNL candidate genes, and their ortholog hits in barley, Brachypodium, rice, maize, and sorghum. Sequence alignment was done using Mega v.7 followed by construction of the phylogenetic trees using the Neighbor-Joining method (Saitou and Nei 1987). Based on amino acid substitutions, the Poisson correction method was used to calculate the evolutionary distances. Also, the *H13* gene belongs to non-TIR NBS LRR subfamily. Sequences of previously identified non-TIR NBS LRR genes from different plant species were obtained, and the phylogenetic analysis was performed for the protein sequence as described above.

2.4 Results

2.4.1 Development of a high-resolution mapping population:

Co-dominant SSR markers *Xcfd132* and *Xgdm36*, flanking the *H13* gene are located proximal to the breakpoint of the deletion bin del6DS-6 (0.99). The F₂ population of 1,368 lines derived from Turcikum57 (PI372129) x Molly (PI562619) (Figure 2.2) was genotyped using polymorphic markers *Xcfd132* and *Xgdm36*.

Sixty-four recombinants between *Xcfd132* and *Xgdm36* markers were selected and utilized for generating a high-resolution map of *H13*. HF infestation screening was done on F₃ families of the high-resolution population (Figure 2.2). Combination of results from genotyping and HF screening resulted in the formation of a linkage map of *H13* region where *Xcfd132* and *Xgdm36* were mapped at 0.4cM proximal and 1.8cM distal to *H13* respectively (Figure 2.2).

2.4.2 Development of a high-resolution map and identification of co-segregating markers:

2.4.2.1 Anchoring of SSR markers on BAC contigs of chromosome 6D

PCR based screening of 3-D BAC pools of chromosomes 1D,4D,6D (CS) was done for SSR markers resulting in the identification of three independent contigs Ctg1153, Ctg269 and Ctg1089 for *Xgdm36*, *Xcfd132* and *Xcfd213* markers respectively (Figure 2.3). As sequence information for BACs was not available, new markers could not be developed. However, one of the genetic markers (AT6D5315) from the genetic and physical map of *Ae. tauschii* (Dvorak et al., 2013) was mapped on Ctg1153. This led to the anchoring of *Xgdm36*, an *H13* flanking

marker, onto the genetic and physical map of *Ae. tauschii* for marker development. None of the *Ae. tauschii* genic markers were mapped to Ctg269 nor Ctg1089 (Figure 2.4).

2.4.2.2 Development of D-genome specific markers:

Locus markers, which were 15cM distal as well as proximal to AT6D5315 in the genetic and physical map of *Ae. tauschii* (Dvorak et al., 2013) were selected to design D-genome specific primers. Primers were designed using polymorphic sites among A, B and D genome homeologs and manually picked with polymorphism at the 3'-end (Figure 2.5). The genome specificity of primers was confirmed by checking on Nullisomic-6D, dt6DS, dt6DL, and deletion lines 6DS-2, 6DS-4, 6DS-6 along with Molly, Newton, and Turcikum57 (Figure 2.6).

2.4.2.3 Genetic mapping of *Ae. tauschii* markers on the high-resolution mapping population

Comparative sequence analysis of D-genome specific markers between Molly and Turcikum57 led to the identification of markers used for high-resolution mapping. Out of a total of 46 genic markers spanning 33cM and 22.84Mbp on the genetic and physical map of *Ae. tauschii* (Dvorak et al., 2013), only seven markers were polymorphic and were selected for mapping on F₂ high-resolution mapping population. Genic marker AT6D5315 was mapped 0.14cM distal to *Xgdm36*, and AT6D5321 was mapped 1.96cM proximal to *Xcfd132*, whereas other genic markers flanked this interval (Figure 2.7).

2.4.3 Radiation hybrid (RH) mapping

As markers between AT6D5316 and AT6D5320 could not be used for mapping because they were monomorphic in Molly and Turcikum57, radiation hybrid panels were used to further fine map the interval. Radiation hybrid panels were obtained from a cross between tetraploid cultivar

Altar84 with irradiated pollens from the hexaploid Chinese Spring (CS). The D-genome specific markers were tested with a positive PCR control on CS and tetraploid cultivar Altar84 (Figure 2.8). Presence and absence of markers were scored on 369 RH lines (Figure 2.8). Nearest collinear markers AT6D5318 and AT6D5319 from the genetic and physical map of *Ae. tauschii* flanked *Xgdm36* and *Xcfd132* markers by 2cR on RH panels (Figure 2.8). The sequence of the 10 kb region harboring each of AT6D5318 and AT6D5319 was used to develop additional D-genome specific markers. Polymorphic markers were identified and used for mapping on F₂ high-resolution mapping population. Genic marker AT6D5318 was mapped 0.07 cM distal to *Xgdm36* and AT6D5319 was mapped 0.55 cM proximal to *Xcfd132* (Figure 2.9) using a high-resolution hexaploid mapping population. *Ae. tauschii* marker AT6D5318 and AT6D5319 were mapped to contig 1251 and contig 826, respectively.

2.4.4 Marker enrichment for *H13* region

The complete sequence of scaffolds containing markers AT6D5318 and AT6D5319 were repeat masked and annotated for gene prediction. This sequence information was used to design D-genome specific primers for predicted genes. Genic markers from scaffold AT6D5318 were mapped distal to *Xgdm36*, whereas those from scaffold AT6D5319 were mapped both distal and proximal to *H13* (Figure 2.10). Thus, markers from scaffold AT6D5319 were used to fine map the *H13* gene.

2.4.5 Fine mapping of *H13* region

Genic marker 19-34 is mapped 0.15 cM distal to *H13*, *CNL-1* cosegregated with *H13*, and 19-124 is mapped 0.26 cM proximal to *H13*. Nine candidate genes were identified between 19-34 and 19-124 markers. The nine genes spanned an interval of 319kb and co-segregated with *H13*.

Out of 9 genes, six genes were predicted to belong to coiled-coil (CC), nucleotide binding (NB), leucine-rich repeat (LRR) – CNL disease resistance genes; one gene was predicted to belong to a family of pentatricopeptide-repeat-containing proteins (PPR); one gene was predicted to belong to the family of serine/threonine kinases (STK); and one gene was predicted to belong to plant mobile domain (PMD) (Figure 2.10; Figure 2.18).

2.4.6 Functional validation of candidate genes by TILLING

For TILLING, the optimal concentration of EMS was used on M₀ seeds (Molly; PI562619) to generate a mutagenized TILLING population. Screening of 2,296 M₃ lines for HF phenotype resulted in 101 susceptible lines (Figure 2.11). The susceptible lines were genotyped using flanking genic markers to rule out deletion mutations in the *H13* region. Furthermore, SSR markers analysis for *Xcfd132* and *Xgdm36* in susceptible lines identified 73 lines which had different marker profile than Molly, suggesting these to be contaminant seeds from field harvest (Figure 2.12). Additional verification of contaminant seeds was performed using sequencing analysis of genic marker *CNL-1*, which revealed 49 SNPs within 500bp region that were consistent among 73 lines and different from Molly (Figure 2.13). The contaminant lines were removed from this study.

For the functional validation of candidate genes, 28 susceptible lines were sequenced. Mis-sense and truncation mutations were present only in *CNL-1*. Three independent susceptible lines had truncation mutations in *CNL-1*. Mutant line M⁶³⁴ had a C698T mutation resulting in a premature stop codon at amino acid position 233. Mutant line M²⁴¹³ had a G2625A mutation resulting in a premature stop codon at amino acid position 875. Mutant line M²³⁷⁸ had a C2634T mutation resulting in a premature stop codon at amino acid position 878. Also, missense mutations

resulting in amino acid change were found in nineteen independent susceptible lines for *CNL-1* (Figure 2.14).

Additionally, the effect of amino acid substitution on protein function was tested using the PROVEAN score, and the corresponding scores for all mutations are listed in Table 2.6. For SIFT score, the cut off probability of less than 0.05 for an amino acid substitution is deleterious on the function of the protein. For the PROVEAN score, prediction of -2.5 or lower for an amino acid substitution is deleterious to the function of the protein. Either silent or intronic mutations were found for *CN-7*, *CNL-4*, and *CNL-2*, and no mutations were observed for other candidate genes. In contrast, by sequencing six resistant mutagenized lines, a missense mutation was observed for *CNL-6*, and a silent mutation in *CNL-1*, which further validated *CNL-1* as the best candidate *H13* gene.

2.4.7 Identification of gene structure for *H13*

Random amplification of cDNA ends (RACE) was done to identify the 5' and 3' ends of cDNA for *H13* from Molly and Newton (Figure 2.15; Figure 2.16). Using the sequence information from cloned fragments of 5'-RACE and 3'-RACE, long-range PCR was done to amplify the full length of *H13* gene from genomic DNA and cDNA (Figure 2.17).

The coding sequence of the *H13* gene from Molly (*H13H13*) is 3192bp consisting of two exons with 618bp 5'-UTR and 2260bp 3'-UTR. It translates into a protein of 1063 amino acids with an N terminal CC, a central NB-ARC, and a C terminal LRR domain (Figure 2.20; Figure 2.21). The coding sequence of the *H13* gene from Newton (*h13h13*) is 4503bp consisting of two exons with 619bp 5'-UTR and 158bp 3'-UTR. It translates into a protein of 1500 amino acids with similar domains as in Molly. However, in Newton, there are differences in number and

organization of LRR repeats in addition to the presence of endonuclease domains towards C terminal (Figure 2.22).

2.4.8 Phylogenetic and comparative analysis

Majority of the candidate genes identified in *H13* region belonged to the family of CNL proteins and were present in clusters which is consistent with the clustering of R genes during evolution and adaptation to selection pressure. Phylogenetic analysis done using the amino acid sequence of all candidate genes showed that CNL-3 and CNL-4 had 80% sequence identity with CNL-1 and were more closely related as compared to other CNLs which had 65% identity with CNL-1 (Figure 2.24). The comparative analysis of CNLs among the different members of grass family including *Ae. tauschii*, *H. vulgare*, *B. distachyon*, *O. sativa*, *Z. mays* and *S. bicolor* identified homologs. Homologs of flanking genic markers were found to be syntenic across *B. distachyon*, *O. sativa*, *Z. mays* and *S. bicolor* but the candidate CNLs identified homologs on other chromosomes and were not orthologous. BLAST searches using the amino acid sequence from CNL-1, CNL-3 and CNL-6 genes resulted in the identification of a single homolog in *H. vulgare*. Similarly, CNL-2 and CNL-4 also resulted in a single hit when searched in *H. vulgare* (Figure 2.19; Table 2.4). Other candidate genes encoding for PPR and STK showed synteny across the different species mentioned above.

The phylogenetic analysis of the full-length amino acid sequence of *H13* with other cloned members of CNL family showed that the *H13* belongs to a separate clade shared by none of the other cloned CNL genes. The sequence identity of the CNLs range between 23 to 27% as compared to the *H13* gene. Thus, *H13* is a unique member of the CNL family and has low sequence similarity with other known CNLs in plants.

2.4.9 Expression analysis

Using two different sets of primers from exonic regions, we observed that the transcript expression of *H13* in Molly was induced at the feeding site during HF infestation by GP biotype (Figure 2.23). Transcription levels increased from 12 hours after egg hatching (HAH) and peaked at 60 HAH, which is consistent with feeding and recognition of potential effectors secreted by Hessian fly larvae, whereas in Newton, over the same period the transcript levels declined rapidly.

2.4.10 Haplotype analysis and allele mining

Using the sequence information from the *H13* gene, full-length coding sequence was amplified and compared among *Ae. tauschii* core collection and different hexaploid cultivars. Sequence alignment identified 16 different haplotypes for the *H13* gene (Table 2.7; Please go to www.krex.k-state.edu and search for AnupamaJoshi2018-Table2.7.Haplotype analysis

Table 2.8; Table 2.9). For two of the haplotypes Hap15 and Hap16, we were unable to amplify the complete gene. A phylogenetic tree using the Neighbor-joining method was constructed indicating the evolutionary relationship between different alleles (Figure 2.25). Among 16 haplotypes, six had a premature stop codon at the beginning of the NBS domain. Haplotype for *H13* gene was designated as Hap1 and four additional haplotypes namely Hap4, Hap5, Hap6, and Hap7 had a stop codon at the same site as Hap1. These haplotypes could be novel functional haplotypes against evolving HF biotypes. *Ae. tauschii* accessions were categorized to all 16 haplotypes, but the hexaploid cultivars used in this study belonged to 2 of the haplotypes only. It is evident that diversity is present in *Ae. tauschii* germplasm. Haplotype derived from Newton

was designated as H3, and it showed the presence of an additional DDE endonuclease towards the C terminus of the protein.

2.5 Discussion

Gall midges (Diptera: Cecidomyiidae) are divided into four major sub families named Catotrichinae, Lestremiinae, Porricondylinae, and Cecidomyiinae. The plant-gall-forming species including HF, Mycetophages, and predators, belong to Cecidomyiinae subfamily (Roskam et al. 2005).

Gall-forming insects have a diverse range of hosts (Gagne et al. 1989, Price et al. 2005) and feed on their host by forming galls on different plant tissues including leaves, stems, flowers, buds and fruits and extracting nutrients (Mamaev et al. 1975, Gagne et al. 1989). Gall midges adopt a biotrophic or hemibiotrophic lifestyle with the localized site of infection on the host which renders the plant susceptible to the insect (Stuart et al. 2012). Gall midges reproduce quickly, adults have a short life span, and host selection is critical for reproductive fitness (Harris et al. 2001; 2003).

HF belongs to a genus of gall midges containing 29 species which feed on grasses (Gagne et al. 2010). HF infestation is associated with different morphological changes which highly correlate with morphological changes in the wheat (Hatchett et al. 1990, Stuart et al. 1987). The life cycle of HF starts with the egg, three larval instars, the pupa, and the adult which gets completed in 28 days. First instar larvae modulate and affect the plant development, second-instar larvae feed on the nutrients provided by the reprogrammed plant. Third-instar larvae and pupae develop within a puparium, and adult flies emerge from the puparia and live for 1-4 days (Harris et al. 2003).

R genes in plants provide resistance against various pests and pathogens, with a limitation of low durability caused by rapidly evolving *Avr* genes in virulent pests and pathogens. Among other

insect resistance genes cloned, *Mi-1.2* gene in tomato (*Solanum lycopersicum*) encodes for an NL domain protein and provides resistance against Potato aphid (*Macrosiphum euphorbiae*) and Silverleaf whitefly (*Bemisia tabaci*) (Rossi et al. 1998, Nombela et al. 2003). *Vat* gene in melon (*Cucumis melo*) encodes for a CNL protein and provides resistance against cotton aphid (*Aphis gossypii*) (Dogimont et al. 2014). Using map-based cloning approach, in rice (*Oryza sativa*) 13 brown planthopper (*Nilaparvata lugens*) (BPH)-resistant genes have been cloned including *Bph14* (Du et al., 2009), *Bph3* (Du et al. 2009, Liu et al. 2015), *Bph15* (Cheng et al. 2013), *Bph26/2* (Tamura et al. 2014), *bph29* (Wang et al. 2015), *Bph18* (Ji et al. 2016), *Bph9/1/7/10/21* (Zhao et al. 2016a) and *Bph32* (Ren et al. 2016). *Bph14*, *Bph18*, and *Bph26* encode for a CNL protein whereas *Bph3* and *Bph15* encode for a membrane-localized lectin receptor kinase (LPK). *Bph32* encodes for a short consensus repeat (SCR) domain protein, and *bph29* encodes for B3 DNA-binding domain. The candidate resistant genes for Asian rice gall midge (*Orseolia oryzae* Wood-Mason) in rice, *Gm2*, *Gm4*, *gm3* belong to CNL family (Sama et al. 2014, Divya et al. 2015). There is a lack of genetic tractability in various plant-insect interactions.

Majority of plant-insect interactions involve effector based mechanisms (Jones et al. 2006, Torto-Alalibo et al. 2009). A multitude of genes in salivary glands of HF encodes for different effector proteins that interact with the R genes of wheat in a gene-for-gene manner. Study of *Avr* cognate pair in the wheat-HF system would help in understanding the mechanism of resistance in wheat.

In our study, we utilized a map-based cloning strategy to clone the *H13* gene. A mapping population was developed by a cross between PI562619 (Molly; *H13H13*; resistant) and PI372129 (Turcikum57; *h13h13*; susceptible). Screening for recombinants between *Xgdm36* and *Xcfd132* resulted in the generation of a high-resolution mapping population for the *H13* region.

The genic markers showed limited polymorphism for genetic mapping, and therefore, Radiation Hybrid (RH) mapping approach (Tiwari et al., 2012, Tiwari et al., 2016) was utilized to map further the *H13* region, followed by fine mapping using high-resolution genetic mapping population.

With the availability of reference genome RefSeq V1.0 assembly for hexaploid Chinese Spring, mapping of gene homeologs to their respective genomes has been possible. Additionally, *Ae. tauschii* genome has been sequenced and provides a resource for comparative mapping of genes. Based on the annotation of scaffold sequences comprising the *H13* gene from *Ae. tauschii* different candidate genes were identified. We used Targeting Induced Local Lesions in Genomes (TILLING) approach for functional validation of candidate genes. EMS is an alkylating agent that reacts with guanine and results in transitions. The G:C base pair gets replaced by A:T pair (Greene et al., 2003). All the mutations found in our study were also G:C to A:T transitions. There are five transitions that can change an amino acid to a stop codon; TGG to TGA, CGA to TGA, CAA to TAA, CAG to TAG and TGG to TAG. The underlined nucleotide represents the transition whereas there are 116 possible transitions, that change a codon so that it encodes for a different amino acid. Krasileva et al. (2017) performed exome capture of 73,895 genes for the EMS-induced TILLING population in hexaploid wheat and reported 23 missense and 1.8 nonsense and splice site mutations per gene.

Similarly, Henry et al. (2014) reported that 32% of the induced mutations are classified as a missense mutation, and 1% is classified as a nonsense mutation in exome capture for the TILLING population in rice. Similar to these reports, in our study, sequencing of the candidate genes from independent susceptible mutant lines identified 24 missense and three nonsense mutations in CNL-1 gene indicating it to be *H13* gene. Based on the mutation frequency

observed in the TILLING population, it can be exploited for characterization of different genes in wheat.

Studies have reported that among all the known gall-forming insects, the interaction between wheat and HF can act as a model system, as the host-pest interactions can be controlled and are based on the gene-for-gene model. The family of CC-NB-ARC-LRR domain-containing proteins are known to be involved in pattern-triggered immunity (PTI), and effector-triggered immunity (ETI) response to a pathogen or pest attack (Jones and Dangl 2006). In monocots, the majority of resistance genes against pathogen infestation belong to CC-NBS-LRR class. Some of the examples of cloned R gene which belong to this family in wheat include: *Lr10*, *Lr21*, *Lr22* and *Lr1* against leaf rust (Feuillet et al., 2003, Huang et al., 2003; Thind et al., 2017; Cloutier et al., 2007); *Sr22*, *Sr45*, *Sr33* and *Sr35* against stem rust (Steuernagel et al., 2016; Saintenac et al., 2013; Periyannan et al., 2013); *Pm3b* and *Pm8* against powdery mildew (Yahiaoui et al., 2004; Hurni et al., 2013), *Yr10* against stripe rust (Liu et al., 2014). *H13* is the first cloned gene specifying resistance against Hessian fly in wheat. Our study has revealed that *H13* encodes for a CC-NB-ARC-LRR protein.

Comparative analysis of amino acid sequences between *H13* and the other cloned CC-NB-LRR proteins from different species has shown sequence similarity between 23-35% with sequence coverage ranging from 40-96%. The results indicate that *H13* is a distinct CC-NB-LRR protein. During evolution, events of gene duplication and divergence lead to clustering and diversity in NB-LRR encoding genes (Michelmore and Meyers, 1998). Grube et al. (2000) have reported the differences in pathogen recognition specificities for R genes, which either arose by duplication or are related by descent.

Increased transcript abundance of the *H13* gene during incompatible interactions at larval feeding stages of GP biotype is consistent with the role of ETI in response to recognition of effectors. Comparative transcriptome analysis between GP infested resistant and susceptible lines would help identify differentially expressed genes and corresponding pathways involved in plant-insect interactions. There exists a wide diversity in sequences of CC-NB-LRR proteins among different R genes or alleles of the same gene. Comparison of the amino acid sequence of *H13* gene between Molly (resistant; *H13H13*) and Newton (susceptible; *h13h13*) lines revealed significant differences in sequence composition and a number of LRR domains in addition to the presence of DDE endonuclease domain towards C terminal of the protein in Newton. The presence of a transposable element in *h13* suggests interallelic recombinations, recent tandem duplications and gene conversions in *CNL* rich region near *H13* locus. R genes are usually present in clusters, which are a result of unequal crossing over during meiosis. These R gene clusters act as a source of novel alleles with new disease resistance specificities (Richter et al. 1995). In our study, the presence of a transposable element, DDE endonuclease, in four out of the six CNLs suggests that this element may play a role in gene duplication. The CNLs arose in common ancestor of Triticeae species and underwent recent duplications in the *Ae tauschii* lineage. DDE endonucleases can transpose from one location to another and can cause insertions and deletions of R genes (Nesmelova et al. 2010). They play an essential role in the evolution of resistance genes such as rice Xa21 homologs and an L6 allele of flax (Michelmore and Meyers, 1998, Ellis et al., 1997). In *Ae. tauschii*, chromosomal ends have a higher abundance of duplicated genes and synteny is not conserved because of the presence of transposable elements (Luo et al. 2017).

A total number of 4,826 genes predicted on chromosome 6D of *Ae. tauschii*. Out of those, the number of predicted resistance gene analogs on chromosome 6D are 6 CC–NBS; 29 CC–NBS–LRR; 6 NBS; 35 NBS–LRR; 119 RLK, receptor-like protein kinase; 42 RLP, receptor-like protein; and 20 TMCC, transmembrane coiled-coil protein (Luo et al. 2017). In our study, out of total nine candidate genes named *CNL-1*, *CNL-2*, *CNL-3*, *CNL-4*, *PMD*, *CNL-6*, *CN-7*, *STK* and *PPR*, six of them encoded for CC-NB-LRR proteins. Microcolinearity in chromosome segments has been observed between different members of grass family including rice, Brachypodium, barley, and *Ae. tauschii* (Liu et al., 2010; Bossolini et al., 2007; Faris et al., 2008; Lu et al., 2006; Griffiths et al., 2006; Valarik et al., 2006), and is often disrupted by duplications and rearrangements (Sorrells et al., 2003; Luo et al., 2013; Bossolini et al., 2007; Lu et al., 2006; Tikhonov et al., 1999; Keller and Feuillet 2000; Dubcovsky et al., 2001; Li and Gill 2002). Synteny analysis among *Ae. tauschii*, *H. vulgare*, *B. distachyon*, *O. sativa*, *Z. mays* and *S. bicolor* using protein sequences of candidate genes from the *H13* region of 319Kb indicates that *CNL-1*, *CNL-3*, and *CNL-6* from *Ae. tauschii* and hexaploid wheat were present as a single copy in *H. vulgare*. Additionally, *CNL-3* and *CNL-6* have 80% and 82% amino acid sequence identity with *H13*. Similarly, *CNL-2* and *CNL-4* were also present as a single copy in *H. vulgare* indicating recent duplication and probable functional adaptation during evolution. Synteny for all 6 *CNL* genes was not observed in Brachypodium, rice, Maize, and Sorghum. These observations are consistent with finding less collinearity for resistance genes as they are constantly evolving (Leister et al., 1998).

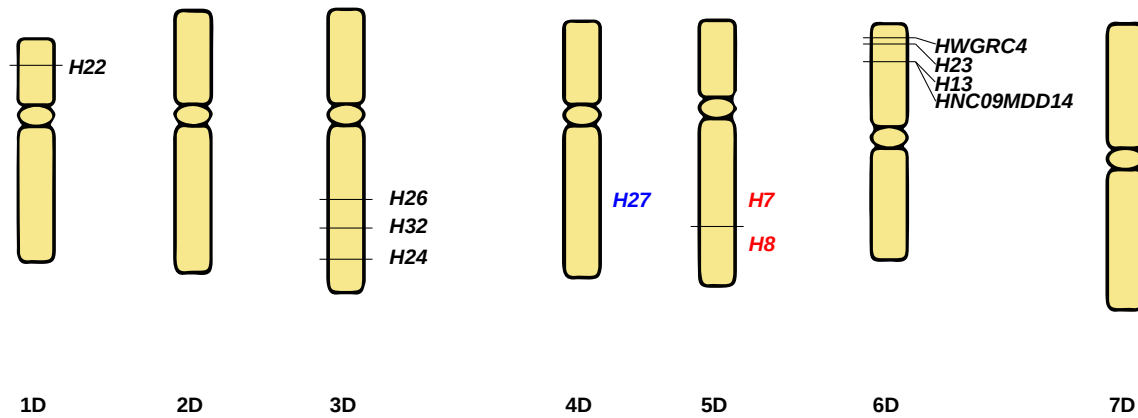
To identify novel diversity and functional alleles of *H13* gene conferring resistance to different biotypes of HF, translated sequences of different haplotypes were compared, and a phylogenetic tree was constructed. Haplotype analysis using *Ae. tauschii* core collection along with hexaploid

cultivars showed most of the sequence diversity was in the LRR region which is consistent with the role of LRR in recognition specificity (Ellis et al., 1999; DeYoung et al., 2006). Based on comparative analysis of conserved domains among different haplotypes, we identified differences in the number of LRR repeats, presence or absence of endonuclease domain, premature stop codon in different haplotypes. Based on similarities in the structural organization of different domains with those present in haplotype Hap1, haplotype 4, 5, 6, seven were selected for further analysis. Similar to Hap1, Hap4 and Hap5 had the same number of LRR repeats and were lacking the DDE endonuclease domain. Further analysis of HF phenotype data for all 11 accessions belonging to Hap4 displayed resistance phenotype indicating Hap4 to be a functional allele of *H13*. Similar HF infestation studies on accessions belonging to Hap5, Hap6, and Hap7 haplotypes would help in the identification of additional functional alleles of *H13*. These different haplotypes of *H13* gene could be tested to have a role in resistance against different variants of the Avr gene, different effector encoding genes or different pathogens. Some of the examples include: alleles of *Mlo* locus provide resistance against variants of Avr for powdery mildew-causing fungus *Blumeria graminis* f. sp. *hordei*. (Halterman and Wise, 2004). RPM1 gene from Arabidopsis recognizes products of different effectors from *Pseudomonas syringae* (Grant et al., 1995). Different alleles namely HRT, RCY1, and RPP8 from the same locus confer resistance towards different genera of viruses, TCV and CMV, and the oomycete *Hyaloperonospora parasitica* respectively (Cooley et al., 2000; Takahashi et al., 2002). Thus, different functional haplotypes observed for *H13* gene should be checked for resistance against different biotypes of HF or different pests.

The cloning of the R-Avr cognate pair for *H13-vH13* would help in understanding the molecular mechanism of gene-for-gene interactions between hexaploid wheat and Hessian fly. Cloning of

H13 also enables us to develop perfect markers useful for breeding programs. Also, this information will be helpful in pyramiding multiple genes for developing a broad-spectrum resistance level for HF in wheat. Furthermore, sequence information on *H13* will enable allele mining from different elite cultivars, that can be useful for combining different alleles for developing durable HF resistance.

2.6 Figures



H22, H23, H26, H32, H13, HWGRC4, HNC09MDD14 : Aegilops tauschii

H7, H8 : Triticum aestivum

H27 : Aegilops ventricosa

Figure 2.1 Chromosomal location of different HFR genes on the D genome of wheat.

Based on their origin, Hessian fly resistance (HFR) genes from *Aegilops tauschii*, *Triticum aestivum* and *Aegilops ventricosa* are labeled in black, blue and red color, respectively. The genetic distances (cM) between different genes is represented on the left side of the chromosomes.

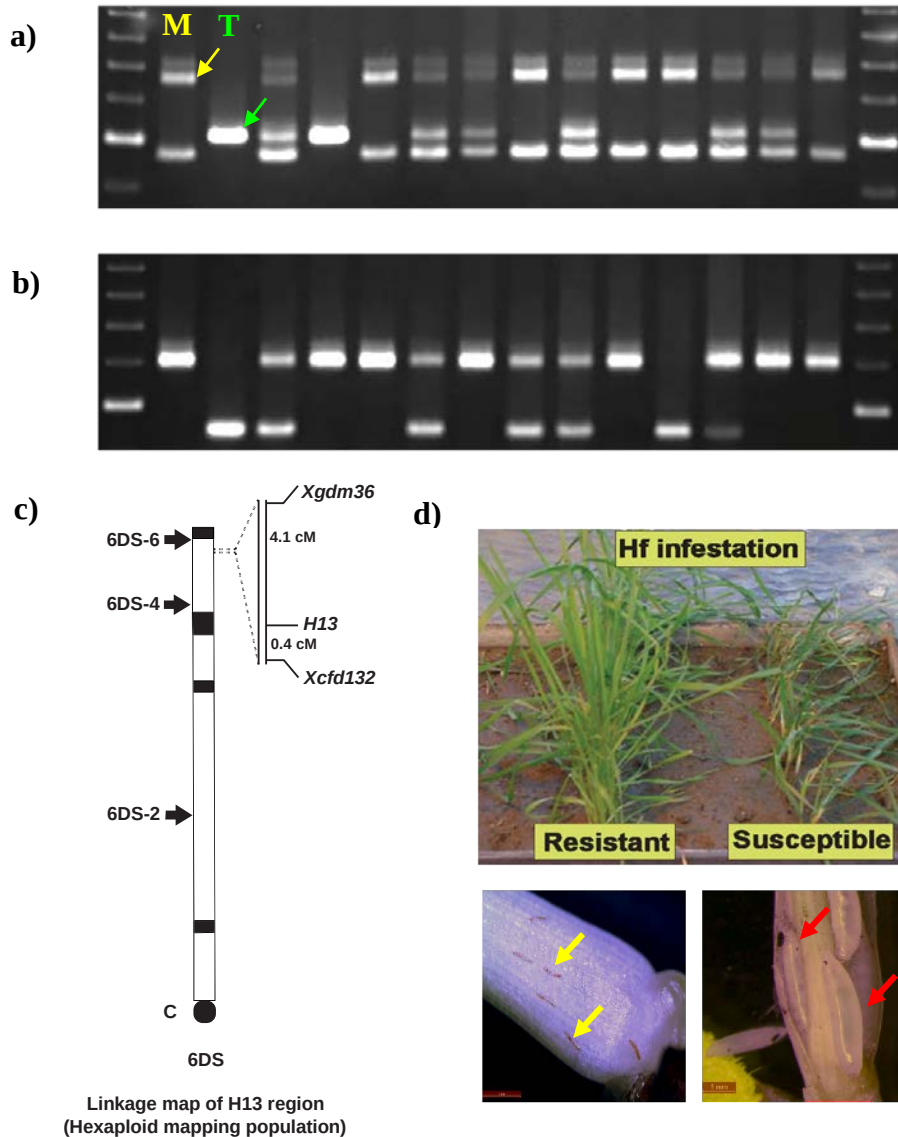


Figure 2.2 Linked SSR marker analysis and HF infestation

a, b) Gel picture showing the genotyping results of individuals from an F₂ population of PI562619 (Molly) x PI372129 (*Dn4*) using linked SSR markers *Xgdm36* (a) and *Xcfd132* (b). Parents of segregating population were used as controls and are labeled as M and T.

c) Linkage map of *H13* region flanked by SSR markers *Xgdm36* and *Xcfd132* on the short arm of chromosome 6D from hexaploid wheat. Different arrows represent breakpoints for each deletion bin.

d) Results from Hessian fly infestation on resistant and susceptible plants. Susceptible plants are darker and stunted as compared to a resistant plant. Three weeks post infestation, HF pupae were visible on susceptible plants (shown by red arrows) whereas dead first-instar HF larvae were present on resistant plants (shown by yellow arrows).

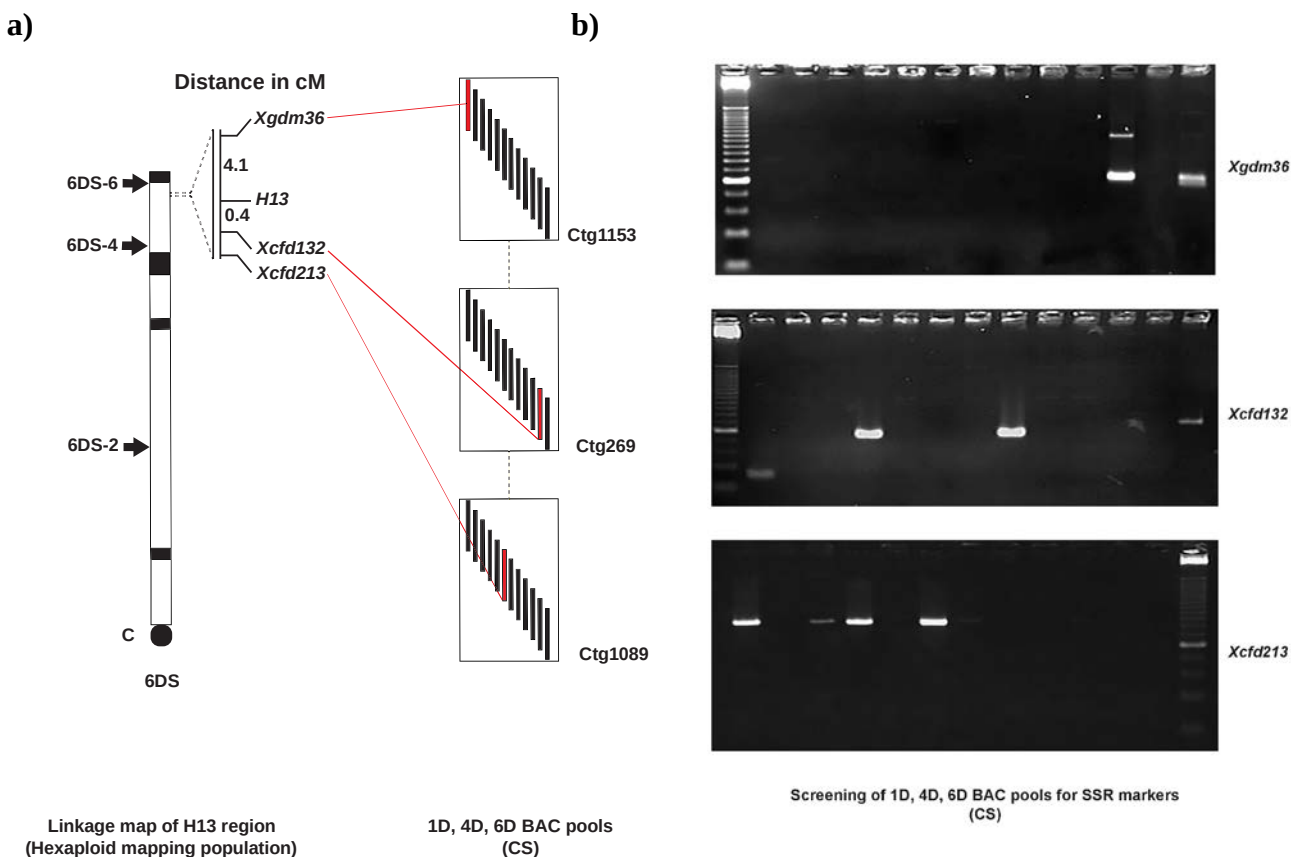


Figure 2.3 Anchoring of SSR markers on Chinese Spring (CS) BAC pools

a) Linked SSR markers *Xgdm36* and *Xcfd132* were anchored onto the BAC pools of D genome (1D, 4D, 6D) assembly for Chinese Spring. Red lines indicate the position of markers *Xgdm36*, *Xcfd132*, *Xcfd213* on BAC pools of corresponding contigs Ctg1153, Ctg269 and Ctg1089, respectively.

b) Individual gel pictures are showing results from the screening of different SSRs (*Xgdm36*, *Xcfd132*, and *Xcfd213*) in BAC pools. The presence of a band indicates the presence of the marker in the BAC pool.

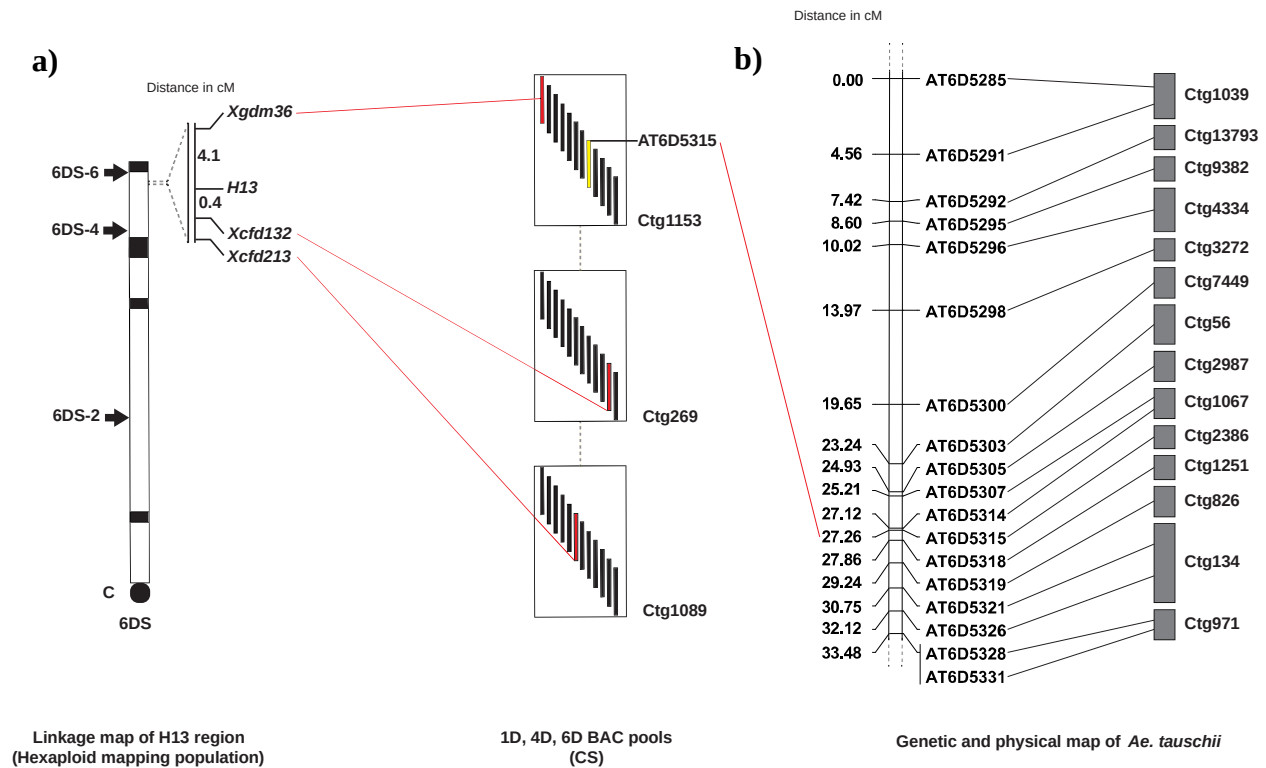


Figure 2.4 Anchoring of SSR markers on *Ae. tauschii* map

a) Linked SSR markers *Xgdm36* and *Xcfd132* were anchored onto the BAC pools of D genome (1D, 4D, 6D) assembly for Chinese Spring. Red lines indicate the position of markers *Xgdm36*, *Xcfd132*, *Xcfd213* on BAC pools of corresponding contigs Ctg1153, Ctg269 and Ctg1089, respectively.

b) *Ae. tauschii* marker AT6D5315 was mapped on Contig Ctg1153 leading to the anchoring of *Xgdm36* on the genetic and physical map of *Ae. tauschii*.

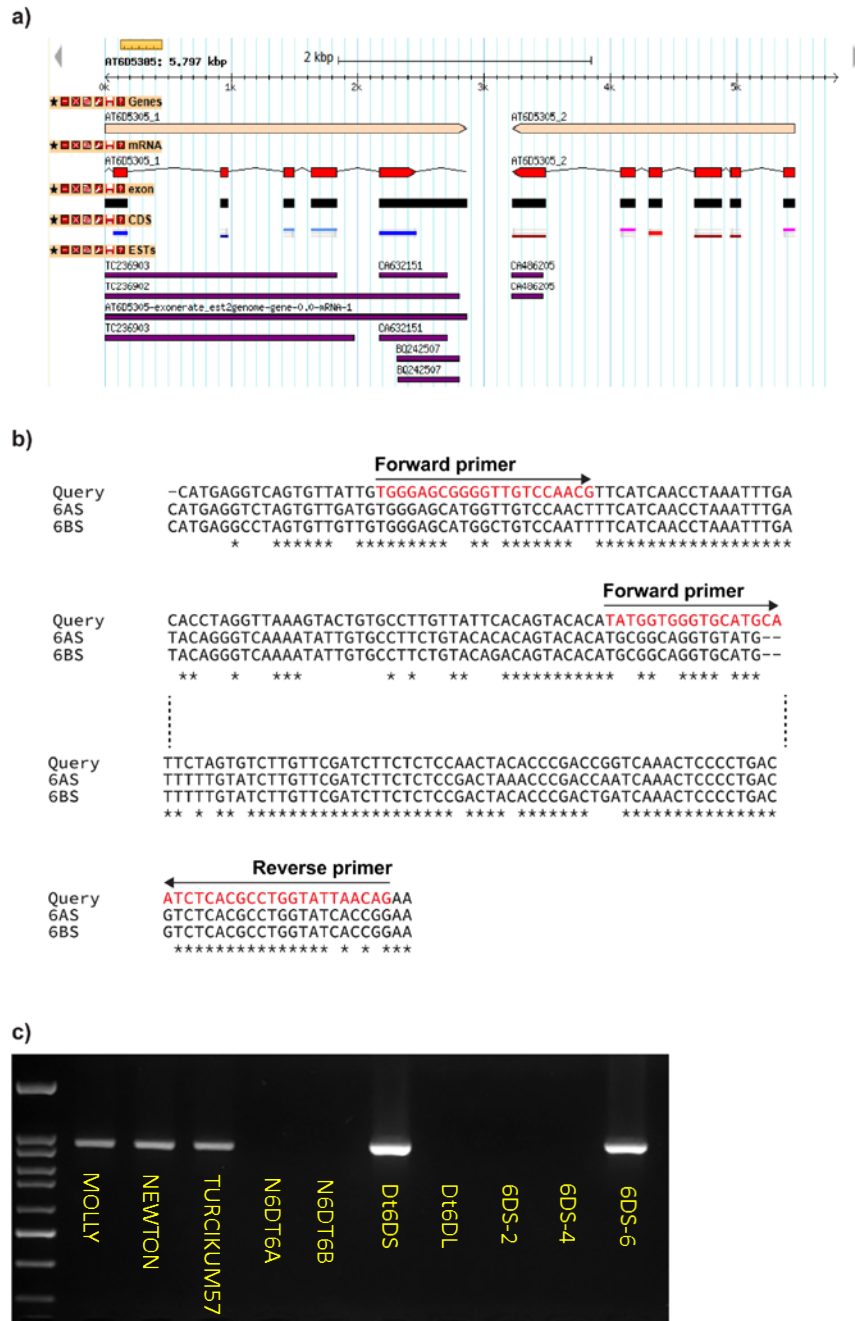


Figure 2.5 Development and validation of D genome-specific markers

a) Genic sequences obtained from Genome browser for *Ae. tauschii* were used to find homeologs for A and B genomes in Chinese Spring.

b) Using sequence information, D genome specific primers were designed using polymorphic sites

c) Gel picture showing testing for genome specificity and marker localization using nullisomic-tetrasomic (NT) lines, ditelosomic (dt) lines and deletion lines.

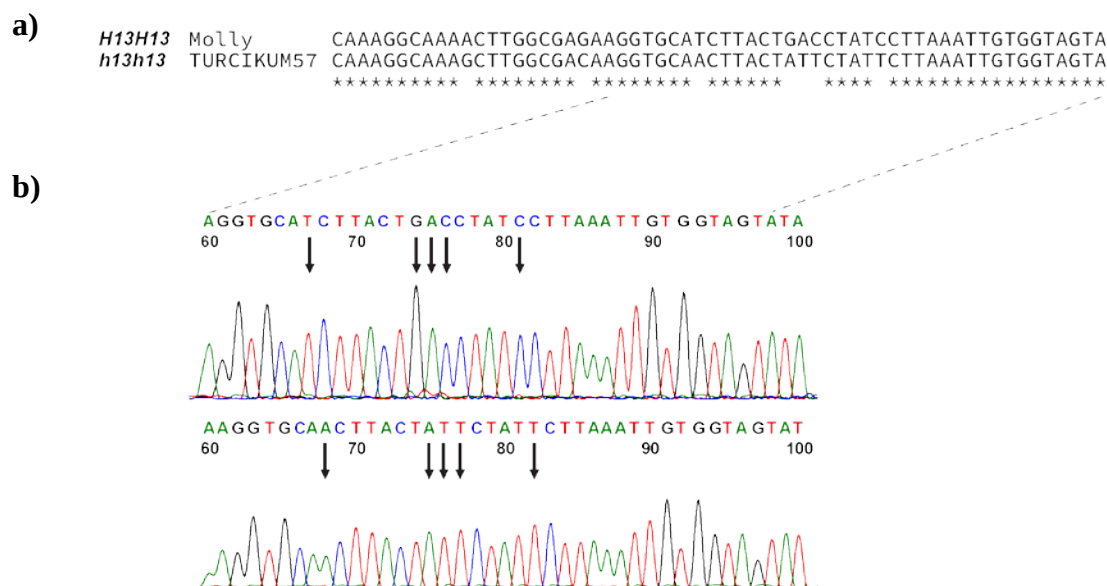


Figure 2.6 Identification of polymorphic markers for genetic mapping

a) Sequence alignment using CLUSTAL omega to identify polymorphic sites between parents of F₂ mapping population: Molly and Turcikum57.

b) Sequence trace files were compared for validation of SNPs using the BioEdit software. Arrows indicate polymorphic sites

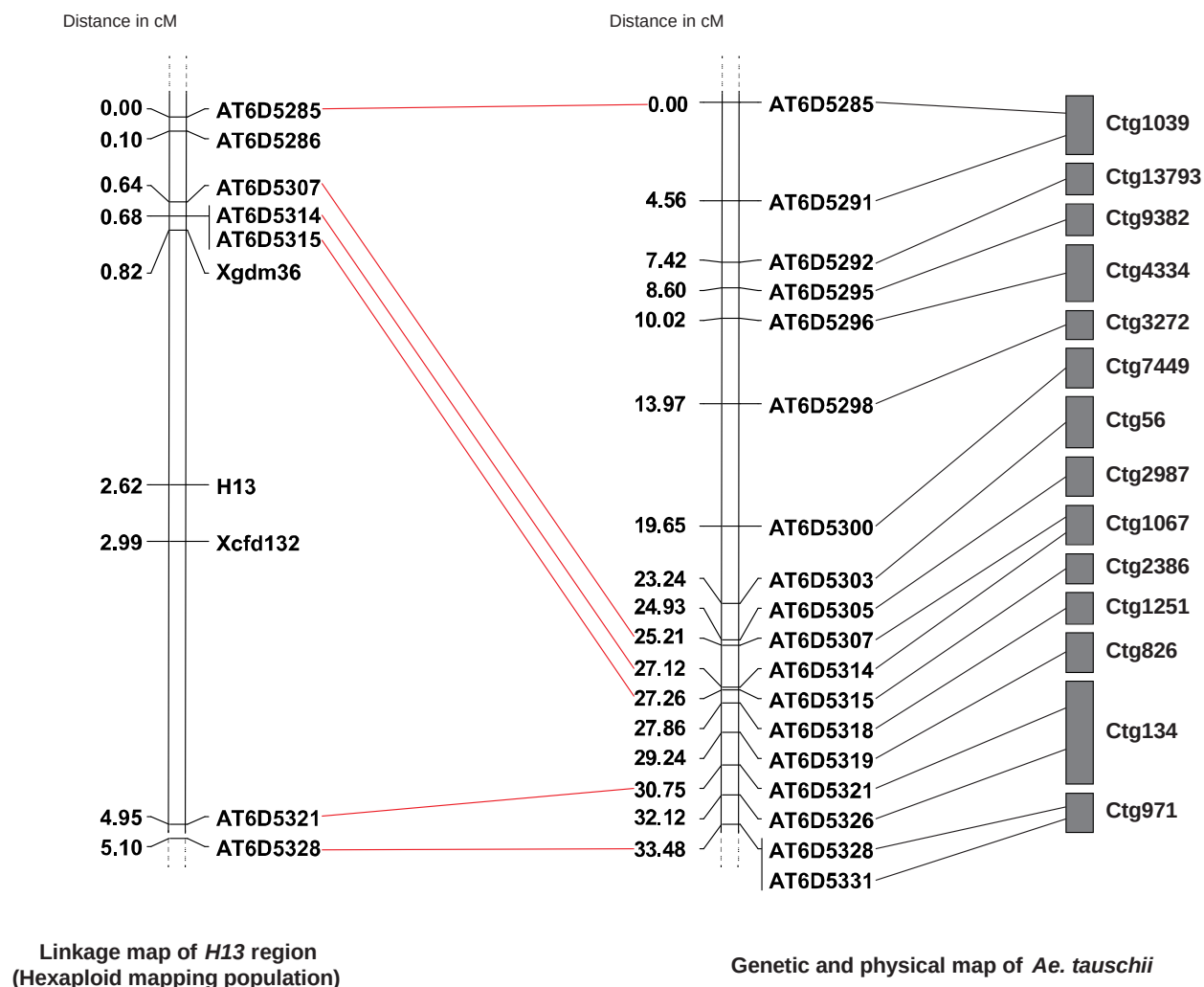


Figure 2.7 Genetic mapping of *Ae. tauschii* markers on high-resolution mapping population

F₂ high-resolution mapping population for Molly x Turcikum57 was genotyped using D genome specific polymorphic markers to generate a linkage map of *H13* region.

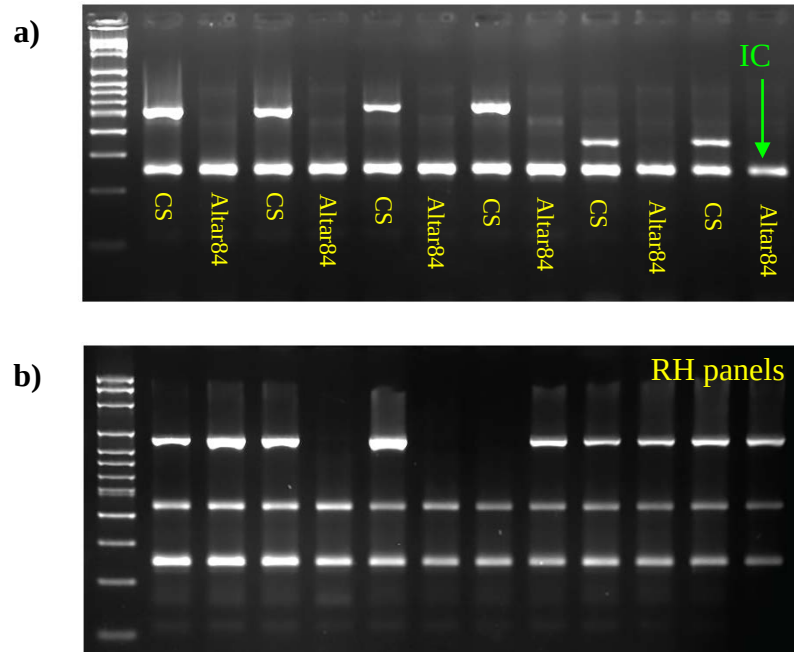


Figure 2.8 Radiation hybrid (RH) mapping

a) Validation of D genome specificity of primers on CS (1) and Altar84 (2) with an internal positive PCR control (IC).

b) D genome polymorphic markers were tested on radiation hybrid (RH) panels.

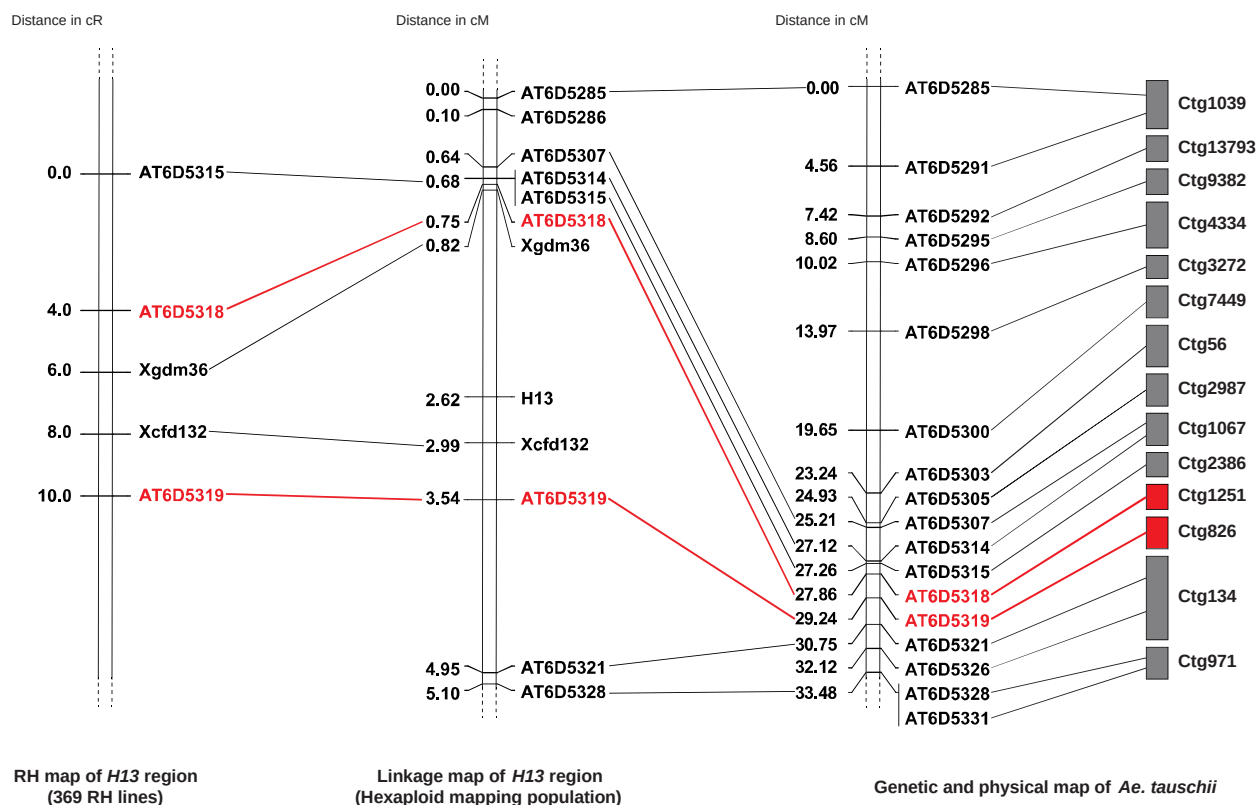


Figure 2.9 Comparative RH and genetic linkage map of *H13* region

Mapping of nearest collinear markers: AT6D5318 and AT6D5319 from the genetic and physical map of *Ae. tauschii* onto RH panels and hexaploid high-resolution mapping population.

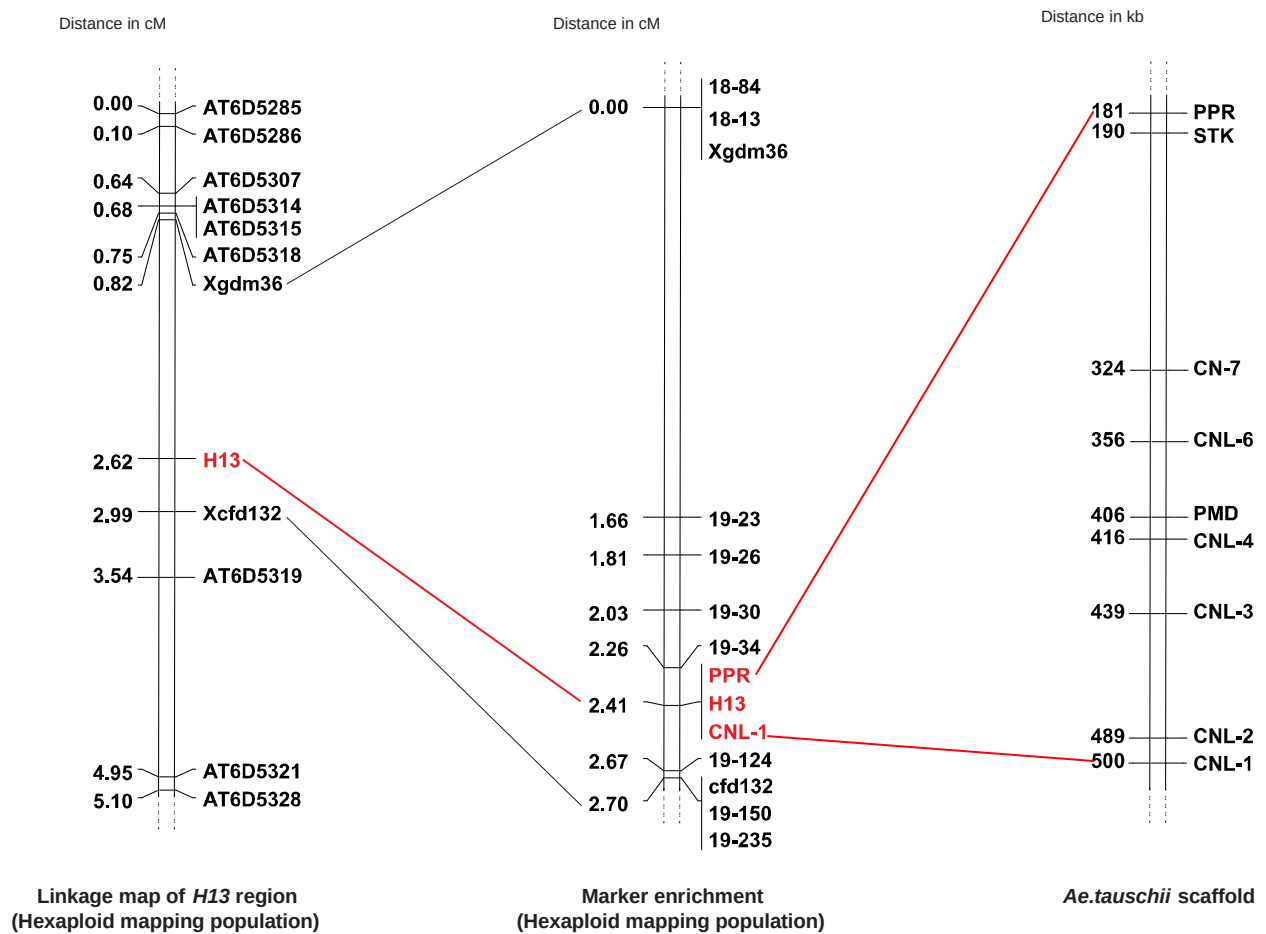
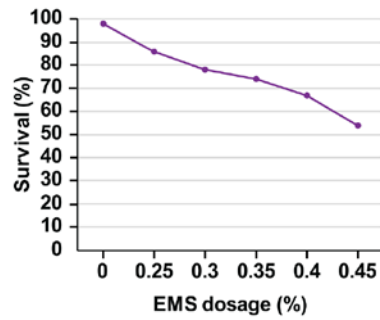


Figure 2.10 Fine mapping and identification of candidate genes co-segregating with *H13*

Genic markers 19-34 and 19-124 mapped 0.15cM distal and 0.26cM proximal to *H13* gene respectively. Nine candidate genes spanning an interval of 319Kb co-segregated with *H13* gene

a)



b)



Figure 2.11 Developing TILLING population and HF screening of M₃ population

a) EMS dosage was optimized, and 0.5% treatment was used to develop a TILLING population in Molly. Screening of the M₃ population was performed for resistant and susceptible plants.

b) The phenotype of wheat seedling 15 days post HF infestation. Molly and Karl92 (as shown in pic) control for resistant and susceptible phenotype respectively.

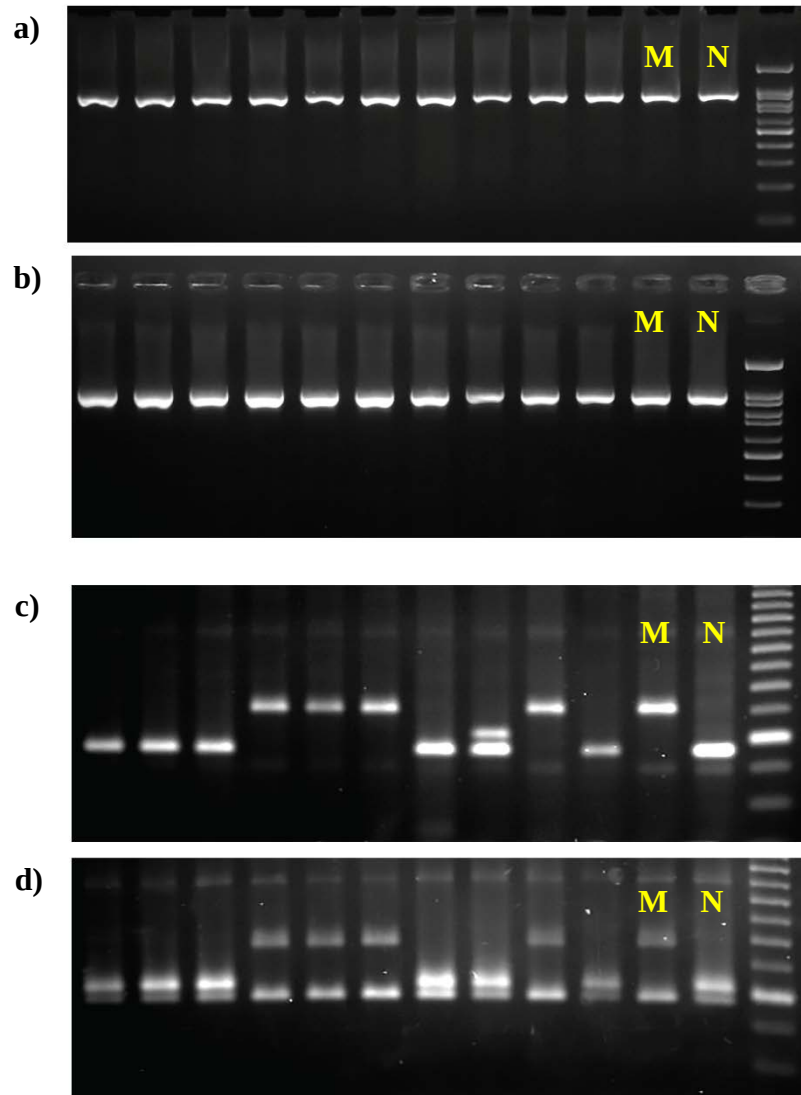


Figure 2.12 Characterization of susceptible lines from M₃ population

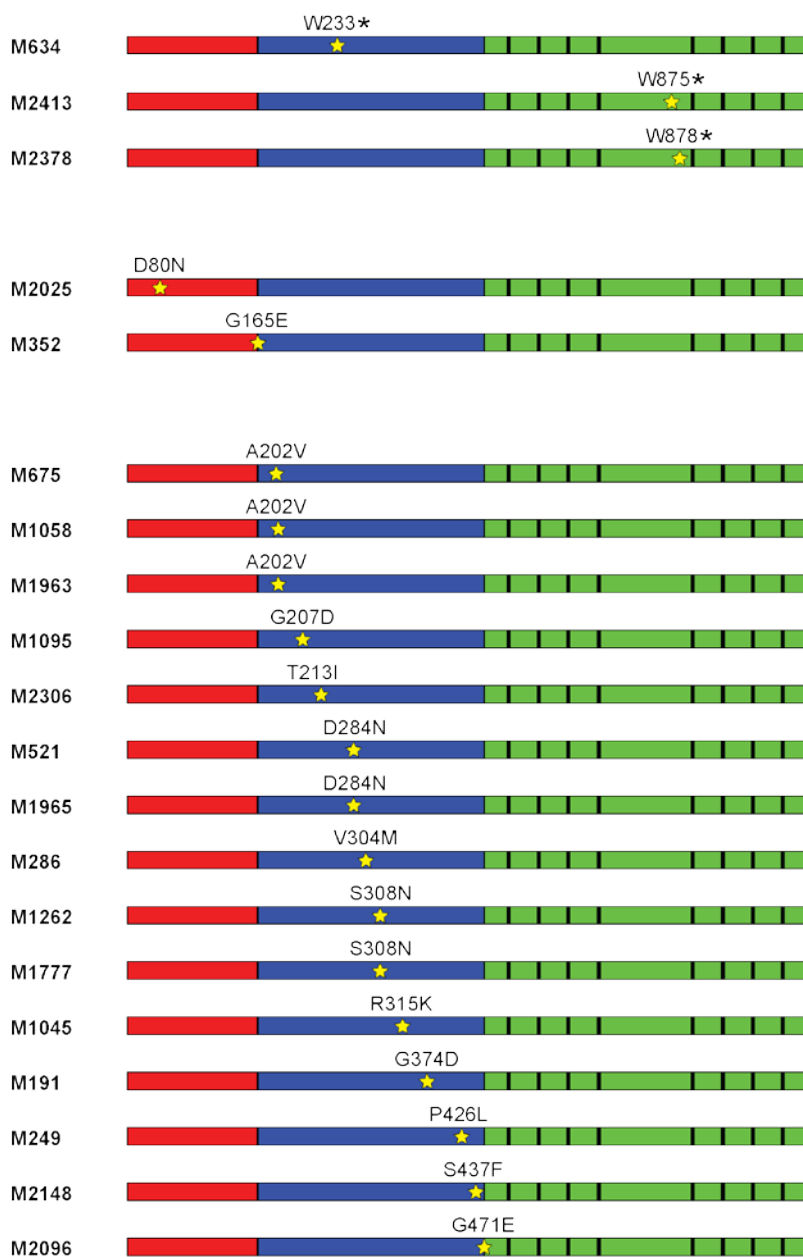
a, b) Susceptible lines were genotyped using genic markers 19-34 and 19-124 flanking the *H13* gene, to test for chromosome/segment deletions

c, d) Susceptible lines were genotyped using SSR markers *Xcfd132* and *Xgdm36* to rule out contaminating lines in M₃ population.

A9-Gene73_A09.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	115
A2-Gene73_A02.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	116
A4-Gene73_A04.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	115
A11-Gene73_A11.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	115
C53-Gene73_E05.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	113
C69-Gene73_F09.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	115
C60-Gene73_E12.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	113
C91-Gene73_H07.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	115
C59-Gene73_E11.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	114
C65-Gene73_F05.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	114
C92-Gene73_H08.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	114
C68-Gene73_F08.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	115
C70-Gene73_F10.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	116
C71-Gene73_F11.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	113
C85-Gene73_H01.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	116
A8-Gene73_A08.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	116
C94-Gene73_H10.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	120
C52-Gene73_E04.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	114
A12-Gene73_A12.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	119
C64-Gene73_F04.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	115
C50-Gene73_E02.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	115
C87-Gene73_H03.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	115
C86-Gene73_H02.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	115
C55-Gene73_E07.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	115
C63-Gene73_F03.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	114
C75-Gene73_G03.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	113
D1-Gene73_A01.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	114
A1-Gene73_A01.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	118
D5-Gene73_A05.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	113
C57-Gene73_E09.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	113
C61-Gene73_F01.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	112
A3-Gene73_A03.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	115
A5-Gene73_A05.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	117
A7-Gene73_A07.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	117
A15-Gene73_B03.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	118
A17-Gene73_B05.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	116
A18-Gene73_B06.abl	TGGCGAAGGTGCACCTTACTATCTATCTTAAATGTGGTAGTATATGGGAGATGA	119
C93-Gene73_H09.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	114
C77-Gene73_G05.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	116
C56-Gene73_E08.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	117
D8-Gene73_A08.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	113
A10-Gene73_A10.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	116
C79-Gene73_G07.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	120
C72-Gene73_F12.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	115
C95-Gene73_H11.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	115
C66-Gene73_F06.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	115
C80-Gene73_G08.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	115
C58-Gene73_E10.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	114
C90-Gene73_H06.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	115
C81-Gene73_G09.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	114
C62-Gene73_F02.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	115
C78-Gene73_G06.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	116
C54-Gene73_E06.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	116
C89-Gene73_H05.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	115
C49-Gene73_E01.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	117
C76-Gene73_G04.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	115
D6-Gene73_A06.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	114
C74-Gene73_G02.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	114
C88-Gene73_H04.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	118
A20-Gene73_B08.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	118
C73-Gene73_G01.abl	TGGCGAGAAGGTGCATCTTACTGACCTATCCTTAAATGTGGTAGTATATGGGAGATGA	113

Figure 2.13 Validation of contaminant lines by sequencing

Sequencing of *CNL-1* gene from susceptible EMS lines identified polymorphism (highlighted in red) in contaminant lines.



Contd

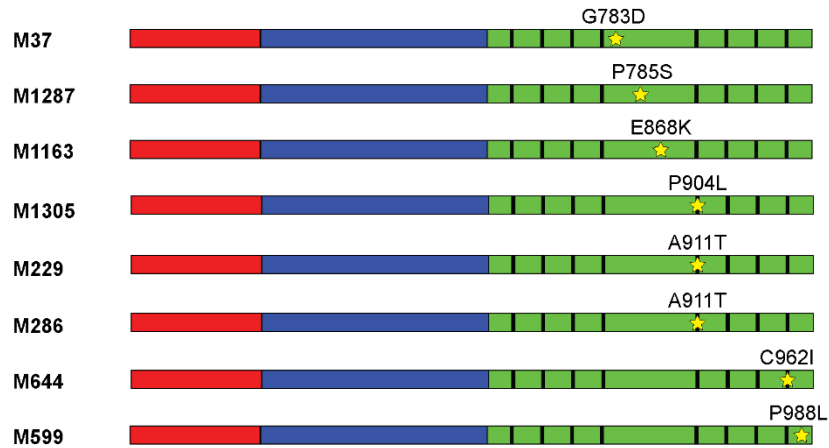


Figure 2.14 Summary of mutations identified in susceptible lines

List of all non-synonymous mutations identified in CC-NBS-LRR domains of *CNL-1* gene. Names of susceptible lines are shown on the left side, and amino acid change is indicated on top of each line. Yellow star indicates the location of the mutation in the protein. The CC, NBS and, LRR domains of the CNL-1 protein are colored red, blue and green, respectively.

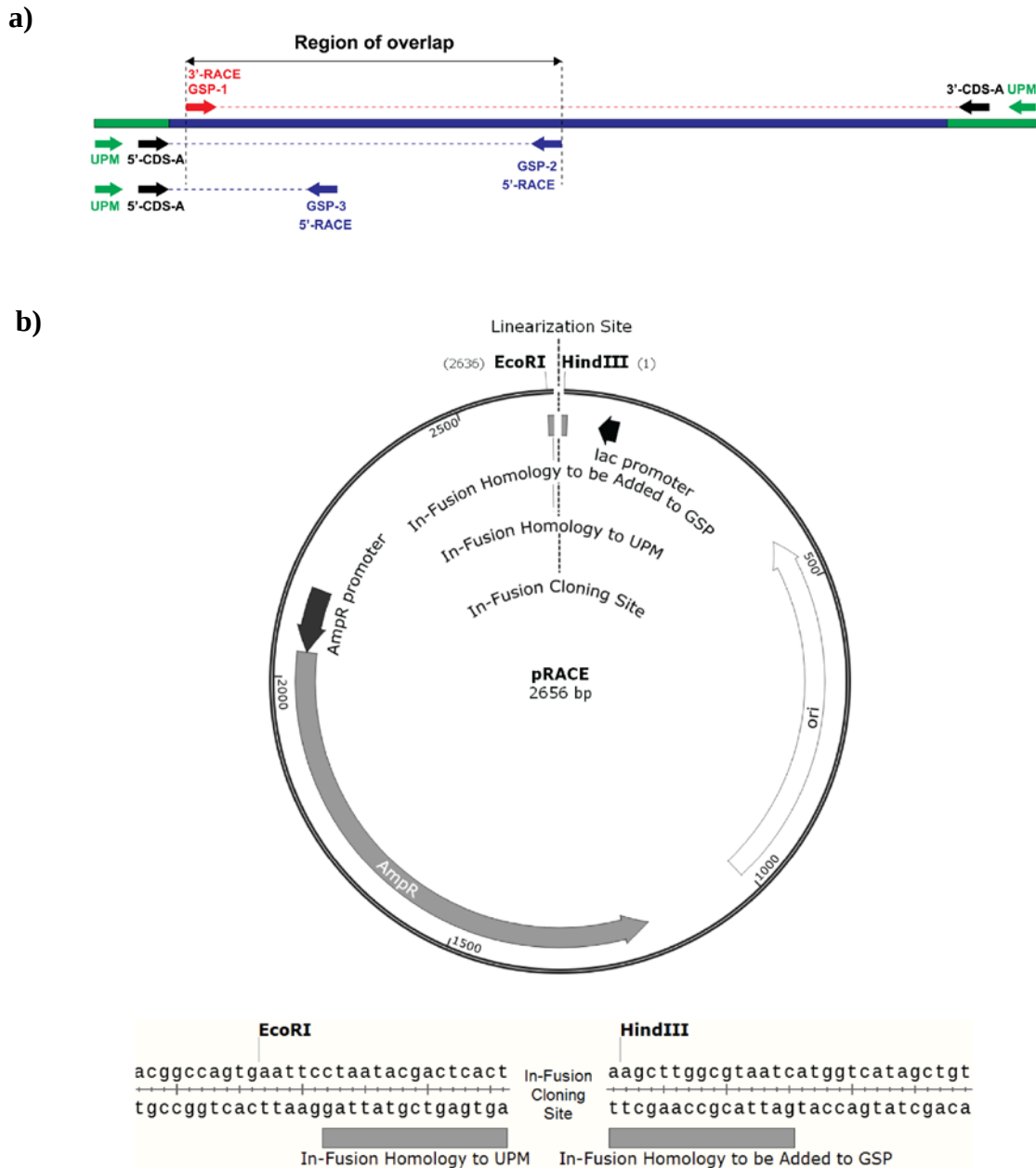


Figure 2.15 Primer locations and vector used for 5'-RACE and 3'-RACE

a) Locations of primers used for amplifying full-length cDNA from Molly and Newton gene are shown. Gene-specific primers (GSP) used for 5'-RACE and 3'-RACE along with primers from universal primers mix (UPM) and CDS primers are colored as blue, red, green and black respectively.

b) Vector map for pRACE vector

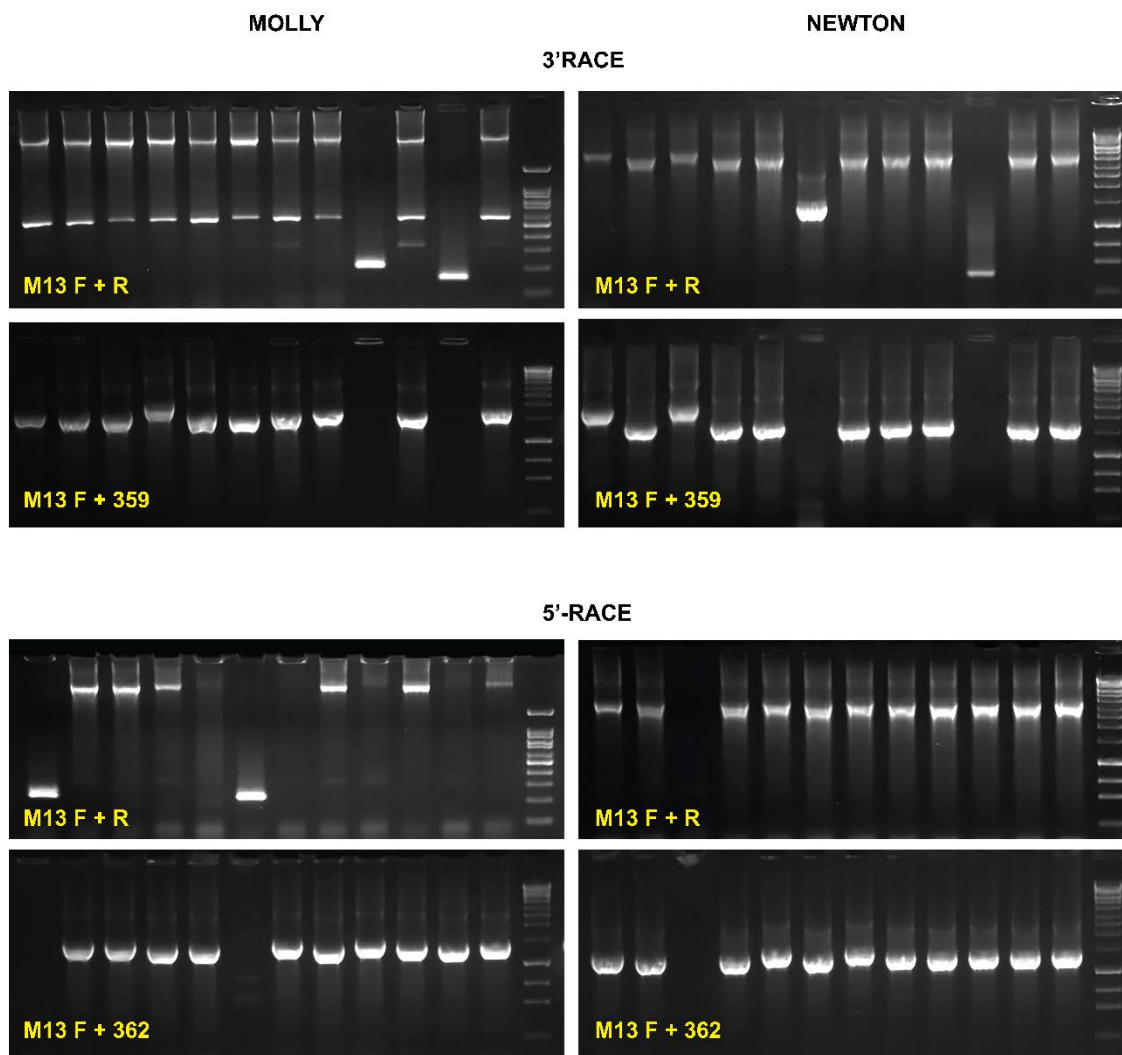


Figure 2.16 Colony screening for positive transformants from 5'-RACE and 3-RACE for *CNL-1*

Screening of transformed colonies for different 5'RACE and 3-RACE products of *CNL-1* from cDNA of Molly and Newton. Primer combinations used are labeled on their respective gel pictures.

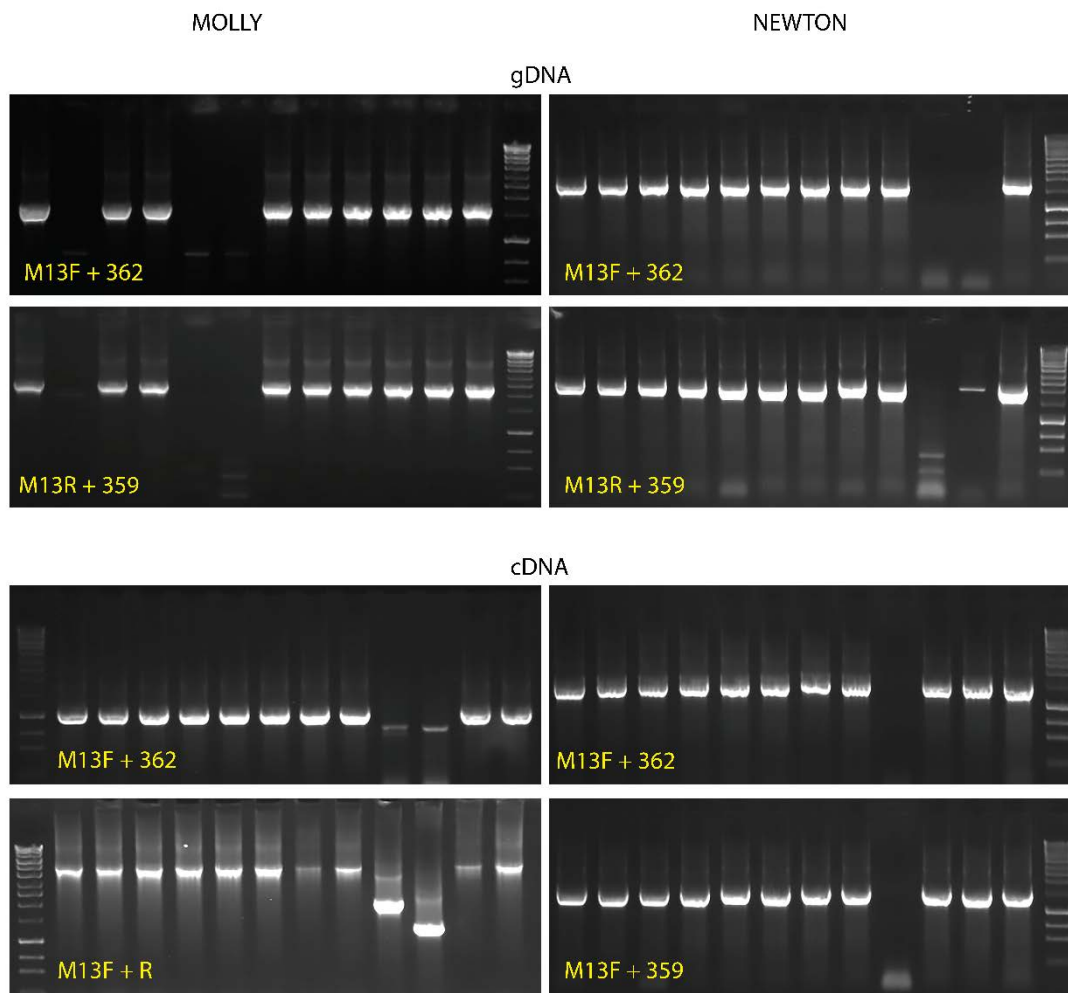


Figure 2.17 Colony screening for positive transformants for full-length gDNA and cDNA for *CNL-1*

Screening of transformed colonies for full-length genomic DNA (gDNA) and complementary DNA (cDNA) of *CNL-1* from Molly and Newton. Primer combinations are labeled on their respective gel pictures. Primer combinations used are labeled on their respective gel pictures.

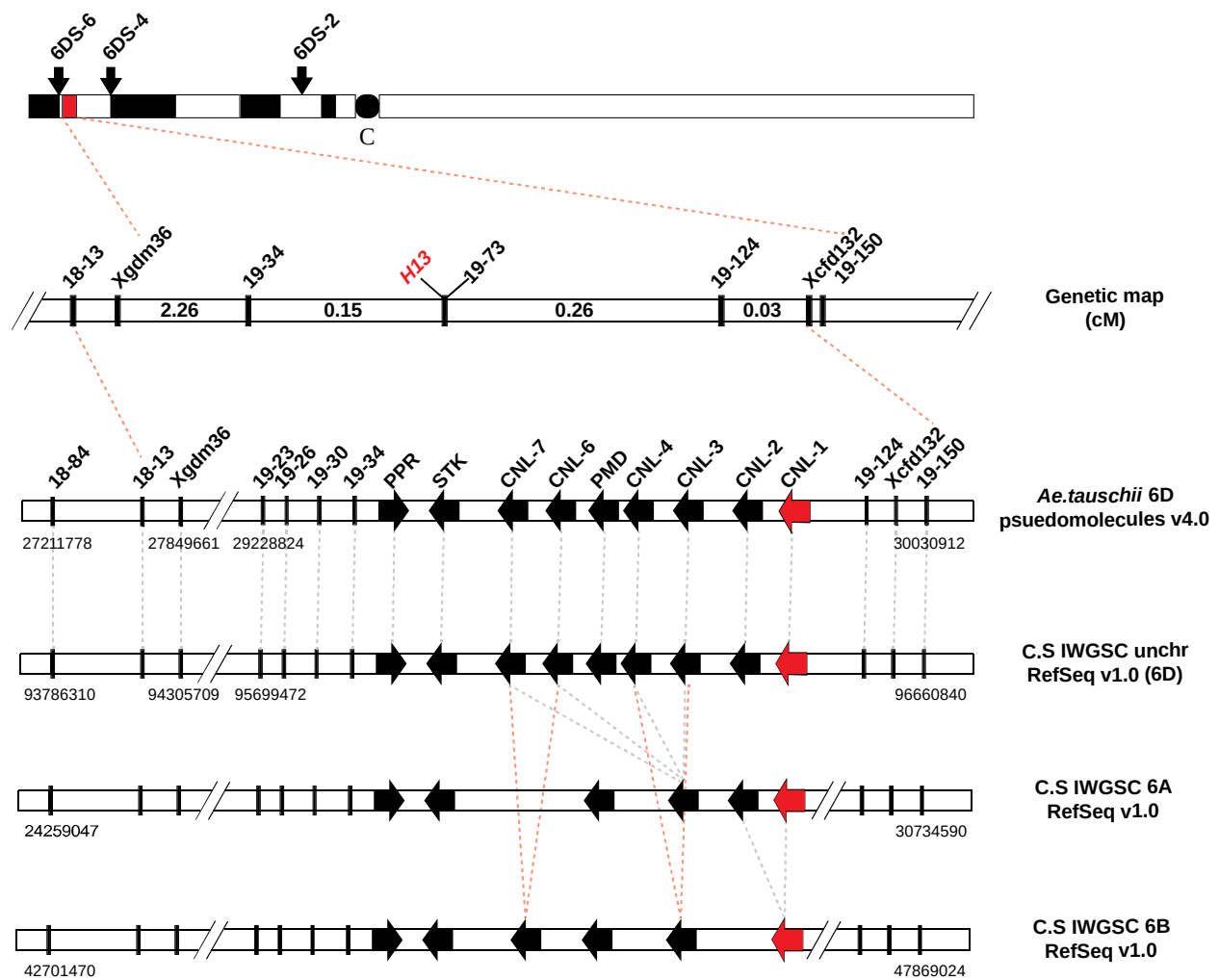


Figure 2.18 Genetic and physical map of *H13* region

A 900-kb region *H13* spanning region was collinear between 6D of *Ae. tauschii* and 6D of Chinese Spring

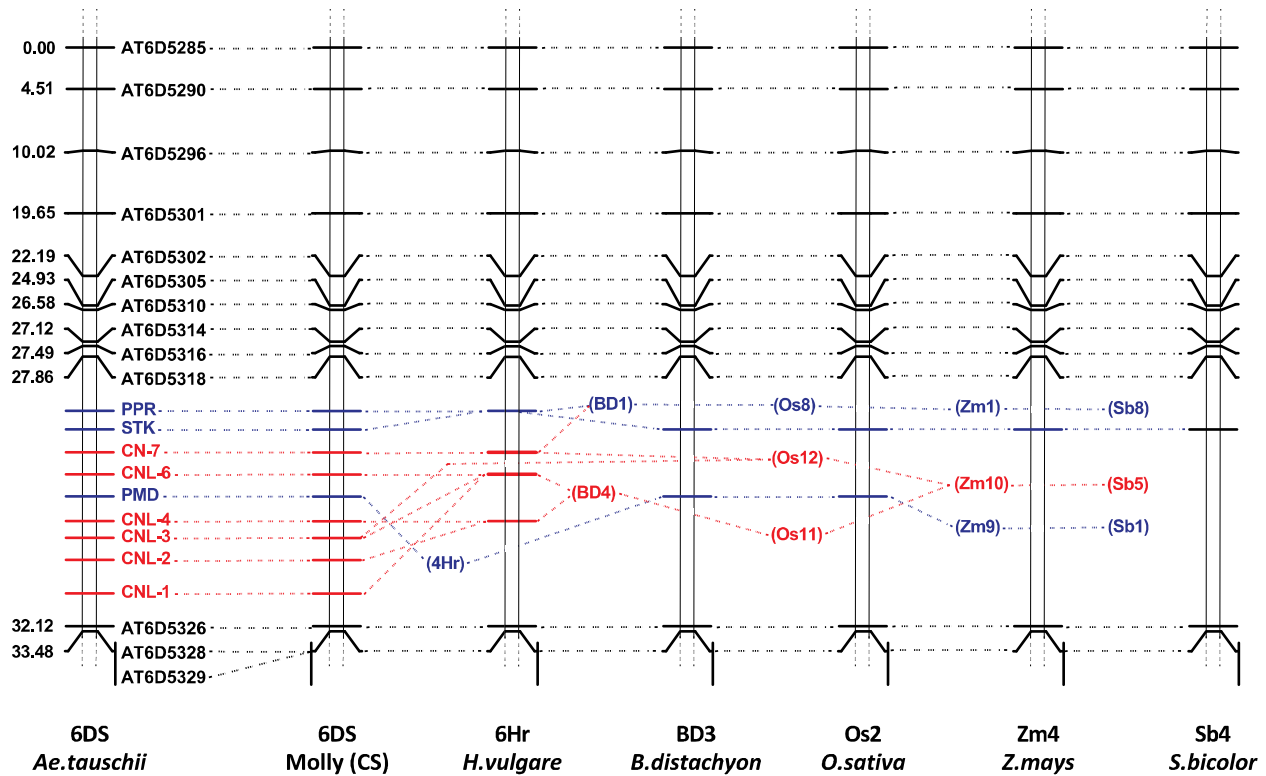
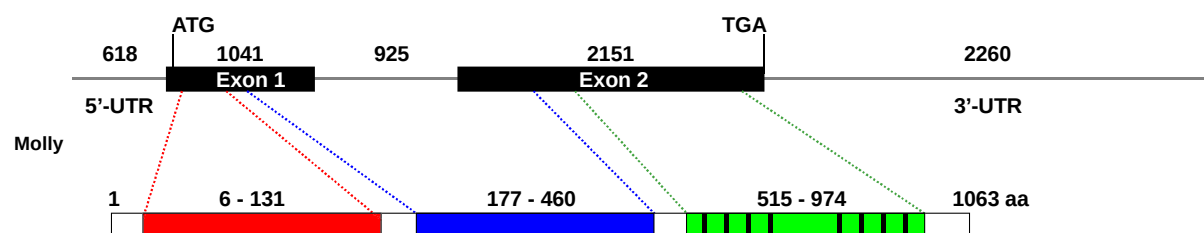


Figure 2.19 Comparative analysis of *H13* region

Homologs from other species were identified for genes near *H13* region. Candidate genes belonging to different protein families are colored. CNL encoding candidate genes, all other candidate genes and flanking genes are colored in red, blue and black respectively. Dotted lines are connecting homologs, and solid lines on syntenic chromosomes in different species indicate collinearity.

Figure 1

A)



B)

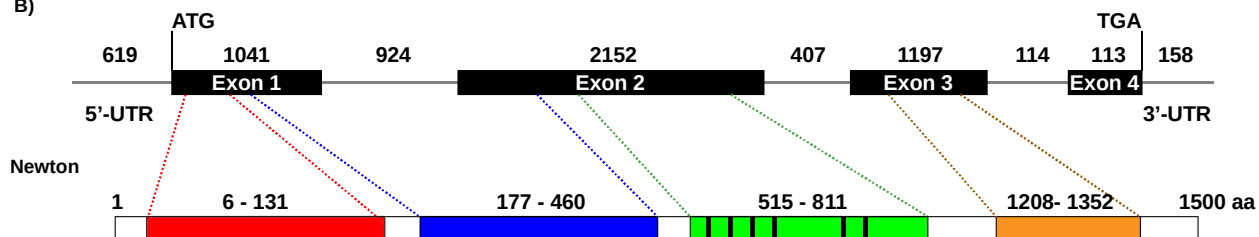


Figure 2.20 Gene structure and protein-domain prediction of *CNL-1*

Black filled boxes, and solid lines represent the coding and non-coding regions of the *H13* gene respectively. Start codon and stop codon are indicated by thin vertical lines. The predicted domains: CC, NBS, LRR, DDE endonuclease; of the *H13* gene from Molly and Newton are colored in red, blue, green and orange, respectively.

atgacgatggctcctggacgcctttgcatcctacgtgcaaaacatgctcacggagatggtg
 M T M V L D A F A S Y V Q N M L T E M V
 agtgaagaggtgcacatgctgcttggagtcacaaacgagatcaagaagatggacgtcaag
 S E E V H M L L G V T N E I K K M D V K
 ctttaaggacctcaagaacttcctcgccgacgctgaccggagaaacatcaccgacgagagc
 L K D L K N F L A D A D R R N I T D E S
 gtgcaggagtggtggcccagctcaagcgtgccatgtacgaagctacagacatccttgac
 V Q E W V A Q L K R A M Y E A T D I L D
 ctctgtcagctcaaggccatggagcgcggtccctccactgcagatgtgggggtccttcaac
 L C Q L K A M E R G P S T A D V G C F N
 cccttgctcttctgcatgcggaatccctcccatgcccacgacatcggcatccgcatcaag
 P L L F C M R N P S H A H D I G I R I K
 ggcgtcaaccagaacctcgacgacatcaaggagcgaagcgtgcttccagcttcatacct
 A L N Q N L D D I K E R S A A F S F I P
 cttgggtcctctgaggtacctagtagcaaggtgcaagcctctcgcttctggcaataagaga
 L G S S E D P S S K V Q A S R S G N K R
 cgtgagacgtcaggggtgtttgaccggctcggtgtagtcggagagaagattgaacaagac
 R E T S G V F D R S A V V G E K I E Q D
 acaagaaagctggttgagatcatgctgagtgaagaggaggcaaccccaacatcatggtg
 T R K L V E I M L S E K E G N P N I M V
 ttgcgcgttgctcggtgtcgccggaatcggaagaccactcttgcccacagggtcttcaat
 F A V V G V G G I G K T T L A H R V F N
 gacgaagccttgaacaacaagttccagaagacgatattggttgagcatcaacaaggacttt
 D E A L N N K F Q K T I W L S I N K D F
 gatgaggttgagttgctcaggacaatcatcacgcaagcggggggagtagcattgtggtcat
 D E V E L L R T I I T Q A G G V H C G H
 aaggcgttggtgtgcttcagccaatccttgctactaccttgaaggggaagaagctattc
 K A L A V L Q P I L V T T L K G K K L F
 ctggtgctggatgatgtgtggaaccatggggcgtgggatgatgtgcttaaaacacccttg
 L V L D D V W N H G A W D D V L K T P L
 gctaattgttggtccaggtagccgagtcctcgtcactactagagatgaaatagtagct
 A N V V A P G S R V L V T T R D E I V A
 cgaatgatgaaagtgtgcttcctaccaccatgtggacaaattagatgaggaggatgcc
 R M M K V V L P Y H H V D K L D E E D A
 tggctcattgtccagaaacagataatctcaagtgagaatgatggacgtgacattgatatg
 W S L L Q K Q I I S S E N D G R D I D M
 ctcaaggacattgggttgcaaatatagcaaatgtgatggtttgcctcttgacgtcaaa
 L K D I G L Q I I A K C D G L P L A V K
 gtaatgggaggtctattgtgtcagaaggacaaaaaccgccgtgagtgaggagatgggttg
 V M G G L L C Q K D K N R R E W E M V L
 gctgattctgtatggtcgatatctggaatgccgaagagctaaaccatgcagtgatttta
 A D S V W S I S G M P E E L N H A V Y L
 agctatgaagatttaccttctactgccaacaatgctttctgcatttctcccttctccct
 S Y E D L P S T A K Q C F L H F S L L P
 aaaactgcaaaagttttttgcaatgaaatcattggcatgtggataagtgaaggatttctt
 K T A K F F C N E I I G M W I S E G F L
 catgggacctcggtgacttagaagaactaggaagcaagtactataaggagcttatattg
 H G T S D D L E E L G S K Y Y K E L I L
 aggaatcttatagagccagatatagagtatatcgatcaatgtgtttgcaacatgcatgat
 R N L I E P D I E Y I D Q C V C N M H D
 gttgtacgctcattgtcagtttgtggctagagatgaggcactagcagctcacagtgagg
 V V R S F A Q F V A R D E A L A A H S G
 gaaactaatattgttagtagacttagtgacatgagtttcttcggttatctctagaaagc
 E T N I V S R L S A H E F L R L S L E S
 aaagcatcagaatcagatgggttagactggagttctttgcaggcacaaaagacactaagg
 K A S E S D G L D W S S L Q A Q K T L R
 gcgctaataatcagttggctatataaatatcaagcctgggtgattccttggttcattttcca
 A L I S V G Y I N I K P G D S L V H F P

```

tgtctacgaactttatatatatgttctgcacatgtttagcattgcatgaatctttgcat
C L R T L Y I C S A H V V A L H E S L H
gaactcaagcacctgaggtacttgtcattggaaaactgatatatctagtctgccggat
E L K H L R Y L S L E N T D I S S L P D
agcattggaagatgaaatttttgcagtacattagccttcatggatgcaacaatttgtg
S I G K M K F L Q Y I S L H G C K Q F V
aaagtctccgcatagctttgttaaagtttagggcagctaaggtaccttagcttcagtgaaca
K V P H S F V K L G Q L R Y L S F S G T
agcatagaaggcataccttaggggtttcatgctctaacgagcatgagggattacatgggt
S I E G I P R G F H A L T S M R V L H G
tttcagctcaggtggatgggtgattggtagtttggaagagttgggtcctctttctcag
F P A Q V D G D W C S L E E L G P L S Q
ctcaagcatcttggattggacggattggagaatataactgcttcctcatctgcagcaaag
L K H L G L D G L E N I T A S S S A A K
gcaaaacttggcgagaaggtgcatcttactgacctatccttaaattgtggtagtatattg
A K L G E K V H L T D L S L N C G S I L
ggagatgatgggctgattaaagaggaagatagtgctcctcgagcaagagcaacagcgaatc
G D D G L I K E E D S V S E Q E Q Q R I
aagaaggtgtttgatgagctctgcctccatccagggttagaataatctttgcatcaatgga
K K V F D E L C P P S R L E Y L C I N G
cattttggccaacgactcccaaggtggatgatgtcatcatcagctgtgtacctcaatagt
H F G Q R L P R W M M S S S A V Y L N S
ttgaggattctgttccctgaagacatggcttgctgcatacaacttccctgatggcttgggt
L R I L F L E D M A C C I Q L P D G L C
caactcccctatttgaattcattcagatcaagcgtgctccagccatcaagcgtgttgga
Q L P Y L E F I Q I K R A P A I K R V G
ccagaattcatgcagtcctaccgtcagcacagtcctcgtccttcccagatgggtgcctgca
P E F M Q S Y R Q H S P R P S Q M V P A
tttcgagattgcatgagttggaatttacagggaatgggtggaatgggaggagtgaggagtg
F P R L H E L E F T G M V E W E E W E W
gaggaggaagtgcgaagcctttctgtcttgcaggaacttatgctaaagcaatgcaaattg
E E E V Q A F P V L Q E L M L K Q C K L
aagtgtcttctcccggttgcctccaggaagggtttgaataaattattcatatac
K C L P P G L A S Q A R A L N K L F I Y
tacgtccagggtctcatctctctagagaactttccatctctgggtgagctggatctggat
Y V Q G L I S L E N F P S L V E L D L D
gaaaaccttgacctagagaggtactaatacttccaagacttcagatgcttagcatcgaa
E N L D L E R I T N L P R L Q M L S I E
gactgccccaaagctgaaggtgttggaagggtgttctctgcactcgagagggtcactacacg
D C P K L K V L D G V P A L E R L I L T
gatgaagacatgaagacactcccagaatacatgggaggtataaaccgaaggcatttggag
D E D M K T L P E Y M G G I N P R H L E
ctatattgtagccttgcactgctcacttccatagccgctgcacagctctggccctgagtg
L Y C S L A L L T S I A A A Q S G P E W
gacaagttcagccatgttaagcatgtcaaagcatatgcatctgaaggagataatacaaaag
D K F S H V K H V K A Y A S E G D N T K
aaatgggtcgtcttgtacacagctgatccctacaacttggagacaaatgtcagccgttct
K W F V L Y T A D P Y N L E T N V S R S
ttcatgtcctga
F M S -

```

Figure 2.21 Sequence of H13 gene and its deduced amino acid sequence

Full-length gene sequence from gDNA and cDNA identified coding sequence and deduced amino acid sequence.

Molly	MTMVLDAFASYVQNMLTEMVSEEVHMLLGVNEIKKMDVKLKDLDKNFLADADRNRNITDES	60	}	
Newton	MTMVLDAFASYVQNMLTEMVSEEVHMLLGVNEIKKMDVKLKDLDKNFLADADRNRNITDES	60		
Molly	VQEWVAQLKRAMYEATDILDLCQLKAMERGPSTADVGCNPLLFMRNPNSHAHDIGIRIK	120	}	Coiled coil (CC) domain (6 - 131 aa)
Newton	VQEWVAQLKRAMYEATDILDLCQLKAMERGPSTADVGCNPLLFMRNPNSHAHDIGIRIK	120		
Molly	ALNQNLDDIKERSAAFSFIPLGSSSEDPSSKVQASRSNGKRRETSGVFDRSAVVGEKIEQD	180	}	
Newton	ALNQNLDDIKERSAAFSFIPLGSSSEDPSSKVQASRSNGKRRETSGVFDRSAVVGEKIEQD	180		
Molly	TRKLVIEIMLSEKEGNPNIMVFAVVGVGIGKTTLAHRVFNDEALNNKFQKTIWLSINKDF	240	}	
Newton	TRKLVIEIMLSEKEGNPNIMVFAVVGVGIGKTTLAHRVFNDEALNNKFQKTIWLSINKDF	240		
Molly	DEVELLRITITQAGGVHCGHKALAVLQPIVLTTLGKKLFLVLDDVWNGAWDDVLKTP	300	}	
Newton	DEVELLRITITQAGGVHCGHKALAVLQPIVLTTLGKKLFLVLDDVWNGAWDDVLKTP	300		
Molly	ANVVPAGSRVLVTTTDEIVARMKVVLPHYHVDKLDEEDAWSLLQKQIISSENDGRDIDM	360	}	NB-ARC domain (177 - 460 aa)
Newton	ANVVPAGSRVLVTTTDEIVARMKVVLPHYHVDKLDEEDAWSLLQKQIISSENDGRDIDM	360		
Molly	LKDIGLQIIAKCDGLPLAVKVMGGLLCQKDKNRREWEMVLADSVWSISGMPEELNHAVYL	420	}	
Newton	LKDIGLQIIAKCDGLPLAVKVMGGLLCQKDKNRREWEMVLADSVWSISGMPEELNHAVYL	420		
Molly	SYEDLPSTAKQCFLHFSLLPKTAKFFCNEIIGMWISGFLHGTSDDEELGSKYYKELIL	480	}	
Newton	SYEDLPSTAKQCFLHFSLLPKTAKFFCNEIIGMWISGFLHGTSDDEELGSKYYKELIL	480		
Molly	RNLIEPDIEYIDQCVCNMHDVVRSAQFVARDEALAAHSGETNIVSRLSAHEFLRLSLES	540	}	
Newton	RNLIEPDIEYIDQCVCNMHDVVRSAQFVARDEALAAHSGETNIVSRLSAHEFLRLSLES	540		
Molly	KASESDGLDWSSLOAQKTLRALISVGYNIKPGDSLHFPCLRTLYICSAHVVALHESLH	600	}	
Newton	KASESDGLDWSSLOAQKTLRALISVGYNIKPGDSLHFPCLRTLYICSAHVVALHESLH	600		
Molly	ELKHLRYLSLENTDISSLPDSIGKMKFLQYISLHGCEQFVKVPHSFVKLGQLRYLSFSGT	660	}	
Newton	ELKHLRYLSLENTDISSLPDSIGKMKFLQYISLHGCEQFVKVPHSFVKLGQLRYLSFSGT	660		
Molly	SIEGIPRGFHALTSMRVLHGFPQVGDWCSLEELGPLSQLKHLGLDLENITASSSAK	720	}	LRR repeat (515 - 974 aa)
Newton	SIEGIPRGFHALTSMRVLHGFPQVGDWCSLEELGPLSQLKHLGLDLENITASSSAK	720		
Molly	AKLGEKVHLTDLNLNCGSILGDDGLIKEEDSVSEQQRIKKVFDELCPPSRLEYLCING	780	}	
Newton	AKLGEKVHLTDLNLNCGSILGDDGLIKEEDSVSEQQRIKKVFDELCPPSRLEYLCING	780		
Molly	HFGQLRPLWMMSSSAVYLNSLRILFLEDMACCIQLPDGLCQLPYLEFIQIKRAPAIKRVG	840	}	
Newton	HFGQLRPLWMMSSSAVYLNSLRILFLEDMACCIQLPDGLCQLPYLEFIQIKRAPAIKRVG	840		
Molly	PEFMQSYRQHSPPSQMVPAFPRLHELEFTGMVEWEWEWEEVQAFPPVLQELMLKQCKL	900	}	
Newton	PEFMQSYRQHSPPSQMVPAFPRLHELEFTGMVEWEWEWEEVQAFPPVLQELMLKQCKL	900		
Molly	KCLPPGLASQARALNKLFIYVQGLISLENFPSLVELDLDENLDLEITNLPRLQMLSIE	960	}	
Newton	KCLPPGLASQARALNKLFIYVQGLISLENFPSLVELDLDENLDLEITNLPRLQMLSIE	960		
Molly	DCPKLKVLGDGPALERLILTDEDMKTLPFYMGGINPRHLELYCSLALLTSIAAAQSGPEW	1020	}	
Newton	DCPKLKVLGDGPALERLILTDEDMKTLPFYMGGINPRHLELYCSLALLTSIAAAQSGPEW	1020		

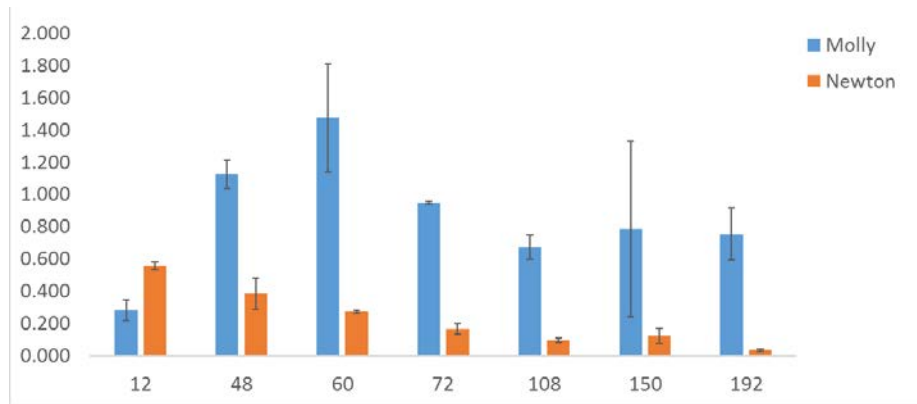
Molly	DKFSHVKKVKAYASEGDNTKKWFVLYTADPYNLETNVSRSFMS-----	1063	
Newton	DKFSHVKKVKAYASEGDNTKKWFVLYTADPYNLETNVSRSFMSRGTLTSFEDTQRFEHVF	1080	

Molly	-----	1063	
Newton	KMTRKTFYSYICSLVMKPSMKDMNSYTFVDGRVLSLEDRAVAVSLKRLYSSEPPETIGSSVG	1140	
Molly	-----	1063	
Newton	VNESTLLLVTESEFVVSVDKRGHHSRWPDSSMDKIKSMFEKIHNMENCCGVICTTHIPF	1200	
Molly	-----	1063	
Newton	GENWDREKND SILVQLVIDPEMRFSIWWGRISSMNL SILHESNLFNECEGGYLNGSK	1260	
Molly	-----	1063	
Newton	LKVALDGSEVG EYIIGDAGFPLLPWLLTPYQEEDLSDSKAEFNRRHNAATTCALKALARF	1320	DDE endonuclease (1208 - 1352aa)
Molly	-----	1063	
Newton	KDTWKYLQGETSCPVNPD TLVETIHACCYLHNIVIDMEEDGAIPSEIEPDSEDDAATSS	1380	
Molly	-----	1063	
Newton	AEERDSEDDAATSSAEFPDSEDDAATSSAEFPDSEDDAATLSVEEPDESEELRQVAN	1440	
Molly	-----	1063	
Newton	EDAVVARDMLSQYFLTTMSESGVSPVDAEEDHEVVALGSGDEDKEQEVQTRRADEEIV	1500	

Figure 2.22 Alignment of amino acid sequences from Molly and Newton

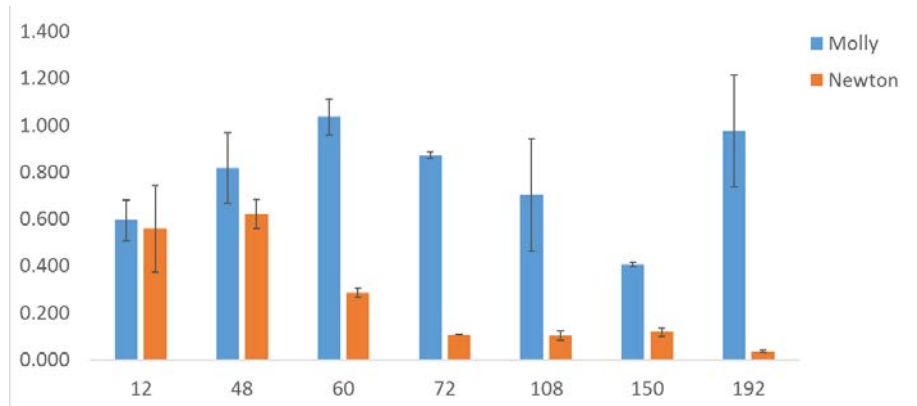
Clustal Omega alignment of the amino acid sequences deduced from the coding sequence. Most of the variation between Molly and Newton was present in LRR repeat sequence. Newton has an additional DDE endonuclease.

Infested (GP) vs uninfested



RT PCR primer 1

Infested (GP) vs uninfested



RT PCR primer 2

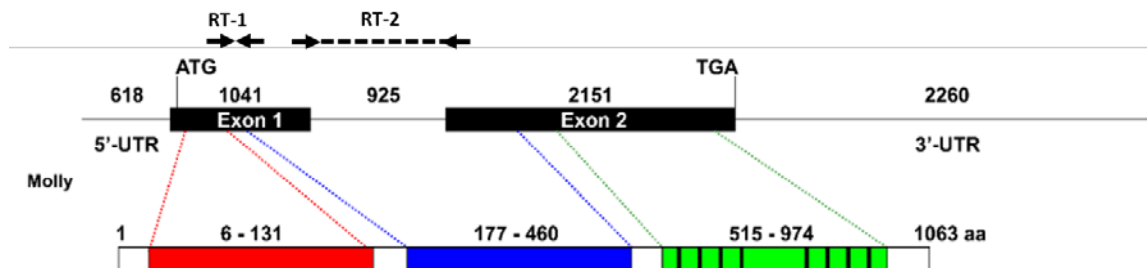


Figure 2.23 Expression profile of *H13* gene

Quantitative PCR from the feeding site of Molly and Newton upon infestation with GP biotype flies.

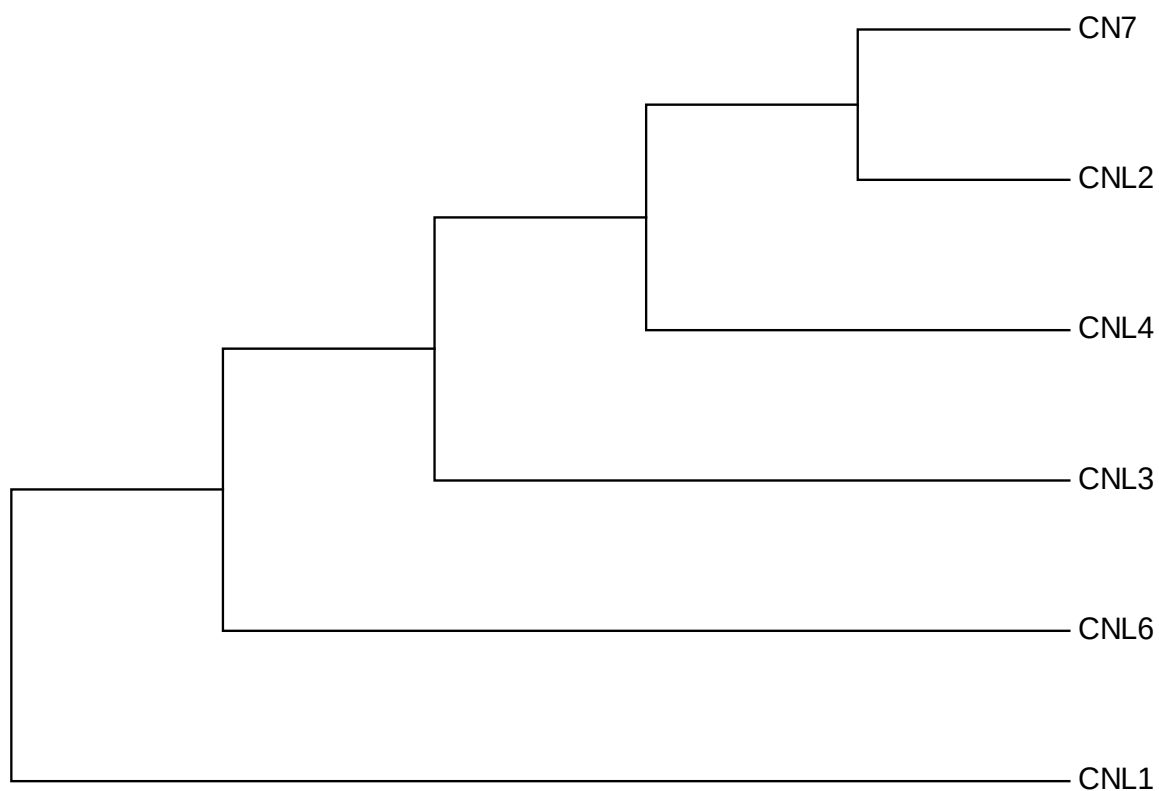


Figure 2.24 Neighbor-joining tree between CNL-1 and other candidate CNLs Full protein sequence was used to construct neighbor-joining tree between CNL-1 and other cloned CNLs identified in *H13* region.

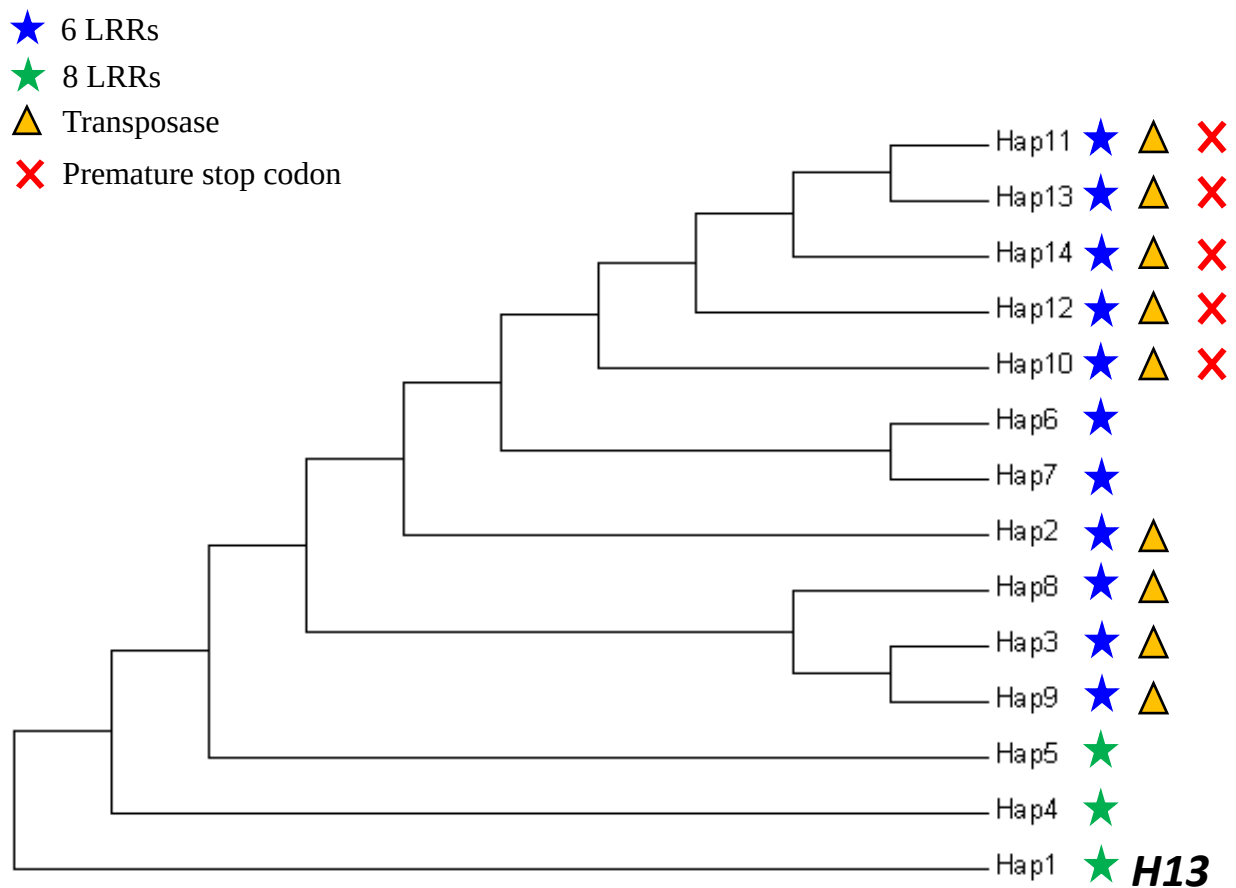


Figure 2.26 Neighbor-joining tree showing relationships between Hap1 and other haplotypes.

Full-length gene was amplified from different accessions and haplotypes were identified from deduced amino acid sequences. A neighbor-joining tree between Hap1 and other haplotypes was constructed.

2.7 Tables

Table 2.1 HFR genes with their source and chromosome locations

R gene Chr loc	Source	Chromosome location
<i>H1</i>	<i>Triticum aestivum</i>	
<i>H2</i>	<i>Triticum aestivum</i>	
<i>H3</i>	<i>Triticum aestivum</i>	1AS
<i>H4</i>	<i>Triticum aestivum</i>	
<i>H5</i>	<i>Triticum aestivum</i>	1AS
<i>H6</i>	<i>Triticum turgidum</i> ssp. <i>durum</i>	1AS
<i>H7</i>	<i>Triticum aestivum</i>	5D
<i>H8</i>	<i>Triticum aestivum</i>	5D
<i>H9</i>	<i>Triticum turgidum</i> ssp. <i>durum</i>	1AS
<i>H10</i>	<i>Triticum turgidum</i> ssp. <i>durum</i>	1AS
<i>H11</i>	<i>Triticum turgidum</i> ssp. <i>durum</i>	1AS
<i>H12</i>	<i>Triticum aestivum</i>	1AS
<i>H13</i>	<i>Aegilops tauschii</i> , KU2076	6DS
<i>H14</i>	<i>Triticum turgidum</i> ssp. <i>durum</i>	5A
<i>H15</i>	<i>Triticum turgidum</i> ssp. <i>durum</i>	5A
<i>H16</i>	<i>Triticum turgidum</i> ssp. <i>durum</i>	1A
<i>H17</i>	<i>Triticum turgidum</i> ssp. <i>durum</i>	1A
<i>H18</i>	<i>Triticum turgidum</i> ssp. <i>durum</i>	2B
<i>H19</i>	<i>Triticum turgidum</i> ssp. <i>durum</i>	5A
<i>H20</i>	<i>Triticum turgidum</i> ssp. <i>durum</i>	2B
<i>H21</i>	<i>Secale cereale</i>	2B
<i>H22</i>	<i>Aegilops tauschii</i> , TA1644	1D
<i>H23</i>	<i>Aegilops tauschii</i> , TA1642	6D
<i>H24</i>	<i>Aegilops tauschii</i> ,	3DL
<i>H25</i>	<i>Secale cereale</i>	4AL

R gene Chr loc	Source	Chromosome location
<i>H26</i>	<i>Aegilops tauschii</i> , TA2473	3DL
<i>H27</i>	<i>Aegilops ventricosa</i>	4D
<i>H28</i>	<i>Triticum turgidum</i> ssp. <i>durum</i>	5A
<i>H29</i>	<i>Triticum turgidum</i> ssp. <i>durum</i>	5A
<i>H30</i>	<i>Aegilops triuncialis</i>	
<i>H31</i>	<i>Triticum turgidum</i> ssp. <i>durum</i>	5BS
<i>H32</i>	<i>Aegilops tauschii</i>	3DL
<i>H33</i>	<i>Triticum turgidum</i> ssp. <i>Durum</i> , PI 134942	3AS
<i>H34</i> (QTL)	<i>Triticum turgidum</i> ssp. <i>durum</i>	6BS
<i>Hdic</i>	<i>Triticum turgidum</i> L. Subsp. <i>dicoccon</i>	1AS
<i>H_{wgrc4}</i>	<i>Aegilops tauschii</i> , TA1695	6D
<i>HNC09MDD14</i>	<i>Aegilops tauschii</i> , TA2492 or TA2377	6DS

Table 2.2 Plant materials used in this study

Accession/TA	Cultivar	Description
PI 562619	Molly	Resistant near-isogenic line, <i>H13H13</i> , hard red winter wheat
TA3002 (Cltr 17715)	Newton	Susceptible near-isogenic line, <i>h13h13</i> , hard red winter wheat
TA2946 (PI 372129)	Turcikum57	Turkmenistan Winter Wheat, <i>Dn4Dn4</i>
TA2962	Chinese Spring	Chinese spring wheat
TA2970	Altar84	Mexican spring durum wheat
	Genetic Stocks	
TA4544	CS del6DS-2	Deletion stocks for Chr 6DS, del6DS-2-0.45
TA4544	CS del6DS-4	Deletion stocks for Chr 6DS, del6DS-4-0.79
TA4544	CS del6DS-6	Deletion stocks for Chr 6DS, del6DS-6-0.99
TA3156	CS N6D-T6A	Nullisomic-Tetrasomic
TA3157	CS N6D-T6B	Nullisomic-Tetrasomic
TA3128	CS Dt6DS	Ditelosomic
TA3129	CS Dt6DL	Ditelosomic
	Mapping population	
		An F ₂ population derived from a cross between Molly and Turcikum57
		RH mapping population generated from Chinese spring (CS) and Altar84
	Hf screening checks	
Cltr 17897	Caldwell	Soft red winter wheat (resistant)
PI 562612	Carol	Hard red winter wheat (resistant)
TA2923/ PI 564245	Karl92	Hard red winter wheat (susceptible)
	<i>H13</i> Association study	

Accession/TA	Cultivar	Description
TA10099		<i>Ae. tauschii</i> Minicore (MC), collected from Armenia
TA10106		<i>Ae. tauschii</i> MC, collected from Kyrgyzstan
TA10108		<i>Ae. tauschii</i> MC, collected from Tajikistan
TA10124		<i>Ae. tauschii</i> MC, collected from Uzbekistan
TA10141		<i>Ae. tauschii</i> MC, collected from PR China
TA10144		<i>Ae. tauschii</i> MC, collected from Syria
TA10162		<i>Ae. tauschii</i> MC, collected from Turkmenistan
TA10179		<i>Ae. tauschii</i> MC, collected from Turkmenistan
TA10210		<i>Ae. tauschii</i> MC, collected from Uzbekistan
TA10212		<i>Ae. tauschii</i> MC, collected from Uzbekistan
TA10330		<i>Ae. tauschii</i> MC, collected from Kyrgyzstan
TA1578		<i>Ae. tauschii</i> MC, H
TA1596		<i>Ae. tauschii</i> MC, H
TA1605		<i>Ae. tauschii</i> MC
TA1631		<i>Ae. tauschii</i> MC, collected from Afghanistan, S
TA1651		<i>Ae. tauschii</i> MC, collected from Iran, R
TA1665		<i>Ae. tauschii</i> MC, collected from Azerbaijan, R
TA1666		<i>Ae. tauschii</i> MC, collected from Azerbaijan, R
TA1667		<i>Ae. tauschii</i> MC, collected from Azerbaijan, R
TA1669		<i>Ae. tauschii</i> MC, collected from Azerbaijan, R
TA1694		<i>Ae. tauschii</i> MC, collected from Turkmenistan, S
TA1707		<i>Ae. tauschii</i> MC, collected from Sweden, R
TA2374		<i>Ae. tauschii</i> MC, collected from Pakistan, S
TA2376		<i>Ae. tauschii</i> MC, collected from Iran, S
TA2378		<i>Ae. tauschii</i> MC, collected from Iran, S
TA2395		<i>Ae. tauschii</i> MC, collected from Afghanistan, S
TA2413		<i>Ae. tauschii</i> MC, collected from Afghanistan, S
TA2431		<i>Ae. tauschii</i> MC, collected from Afghanistan, R
TA2435		<i>Ae. tauschii</i> MC, collected from Afghanistan, R

Accession/TA	Cultivar	Description
TA2448		<i>Ae. tauschii</i> MC, collected from Iran, R
TA2458		<i>Ae. tauschii</i> MC, collected from Iran, R
TA2468		<i>Ae. tauschii</i> MC, collected from Iran, R
TA2474		<i>Ae. tauschii</i> MC, collected from Iran, R
TA2485		<i>Ae. tauschii</i> MC, collected from Iran, S
TA2488		<i>Ae. tauschii</i> MC, collected from Iran, R
TA2508		<i>Ae. tauschii</i> MC, collected from Turkey, R
TA2514		<i>Ae. tauschii</i> MC, collected from Iran, R
TA2536		<i>Ae. tauschii</i> MC, collected from Afghanistan, H
TA2545		<i>Ae. tauschii</i> MC, collected from Afghanistan, R
TA2586		<i>Ae. tauschii</i> MC, collected from Georgia, H
TA10132		<i>Ae. tauschii</i> collected from Armenia
TA1594		<i>Ae. tauschii</i> collected from Turkey, R
TA1604		<i>Ae. tauschii</i> collected from Afghanistan
TA1622		<i>Ae. tauschii</i> collected from Azerbaijan, R
TA1642		<i>Ae. tauschii</i> collected from Iran, R
TA1644		<i>Ae. tauschii</i> collected from Iran, R
TA1645		<i>Ae. tauschii</i> collected from Iran, R
TA1660		<i>Ae. tauschii</i> collected from Azerbaijan, R
TA1664		<i>Ae. tauschii</i> collected from Azerbaijan, R
TA1668		<i>Ae. tauschii</i> collected from Azerbaijan, R
TA1670		<i>Ae. tauschii</i> collected from Azerbaijan, R
TA1671		<i>Ae. tauschii</i> collected from Azerbaijan, R
TA1672		<i>Ae. tauschii</i> collected from Azerbaijan, R
TA1674		<i>Ae. tauschii</i> collected from Russian Federation, R
TA1678		<i>Ae. tauschii</i> collected from Azerbaijan, R
TA1680		<i>Ae. tauschii</i> collected from Azerbaijan, R
TA1687		<i>Ae. tauschii</i> collected from Azerbaijan, R
TA1695		<i>Ae. tauschii</i> collected from Japan, R

Accession/TA	Cultivar	Description
TA1704		<i>Ae. tauschii</i> collected from Tajikistan, H
TA1706		<i>Ae. tauschii</i> collected from Iran, H
TA1715		<i>Ae. tauschii</i> collected from Iran, R
A2377		<i>Ae. tauschii</i> collected from Iran, R
TA2418		<i>Ae. tauschii</i> collected from Afghanistan, R
TA2419		<i>Ae. tauschii</i> collected from Afghanistan, R
TA2434		<i>Ae. tauschii</i> collected from Afghanistan, R
TA2438		<i>Ae. tauschii</i> collected from Afghanistan, R
TA2440		<i>Ae. tauschii</i> collected from Afghanistan, R
TA2442		<i>Ae. tauschii</i> collected from Afghanistan, R
TA2443		<i>Ae. tauschii</i> collected from Afghanistan, R
TA2449		<i>Ae. tauschii</i> collected from Iran, R
TA2450		<i>Ae. tauschii</i> collected from Iran
TA2452		<i>Ae. tauschii</i> collected from Iran
TA2453		<i>Ae. tauschii</i> collected from Iran, R
TA2454		<i>Ae. tauschii</i> collected from Iran, R
TA2455		<i>Ae. tauschii</i> collected from Iran, R
TA2463		<i>Ae. tauschii</i> collected from Iran, R
TA2464		<i>Ae. tauschii</i> collected from Iran, R
TA2465		<i>Ae. tauschii</i> collected from Iran, R
TA2466		<i>Ae. tauschii</i> collected from Iran, R
TA2467		<i>Ae. tauschii</i> collected from Iran, R
TA2469		<i>Ae. tauschii</i> collected from Iran, R
TA2470		<i>Ae. tauschii</i> collected from Iran, R
TA2472		<i>Ae. tauschii</i> collected from Iran, R
TA2473		<i>Ae. tauschii</i> collected from Iran, R
TA2475		<i>Ae. tauschii</i> collected from Iran, R
TA2476		<i>Ae. tauschii</i> collected from Iran, R
TA2477		<i>Ae. tauschii</i> collected from Iran, R

Accession/TA	Cultivar	Description
TA2481		<i>Ae. tauschii</i> collected from Iran, R
TA2492		<i>Ae. tauschii</i> collected from Iran, R
TA2507		<i>Ae. tauschii</i> collected from Turkey, R
TA2512		<i>Ae. tauschii</i> collected from Iran, H
TA2561		<i>Ae. tauschii</i> collected from Azerbaijan, R
TA2562		<i>Ae. tauschii</i> collected from Azerbaijan, R
TA2563		<i>Ae. tauschii</i> collected from Azerbaijan, R
TA2569		<i>Ae. tauschii</i> collected from Armenia, R
	9346A1-2-5-5-2-1	<i>H13H13</i>
	INW9811	<i>H13H13</i>
	Dawson	<i>H1-2H1-2</i>
	Carol	<i>H3H3</i>
	Monon	<i>H3H3</i>
	Java	<i>H4H4</i>
	Erin	<i>H5H5</i>
	Magnum	<i>H5H5</i>
	Flynn	<i>H6H6</i>
	Caldwell	<i>H6H6</i>
	Seneca	<i>H7H7</i>
	Iris	<i>H9H9</i>
	Joy	<i>H10H10</i>
	Karen	<i>H11H11</i>
	Lola	<i>H12H12</i>
	92167A3-5	<i>H14H14</i>
	81602	<i>H15H15</i>
	921682A4-6	<i>H16H16</i>
	921680D1-7	<i>H17H17</i>
	Redland	<i>H18H18</i>
	84702B14	<i>H19H19</i>

Accession/TA	Cultivar	Description
	Jori	<i>H20H20</i>
	Hamlet	<i>H21H21</i>
TA5005/ PI 499691	KS85WGRC01	<i>H22H22</i>
TA5013/ PI 535766	KS89WGRC03	<i>H23H23</i>
TA5016/ PI 535769	KS89WGRC06	<i>H24H24</i>
TA5033/ PI 592732	KS92WGRC20	<i>H25H25</i>
TA5039/ PI 572542	KS93WGRC26	<i>H26H26</i>
	Sincape90	<i>H29H29</i>
	P921696A1-15	<i>H31H31</i>
TA5062	KS99WGRC42	<i>HdicHdic</i>
TA5014/ PI 535767	KS89WGRC04	
TA5018/ PI 549276	KS89WGRC08	<i>H21H21</i>
TA5030/ PI 592729	KS92WGRC17	<i>H25H25</i>
TA5031/ PI 592730	KS92WGRC18	<i>H25H25</i>
TA5044/ PI 586956	KS94WGRC31	
TA5062	KS99WGRC42	
	PIONEER-2555	
	Danby	
TA9111	Tam111	U.S. Hard Red Winter Wheat

Accession/TA	Cultivar	Description
TA9187	TAM112	U.S. Hard Red Winter Wheat
TA2909	Jagger	U.S. Hard Red Winter Wheat
TA2959	Chris	U.S. Hard Red Spring Wheat
TA9121	Everest	U.S. Hard Red Winter Wheat
	Overland	
	Lyman	
TA2908/PI 574488	Ike	U.S. Hard Red Winter Wheat
TA2919/Citr 6887	Marquillo	U.S. Hard Red Spring Wheat

Table 2.3. Haplotype of *H13* region in high resolution mapping population

POPULATION NUMBER	2011-214-79	2011-214-116	2011-214-127	2011-214-287	2011-214-324	2011-214-377	2011-214-392	2011-214-556	2011-214-588	2011-214-831	2011-214-969	2011-214-975	2011-214-978	2011-214-1072	2011-214-1166	2011-214-1192	2011-214-1208	2011-214-1223	2011-214-1291	2011-214-1353
MARKER																				
18-84	B	H	H	A	A	H	H	H	B	B	A	H	H	H	A	H	H	B	H	B
18-13	B	H	H	A	A	H	H	H	B	B	A	H	H	H	A	H	H	B	H	B
GDM 36	B	H	H	A	A	H	H	H	B	B	A	H	H	H	A	H	H	B	H	B
19-23	B	H	H	A	A	H	H	H	B	B	A	H	H	H	A	H	H	B	H	B
19-26	B	B	H	A	A	H	H	H	B	B	A	H	A	H	A	H	B	B	H	B
19-30	B	B	H	A	A	H	H	H	B	B	A	H	A	H	A	H	B	B	B	B
19-34	B	B	B	H	A	H	B	H	B	B	H	H	A	H	A	H	B	B	B	H
PPR	B	B	B	H	A	H	B	A	H	B	H	H	A	H	H	H	B	H	B	H
STK	B	B	B	H	A	H	B	A	H	B	H	H	A	H	H	H	B	H	B	H
CN-7	B	B	B	H	A	H	B	A	H	B	H	H	A	H	H	H	B	H	B	H
<i>H13</i>	B	B	B	H	A	H	B	A	H	B	H	H	A	NA	H	H	B	H	B	H
CNL-1	B	B	B	H	A	H	B	A	H	B	H	H	A	H	H	H	B	H	B	H
19-124	H	B	B	H	A	A	B	A	H	H	H	B	A	A	H	B	B	H	B	H
CFD132	H	B	B	H	H	A	B	A	H	H	H	B	A	A	H	B	B	H	B	H
19-150	H	B	B	H	H	A	B	A	H	H	H	B	A	A	H	B	B	H	B	H

Table 2.4 Comparative analysis of candidate genes

Please go to www.krex.k-state.edu and search for **AnupamaJoshi2018-Table2.4.Comparative analysis**

Table 2.5 Phenotypic data of HS infestation on EMS M₃ lines

BATCH 1 - TRAY 1 TO TRAY 28			
Sr.no	Tray	Susceptible EMS Molly M ₃ Line	Phenotype segregation
1	Tray2	2014-635-37	10 R, 10 S
2	Tray3	2014-635-49	20 S
3	Tray3	2014-635-65	20 S
4	Tray5	2014-635-93	20 S
5	Tray5	2014-635-96	7 S, 13 R
6	Tray8	2014-635-191	7 S, 13 R
7	Tray9	2014-635-215	18 S
8	Tray10	2014-635-229	20 S
9	Tray10	2014-635-231	20 S
10	Tray10	2014-635-249	1 S, 29 R
11	Tray12	2014-635-288	20 S
12	Tray13	2014-635-316	20 S
13	Tray14	2014-635-336	20 S
14	Tray14	2014-635-352	20 S
15	Tray16	2014-635-401	18 S, 2 R
16	Tray17	2014-635-431	20 S
17	Tray19	2014-635-480	20 S
18	Tray21	2014-635-521	19 S
19	Tray21	2014-635-525	20 S
20	Tray21	2014-635-537	20 S
21	Tray22	2014-635-542	18 S
22	Tray22	2014-635-565	7 S, 13 R
23	Tray24	2014-635-596	20 S
24	Tray24	2014-635-599	20 S
25	Tray25	2014-635-634	6S, 14 R
26	Tray26	2014-635-644	4S, 16 R
27	Tray27	2014-635-666	20 S
28	Tray27	2014-635-675	7 S, 13 R

BATCH 2 - TRAY 29 TO TRAY 56			
Sr.no.	Tray	Susceptible EMS Molly M ₃ line	Phenotype segregation
1	Tray31	2014-635-772	20 S
2	Tray39	2014-635-1015	17 S
3	Tray40	2014-635-1040	20 S
4	Tray40	2014-635-1045	10 S, 10 R
5	Tray41	2014-635-1058	4 S, 16 R
6	Tray42	2014-635-1095	8 S, 12 R
7	Tray43	2014-635-1113	20 S
8	Tray44	2014-635-1131	4 S, 16 R
9	Tray44	2014-635-1138	20 S
10	Tray45	2014-635-1145	20 S
11	Tray46	2014-635-1163	10 S, 10 R
12	Tray46	2014-635-1175	20 S
13	Tray47	2014-635-1196	20 S
14	Tray49	2014-635-1240	20 S
15	Tray50	2014-635-1258	20 S
16	Tray50	2014-635-1262	20 S
17	Tray50	2014-635-1266	19 S, 1 R
18	Tray51	2014-635-1287	20 S
19	Tray53	2014-635-1360	20 S

BATCH 3 - TRAY 57 TO TRAY 84			
Sr.no.	Tray	Susceptible EMS Molly M ₃ line	Phenotype segregation
1	Tray 57	2014-635-1464	20 S
2	Tray 61	2014-635-1558	20 S
3	Tray 65	2014-635-1640	20 S
4	Tray 65	2014-635-1649	6 S + 14 R
5	Tray 65	2014-635-1653	19 S
6	Tray 66	2014-635-1676	20 S
7	Tray 67	2014-635-1710	4 S + 16 R
8	Tray 68	2014-635-1728	20 S
9	Tray 68	2014-635-1736	18 S
10	Tray 70	2014-635-1777	20 S
11	Tray 73	2014-635-1848	18 S
12	Tray 75	2014-635-1903	20 S
13	Tray 75	2014-635-1912	18 S + 1R
14	Tray 75	2014-635-1915	20 S
15	Tray 77	2014-635-1963	10 S +10 R
16	Tray 77	2014-635-1965	5 S + 15 R
17	Tray 78	2014-635-1975	20 S
18	Tray 79	2014-635-2025	9 S + 11 R
19	Tray 80	2014-635-2033	1 S + 0 R
20	Tray 80	2014-635-2039	20 S
21	Tray 82	2014-635-2096	20 S
22	Tray 84	2014-635-2148	8 S + 12 R
23	Tray 84	2014-635-2150	18 S + 2 R

BATCH 4 - TRAY 85 TO TRAY 112			
Sr.no.	Tray	Susceptible EMS Molly M ₃ line	Phenotype segregation
1	Tray 88	2014-635-471	only 4 plants S
2	Tray 91	2014-635-2306	13 S + 1R
3	Tray 91	2014-635-2327	4 S + 12 R
4	Tray 92	2014-635-2333	7 S + 12 R
5	Tray 92	2014-635-2355	20 S
6	Tray 93	2014-635-2378	5 S + 14 R
7	Tray 94	2014-635-2413	20 S
8	Tray 97	2014-635-201	18 S
9	Tray 98	2014-635-253	19 S
10	Tray 98	2014-635-286	20 S
11	Tray 98	2014-635-309	20 S
12	Tray 99	2014-635-2060	14 S
13	Tray 99	2014-635-375	12 S +2 R
14	Tray 100	2014-635-510	20 S
15	Tray 101	2014-635-594	17 S
16	Tray 101	2014-635-717	20 S
17	Tray 102	2014-635-736	20 S
18	Tray 104	2014-635-937	20 S
19	Tray 104	2014-635-1421	15 S
20	Tray 105	2014-635-1216	20 S
21	Tray 106	2014-635-1270	18 S
22	Tray 106	2014-635-1305	5 S + 12 R
23	Tray 107	2014-635-1537	20 S
24	Tray 107	2014-635-1603	20 S
25	Tray 108	2014-635-155	only 1 S + 15 R
26	Tray 109	2014-635-1883	20 S
27	Tray 110	2014-635-2065	18 S
28	Tray 110	2014-635-2067	20 S
29	Tray 111	2014-635-2365	20 S
30	Tray 111	2014-635-2408	20 S
31	Tray 112	2014-635-1214	20 S

Table 2.6 Details of mutations found in candidate genes

Gene	Primer Set	Plant ID	Type of mutation	DNA change	Amino Acid change =silent, xstop codon and > substitution	PROVEAN score	SIFT SCORE	SIFT PREDICTION	HF infestation(R/S)
<i>CNL-1</i>	305F1-R1	352	Substitution	C 494 T	G 165 E	-1.372	0.38	Tolerated	S
		521	Substitution	C 850 T	D 284 N	-3.755	0.00	Damaging	S
		634	Truncation	C 698 T	W 233 X	-	-		S
		675	Substitution	G 605 A	A 202 V	-2.615	0.24	Tolerated	S
		1058	Substitution	G 605 A	A 202 V	-2.615	0.24	Tolerated	S
		1095	Substitution	C 620 T	G 207 D	-4.609	0.00	Damaging	S
		1963	Substitution	G 605 A	A 202 V	-2.615	0.24	Tolerated	S
		1965	Substitution	C 850 T	D 284 N	-3.755	0.00	Damaging	S
		2025	Substitution	C 238 T	D 80 N	-3.669	0.00	Damaging	S
	348F-305R1	2306	Substitution	C 638 T	T 213 I	-4.1	0.00	Damaging	S
		1045	Substitution	G 944 A	R 315 K	-2.203	0.00	Damaging	S
		1262	Substitution	G 923 A	S 308 N	-2.086	0.07	Tolerated	S
		1777	Substitution	G 923 A	S 308 N	-2.086	0.07	Tolerated	S
		286	Substitution	G 910 A	V 304 M	-0.303	0.07	Tolerated	S
		37	Substitution	G 2348 A	G 783 D	-5.819	0.00	Damaging	S
	275F1-R2	229	Substitution	C 2731 T	A 911 T	-2.211	0.59	Tolerated	S
		644	Substitution	C 2885 T	C 962 I	-7.812	0.15	Tolerated	S
		1163	Substitution	G 2602 A	E 868 K	-0.265	0.51	Tolerated	S
		1287	Substitution	C 2355 T	P 787 S	-7.231	0.00	Damaging	S

Gene	Primer Set	Plant ID	Type of mutation	DNA change	Amino Acid change =silent, xstop codon and > substitution	PROVEAN score	SIFT SCORE	SIFT PREDICTION	HF infestation(R/S)
<i>CNL-1</i>		2413	Truncation	G 2625 A	W 875 X	-	-		S
		1305	Substitution	G 2711 A	P 904 L	-8.467	0.00	Damaging	S
		2378	Truncation	C 26334 T	W 878 X	-	-		S
	352F-R	286	Substitution	C 2731 T	A 911 T	-2.211	0.59	Tolerated	S
	481F-482R	191	Substitution	G 1121 A	G 374 D	-5.537	0.00	Damaging	S
		249	Substitution	C 1277 T	P 426 L	-6.544	0.00	Damaging	S
		599	Substitution	C 2963 T	P 988 L	-8.252	0.00	Damaging	S
		2096	Substitution	G 1412 A	G 471 E	-6.096	0.00	Damaging	S
		2148	Substitution	C 1310 T	S 437 F	-4.732	0.02	Damaging	S
	274F2-R2	482	Silent	C 1308 T	F 436 F	-	1.0	Tolerated	R
<i>CNL-2</i>	285F-R1	2025	Intronic	G 1433 A	-				S
<i>CNL-4</i>	341F-R1	644	Silent	C 249 T	K 83 K				S
<i>CNL-6</i>	337F1-R2	567	Substitution	G 2536 A	A 846 T	-0.239	0.23	Tolerated	R
<i>CNL-7</i>	297F2-R	1163	Silent	C 786 T	A 262 A				S
	334F1-R1	2413	Silent	C 2463 T	R 821 R				S

Table 2.7 Full-length haplotypes of *H13* gene

Please go to www.krex.k-state.edu and search for **AnupamaJoshi2018-Table2.7.Haplotype analysis**

Table 2.8 Sites of variation in the haplotypes of *H13* gene

Haplotype	32	33	51	90	117	147	161	182	217	229	247	252	258	292	317	394	444	472	487	494	495	503
Hap-16	N	E	A	G	T	R	R	R	K	Q	R	Q	G	C	E							
Hap-8	N	E	A	G	T	R	R	R	K	Q	R	Q	G	W	E	R	K	S	D	C	V	R
Hap-9	N	Q	T	S	T	P	R	R	K	Q	K	Q	G	W	E	H	K	R	N	N	L	R
Hap-2	N	E	A	G	T	P	R	G	K	Q	R	Q	C	W	E	R	M	S	D	C	V	R
Hap-3	N	E	A	G	T	P	R	R	K	Q	R	Q	C	W	E	R	K	R	N	N	L	R
Hap-6	N	E	A	G	T	P	R	R	R	Q	R	Q	C	W	E	R	M	S	D	C	V	R
Hap-5	N	E	A	G	T	P	R	R	R	Q	R	Q	C	W	E	R	K	S	D	C	V	R
Hap-4	N	E	A	G	T	P	R	R	R	Q	R	Q	C	W	E	R	K	S	D	C	V	R
Hap-1	N	E	A	G	I	P	R	R	R	Q	R	Q	C	W	E	R	K	S	D	C	V	R
Hap-7	N	E	A	G	T	P	S	R	R	Q	R	Q	C	W	E	R	M	S	D	C	V	R
Hap-10	T	E	A	G	T	P	R	R	R	-	R	Q	C	W	E	R	M	S	D	C	V	R
Hap-15	T	E	A	G	T	P	R	R	R	-	R	R	C	W	K							
Hap-11	T	E	A	G	T	P	R	R	R	-	R	Q	C	W	K	R	M	S	D	C	V	C
Hap-12	T	E	A	G	T	P	R	R	R	-	R	Q	C	W	K	R	M	S	D	C	V	R
Hap-14	T	E	A	G	T	P	R	R	R	-	R	Q	C	W	K	R	M	S	D	C	V	C
Hap-13	T	E	A	G	T	P	R	R	R	-	R	Q	C	W	K	R	M	S	D	C	V	C

Haplotype	586	588	611	617	623	634	637	641	642	646	647	651	654	656	657	658	663	667	670	674	675	679	683
Hap-16																							
Hap-8	C	Y	I	S	G	D	E	K	V	F	V	Q	Y	S	F	S	E	K	H	N	L	T	A
Hap-9	C	Y	I	S	G	H	E	K	V	F	V	L	Y	S	F	S	E	K	H	N	L	T	A
Hap-2	Y	Y	E	R	A	Q	Q	E	L	I	I	Q	F	N	L	T	K	R	C	N	L	I	T
Hap-3	C	Y	I	S	G	H	E	K	V	F	V	L	Y	S	F	S	E	K	H	N	L	T	A
Hap-6	Y	Y	E	R	A	Q	Q	E	L	I	I	Q	F	N	L	T	K	R	C	N	L	I	T
Hap-5	Y	C	E	S	G	H	K	K	V	F	V	Q	Y	S	F	S	E	R	H	S	M	H	A
Hap-4	Y	C	E	S	G	H	K	K	V	F	V	Q	Y	S	F	S	E	R	H	S	M	H	A
Hap-1	Y	C	E	S	G	H	K	K	V	F	V	Q	Y	S	F	S	E	R	H	S	M	H	A
Hap-7	Y	Y	E	R	A	Q	Q	E	L	I	I	Q	F	N	L	T	K	R	C	N	L	I	T
Hap-10	Y	Y	E	R	A	Q	Q	E	L	I	I	Q	F	N	L	T	K	R	C	N	L	I	T
Hap-15																							
Hap-11	Y	Y	E	R	A	Q	Q	E	L	I	I	Q	F	N	L	T	K	R	C	N	L	I	T
Hap-12	Y	Y	E	R	A	Q	Q	E	L	I	I	Q	F	N	L	T	K	R	C	N	L	I	T
Hap-14	Y	Y	E	R	A	Q	Q	E	L	I	I	Q	F	N	L	T	K	R	C	N	L	I	T
Hap-13	Y	Y	E	R	A	Q	Q	E	L	I	I	Q	F	N	L	T	K	R	C	N	L	I	T

Haplotype	685	700	705	706	707	708	712	714	719	725	728	731	733	744	746	751	753	755	756	758	761	765	771
Hap-16																							
Hap-8	V	Q	E	L	D	G	I	A	T	D	H	I	S	V	V	G	P	E	E	K	E	D	P
Hap-9	V	Q	E	L	D	G	I	A	T	N	Q	I	F	G	I	G	S	E	K	Q	E	E	H
Hap-2	E	L	G	M	Y	E	V	T	A	E	H	D	S	G	I	S	S	Q	E	Q	K	E	P
Hap-3	V	Q	E	L	D	G	I	A	T	D	Q	I	F	G	I	G	S	E	K	Q	E	E	H
Hap-6	E	L	R	M	Y	E	V	T	A	E	H	D	S	G	I	S	S	Q	E	Q	K	E	P
Hap-5	V	Q	G	L	D	G	I	A	A	E	H	D	S	G	I	S	S	Q	E	Q	K	D	S
Hap-4	V	Q	G	L	D	G	I	A	A	E	H	D	S	G	I	S	S	Q	E	Q	K	D	S
Hap-1	V	Q	G	L	D	G	I	A	A	E	H	D	S	G	I	S	S	Q	E	Q	K	D	S
Hap-7	E	L	R	M	Y	E	V	T	A	E	H	D	S	G	I	S	S	Q	E	Q	K	E	P
Hap-10	E	L	R	M	Y	E	V	T	A	E	H	D	S	G	I	S	S	Q	E	Q	K	E	P
Hap-15																							
Hap-11	E	L	R	M	Y	E	V	T	A	E	H	D	S	G	I	S	S	Q	E	Q	K	E	P
Hap-12	E	L	R	M	Y	E	V	T	A	E	H	D	S	G	I	S	S	Q	E	Q	K	E	P
Hap-14	E	L	R	M	Y	E	V	T	A	E	H	D	S	G	I	S	S	Q	E	Q	K	E	P
Hap-13	E	L	R	M	Y	E	V	T	A	E	H	D	S	G	I	S	S	Q	E	Q	K	E	P

Haplotype	775	777	779	781	783	784	785	788	797	799	803	806	807	809	813	820	824	831	832	839	840	844	848
Hap-16																							
Hap-8	I	S	N	Y	S	R	Q	R	P	K	N	I	T	L	I	C	Y	D	R	V	G	M	H
Hap-9	T	V	K	Y	G	Q	R	L	P	K	N	I	T	L	T	S	F	N	H	I	R	I	H
Hap-2	N	D	K	Y	G	R	W	R	A	K	I	I	T	L	T	C	F	N	R	I	G	I	H
Hap-3	T	V	K	Y	G	Q	R	L	P	K	N	I	T	L	T	S	F	N	H	I	G	I	H
Hap-6	N	D	K	Y	G	R	W	R	A	K	I	I	T	L	T	C	F	N	R	I	G	I	H
Hap-5	Y	C	N	H	G	Q	R	R	Y	N	I	L	K	M	I	C	Y	K	R	V	G	M	R
Hap-4	Y	C	N	H	G	Q	R	R	Y	N	I	L	E	M	I	C	Y	K	R	V	G	M	R
Hap-1	Y	C	N	H	G	Q	R	R	Y	N	I	L	E	M	I	C	Y	K	R	V	G	M	R
Hap-7	N	D	K	Y	G	R	W	R	A	K	I	I	T	L	T	C	F	N	R	I	G	I	H
Hap-10	N	D	K	Y	G	R	W	R	A	K	I	I	T	L	T	C	F	N	R	I	G	I	H
Hap-15																							
Hap-11	N	D	K	Y	G	R	W	R	A	K	I	I	T	L	T	C	F	N	R	I	G	I	H
Hap-12	N	D	K	Y	G	R	W	R	A	K	I	I	T	L	T	C	F	N	R	I	G	I	H
Hap-14	N	D	K	Y	G	R	W	R	A	K	I	I	T	L	T	C	F	N	R	I	G	I	H
Hap-13	N	D	E	Y	G	R	W	R	A	K	I	I	T	L	T	C	F	N	R	I	G	I	H

Haplotype	849	850	853	857	859	866	867	868	869	870	883	892	894	895	896	897	905	909	913	916	918	920	921
Hap-16																							
Hap-8	H	Q	H	V	A	K	I	E	L	I	Q	E	R	I	K	Q	S	F	A	S	S	T	S
Hap-9	Q	H	C	M	A	V	M	N	L	I	Q	Q	M	L	E	Q	P	S	A	S	S	T	S
Hap-2	H	H	C	M	A	E	M	N	L	I	Q	Q	M	L	E	H	P	S	T	S	S	T	S
Hap-3	H	H	C	M	A	E	M	N	L	I	Q	Q	M	L	E	Q	P	S	A	S	S	T	S
Hap-6	H	H	C	M	A	E	M	N	L	I	Q	Q	M	L	E	H	P	S	T	S	S	T	S
Hap-5	Q	H	R	M	P	E	L	E	F	T	E	E	M	L	K	Q	P	S	A	K	F	Y	Y
Hap-4	Q	H	R	M	P	E	L	E	F	T	E	E	M	L	K	Q	P	S	A	K	F	Y	Y
Hap-1	Q	H	R	M	P	E	L	E	F	T	E	E	M	L	K	Q	P	S	A	K	F	Y	Y
Hap-7	H	H	C	M	A	E	M	N	L	I	Q	Q	M	L	E	H	P	S	T	S	S	T	S
Hap-10	H	H	C	M	A	E	M	N	L	I	Q	Q	M	L	E	H	P	S	T	S	S	T	S
Hap-15																							
Hap-11	H	H	C	M	A	E	M	N	L	I	Q	Q	M	L	E	H	P	S	T	S	S	T	S
Hap-12	H	H	C	M	A	E	M	N	L	I	Q	Q	M	L	E	H	P	S	T	S	S	T	S
Hap-14	H	H	C	M	A	E	M	N	L	I	Q	Q	M	L	E	Q	P	S	A	S	S	T	S
Hap-13	H	H	C	M	A	E	M	N	L	I	Q	Q	M	L	E	H	P	S	T	S	S	T	S

Haplotype	926	927	942	956	958	967	969	975	1064
Hap-16									
Hap-8	V	S	I	K	T	T	E	Q	R
Hap-9	V	S	I	K	T	T	E	Q	R
Hap-2	V	S	I	K	T	T	E	Q	R
Hap-3	V	S	I	K	T	T	E	Q	R
Hap-6	V	S	I	K	T	T	E	Q	-
Hap-5	I	S	N	M	S	V	D	E	-
Hap-4	I	S	N	M	S	V	D	E	-
Hap-1	I	S	N	M	S	V	D	E	-
Hap-7	V	S	I	K	T	T	E	Q	-
Hap-10	V	S	I	K	T	T	E	Q	R
Hap-15									
Hap-11	V	S	I	K	T	T	E	Q	R
Hap-12	V	S	I	K	T	T	E	Q	R
Hap-14	V	F	I	K	T	T	E	Q	R
Hap-13	V	S	I	K	T	T	E	Q	R

Table 2.9 List of accessions belonging to each haplotype

HAPLOTYPE	Hap1	Hap2	Hap4	Hap5	Hap6
LINES	<i>H13-9346A1</i>	TA2431	TA1622	1605	TA1706
	<i>H13-INW9811</i>	TA1631	TA1642		TA1596
	TA3386	TA2374	TA1660		TA2435
	TA3387	TA10141	TA1665		TA10106
	TA5011	TA10144	TA1666		TA10330
	TA5012	TA10210	TA1667		
	<i>H13-Molly</i>	TA2418	TA1669		
	TA2377	TA2419	TA1680		
	TA1594	TA2438	TA1687		
	TA2473	TA1664	TA2561		
	TA2508		TA2562		
	TA2445				
	TA2448				
	TA2488				
HAPLOTYPE	Hap7	Hap8	Hap9	Hap10	Hap11
LINES	TA1694	TA2453	TA2450	TA2492	TA2413
		TA2468	TA2512		TA2536
		TA2474			TA10162
		TA2485			TA10212
		TA2466			
		TA2470			
		TA2472			
		TA2476			
		TA2481			
HAPLOTYPE	Hap12	Hap13	Hap14	Hap15	Hap16
LINES	TA2395	TA10108	TA1578	TA10179	TA1715
	TA2454				
	TA2514				
	TA2545				
	TA2586				
	TA10099				
EXON2 DATA IS NOT COMPLETE FOR THESE LINES					

HAPLOTYPE	Hap3	
LINES	Newton	TAM112
	TA1695	PIONEER-2555
	TA1651	Danby
	TA10124	TA1644
	BOBWHITE	TA1645
	H3-Carol	TA2443
	H3-Monon	TA2449
	H4-Java	TA2463
	H5-Erin	TA2464
	H5-Magnum	TA2467
	H6-Flynn	TA2469
	H6-Caldwell	TA2477
	H7-8-Seneca	TA2507
	H9-Iris	TA2908
	H10-Joy	TA2919
	H11-Karen	TA5005
	H12-Lola	TA5016
	H14-92167A3-5	TA5018
	H15-81602	TA5030
	H16-921682A4-6	TA5031
	H17-921680D1-7	TA5033
	H18-Redland	TA5039
	H19-84702B14	TA5044
	H21-Hamlet	TA5062
	H22-KS85WGRC01	
	H23-KS89WGRC03	
	H24-KS89WGRC06	
	H25-KS92WGRC20	
	H26-KS93WGRC26	
	H31-P921696A1-15	
	Hdic-KS99WGRC42	
	TA1707	
	Tam111	

Table 2.10 List of primers used in this study

Order name	Details	Sequence 5' to 3'
H13_302_F	Gene42_masked scffold5319_7TH EXON	ACTGCTCATCCAGCCCTCCAC
H13_302_R1	Gene42_masked scffold5319_B4 8TH EXON	CGTGACATTTTCGGCCGAACG
H13_302_R2	Gene42_masked scffold5319_8TH EXON	TCAAGAGGTAGCAGCGCATA
H13_303_F1	Gene42_masked scffold5319_5TH EXON	TGACGGCTGGGTCCACGG
H13_303_F2	Gene42_masked scffold5319_5TH EXON	CTGACAGGTGGGCCACGA
H13_303_R1	Gene42_masked scffold5319_6TH EXON	GGCACTTCCTTGCTGCGTCT
H13_303_R2	Gene42_masked scffold5319_6TH EXON	CAGCTATCAGCCTCTCCACG
H13_304_F	Gene42_masked scffold5319_6TH EXON	GGAAGTGCCTCCTTAATACGC
H13_304_R1	Gene42_masked scffold5319_b4 7TH EXON	ATCGGCTGTCCATCCAAGC
H13_304_R2	Gene42_masked scffold5319_7TH EXON	ATCCCGTTTCGAACGTGACC
H13_305_F1	Gene73_masked scaffold5319_1st exon	ATGACGATGGTCCTGGATGCC
H13_305_F2	Gene73_masked scaffold5319_1st exon	TTCGCGTCCTATGTCGCCG
H13_305_F3	Gene73_masked scaffold5319_1st exon	GCTGCTTGGAGTCACAAACC
H13_305_F4	Gene73_masked scaffold5319_1st exon	GCTTCTTGGAGTCGGAGACG
H13_305_R1	Gene73_masked scaffold5319_1st exon	TAGGGAAGCACCACCTTTCATCA
H13_305_R2	Gene73_masked scaffold5319_1st exon	ACGGGCAACTCTTGTATCTCTC
H13_306_F1	Gene73_masked scaffold5319_2ndexon end	GCACAGGTTAGAACTCTTGTC
H13_306_F2	Gene73_masked scaffold5319_2ndexon end	TGATGTCAAGATCAGCACCGG
H13_306_R1	Gene73_masked scaffold5319_2ndexon end	GGATCAGCTGTGTACAAGACGA
H13_306_R2	Gene73_masked scaffold5319_2ndexon end	CATGCTGCAGGGCTCCCTC
H13-137-F1	Pool5751.3_AT6D_5319_gene1	CAGCAGAAAACAAAAAATTTCCA
H13-137-F2	Pool5751.3_AT6D_5319_gene1	AACAGATTTTCATCTGTCACCC
H13-137-R	Pool5751.3_AT6D_5319_gene1	GGTCCTTGTGGAAGGACTTCA
H13-138-F	Pool5751.3_AT6D_5319_gene1	GGTATCATCGAACTTATCCG
H13-138-R	Pool5751.3_AT6D_5319_gene1	ATCCATAAAGGCTCTGGGGAG
H13-139-F1	Pool5751.3_AT6D_5319_gene2	CCTTGGCAATAAGAGCAGA
H13-139-F2	Pool5751.3_AT6D_5319_gene2	TTGGTACCATGCTTGGCATGG

Order name	Details	Sequence 5' to 3'
H13-139-R	Pool5751.3_AT6D_5319_gene2	ATCACCTGCAAGGTTGTTCAAG
H13-140-F	Pool5751.3_AT6D_5319_gene2	CAGTTTCTGAGCTGTCTGTGG
H13-140-R	Pool5751.3_AT6D_5319_gene2	AAGGGTCCCCAACTCATACGACT
H13-141-F	Pool5751.3_AT6D_5319_gene2	AGTCGTATGAGTTGGGGACCCTT
H13-141-R	Pool5751.3_AT6D_5319_gene2	TAAAGTTGGCATTTCGCGGGGG
H13-142-F	Pool5751.3_AT6D_5319_gene2	GCAACTGCCAACCAAGCACG
H13-142-R	Pool5751.3_AT6D_5319_gene2	ACCACTGCAATAAGGCAAGT
H13-143-F	HfrDrd_F1a	CTCTCTTTTCCCCTTCTTGCAAAC
H13-143-R	HfrDrd_R1	CCTTGCTCAAACCTCTATGTTATGAGTC
H13-144-F	HfrDrd_F2	GATGGTACATGACTCATAACATAG
H13-144-R	HfrDrd_R2	CCCTGTGCCATAATGAATTGACC
H13-145-F	HfrDrd_F3	GGGTATAATAAGTTTCACAAATC
H13-145-R	HfrDrd_R3a	AGACACCAACAAACAGATGGG
H13-146-F	HfrDrd	ATAGATAGGTTTGCAGGGTCC
H13-146-R	HfrDrd	CTGGATATCCATTATGCTAC
H13-147-F	Pool5751.3_AT6D_5319_gene3	ACAAAGCCGATCCAAATCAG
H13-147-R1	Pool5751.3_AT6D_5319_gene3	CTCTTGACGCTCTTGTCTG
H13-147-R2	Pool5751.3_AT6D_5319_gene3	CCACCCACTCTTGACACGC
H13-148-F	Pool5751.3_AT6D_5319_gene3	CATGCTCGAGGACGACTACA
H13-148-R1	Pool5751.3_AT6D_5319_gene3	CAAGACGGTTGACTCGTTC
H13-148-R2	Pool5751.3_AT6D_5319_gene3	CTCAGTTACCAACAAGACGG
H13-149-F1	Pool5751.3_AT6D_5319_gene5	CCCTAGCCCTGATGATTCAA
H13-149-F2	Pool5751.3_AT6D_5319_gene5	GTTTGACACCCTAGCCCTG
H13-149-R1	Pool5751.3_AT6D_5319_gene5	TTTCCACATCCAACAGGTCA
H13-149-R2	Pool5751.3_AT6D_5319_gene5	GAGATCGGGTCATTTCCACA
H13-150-F1	Pool5751.3_AT6D_5319_gene5	CACGCGATGCTTACTTTCAA
H13-150-F2	Pool5751.3_AT6D_5319_gene5	GTAAACTCACGCGATGCTTAC

Order name	Details	Sequence 5' to 3'
H13-150-R1	Pool5751.3_AT6D_5319_gene5	AGGCACCCAGATCTTCTTCA
H13-150-R2	Pool5751.3_AT6D_5319_gene5	CACCTTTTAGGCACCCAGATC
H13-151-F	Pool5751.3_AT6D_5319_gene9	GCACTCGGAAAAGTGGAGAC
H13-151-R	Pool5751.3_AT6D_5319_gene9	GGATGGATATCACCGACACC
H13-152-F	Pool5751.3_AT6D_5319_gene9	GTGGTTGCGTTGTACGTGTC
H13-152-R	Pool5751.3_AT6D_5319_gene9	GGGTCATGCTATCTGCCAAT
H13-153-F	Pool5751.3_AT6D_5319_gene9	CTTGGTGTACGTGCCATCAC
H13-153-R	Pool5751.3_AT6D_5319_gene9	TGCTGGTCAAGAAGTGTTC
H13-154-F	Pool5751.3_AT6D_5319_gene9	AGACAGATGAACGTGAAGTC
H13-154-R	Pool5751.3_AT6D_5319_gene9	CTGGCAGACTAGATATATCTGT
H13-155-F	Pool5751.3_AT6D_5319_gene12	AGGTTTGTGCTGCTGTGTG
H13-155-R	Pool5751.3_AT6D_5319_gene12	GCTGTTCTTGTCTGCCCTTC
H13-156-F	Pool5751.3_AT6D_5319_gene12	TGGACTCCAGAGGTTCTGTTT
H13-156-R	Pool5751.3_AT6D_5319_gene12	GTGATCTGCTTTCTGCCACA
H13-157-F	Pool5751.3_AT6D_5319_gene12	AATCAGATGGGTTGGACTGG
H13-157-R	Pool5751.3_AT6D_5319_gene12	TGTGCTGATGGTAGGACTGC
H13-158-F	Pool5751.3_AT6D_5319_gene12	AAGTGCGCCAGTTAGCAAAT
H13-158-R	Pool5751.3_AT6D_5319_gene12	TTGCACCAGCTTACTGATGG
H13-159-F	Pool5751.3_AT6D_5319_gene12	TTGAGTTTTCTGCATTACG
H13-159-R	Pool5751.3_AT6D_5319_gene12	AGGTCCCGCTTGATAGCTT
H13-160-F	Pool5751.3_AT6D_5319_gene14	ATGGTGGAAGCAGAGGAATG
H13-160-R	Pool5751.3_AT6D_5319_gene14	GTA CTGCTGGGAGGACAAGG
H13-161-F	Pool5751.3_AT6D_5319_gene14	CTCATACCCCTGCATGCTCT
H13-161-R	Pool5751.3_AT6D_5319_gene14	ACCTCCGTTTCTGTCGTCTG
H13-162-F	Pool5751.3_AT6D_5319_gene14	TGCCTGCCTAATAGTAGCATTC
H13-162-R	Pool5751.3_AT6D_5319_gene14	AGTGTGTTAACGTGTTGCTCA
H13-163-F	Pool5751.3_AT6D_5319_gene28	CACCCAGTTACGTTGTGACG
H13-163-R	Pool5751.3_AT6D_5319_gene28	CGTCAGAATCTGCAACCTCA

Order name	Details	Sequence 5' to 3'
H13-164-F	Pool5751.3_AT6D_5319_gene28	TTGATTGCCCAGTGGATGTA
H13-164-R	Pool5751.3_AT6D_5319_gene28	CTGTTAACGGCCCATTCACT
H13-165-F	Pool5751.3_AT6D_5319_gene28	GCCGAAATATCATCGGGCCG
H13-165-R	Pool5751.3_AT6D_5319_gene28	GTACACTTCCGTGATGATACG
H13-166-F	Scaffold5751_AT6D_5319_gene8	GGCTCTTATCCCAGTGGTCA
H13-166-R	Scaffold5751_AT6D_5319_gene8	GAAGGCATGCGTGAAATCTT
H13-167-F	Scaffold5751_AT6D_5319_gene8	AGCAGCGCAAAATCTCTTGT
H13-167-R	Scaffold5751_AT6D_5319_gene8	GCGAGCATCAAAGCAGTACA
H13-168-F	Scaffold5751_AT6D_5319_gene8	TTGGCAGGGTAAGTGTGTG
H13-168-R	Scaffold5751_AT6D_5319_gene8	CGGCAATCATGTGTGCTTAC
H13-169-F	Scaffold5751_AT6D_5319_gene8	GCATCTGGACGATGCATGCA
H13-169-R	Scaffold5751_AT6D_5319_gene8	GGGCTCCTACGATGACACTGC
H13-170-F1	Scaffold5751_AT6D_5319_gene20	CAAGATCCCCTCCGTGTAGA
H13-170-F2	Scaffold5751_AT6D_5319_gene20	CCATAGGAACACTTTGGAGCA
H13-170-R	Scaffold5751_AT6D_5319_gene20	GGCCTCTCACAAGTCTGGAA
H13-171-F	Scaffold5751_AT6D_5319_gene34	TGCCATGCTCTTGTTCATAGC
H13-171-R	Scaffold5751_AT6D_5319_gene34	TGTCATACCCCTGTCCACGA
H13-172-F	Scaffold5751_AT6D_5319_gene34	GCCGTTGAGGAAAATCATGT
H13-172-R	Scaffold5751_AT6D_5319_gene34	TTTGTCTTGTATGGCTCACG
H13-173-F	Scaffold5751_AT6D_5319_gene34	TTGGAGGCTTTCGTTCAACT
H13-173-R	Scaffold5751_AT6D_5319_gene34	TGGAGCCATTGTCTTCAGTG
H13-174-F	Scaffold5751_AT6D_5319_gene34	TCTGCTCCCACCTACTGCTT
H13-174-R	Scaffold5751_AT6D_5319_gene34	AGGCACCCAGATCTTCTTCA
H13-175-F	Scaffold5751_AT6D_5319_gene34	AAGTTCACTAAGATGAAAGGC
H13-175-R	Scaffold5751_AT6D_5319_gene34	AGTAGCAGTATCACGCTCAAG
H13-176-F	Scaffold5751_AT6D_5319_gene34	TCACAACGAAGATTATGATTC
H13-176-R	Scaffold5751_AT6D_5319_gene34	AAGACTATTGGCAGCTAGGC

Order name	Details	Sequence 5' to 3'
H13-177-F	Scaffold5751_AT6D_5319_gene50	GGATCAAGTCAAGCTGCACA
H13-177-R	Scaffold5751_AT6D_5319_gene50	ATGGTCTCCACATCCGAGTC
H13-178-F	Scaffold5751_AT6D_5319_gene50	CTTGCTATCCGATTCGTCCG
H13-178-R	Scaffold5751_AT6D_5319_gene50	CTTTGTGTATTCATGCCCATTC
H13-179-F	Scaffold5751_AT6D_5319_gene71	GAACGCTATGGGTGTTTCGAT
H13-179-R	Scaffold5751_AT6D_5319_gene71	GTTTCGGAAGACATGCAGACA
H13-180-F	Scaffold5751_AT6D_5319_gene71	ACGGAAGACCGAAAGGTCGAG
H13-180-R	Scaffold5751_AT6D_5319_gene71	CCGGCAACTGGCTTACCG
H13-181-F	Scaffold5751_AT6D_5319_gene71	ATTGCAAGGGAAGTGGTCACTG
H13-181-R	Scaffold5751_AT6D_5319_gene71	CTCCGCGAGCCTCAACAC
H13-182-F	Scaffold5319_masked_gene373	GTAAAGTGGTTGTGGCAGCA
H13-182-R	Scaffold5319_masked_gene373	TACCAGAACCGGAACACCTC
H13-183-F	Scaffold5319_masked_gene373	TCGTTGGAATTGAACTGCTG
H13-183-R	Scaffold5319_masked_gene373	ACAGGTATGGTAGCGGATGG
H13-184-F	Scaffold5319_masked_gene373	GGTCTGGAAGAACAGGGTGA
H13-184-R	Scaffold5319_masked_gene373	TGGCCAGTTCATTGTCTCTG
H13-185-F	Scaffold5319_masked_gene373	AAAACCGGCAGAGAGTCTGA
H13-185-R	Scaffold5319_masked_gene373	GCTGCCACAACCACTTTACA
H13-186-F	Scaffold5319_masked_gene373	CGCCATGGCCTTGTGACG
H13-186-R	Scaffold5319_masked_gene373	GCCATTGGGTATTGGGATGCG
H13-187-F	Scaffold5319_masked_gene26	CATGCTCGAGGACGACTACA
H13-187-R	Scaffold5319_masked_gene26	CAAGACGGTTGACTCGTTCA
H13-188-F	Scaffold5319_masked_gene26	GACTGAAAAAGAGGGCAACG
H13-188-R	Scaffold5319_masked_gene26	TCCTATGCTGCCATCAAACA
H13-189-F	Scaffold5319_masked_gene26	CGTGTCTTGCTATCGTCGAA
H13-189-R	Scaffold5319_masked_gene26	ACAGAGCGGGCTGTCTTAAA
H13-190-F	Scaffold5319_masked_gene26	GAGATGACATCGACAAGATGGAT

Order name	Details	Sequence 5' to 3'
H13-190-R	Scaffold5319_masked_gene26	ATGCATGGAGTTTGGCTATGT
H13-191-F	Scaffold5319_masked_gene26	CAATAACCAAGTCGGTTGGCC
H13-191-R	Scaffold5319_masked_gene26	CAGTACAGAGCGGGCTGTCT
H13-192-F	Scaffold5319_masked_gene30	GCACTCGGAAAAGTGGAGAC
H13-192-R	Scaffold5319_masked_gene30	GGATGGATATCACCGACACC
H13-193-F	Scaffold5319_masked_gene30	GTGGTTGCGTTGTACGTGTC
H13-193-R	Scaffold5319_masked_gene30	GGGTCATGCTATCTGCCAAT
H13-194-F	Scaffold5319_masked_gene30	CTTGGTGTACGTGCCATCAC
H13-194-R	Scaffold5319_masked_gene30	TGCTGGTCAAGAAGTGTTC
H13-195-F	Scaffold5319_masked_gene30	TCCGCCCATCCATCTATTTA
H13-195-R	Scaffold5319_masked_gene30	AATGGCGACCACCATAATGT
H13-196-F	Scaffold5319_masked_gene30	CGAAGAGCTAAACCATGCAGTG
H13-196-R	Scaffold5319_masked_gene30	TCCTTGAGCCGAGAAAGAGG
H13-197-F	Scaffold5319_masked_gene34	AGGTTTGTGTGCTGCTGTGTG
H13-197-R	Scaffold5319_masked_gene34	GCTGTTCTTGTCTGCCCTTC
H13-198-F	Scaffold5319_masked_gene34	AAGTGCGCCAGTTAGCAAAT
H13-198-R	Scaffold5319_masked_gene34	TTGCACCAGCTTACTGATGG
H13-199-F	Scaffold5319_masked_gene34	TGGACTCCAGAGGTTCGTTT
H13-199-R	Scaffold5319_masked_gene34	GTGATCTGCTTTCTGCCACA
H13-200-F	Scaffold5319_masked_gene34	CCAACATAGTGGCTGGGAGT
H13-200-R	Scaffold5319_masked_gene34	AAAGCCATAGCCTTGCAGAA
H13-201-F	Scaffold5319_masked_gene34	GGTTAAACCTGCTTGCAGATT
H13-201-R1	Scaffold5319_masked_gene34	GGCAGACTAGATATATCAGTG
H13-201-R2	Scaffold5319_masked_gene34	CAAGTACCTCAAGTGCTTGAG
H13-202-F	Scaffold5319_masked_gene332	TCTCCAAGCTTGACCTTCGT
H13-202-R	Scaffold5319_masked_gene332	ATTTCGACCCATCGAGTCTG
H13-203-F	Scaffold5319_masked_gene332	CTCAAGGCTTTCTTGCCAAC

Order name	Details	Sequence 5' to 3'
H13-203-R	Scaffold5319_masked_gene332	ATTTCCACGCTTGGACTTTG
H13-204-F	Scaffold5319_masked_gene332	TCTCCAAGCTTGACCTTCGT
H13-204-R	Scaffold5319_masked_gene332	CATTTTCGACCCATCGAGTCT
H13-205-F	Scaffold5319_masked_gene332	TTTCGTTCCATTTGGAGGTC
H13-205-R	Scaffold5319_masked_gene332	CCAGAGATTTGGGCATCAAT
H13-206-F	Scaffold5319_masked_gene332	TCTAGCCGTAGCTCGGTGTT
H13-206-R	Scaffold5319_masked_gene332	GGCTTAGGTGGCAAATGGTA
H13-207-F	Scaffold5319_masked_gene332	CGTCAGTTGAAGCCCAACGAG
H13-207-R	Scaffold5319_masked_gene332	GGCCACTTTAGAAAGCTTGTCG
H13-208-F	Scaffold5319_masked_gene320	TTGTCTGCAACGGTGTTCCTC
H13-208-R	Scaffold5319_masked_gene320	TGCCCTTTTGGTACTTTTGC
H13-209-F	Scaffold5319_masked_gene320	CTCCGGGGTTACAGTTTCAA
H13-209-R	Scaffold5319_masked_gene320	TGATGTTTGGGGCTTTGATT
H13-210-F	Scaffold5319_masked_gene320	GATGACCAAAGCTATGTTTGC
H13-210-R	Scaffold5319_masked_gene320	CCATTGAGATGCGTGATGGGT
H13-211-F	Scaffold5319_masked_gene320	TAGCAGGCATGCCTACCA
H13-211-R	Scaffold5319_masked_gene320	GAGATGCATCGACATGGTATT
H13-212-F	Scaffold5319_masked_gene301	CCATATCAGCAGGGCTTGTT
H13-212-R	Scaffold5319_masked_gene301	GTTGGAGCATCAAACGACCT
H13-213-F	Scaffold5319_masked_gene301	ATGAGGCAGATGGGAATTTG
H13-213-R	Scaffold5319_masked_gene301	CACACTCCTCCGTCGTAACA
H13-214-F	Scaffold5319_masked_gene301	CTTAGCACCGTTTCCTCCTG
H13-214-R	Scaffold5319_masked_gene301	TGAAACCCACTAGGCCAAAC
H13-215-F	Scaffold5319_masked_gene301	ATCTATCCATCCATGCCCCTG
H13-215-R	Scaffold5319_masked_gene301	ATATGCGAGATGTGATATGCG
H13-216-F	Scaffold5319_masked_gene187	AGCGGAGCAGCAAGTAAATC
H13-216-R	Scaffold5319_masked_gene187	ATCCCTCCCTCTTCAAAGGA

Order name	Details	Sequence 5' to 3'
H13-217-F	Scaffold5319_masked_gene187	TCGCGAAGCAACATGAATAG
H13-217-R	Scaffold5319_masked_gene187	GAGGGAGGGAGAAGCAAGAT
H13-218-F	Scaffold5319_masked_gene187	AGGGTGCCAGAGTCGTTCTA
H13-218-R	Scaffold5319_masked_gene187	CCCGCAGCTAGGATTTACAA
H13-219-F	Scaffold5319_masked_gene187	TCACGAGTCTGCAATATCCCT
H13-219-R	Scaffold5319_masked_gene187	GACATGACCCGGAACCCCG
H13-220-F	Scaffold5319_masked_gene154	ATCGTGTGGATGAACTTAAAG
H13-220-R	Scaffold5319_masked_gene154	TAGCCAGTACATATATCACCA
H13-221-F	Scaffold5319_masked_gene154	GTGCATGCCTTGAATCATTG
H13-221-R	Scaffold5319_masked_gene154	CTTCGTCCATCCTTCTCTCG
H13-222-F	Scaffold5319_masked_gene154	CAAATGGAACGACTGTGGTG
H13-222-R	Scaffold5319_masked_gene154	GAGGTGAGCTCCTTGCATTC
H13-223-F1	Scaffold5319_masked_gene132	TTGGTTGGATCCTTCTGGAC
H13-223-F2	Scaffold5319_masked_gene132	GGAGGACTAGACGGGGACTC
H13-223-F3	Scaffold5319_masked_gene132	ACTGCCCCCTCTCTTCCTGTT
H13-223-F4	Scaffold5319_masked_gene132	CTGCATGGAGACATCCTTCC
H13-223-R1	Scaffold5319_masked_gene132	GAACTCCAGGATGTGGTGGT
H13-223-R2	Scaffold5319_masked_gene132	GCTAGGGTTCACACCGAAAA
H13-223-R3	Scaffold5319_masked_gene132	GAAGCAATGATGGCAGAGGT
H13-223-R4	Scaffold5319_masked_gene132	CTGCGTCGAGTTCAGGCATG
H13-224-F1	Scaffold5319_masked_gene150	TTTTTGATGCTCCTTTTATCA
H13-224-F2	Scaffold5319_masked_gene150	TTTCATGAGATGGCACATAAC
H13-224-R1	Scaffold5319_masked_gene150	ATGATGCCCAACCTCAAGAC
H13-224-R2	Scaffold5319_masked_gene150	CCAACGCATACCAAAAGAAAA
H13-225-F1	Scaffold5319_masked_gene150	ACCACTGGGTTCACCTGGTA
H13-225-F2	Scaffold5319_masked_gene150	GGCATGTAAAAGGAGGTGCT
H13-225-F3	Scaffold5319_masked_gene150	ACTGTATCAGGCATGTAAAAG
H13-225-R1	Scaffold5319_masked_gene150	AACGGGCAACAAATCAAAAC

Order name	Details	Sequence 5' to 3'
H13-225-R2	Scaffold5319_masked_gene150	ACAAGCAGATGCATCACAGC
H13-226-F	Scaffold5319_masked_gene172	GTCTCCGACCCGTGATTCGTC
H13-226-R	Scaffold5319_masked_gene172	ACCCGAGAATACCAACCAAA
H13-227-F	Scaffold5319_masked_gene172	ACAGAGAGTGGTGGCCTTGGT
H13-227-R1	Scaffold5319_masked_gene172	GGAATGGAGTGTTCGTTTCAGG
H13-227-R2	Scaffold5319_masked_gene172	AACGGGGCAATCATCTTATG
H13-228-F	Scaffold5319_masked_gene172	TTGCCTTCAGTTACCCCAT
H13-228-R1	Scaffold5319_masked_gene172	AAACATCAATGCTGGCAAGA
H13-228-R2	Scaffold5319_masked_gene172	TTAGTCTGCATAGCATCACAA
H13-229-F	Scaffold5319_masked_gene172	CTATGCTTTGTGTGTGTGGTC
H13-229-R1	Scaffold5319_masked_gene172	CCATTTTCACAGGTGACCGG
H13-229-R2	Scaffold5319_masked_gene172	CTTCAACCTTTGTGCCAGTTG
H13-230-F	Scaffold5319_masked_gene172	ATATGCCCTCATGTGTTCG
H13-230-R	Scaffold5319_masked_gene172	ATCGACCTTAACAGTATCTCC
H13-231-F	Scaffold5319_masked_gene172	TCGACGTGGCATAACTGTTC
H13-231-R	Scaffold5319_masked_gene172	TGCTCTTATTGGCACTGCTC
H13-232-F	Scaffold5319_masked_gene172	GCGGCGTCTAGTAAGTTAACC
H13-232-R	Scaffold5319_masked_gene172	TCACTGTCAAATTCACACTACAG
H13-233-F	Scaffold5319_masked_gene180	AACTCAAGTGGCGTGAAGGT
H13-233-R	Scaffold5319_masked_gene180	AGGTGCACATGCTCACAACAT
H13-234-F	Scaffold5319_masked_gene180	GCCACATTTCTCACAATGGTT
H13-234-R	Scaffold5319_masked_gene180	CGGTCTTAGTAGGTGCGCTCC
H13-235-F	Scaffold5319_masked_gene180	AGGGCTTGAACAAGACTCCA
H13-235-R	Scaffold5319_masked_gene180	CCACTTGGTTCCAATGACAT
H13-236-F1	Scaffold5319_masked_gene235	TGATTTCCTTGCCCGCCATCG
H13-236-F2	Scaffold5319_masked_gene235	CATAACCCACCTCAATTACTA
H13-236-R	Scaffold5319_masked_gene235	GTTCCAACCCTCAAATCCC

Order name	Details	Sequence 5' to 3'
H13-237-F1	Scaffold5319_masked_gene235	TGCATCTTAATGGAATGTCTA
H13-237-F2	Scaffold5319_masked_gene235	GGATCTTGAGGCACAACAC
H13-237-R	Scaffold5319_masked_gene235	TCCTCTTGGCTGAACTGG
H13-238-F	Scaffold5319_masked_gene235	CCCTTCTCCTTGCAGTTCTG
H13-238-R	Scaffold5319_masked_gene235	TGATATGAACCGGCCCTA
H13-239-F	Scaffold5319_masked_gene235	CCCAAGTACAAGGCACAACC
H13-239-R	Scaffold5319_masked_gene235	TGGACCTGCAAATAGGTTCC
H13-240-F1	Scaffold5318_masked_gene_9	CAATGACAAAAGATCTTCTAG
H13-240-F2	Scaffold5318_masked_gene_9	TTCCAAGTCAAGGTCCTAC
H13-240-F3	Scaffold5318_masked_gene_9	CTTCTAGTCGTGAGCGTAAG
H13-240-R1	Scaffold5318_masked_gene_9	ACATTGGATTTAGCATGATTG
H13-240-R2	Scaffold5318_masked_gene_9	CAAGAGAACTACTAGAAGCAT
H13-240-R3	Scaffold5318_masked_gene_9	TGCTCAAGTGTAAGTAGGC
H13-241-F	Scaffold5318_masked_gene_9	TCCGCCAACATACTTCACTCC
H13-241-R	Scaffold5318_masked_gene_9	CAACTCGGTGTGACCGAAAT
H13-242-F1	Scaffold5318_masked_gene_9	CCGCCGCCGTCAATGGAAA
H13-242-F2	Scaffold5318_masked_gene_9	GTCCTCTCGTCGCCAGCGT
H13-242-R1	Scaffold5318_masked_gene_9	GTCTTCCTGCTGGCCTCTTTG
H13-242-R2	Scaffold5318_masked_gene_9	TGTGGCAACTTTTCTTGGCTTA
H13-243-F	Scaffold5318_masked_gene_13	CACCGGTGGACTGCATTTTC
H13-243-R	Scaffold5318_masked_gene_13	CAGTCCATCCACCACACCAT
H13-244-F	Scaffold5318_masked_gene_17	GCGTATTCTTGTGGTGAAC
H13-244-R	Scaffold5318_masked_gene_17	AGGAGCTGACTTGCACCTT
H13-245-F	Scaffold5318_masked_gene_25	GATCGCCTCTGATACGTGA
H13-245-R1	Scaffold5318_masked_gene_25	AATGATAGTGCAATTATGTCG
H13-245-R2	Scaffold5318_masked_gene_25	CGCATGGCCTATAGTCCGG
H13-246-F	Scaffold5318_masked_gene_25	TACGTTGTTGAATATTTCA
H13-246-R1	Scaffold5318_masked_gene_25	GGAAAGAGAAAAGTGGATTGC

Order name	Details	Sequence 5' to 3'
H13-246-R2	Scaffold5318_masked_gene_25	GCGTGTCCGAGGTCCCTGT
H13-247-F1	Scaffold5318_masked_gene_35	ATCTCGATGCGCCCTTCG
H13-247-F2	Scaffold5318_masked_gene_35	AGGTCAAGTGCTCGGATGATA
H13-247-R	Scaffold5318_masked_gene_35	CTCCAGGCTCTCCGCGTC
H13-248-F1	Scaffold5318_masked_gene_35	GACGCGGAGAGCCTGGAG
H13-248-F2	Scaffold5318_masked_gene_35	AGCTGCTGGCACCGGTGCTA
H13-248-F3	Scaffold5318_masked_gene_35	GTGGGGTGGGCTAAGTTAGT
H13-248-R1	Scaffold5318_masked_gene_35	ACCTTCTCTTCCTCATCTGT
H13-248-R2	Scaffold5318_masked_gene_35	TCTAGCAAGACCTGGGCGG
H13-248-R3	Scaffold5318_masked_gene_35	ACTAACTTAGCCCACCCAC
H13-249-F1	Scaffold5318_masked_gene_43	GTTCAAACCTCAGTTAGAACAA
H13-249-F2	Scaffold5318_masked_gene_43	TGCATGACTTTGCTGTTAGAG
H13-249-R1	Scaffold5318_masked_gene_43	AGGTGTGAGTTTATTCAAGTA
H13-249-R2	Scaffold5318_masked_gene_43	CAAATGGCTGATAAATCAAC
H13-250-F1	Scaffold5318_masked_gene_43	TATCGGTATCTGTGAGGATC
H13-250-F2	Scaffold5318_masked_gene_43	TTGGTAGAACCTTCTTGAAC
H13-250-R1	Scaffold5318_masked_gene_43	CACAACCCACCAGGGCA
H13-250-R2	Scaffold5318_masked_gene_43	TCCTAGCGTGCCCAGGTA
H13-251-F	Scaffold5318_masked_gene_48	GGTAGAAGTGTTGGTCCAGT
H13-251-R1	Scaffold5318_masked_gene_48	TATCTTGCTTCCATACTTGT
H13-251-R2	Scaffold5318_masked_gene_48	GTGCATTATTATGCCCACC
H13-252-F1	Scaffold5318_masked_gene_48	CATGACATATCACAAGTATGG
H13-252-F2	Scaffold5318_masked_gene_48	TGATAGAATGGGCATAAGGT
H13-252-F3	Scaffold5318_masked_gene_48	GAAGAGATGCGTGTAGTTCG
H13-252-R	Scaffold5318_masked_gene_48	TTTGACAACCCATTGAAC TTG
H13-253-F	Scaffold5318_masked_gene_57	CGTTATGTCTAGACGCCGATC
H13-253-R	Scaffold5318_masked_gene_57	ACATGCACATGCGCATTGTAC

Order name	Details	Sequence 5' to 3'
H13-254-F	Scaffold5318_masked_gene_57	CTGAGTGTGTACAATGCGCAT
H13-254-R	Scaffold5318_masked_gene_57	GAGAGGAGCGTCTCTGTCCT
H13-255-F	Scaffold5318_masked_gene_57	GGACAGAGACGCTCCTCTCG
H13-255-R	Scaffold5318_masked_gene_57	GCGGTCTCCCTTATCACCTC
H13-256-F	Scaffold5318_masked_gene_75	ATGGCACCCCCAAGATAATA
H13-256-R	Scaffold5318_masked_gene_75	TGGCTGTTTTGAGGAAGGTT
H13-257-F	Scaffold5318_masked_gene_75	CATGGCAATACCCTACTCACC
H13-257-R	Scaffold5318_masked_gene_75	GGGTCCAAGTCTCCTGGATC
H13-258-F1	Scaffold5318_masked_gene_84	CATCCCGCAACCGCCATTC
H13-258-F2	Scaffold5318_masked_gene_84	TGCAAGAGCCTCAACTAGTG
H13-258-R1	Scaffold5318_masked_gene_84	ATCGATCGCCACCAGCCG
H13-258-R2	Scaffold5318_masked_gene_84	CCAAGGGAAGCTGGTCTGTG
H13-259-F	Scaffold5318_masked_gene_84	GATGTTGCAAGCTCCTTCGA
H13-259-R	Scaffold5318_masked_gene_84	CGTTGTGATCAAACCACTCC
H13-260-F1	Scaffold5319_masked_Rgene_26	ACGGAGATGGTGAAGGAAGAA
H13-260-F2	Scaffold5319_masked_Rgene_26	CTGCTTGGAGTCAGAGATGAC
H13-260-F3	Scaffold5319_masked_Rgene_26	ACATCGACAAGATGGATGTC
H13-260-R1	Scaffold5319_masked_Rgene_26	GCCAACGCTTTTTCACCACA
H13-260-R2	Scaffold5319_masked_Rgene_26	TCACCCCATGCTTCGTGACT
H13-261-F1	Scaffold5319_masked_Rgene_26	CTTGCTGGGGGAAAACACTG
H13-261-F2	Scaffold5319_masked_Rgene_26	ATTCTCACTACTGCCCTGACA
H13-261-F3	Scaffold5319_masked_Rgene_26	GGTGCTAGATGATGTGTGGA
H13-261-R1	Scaffold5319_masked_Rgene_26	TTCATGGGATTTCATAGGAT
H13-261-R2	Scaffold5319_masked_Rgene_26	TGCTCACCCAAACAAGGTGCA
H13-261-R3	Scaffold5319_masked_Rgene_26	TCAAAGTCTGGTTTCGCCAG
H13-262-F	Scaffold5319_masked_Rgene_26	GCCCGCTCTGTACTGAATCG
H13-262-R	Scaffold5319_masked_Rgene_26	GATTCAACCAACGCGGCATGT
H13-263-F1	Scaffold5319_masked_Rgene_26	AATGTAAGTCTTCCTCGTTCT

Order name	Details	Sequence 5' to 3'
H13-263-F2	Scaffold5319_masked_Rgene_26	AAGGCAAAGCTTGTTGAGAAGT
H13-263-F3	Scaffold5319_masked_Rgene_26	TAGTAGATTGGGGGATGACGG
H13-263-R1	Scaffold5319_masked_Rgene_26	CTCAAGTCCAGATTGTGCCAC
H13-263-R2	Scaffold5319_masked_Rgene_26	ACTTCCCATGTATTCTGGGAG
H13-264-F1	Scaffold5319_masked_Rgene_26	CACTCCCAGAATACATGGGAA
H13-264-F2	Scaffold5319_masked_Rgene_26	GGCATATGCACATGAAGGAGAC
H13-264-R1	Scaffold5319_masked_Rgene_26	CACACCAACAGTGGATCCTAT
H13-264-R2	Scaffold5319_masked_Rgene_26	CAAGTGATGCTTTGCTCGATT
H13-265-F1	Scaffold5319_masked_Rgene_26	CGGTGTTATATGTACAACCTCG
H13-265-F2	Scaffold5319_masked_Rgene_26	ACCAAACCGCGACAGTGAGT
H13-265-F3	Scaffold5319_masked_Rgene_26	GGTTGGCATCGACAGGTAGC
H13-265-R1	Scaffold5319_masked_Rgene_26	GCCTCTCTCCTCTGCTGATT
H13-265-R2	Scaffold5319_masked_Rgene_26	GTCCTCCTCTGCATCCACTGT
H13-266-F1	Scaffold5319_masked_Rgene_30	TTGTTGCAGCAAAAAGCAGCG
H13-266-F2	Scaffold5319_masked_Rgene_30	GCTGACGGAGATGGCGAGT
H13-266-F3	Scaffold5319_masked_Rgene_30	TTTGCATCCTACGTGCATAAC
H13-266-R1	Scaffold5319_masked_Rgene_30	AGGGTGTTTCAAGCACATCAC
H13-266-R2	Scaffold5319_masked_Rgene_30	CGGCTACCTTGAGCCACAGT
H13-266-R3	Scaffold5319_masked_Rgene_30	TTCACATGGTGGTGAAGCA
H13-267-F1	Scaffold5319_masked_Rgene_30	CAGCTACAACCTGTGCGCTG
H13-267-F2	Scaffold5319_masked_Rgene_30	CTGAAGGCGCTAGCAAGGT
H13-267-F3	Scaffold5319_masked_Rgene_30	CATGGTGGCCAGTTGATCTG
H13-267-R1	Scaffold5319_masked_Rgene_30	GCCACAGCTACTCTGAATTCG
H13-267-R2	Scaffold5319_masked_Rgene_30	ACTGCAGGATGGAGGGCTTC
H13-267-R3	Scaffold5319_masked_Rgene_30	GATGAACGACCCACACCACA
H13-268-F1	Scaffold5319_masked_Rgene_30	GGGACAAAGTTCAGACATGTTA
H13-268-F2	Scaffold5319_masked_Rgene_30	ATGTCAAAGCGTACTCAAATG

Order name	Details	Sequence 5' to 3'
H13-268-F3	Scaffold5319_masked_Rgene_30	GTCAGCCGCTCTTTCCTGTCT
H13-268-R1	Scaffold5319_masked_Rgene_30	CTCCCATCAACAAAGGTATAG
H13-268-R2	Scaffold5319_masked_Rgene_30	ATCTTTCAGTGATGGCACGTA
H13-269-F1	Scaffold5319_masked_Rgene_34	AGCAAATGAAGACGCTGTCAA
H13-269-F2	Scaffold5319_masked_Rgene_34	TGAATCGGGAGGTGAATTAGT
H13-269-F3	Scaffold5319_masked_Rgene_34	TGCTGTCGCGACACTTCATGT
H13-269-R1	Scaffold5319_masked_Rgene_34	CATATATACACCTGACCTGAC
H13-269-R2	Scaffold5319_masked_Rgene_34	TGGATGGCAAAATACACAATC
H13-269-R3	Scaffold5319_masked_Rgene_34	GCCTATGCATCTGCGCGAC
H13-270-F1	Scaffold5319_masked_Rgene_34	AGCTATACTGCAGCCTAGCAT
H13-270-F2	Scaffold5319_masked_Rgene_34	CCCTGAGTGGGACAAGTTCG
H13-270-R1	Scaffold5319_masked_Rgene_34	TCTGAATCTTCTCATGGTCGC
H13-270-R2	Scaffold5319_masked_Rgene_34	GCGGTAGTGGTGATCTGCTTT
H13-271-F1	Scaffold5319_masked_Rgene_52	TCAGATGAACGACCAAGCACC
H13-271-F2	Scaffold5319_masked_Rgene_52	TCAGATGAACGACCAAGCACC
H13-271-R	Scaffold5319_masked_Rgene_52	TCTGTGCTACTCCTATCCAG
H13-272-F1	Scaffold5319_masked_Rgene_73	TGGATGCCTTTGCATCCTACG
H13-272-F2	Scaffold5319_masked_Rgene_73	TGCTGCTTGGAGTCACAAACC
H13-272-R1	Scaffold5319_masked_Rgene_73	GACAAGGATTGGCTGAAGCAT
H13-272-R2	Scaffold5319_masked_Rgene_73	CCATGGTTCCACACATCATCC
H13-272-R3	Scaffold5319_masked_Rgene_73	ACGAGGACTCGGCTACCTG
H13-273-F1	Scaffold5319_masked_Rgene_73	CCACCATGTGGACAAATTAGAT
H13-273-F2	Scaffold5319_masked_Rgene_73	AGGATGCCTGGTCATTGCTCC
H13-273-R1	Scaffold5319_masked_Rgene_73	GTTAACCGCTAGTGTCGAAGC
H13-273-R2	Scaffold5319_masked_Rgene_73	TGCAAGCAGTTGAACTGATGC
H13-274-F1	Scaffold5319_masked_Rgene_73	ATATGCTCAAGGACATTGGGT
H13-274-F2	Scaffold5319_masked_Rgene_73	CAGTCAAAGTAATGGGAGGTC
H13-274-F3	Scaffold5319_masked_Rgene_73	GAGTGGGAGATGGTTCTGGC

Order name	Details	Sequence 5' to 3'
H13-274-R1	Scaffold5319_masked_Rgene_73	CTATGCTTGTTCCTGAAGC
H13-274-R2	Scaffold5319_masked_Rgene_73	CCTGGGCTGGAAACCCAGT
H13-275-F1	Scaffold5319_masked_Rgene_73	GGTCCTCTTTCTCAGCTCAAG
H13-275-F2	Scaffold5319_masked_Rgene_73	CAACAAAGGCAAAGCTTGGCA
H13-275-F3	Scaffold5319_masked_Rgene_73	AAGCTTGGCAACAAGGTGCAA
H13-275-R1	Scaffold5319_masked_Rgene_73	GTCTTCATGTCTTCATCCGTG
H13-275-R2	Scaffold5319_masked_Rgene_73	GTGAGCAGTGAAGGCTACAA
H13-275-R3	Scaffold5319_masked_Rgene_73	TGTGCAGCGGTATGGAAGT
H13-276-F1	Scaffold5319_masked_Rgene_73	AGCATATGCATCTGAAGGAGA
H13-276-F2	Scaffold5319_masked_Rgene_73	ATGGTTCGTCTTGTACACAGC
H13-276-F3	Scaffold5319_masked_Rgene_73	TACACAGCTGATCCCTACAAC
H13-276-R1	Scaffold5319_masked_Rgene_73	GAGTTGTACATATAACACCGCAA
H13-276-R2	Scaffold5319_masked_Rgene_73	TCTCACGGTCCCAGTTTGGC
H13-277-F1	Scaffold5319_masked_Rgene_73	CATCCCATTGGGGCCAACTG
H13-277-F2	Scaffold5319_masked_Rgene_73	GAATGACAGCATCCTAGTACAG
H13-277-F3	Scaffold5319_masked_Rgene_73	ATCCAGAGATGAGGTTTCAGGT
H13-277-R1	Scaffold5319_masked_Rgene_73	TCTTCATCCCCTGAGCCCGA
H13-277-R2	Scaffold5319_masked_Rgene_73	CCTCTGCATCCACCGGGCT
H13-278-F	Scaffold5319_masked_Rgene_56	CAGTCTGGCCAGCCTGT
H13-278-R	Scaffold5319_masked_Rgene_56	CCTTTGAATACCCCTTCGAGAG
H13-279-F	Scaffold5319_masked_Rgene_56	GGACGTAACCTAGGGGGG
H13-279-R	Scaffold5319_masked_Rgene_56	CCTGTCTTGCAGCAACTTTTGC
H13-280-F	Scaffold5319_masked_Rgene_56	GCTAAATTACCTCTGGACATGAAAA
H13-280-R	Scaffold5319_masked_Rgene_56	GCTTGAGGATTATGTTCAATTGC
H13-281-F	Scaffold5319_masked_Rgene_56	GTTCTCTAGGGAGATGAGACTG
H13-281-R	Scaffold5319_masked_Rgene_56	CCTCTTTCTCGGCTCAAGAAA
H13-282-F1	Scaffold5319_masked_Rgene_70	GTGTCTCCGTGACTTTATCCGC

Order name	Details	Sequence 5' to 3'
H13-282-F2	Scaffold5319_masked_Rgene_70	CTTTGTGCAGTCTTCCATGGA
H13-282-R1	Scaffold5319_masked_Rgene_70	AGAAACTTCAATTGGGCGGAC
H13-282-R2	Scaffold5319_masked_Rgene_70	GAGTTGGGTCCTCTTTCTCGT
H13-283-F1	Scaffold5319_masked_Rgene_70	CTTGGA CTGGAAGTCAAGTAC
H13-283-R1	Scaffold5319_masked_Rgene_70	TGGTCAATATCTGGACTGCCT
H13-283-R2	Scaffold5319_masked_Rgene_70	CATGAGTGGGAGATGGTTCTG
H13-284-F1	Scaffold5319_masked_Rgene_70	CCAGAACCATCTCCCACTCAT
H13-284-F2	Scaffold5319_masked_Rgene_70	ACTGCATGGTTTAGCTCTTCA
H13-284-R	Scaffold5319_masked_Rgene_70	TTCAAAC TATTCGTAGTCGTCAC
H13-285-F	Scaffold5319_masked_Rgene_70	CTGGTGTGTTGTGGGAGTGC
H13-285-R1	Scaffold5319_masked_Rgene_70	ACACCATCAAGGAGCGAAGCG
H13-285-R2	Scaffold5319_masked_Rgene_70	GCTCAACAAGAGGCTTGACAC
H13-286-F1	Scaffold5319_masked_Rgene_70	GATGCGGGTGCCGATGTCAC
H13-286-F2	Scaffold5319_masked_Rgene_70	GGATGGACCCCGCTCCATG
H13-286-R1	Scaffold5319_masked_Rgene_70	AAATCCGAGCTTGACTGGTG
H13-286-R2	Scaffold5319_masked_Rgene_70	AGCTGATATGTTGGTTTACCC
H13-287-F	Scaffold5319_masked_Rgene_60	GGCACGCTTGAGCTGGGCCG
H13-287-R	Scaffold5319_masked_Rgene_60	GACGATGGTCCTGGATGCCTTT
H13-288-F1	Scaffold5319_masked_Rgene_63	CCATTGCCAGGGCAGCTCA
H13-288-F2	Scaffold5319_masked_Rgene_63	CAATGAAGGCTTCCCTGCCT
H13-288-R1	Scaffold5319_masked_Rgene_63	GAAGTGCGCCAGTTAGCAAAA
H13-288-R2	Scaffold5319_masked_Rgene_63	GCCCCAAGCGGAGGAATA
H13-289-F1	Scaffold5319_masked_Rgene_63	CCACGGGAGAAGAGGGAAC
H13-289-F2	Scaffold5319_masked_Rgene_63	TCCCCGACTTCTGATCCATCTAA
H13-289-R	Scaffold5319_masked_Rgene_63	CAAAAGAAAATAGATGACCGGCC
H13-290-F1	Scaffold5319_masked_Rgene_63	TTCAAAC TCAACATACTGACAG
H13-290-F2	Scaffold5319_masked_Rgene_63	ACTCAGGGCCAGATTGTCT
H13-290-R1	Scaffold5319_masked_Rgene_63	CCATCAAGCTTGTTGGACCC

Order name	Details	Sequence 5' to 3'
H13-290-R2	Scaffold5319_masked_Rgene_63	CACAGCTTCCTATTGGTTTGTGTT
H13-291-F1	Scaffold5319_masked_Rgene_63	GGGCAGAGCTCATCAAATAG
H13-291-F2	Scaffold5319_masked_Rgene_63	CTCGATTGTGTTGTCGCTTCC
H13-291-R1	Scaffold5319_masked_Rgene_63	GGAAGCATGTACTACAAGGAGT
H13-291-R2	Scaffold5319_masked_Rgene_63	CCTCGGATGACTTAGAAGAAGT
H13-292-F1	Scaffold5319_masked_Rgene_63	AAATCCTTCACTAATCCACATCG
H13-292-F2	Scaffold5319_masked_Rgene_63	CCACATCGCAATGATTTTCATTTT
H13-292-R1	Scaffold5319_masked_Rgene_63	GTCCAACCAAAAATTAGTAACTAGG
H13-292-R2	Scaffold5319_masked_Rgene_63	AAATGCAAGTCCAACCAAAAATTAG
H13-293-F1	Scaffold5319_masked_Rgene_63	TGAGTCACACTAGCAATTTCTTTT
H13-293-F2	Scaffold5319_masked_Rgene_63	ACTCCTAGGTACAATAATCAA
H13-293-R1	Scaffold5319_masked_Rgene_63	TGAGCATCAACAAGGACTTCA
H13-293-R2	Scaffold5319_masked_Rgene_63	TGAATGCAGAGTTCGAGAAGA
H13-294-F	Scaffold5319_masked_Rgene_63	CCTGATCATCACCATGTACTC
H13-294-R	Scaffold5319_masked_Rgene_63	TGGTCCTGGATGCCTTTGCG
H13-295-F1	Scaffold5319_masked_Rgene_52	CGAAGAGTCAACAAGGAGCGG
H13-295-F2	Scaffold5319_masked_Rgene_52	CTCGTGAAGAGAGACTCACGG
H13-295-R1	Scaffold5319_masked_Rgene_52	TGATGAACGTTGGACAACCCC
H13-295-R2	Scaffold5319_masked_Rgene_52	ACAAATGCATGCACCCACCAT
H13-296-F1	Scaffold5319_masked_Rgene_52	TGGTGGGTGCATGCATTTGTG
H13-296-F2	Scaffold5319_masked_Rgene_52	AATAGTGTCCGCGCTGTATGG
H13-296-R1	Scaffold5319_masked_Rgene_52	AAAGATGTGGGGTGCTTCATC
H13-296-R2	Scaffold5319_masked_Rgene_52	GCTGACATCCTCGACCTATGC
H13-297-F1	Scaffold5319_masked_Rgene_52	GATGGTAGGGAAGCACAGCC
H13-297-F2	Scaffold5319_masked_Rgene_52	CGAGCCACGACATTAGCCAC
H13-297-R	Scaffold5319_masked_Rgene_52	CGATGGTCCCGGATGCCT
H13-298-F1	Scaffold5319_masked_Rgene_52	CCAACCTTCTGATCCATCTACCG

Order name	Details	Sequence 5' to 3'
H13-298-F2	Scaffold5319_masked_Rgene_52	GCTACCTTCAACTTGATGCCG
H13-298-R	Scaffold5319_masked_Rgene_52	AGTTGGGCTCTTGCCTTGC
H13-299-F	Scaffold5319_masked_Rgene_52	TGTCTCCAAGTTGAAGGGGTT
H13-299-R	Scaffold5319_masked_Rgene_52	TGATGAGCTCTGCCCTCCG
H13-300-F1	Scaffold5319_masked_Rgene_65	TTTGGTCCAAATGGGATGTGA
H13-300-F2	Scaffold5319_masked_Rgene_65	AAACTTGGATTTGATTTTCATCCTTC
H13-300-R1	Scaffold5319_masked_Rgene_65	TCAGCCACTCTTTCATGCTC
H13-300-R2	Scaffold5319_masked_Rgene_65	GTCTTCTACACAGCTAAGCCATAT
H13-301-F1	Scaffold5319_masked_Rgene_65	ATTTCTTGAATTACCTCCTTCG
H13-301-F2	Scaffold5319_masked_Rgene_65	TGACATGCTCAACATGGCC
H13-301-R	Scaffold5319_masked_Rgene_65	CCTCCACCCAGGTTAGAAA
H13-307F	Cell division control prot., AAA-superfamily of ATPases	CAAATACGCCATCAGGGAGAACATC
H13-307R	Cell division control prot., AAA-superfamily of ATPases	CGCTGCCGAAACCACGAGAC
H13-308F	ADP-ribosylation factor	GCTCTCCAACAACATTGCCAAC
H13-308R	ADP-ribosylation factor	GCTTCTGCCTGTCACATACGC
H13-309F	RNase L inhibitor-like protein	CGATTTCAGAGCAGCGTATTGTTG
H13-309R	RNase L inhibitor-like protein	AGTTGGTCGGGTCTCTTCTAAATG
H13-310F	α -tubulin	GCCATCTACGACATCTGC
H13-310R	α -tubulin	GGTCTGGAACTCGGTTATG
H13-311F	β -tubulin	GCTTGCTGTCAATCTCATC
H13-311R	β -tubulin	CTTGGCATCCCACATTTG
H13-312F	Ubiquitin	GCACCTTGGCGGACTACAACATTC
H13-312R	Ubiquitin	GACACCGAAGACGAGACTTGTGAACC
H13-313F	Actin	TGACCGTATGAGCAAGGAG
H13-313R	Actin	CCAGACAACCTCGCAACTTAG
H13-314F	GAPDH	TTGCTCTGAACGACCATTTC
H13-314R	GAPDH	GACACCATCCACATTTATTCTTC
H13-315F	Ribosomal protein	CAAGGAGTACCGTGACAC

Order name	Details	Sequence 5' to 3'
H13-315R	Ribosomal protein	GCGGGAACCTTGATCTTCG
H13-316F	Histone	GTCACCATCATGCCCAAG
H13-316R	Histone	CAACACATTCCACTTCCG
H13-317F	TaWIN1 ,14-3-3 like protein	CATCACTACCTCTGTTTAATTATGCACTT
H13-317Pr	TaWIN1 ,14-3-3 like protein	AAGTTTCTGGGCTATGTTTTTCTGGTGCTGAT
H13-317R	TaWIN1 ,14-3-3 like protein	CACGAATCCAGCCAGCATT
H13-318F	TaFNRII (ferredoxin-NADP(H) oxidoreductase	CACCGGCCAGTGATCTT
H13-318Pr	TaFNRII (ferredoxin-NADP(H) oxidoreductase	AACCCGGCTTCAGGTCACAGAGG
H13-318R	TaFNRII (ferredoxin-NADP(H) oxidoreductase	AAGGGCGTCTGCTCCAAC
H13-319F	rrn26 ,a putative homologue to RNA 26S gene	TGAGAAGGAAGGACGCTTTCATG
H13-319Pr	rrn26 ,a putative homologue to RNA 26S gene	CGGAGATCCATGGTATCACAGAAAA
H13-319R	rrn26 ,a putative homologue to RNA 26S gene	ACCATGGCCTTTTCGCC
H13-320F	CYP18-2 ,Cyclophilin A	GATCTCCGTGGTTGGTTTAGGA
H13-320Pr	CYP18-2 ,Cyclophilin A	TTTAGGTGGACTCTTTTGCC
H13-320R	CYP18-2 ,Cyclophilin A	CGCCGGACACAGATCCA
H13-321F1	RT-PCR PRIMER, NEAR END OF 1ST EXON	TGATGAAAGTGGTGCTTCCCTAC
H13-321F2	RT-PCR PRIMER, NEAR END OF 1ST EXON	CCACCATGTGGACAAATTAGAT
H13-321F3	RT-PCR PRIMER, NEAR END OF 1ST EXON	AGGATGCCTGGTCATTGCTCC
H13-321R1	RT-PCR PRIMER, NEAR BEGINNING OF 2ND EXON	CACGTCCATCATTCTCACTTGAGAT
H13-321R2	RT-PCR PRIMER, NEAR BEGINNING OF 2ND EXON	GCTATAATTTGCAACCCAATGTCCT
H13-321R3	RT-PCR PRIMER, NEAR BEGINNING OF 2ND EXON	CAATAGACCTCCCATTACTTTGACT
H13-321R4	RT-PCR PRIMER, NEAR BEGINNING OF 2ND EXON	TTTTGCTATAATTTGCAACCCAATG
H13-322F	RT-PCR PRIMER , WITHIN 1ST EXON	ATGACGATGGTCCTGGATGCC
H13-322R	RT-PCR PRIMER , WITHIN 1ST EXON	CTCTCGTGGTGATGTTTCTCC
H13-323F	RT-PCR PRIMER , WITHIN 1ST EXON	TATTCCTGGTGCTGGATGATGTG
H13-323R	RT-PCR PRIMER , WITHIN 1ST EXON	TAGGGAAGCACCCTTTCATCA
H13-324F	RT-PCR PRIMER WITHIN 2ND EXON	GGTCCTCTTCTCAGCTCAAG

Order name	Details	Sequence 5' to 3'
H13-324R	RT-PCR PRIMER WITHIN 2ND EXON	CGATTTCGTTGTTGCTTTTCCTCA
H13-325F1	RT-PCR PRIMER WITHIN 2ND EXON	GTTGAGCTGGATCTGGATGAAAT
H13-325F2	RT-PCR PRIMER WITHIN 2ND EXON	GGATGAAATCCTTGACCTGGAGAG
H13-325R	RT-PCR PRIMER WITHIN 2ND EXON	GTGAGCAGTGCAAGGCTACAA
H13-326F	,2nd exon, replace 325f2 with M	GGATGAAAACCTTGACCTAGAGAG
H13-327F1	gene35_masked_scaffold_5319	CGCAAGAGAAATGCCTGCGCA
H13-327F2	gene35_masked_scaffold_5319	CTGATCAAGTCCGCGCTCACT
H13-327R1	gene35_masked_scaffold_5319	AGGCGTGACTCCTTCGAAACC
H13-327R2	gene35_masked_scaffold_5319	GCAATTAGAGCATTGCAAACAGGC
H13-327R3	gene35_masked_scaffold_5319	GGAGAAGGTGTCCTTAAGAAGCAG
H13-328F1	gene35_masked_scaffold_5319	TGGTTTGGTGATCAAGAAGGGT
H13-328F2	gene35_masked_scaffold_5319	GGTATTCACAGCACGGTTTGAGT
H13-328R1	gene35_masked_scaffold_5319	TCCACCTGTTTCTTGATGCACAG
H13-328R2	gene35_masked_scaffold_5319	GTTTCTTGATGCACAGGCTGCATA
H13-329F1	gene36_masked_scaffold_5319	TGAGGAGGAATCTGCTTCTCCAG
H13-329F2	gene36_masked_scaffold_5319	GTCCAGCCCAGGTGAGGAGG
H13-329F3	gene36_masked_scaffold_5319	CCTCCGTTTCTGTCTGCTGCT
H13-329R1	gene36_masked_scaffold_5319	CAGGATCTGGCCCTCGAGGAAG
H13-329R2	gene36_masked_scaffold_5319	GAACCCTCCCTCCCCGAGG
H13-329R3	gene36_masked_scaffold_5319	CTCATACCCCTGCATGCTCT
H13-330F1	gene36_masked_scaffold_5319	TGCTACTATTAGGCAGGCAGACA
H13-330F2	gene36_masked_scaffold_5319	TTTCCAAAGGTTGCCGCAGCTA
H13-330R1	gene36_masked_scaffold_5319	GTTTGTACCATGGCTGGCACAA
H13-330R2	gene36_masked_scaffold_5319	AACCTATCAGGCATGTGTTTATGC
H13-331F1	gene36_masked_scaffold_5319	TGCCAGCCATGGTACAACTGAAG
H13-331F2	gene36_masked_scaffold_5319	TGTCCTTATGCCTGTGAACCACAG
H13-331R1	gene36_masked_scaffold_5319	TCACCGTGCTTAGTGCGGTG
H13-331R2	gene36_masked_scaffold_5319	CGCTGCCTCCACGTGGCGAC

Order name	Details	Sequence 5' to 3'
H13-331R3	gene36_masked_scaffold_5319	AGCAGAGCGTCCGTGGGCG
H13-332F1	gene52_masked_scaffold_5319	GATTCAAGCAATGCTACGACATC
H13-332F2	gene52_masked_scaffold_5319	AATGCTACGACATCTGCAGAGCAT
H13-332R	gene52_masked_scaffold_5319	AGTTGACACTGGAACACTAGATC
H13-333F1	gene52_masked_scaffold_5319	CCCTTGCCTGGGAGGCAAGG
H13-333F2	gene52_masked_scaffold_5319	GAAGACACTTCAATTTGCATTGGC
H13-333R1	gene52_masked_scaffold_5319	GCGAACAAATTTGTGAAAGTTCCA
H13-333R2	gene52_masked_scaffold_5319	TGAAGTTAGGTCAGCTAAGGCAT
H13-334F1	gene52_masked_scaffold_5319	TGCTGACATTGTGTCAACTAGTC
H13-334F2	gene52_masked_scaffold_5319	AGGGTGTATCCAAAAGGAATGAG
H13-334R1	gene52_masked_scaffold_5319	GCGATATGTCCAGGGTCTCATA
H13-334R2	gene52_masked_scaffold_5319	GCTTTGAATAAATTATCCGTGCG
H13-335F	gene52_masked_scaffold_5319	AGT GGA CGC AGC GGC GAC TCG
H13-335R	gene52_masked_scaffold_5319	GCGGCAGCCCATCACAGCCG
H13-336F1	gene56_masked_scaffold_5319	AATGACCAGGCATCCTCCTCCC
H13-336F2	gene56_masked_scaffold_5319	TCCCCTAATTTGTCCACATGGTGG
H13-336R1	gene56_masked_scaffold_5319	CATTGTCATCCTACTTGCAAAACG
H13-336R2	gene56_masked_scaffold_5319	TGACGATGGTCCTGGATGCA
H13-337F1	gene56_masked_scaffold_5319	GAGGACCCAACTCTTCCAAACTG
H13-337F2	gene56_masked_scaffold_5319	AACTGCACCAATCACCATCCACT
H13-337F3	gene56_masked_scaffold_5319	TTGGGCTGGAAAGTCAAGTACTA
H13-337R1	gene56_masked_scaffold_5319	GATTATAGCAAAATGTGATGGTC
H13-337R2	gene56_masked_scaffold_5319	CTCAAGGATATTGGAGTGCAGATTA
H13-338F1	gene56_masked_scaffold_5319	TTAATTCACCTCCCAGTTTGGAC
H13-338F2	gene56_masked_scaffold_5319	ACGACCCGCTATCCAAGAAGTGT
H13-338R1	gene56_masked_scaffold_5319	GTCTTTAGAAGATCGGGTAGCTG
H13-338R2	gene56_masked_scaffold_5319	GAACAGCTACACCTTTGTTGAT

Order name	Details	Sequence 5' to 3'
H13-339F	gene61_masked_scaffold_5319	TGAGCGGTATACACGGTACATC
H13-339R1	gene61_masked_scaffold_5319	AAACCTTGTCCCACCTGTACC
H13-339R2	gene61_masked_scaffold_5319	ACATTCTCTTTTGGTCGCCGT
H13-340F1	gene61_masked_scaffold_5319	CTACGACAGCCTACATGGGG
H13-340F2	gene61_masked_scaffold_5319	CCGAAAGTGATCCGTTACGG
H13-340R1	gene61_masked_scaffold_5319	CGGGCATAACTCCACACGA
H13-340R2	gene61_masked_scaffold_5319	CAGCTTCATCATGCACTTGGC
H13-341F	gene63_masked_scaffold_5319	ATTCCATGCTCCATGGTTCCAC
H13-341R1	gene63_masked_scaffold_5319	TACGTGCTACGCATGCTGACG
H13-341R2	gene63_masked_scaffold_5319	CTGACGGAGATGGCAAGCG
H13-342F1	gene65_masked_scaffold_5319	CACAGAAGGTGTCACGGCGT
H13-342F2	gene65_masked_scaffold_5319	CTCCGGGTGACCTGGCCGA
H13-342F3	gene65_masked_scaffold_5319	GCATAAGAAGGACGACCCACCA
H13-342R	gene65_masked_scaffold_5319	CGGGCTATGTCACCGTTCCTT
H13-343F1	gene65_masked_scaffold_5319	CAATGACCAGGCATCCTCCTCC
H13-343F2	gene65_masked_scaffold_5319	ACAACCTTCATCCTACGAGCTAG
H13-343F3	gene65_masked_scaffold_5319	CACATGGTGGTAGGGAAGCACAA
H13-343R1	gene65_masked_scaffold_5319	GAAGCTGCTGACATCCTCGACA
H13-343R2	gene65_masked_scaffold_5319	CGTGCCATGTATGAAGCTGCT
H13-344F1	gene65_masked_scaffold_5319	CCTCTTTAATCAGCCCATCATCTT
H13-344F2	gene65_masked_scaffold_5319	TGGAGGGCAGAGCTCCTCAAA
H13-344R1	gene65_masked_scaffold_5319	ATTGAGCCAGATGCAGAGTTTA
H13-344R2	gene65_masked_scaffold_5319	GATGAAATCATTGGCATGTGGATC
H13-345F	gene65_masked_scaffold_5319	TGCTAGCCAAGAAGTATTGTGAC
H13-345R	gene65_masked_scaffold_5319	TGACATCGCTGTAATCCAATTCA
H13-346F1	gene70_masked_scaffold_5319	TGAACCTTAGCACCCCTTAGTGACA
H13-346F2	gene70_masked_scaffold_5319	GCTTGCAAAGAAGTCCAGTCTAAA
H13-346R1	gene70_masked_scaffold_5319	AATATTGAGTTGACACTAGTAGCG

Order name	Details	Sequence 5' to 3'
H13-346R2	gene70_masked_scaffold_5319	CAAGAAAAACCCTCGTTATTGACC
H13-347R	GENE73-MASKED SCAFFOLD5319	TGTCGTGGGCATGGGAGGGA
H13-348F	GENE73-MASKED SCAFFOLD5319	CAAGTTCCAGAAGACGATAT
H13-349R	GENE73-MASKED SCAFFOLD5319	GCAAACCATACACGAGGTCT
H13-350F	GENE73-MASKED SCAFFOLD5319	GCAGAGCCAATCCTTGGG
H13-351F	GENE73-MASKED SCAFFOLD5319	TACTTGTCGTTAATAAACACGGAC
H13-351R	GENE73-MASKED SCAFFOLD5319	TTGGACAGATGTTATGGATAACGA
H13-352F	GENE73-MASKED SCAFFOLD5319	ATATAACTGCTTCCTCATCTGCAA
H13-352R	GENE73-MASKED SCAFFOLD5319	GTGTCTTCATGTCTTCATCCGT
H13-353	CLONTECH-UNIVERSAL PRIMER A	CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT
H13-354	CLONTECH-UNIVERSAL PRIMER SHORT	CTAATACGACTCACTATAGGGC
H13-355	305F1-Gene73 - 3' race primer-1	GATTACGCCAAGCTTATGACGATGGTCCTGGATGCC
H13-356	NEAR 348F-Gene73 - 3' race primer-2	GATTACGCCAAGCTTGACGAAGCCTTGAACAACAAGTTCC
H13-357	275F1-Gene73 - 3' race primer-3	GATTACGCCAAGCTTGGTCCTCTTTCTCAGCTCAAG
H13-358	NEAR 275F1-Gene73 - 3' race primer-4	GATTACGCCAAGCTTCAGCTCAAGCATCTTGAATTGGACG
H13-359	276F1-Gene73 -p 3' race primer-5	GATTACGCCAAGCTTAGCATATGCATCTGAAGGAGA
H13-360	NEAR 351R-Gene73 - 3' race primer-6	GATTACGCCAAGCTTCCATAACATCTGTCCAAGGTCTTG
H13-361	NEAR 352R-Gene73 - 3' race primer-7	GATTACGCCAAGCTTCTCACGGATGAAGACATGAAGACAC
H13-362	NEAR 305R1-Gene73 - 5' race primer-1	GATTACGCCAAGCTTGGTAGGGAAGCACCACCTTTCATCA
H13-363	321R2-Gene73 - 5' race primer-2	GATTACGCCAAGCTTGCTATAATTTGCAACCCAATGTCCT
H13-364	NEAR 321R2-Gene73 - 5' race primer-3	GATTACGCCAAGCTTGAGCATATCAATGTCACGTCCATCAT
H13-365	276R2-Gene73 - 5' race primer-4	GATTACGCCAAGCTTTCTCACGGTCCCAGTTTGGC
H13-366	275R2-Gene73 - 5' race primer-5	GATTACGCCAAGCTTGTGAGCAGTGCAAGGCTACAA
H13-367	274R2-Gene73 - 5' race primer-6	GATTACGCCAAGCTTCCTGGGCTGGAACCCAGT
H13-368	M13-F-20	GTAACGACGGCCAGT
H13-369	M13-R-27	CAGGAAACAGCTATGAC
H13-370	M13-F-41	CGCCAGGGTTTCCCAGTCACGAC

Order name	Details	Sequence 5' to 3'
H13-371	M13-R-48	AGCGGATAACAATTTACACAGG
H13-372	First 26 bases from start codon	ATGACGATGGTCCTGGACGCCTTTGC
H13-373	First 30 bases from start codon	ATGACGATGGTCCTGGACGCCTTTGCATCC
H13-374	15bp upstream of start codon	CAAAAGGAAGCGGAAATGACGATGGTC
H13-375	15bp upstream of start codon	CAAAAGGAAGCGGAAATGACGATGGTCCTGG
H13-376	424bp upstream of start codon	ACCCTCTCTCTCAATCATTTCATCGCTCC
H13-377	408bp upstream of start codon	ATTCATCGCTCCGCCCCACTCCTCACT
H13-378	500bp upstream of start codon	GGGCCGCAAACCCACCGGCGGC
H13-379	485bp upstream of start codon	ACCGGCGGCGGCTCCGACAGCA
H13-380	470bp upstream of start codon	GACAGCACCATGGCGCTGCGCT
H13-381	445bp upstream of start codon	TGTCGAGAGGCTTTCCGCGCGGTGT
H13-382	~ 541bp upstream of start codon	AGCGCACTGTGTACTAATTGGCCT
H13-383	370bp upstream of start codon	CAGGCGAAAGAAGCGCAAGAAGT
H13-384	245bp upstream of start codon	AGGCGTGATCTGGAGAGGATGTTGA
H13-385	at stop codon	TTAGGTAATGATATCATGGTCACAATCACGT
H13-386	3'UTR	GCTCCATTCCGACAGAACAATGAAGGCT
H13-387	3'UTR	AGTGCAAACCTACACACATAACACTAGGC
H13-388	3'UTR	CGACACATAAAAGCCTTACTCCACATAC
H13-389	3'UTR	CTCCACATACATAAAGTGCAAACCTACACAC
H13-390	3'UTR	CAAGATGTGGCAAAAAGAAATCCTGT
H13-391	last exon?	GACATGCTGGTCAAGAAGTATTGTGAC
H13-392	last exon?	CAACTCCAGATTCGGATGACATGCTG
H13-393	3'UTR	GATTCATGCTGCAGCAACAAAGTGCTG
H13-394	3'UTR	CATCGTGATTTTCTCGACACATAAAAGC
H13-395	3'UTR	GACAGAACAATGAAGGCTTCCCTGCCAC
H13-396	last exon?	TGGGGCTCGGCATGGCCCCATCC
H13-397	last exon?	CTAACTGGCGCACTTCCTTGAGCT
H13-398	3'UTR	AAGTCTGCACCCAACCGACGGATGG

Order name	Details	Sequence 5' to 3'
H13-399	3'UTR	GTCTCCGAGCGGCTACCCTACATCG
H13-400	3rd exon end	CTCGGCATGGCCCCATCCTCCTCC
H13-401	for cloning	CTAATACGACTCACTATGACGATGGTCCTGGACGCCTTTGC
H13-402	for cloning	CTAATACGACTCACTATGACGATGGTCCTGGACGCCTTTGCATCC
H13-403	for cloning	CTAATACGACTCACTCAAAAGGAAGCGGAAATGACGATGGTC
H13-404	for cloning	CTAATACGACTCACTCAAAAGGAAGCGGAAATGACGATGGTCCTGG
H13-405	for cloning	CTAATACGACTCACTACCCTCTCTCTCAATCATTCATCGCTCC
H13-406	for cloning	CTAATACGACTCACTATTTCATCGCTCCGCCCACTCCTCACT
H13-407	for cloning	CTAATACGACTCACTGGGCCGCAAACCCCAACGGCGGC
H13-408	for cloning	CTAATACGACTCACTACCGGCGGCGGCTCCGACAGCA
H13-409	for cloning	CTAATACGACTCACTGACAGCACCATGGCGCTGCGCT
H13-410	for cloning	CTAATACGACTCACTTGTCGAGAGGCTTTCCGGCCGGTGT
H13-411	for cloning	CTAATACGACTCACTAGCGCACTGTGTACTAATTGGCCT
H13-412	for cloning	CTAATACGACTCACTCAGGCGAAAGAAGCGCAAGAAGT
H13-413	for cloning	CTAATACGACTCACTAGGCGTGATCTGGAGAGGATGTTGA
H13-414	for cloning	GATTACGCCAAGCTTTTAGGTAATGATATCATGGTCACAATCACGT
H13-415	for cloning	GATTACGCCAAGCTTGCTCCATTCCGACAGAACAATGAAGGCT
H13-416	for cloning	GATTACGCCAAGCTTAGTGCAAACCTACACACATAACACTAGGC
H13-417	for cloning	GATTACGCCAAGCTTCGACACATAAAAGCCTTACTCCACATAC
H13-418	for cloning	GATTACGCCAAGCTTCTCCACATACATAAAGTGCAAACCTACACAC
H13-419	for cloning	GATTACGCCAAGCTTCAAGATGTGGCAAAAAGAAATCCTGT
H13-420	for cloning	GATTACGCCAAGCTTGACATGCTGGTCAAGAAGTATTGTGAC
H13-421	for cloning	GATTACGCCAAGCTTCAACTCCAGATTTCGGATGACATGCTG
H13-422	for cloning	GATTACGCCAAGCTTGATTCATGCTGCAGCAACAAAGTGCTG
H13-423	for cloning	GATTACGCCAAGCTTCATCGTGATTTTCTCGACACATAAAAGC
H13-424	for cloning	GATTACGCCAAGCTTGACAGAACAAATGAAGGCTTCCCTGCCAC
H13-425	for cloning	GATTACGCCAAGCTTTGGGGCTCGGCATGGCCCCATCC

Order name	Details	Sequence 5' to 3'
H13-426	for cloning	GATTACGCCAAGCTTCTAACTGGCGCACTTCCTTGGAGCT
H13-427	for cloning	GATTACGCCAAGCTTAAGTCTGCACCCAACCGACGGATGG
H13-428	for cloning	GATTACGCCAAGCTTGTCTCCGAGCGGCTACCCTACATCG
H13-429	for cloning	GATTACGCCAAGCTTCTCGGCATGGCCCCATCCTCCTCC
H13-430	M13-F-20	GTAAAACGACGGCCAGT
H13-431	M13-R-27	CAGGAAACAGCTATGAC
H13-432	M13-F-41	CGCCAGGGTTTCCCAGTCACGAC
H13-433	M13-R-48	AGCGGATAACAATTCACACAGG
H13-434	Gene73-N-start-A-F	ATGACGATGGTCCTGGACGCCTTTGC
H13-435	Gene73-N-start-B-F	ATGACGATGGTCCTGGACGCCTTTG
H13-436	Gene73-N-Up485-F	GGAGCACGTTGCAGTTGGCGAATAGTG
H13-437	Gene73-N-Up424-F	ACCCTCTCTCTCAATCATTCATCGCTCC
H13-438	Gene73-N-stop-R	TCAGTTAACGATCTCTTCATCTGCTCTTCT
H13-439	Gene73-N-3'UTR-A	ATAATGTGTTGTATCAGTTCCTTTCTTTTCG
H13-440	Gene73-N-3'UTR-B	ATTAGACAGACACTTTCAGCCATGCTCAAG
H13-441	Gene73-N-3'UTR-C	ACTACAGAGAGTCTCCATGCATCTTCTAC
H13-442	Gene73-N-3'UTR-D	GTACACTGACTGGGAACTACAGAGAGTC
H13-443	TAG-H13-434	CTAATACGACTCACTATGACGATGGTCCTGGACGCCTTTGC
H13-444	TAG-H13-435	CTAATACGACTCACTATGACGATGGTCCTGGACGCCTTTG
H13-445	TAG-H13-436	CTAATACGACTCACTGGAGCACGTTGCAGTTGGCGAATAGTG
H13-446	TAG-H13-437	CTAATACGACTCACTACCCTCTCTCTCAATCATTCATCGCTCC
H13-447	TAG-H13-438	GATTACGCCAAGCTTTCAGTTAACGATCTCTTCATCTGCTCTTCT
H13-448	TAG-H13-439	GATTACGCCAAGCTTATAATGTGTTGTATCAGTTCCTTTCTTTTCG
H13-449	TAG-H13-440	GATTACGCCAAGCTTATTAGACAGACACTTTCAGCCATGCTCAAG
H13-450	TAG-H13-441	GATTACGCCAAGCTTACTACAGAGAGTCTCCATGCATCTTCTAC
H13-451	TAG-H13-442	GATTACGCCAAGCTTGTACACTGACTGGGAACTACAGAGAGTC
H13-452	Gene73-3RD EXON-F	TAGATATGGAGGAGGATGGGGCCAT
H13-453	in the seq reads of 274F2 primer, for covering to get overlap	GGTTAGACTGGAGTTCTTTGCAGGCACA

Order name	Details	Sequence 5' to 3'
H13-454	in the seq reads of 274F2 primer, for covering to get overlap	GTAGACTTAGTGACATGAGTTTCTTCGG
H13-455	in the seq reads of 274F2 primer, for covering to get overlap	GGATGCAAACAATTTGTGAAAGTTCCGCA
H13-456	in the seq reads of 275F1 primer, for covering to get overlap	TTACAGGAATGGTGAATGGGAGGAGTG
H13-457	5'UTR – F	GTTTGCTTGCTTCTATCGATCCGGTATCAG
H13-458	1st exon start- F	ATGACGATGGTCCTGGACGCCTTTG
H13-459	1st exon end –R	CTGTTTCTGGAGCAATGACCAGGCATCC
H13-460	2nd exon start – R	GTCACGTCCATCATTCTCACTTGAGATTAT
H13-461	1st exon end – F	AGGATGCCTGGTCATTGCTCCAGAAACAG
H13-462	2nd exon start- F	ATAATCTCAAGTGAGAATGATGGACGTGACAT
H13-463	2nd exon middle1 – F	GCACTAGCAGCTCACAGTGGGGAAAC
H13-464	2nd exon middle2 – F	GGTGTAGTTTGGAAGAGTTGGGTCCTC
H13-465	2nd exon middle3 – F	GCATGAGTTGGAATTTACAGGAATGGTGG
H13-466	2nd exon-end- R	CTCAGGACATGAAAGAACGGCTGACA
H13-467	3rd exon - start- R	CGTTGTGTGTCCTCAAAAGAGGTTAAGGTTT
H13-468	2nd exon end –F	CAAATGTCAGCCGTTCTTTTCATGTCCTGAG
H13-469	3rd exon start – F	GAACCTTAACCTCTTTTGAGGACACACAACG
H13-470	3rd exon middle – F	TGGGGACGGATAAGCAGTATGAACATG
H13-471	3rd exon end – F	CAGTTAGCAACAGAGGATGCTGTGAG
H13-472	4th exon end – R	CTAGAATCAAGATGTGGCAGAAAGAAATCCTG

2.8 References

- Aggarwal et al., 2014. Avirulence effector discovery in a plant galling and plant parasitic arthropod, the Hessian fly (*Mayetiola destructor*). PLoS One, 9: e100958.
- Alexandratos, N. and J. Bruinsma. 2012. World agriculture towards 2030/2050: the 2012 revision. ESA Working paper No. 12-03. Rome, FAO
- Allan et al., 1959. Genetic analysis of ten sources of Hessian fly resistance, their interrelationships, and association with leaf rust reaction in wheat. Kansas Agri Exp Station Tech Bull. 104: 54.
- Amri et al., 1990. Chromosomal location of the Hessian fly resistance gene H20 in 'Jori' durum wheat. J Heredity, 81: 71–72.
- Amri et al., 1990. Complementary action of genes for Hessian fly resistance in the wheat cultivar 'Seneca.' J Heredity, 81: 1224–1227.
- Barnes, H.F. 1956. Gall midges of economic importance: Vol. VII. Gall midges of cereal crops. Crosby Lockwood & Son, London.
- Bossolini et al., 2007. Comparison of orthologous loci from small grass genomes Brachypodium and rice: implications for wheat genomics and grass genome annotation. Plant J, 49: 704–717.
- Buntin GD, Chapin JW. 1990. Biology of Hessian fly (Diptera:Cecidomyiidae) in the southeastern United States: geographic variation and temperature-dependent phenology. J Econ Entomol, 83: 1015-1024.
- Buntin GD. 1999. Hessian fly (Diptera: Cecidomyiidae) injury and loss of winter wheat grain yield and quality. J Econ Entomol., 92: 1190–1197.
- Byers RA and Gallun RL. 1971. The ability of the Hessian fly to stunt winter wheat. I. Effect of larval feeding on the elongation of leaves. J Econ Entomol., 65: 955-958.
- Cainong et al., 2010. Wheat-rye T2BS2BL-2RL recombinants with resistance to Hessian fly (H21). Crop Sci, 50: 920–925.
- Caldwell et al., 1946. RM, Cartwright WB, Compton LE: Inheritance of Hessian fly resistance derived from W38 and durum P.I.94587. J Am Soc Agro, 38: 398–409.
- Caldwell et al., 1957. Utilization of the W38 resistance of wheat to Hessian fly. Agron J, 49: 520.
- Cambron et al., 2010. Virulence in Hessian fly (Diptera: Cecidomyiidae) field collections from the Southeastern United States to 21 resistance genes in wheat. J Econ Entomol, 103: 2229–2235.
- Cebert et al., 1996. Genetic analysis of Hessian fly resistance in durum wheat. Agro Abstracts, 88: 88.

- Chen et al., 2009a. Registration of ‘Cataldo’ wheat. *J Plant Regist*, 3: 264–268.
- Chen et al., 2009b. Virulence analysis of Hessian fly populations from Texas, Oklahoma, and Kansas. *J Econ Entomol*, 102: 774–780.
- Cheng X, Zhu L, He G. 2013. Towards understanding of molecular interactions between rice and the brown planthopper. *Mol Plant*, 6: 621–634.
- Choi Y. et al., 2015. PROVEAN we server: a toll to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*, 31(16): 2745–2747.
- Cloutier et al., 2007. Leaf rust resistance gene *Lr1*, isolated from bread wheat (*Triticum aestivum* L.) is a member of the large *psr567* gene family. *Plant Mol. Biol.*, 65: 93–106.
- Cooley et al., 2000. Members of the Arabidopsis HRT/RPP8 family of resistance genes confer resistance to both viral and oomycete pathogens. *The Plant Cell*, 12: 663–676.
- Delibes et al., 1997. Hessian fly-resistance gene transferred from chromosome 4Mv of *Aegilops ventricosa* to *Triticum aestivum*. *Theor App Gen*, 94: 858–864.
- DeYoung BJ, Innes RW. 2006. Plant NBS-LRR proteins in pathogen sensing and host defense. *Nat Immunol*, 7: 1243–1249.
- Divya et al. 2015. Cloning of a gene encoding LRR protein and its validation as candidate gall midge resistance gene, *Gm4*, in rice. *Euphytica*. 203: 185–195.
- Dogimont et al. 2014. The *Vat* locus encodes for a CC-NBS-LRR protein that confers resistance to *Aphis gossypii* infestation and *A. gossypii*-mediated virus resistance. *Plant J*. 80: 6: 993–1004.
- Du et al., 2009. Identification and characterization of *Bph14*, a gene conferring resistance to brown planthopper in rice. *Proc. Natl Acad. Sci.*, 106: 22163–22168.
- Dubcovsky et al., 2001. Comparative sequence analysis of colinear barley and rice bacterial artificial chromosomes. *Plant Physiol.*, 125: 1342–1353.
- El Bouhssini M, Hatchett JH, Wilde GE. 1999. Hessian fly (Diptera: Cecidomyiidae) larval survival as affected by wheat resistance alleles, temperature and larval density. *J Agric Urban Entomol*, 16: 245–254.
- Ellis et al., 1997. Advances in the molecular genetic analysis of the flax–flax rust interaction. *Annu. Rev. Phytopathol.* 35: 271–291.
- Ellis JG, Lawrence GJ, Luck JE, Dodds PN. 1999. Identification of regions in alleles of the flax rust resistance gene *L* that determine differences in gene-for-gene specificity. *Plant Cell*, 11: 495–506.
- Endo TR, Gill BS. 1996. The deletion stocks of common wheat. *J Hered*, 87: 295–307.
- Endo TR. 1988. Induction of chromosomal structural changes by a chromosome of *Aegilops cylindrica* L. in common wheat. *J Hered*, 79: 366–370.

- F. Gould. 1998. Sustainability of transgenic insecticidal cultivars: Integrating pest genetics and ecology. *Annu. Rev. Entomol.*, 43: 701–726.
- Faris et al., 2008. Micro-colinearity between rice, *Brachypodium*, and *Triticum monococcum* at the wheat domestication locus Q. *Funct Integr Genomics*, 8: 149–164.
- Feuillet et al., 2003. Map-based isolation of the leaf rust disease resistance gene Lr10 from the hexaploid wheat (*Triticum aestivum* L.) genome. *Proc. Natl Acad. Sci.*, 100: 15253–15258.
- Flanders KL, Buntin GD, Mask PL. 2008. Biology and management of Hessian fly in wheat. *Ala. Coop. Ext. Serv. Bull. ANR*, 1069: 4.
- Foster et al., 1991. Effectiveness of deploying single gene resistances in wheat for controlling damage by the Hessian fly (Diptera: Cecidomyiidae). *Environ. Entomol.*, 20: 964–969.
- Friebe et al., 1991. Transfer of Hessian fly resistance from rye to wheat via radiation-induced terminal and intercalary chromosome translocations. *Theor App Gen*, 83: 33–40.
- Gagne RJ. 1989. *The Plant-Feeding Gall Midges of North America*. Ithaca, NY: Cornell Univ. Press. 340 pp.
- Gagne RJ. 1994. *The Gall Midges of the Neotropical Region*. Ithaca, NY: Comstock Publ. Assoc. 352 pp
- Gagne RJ. 2010. *Update for a Catalog of the Cecidomyiidae (Diptera) of the World*. Washington, DC: Entomol. Soc. Wash. 544 pp. http://www.ars.usda.gov/SP2UserFiles/Place/12754100/Gagne_2010_World_Catalog_Cecidomyiidae.pdf
- Gallun RL, Patterson FL 1977. Monosomic analysis of wheat for resistance to Hessian fly. *J Heredit*, 68: 223–226.
- Gill BS, Hatchett JH, Raupp WJ. 1987. Chromosomal mapping of Hessian fly resistance gene *H13* in the D genome of wheat. *J Heredity*, 78: 97–100.
- Gill et al., 1991. Registration of KS89WGRC3 and KS89WGRC6 Hessian fly-resistant hard red winter wheat germplasm. *Crop Sci*, 31: 245.
- Givry et al., 2005. CARTHAGENE: multipopulation integrated genetic and radiated hybrid mapping. *Bioinformatics*, 21(8): 1703–1704.
- Grant et al., 1995. Structure of the Arabidopsis RPM1 gene enabling dual specificity disease resistance. *Science*, 269: 843–846.
- Greene et al., 2003. Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in Arabidopsis. *Genetics* 164, 731–740.
- Griffiths et al., 2006. Molecular characterization of Ph1 as a major chromosome pairing locus in polyploid wheat. *Nature*, 439: 749–752.

- Grube, R.C., Radwanski, E.R. and Jahn, M. 2000. Comparative genetics of disease resistance within the Solanaceae. *Genetics*, 155: 873–887.
- Hall et al., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium series* 41: 95-98.
- Halterman, D.A. and Wise, R.P. 2004. A single-amino acid substitution in the sixth leucinerich repeat of barley MLA6 and MLA13 alleviates dependence on RAR1 for disease resistance signaling. *The Plant Journal*, 38: 215–226.
- Harris MO, Sandanayake M, Griffin W. 2001. Oviposition preferences of the Hessian fly and their consequences for the survival and reproductive potential of offspring. *Ecol. Entomol.* 26:1–14
- Harris MO, Stuart JJ, Mohan M, Nair S, Lamb RJ, Rohfritsch O. 2003. Grasses and gall midges: plant defense and insect adaptation. *Annu. Rev. Entomol.* 48:549–77
- Harris MO, Stuart JJ, Mohan M, Nair S, Lamb RJ, Rohfritsch O. 2003. Grasses and gall midges: plant defense and insect adaptation. *Annu. Rev. Entomol.* 48:549–77
- Hatchett JH, Gill BS. 1981. D-genome sources of resistance in *Triticum tauschii* to Hessian fly. *J Hered*, 72: 126–127.
- Hatchett JH, Kreitner GL, Elzinga RJ. 1990. Larval mouthparts and feeding mechanism of the Hessian fly (Diptera: Cecidomyiidae). *Ann. Entomol. Soc. Am.* 83:1137–47
- Hatchett JH, Martin TJ, Livers RW. 1981. Expression and inheritance of resistance to Hessian fly in hexaploid wheats derived from *Triticum tauschii* (Coss.) Schmal. *Crop Sci*, 21: 731–734.
- Henry et al., 2014. Efficient genome-wide detection and cataloging of EMS-induced mutations using exome capture and next-generation sequencing. *Plant Cell*. 26(4):1382-1397.
- <http://www.fao.org/resources/infographics/infographics-details/en/c/240943/>
- Huang et al., 2003. Map-based cloning of leaf rust resistance gene Lr21 from the large and polyploid genome of bread wheat. *Genetics*, 164: 655–664.
- Hurni et al., 2013. Rye Pm8 and wheat Pm3 are orthologous genes and show evolutionary conservation of resistance function against powdery mildew. *Plant J.*, 76: 957–969.
- IciMapping 3.3: <https://www.integratedbreeding.net/supplementary-toolbox/geneticmapping-and-qt/ici-mapping>
- Ji et al., 2016. Map-based cloning and characterization of the *BPH18* gene from wild rice Conferring Resistance to Brown Planthopper (BPH) Insect Pest. *Sci Rep*, 6: 34376.
- Johnson J, Buck J, Buntin G, Chen Z. 2009. Contributions by Georgia experiment station, University of Georgia. *Annu Wheat Newsl*, 55: 196–198.
- Jones and Dangl .2006. The plant immune system. *Nature*, 444: 323–329.

- Jones JDG, Dangl JL. 2006. The plant immune system. *Nature* 444:323–29
- Keller, B., and C. Feuillet, 2000. Colinearity and gene density in grass genomes. *Trends Plant Sci.*, 5: 246–251.
- Krasileva et al., 2017. Uncovering hidden variation in polyploid wheat. *PNAS*. 114, 6: E913-E921.
- Leister et al., 1998 genomes. Rapid reorganization of resistance gene homologues in cereal genomes. *Proc. Natl. Acad. Sci.*, 95: 370–375.
- Leroy P. et al., 2012. TriAnnot: a versatile and high-performance pipeline for the automated annotation of plant genomes. *Front Plant Sci.*, 3: 5.
- Li, W., and B. S. Gill. 2002. The colinearity of the Sh2/A1 orthologous region in rice sorghum and maize is interrupted and accompanied by genome expansion in the Triticeae. *Genetics*, 160: 1153–1162.
- Liu et al., 2005. Genetic characterization and molecular mapping of a Hessian fly-resistance gene transferred from *T. turgidum* ssp. *dicoccum* to common wheat. *Theor Appl Genet*, 111: 1308–1315.
- Liu et al., 2005. H9, H10, and H11 compose a cluster of Hessian fly resistance genes in the distal gene-rich region of wheat chromosome 1AS. *Theor Appl Genet*, 110: 143–148.
- Liu et al., 2014. The Stripe Rust Resistance Gene Yr10 Encodes an Evolutionary-Conserved and Unique CC–NBS–LRR Sequence in Wheat. *Molecular Plant*, 7: 1740–1755.
- Liu et al., 2015. A gene cluster encoding lectin receptor kinase confers broad-spectrum and durable insect resistance in rice. *Nat Biotechnol*, 33: 301–305.
- Liu X, Gill B, Chen MS. 2005b. Hessian fly resistance gene H13 is mapped to a distal cluster of resistance genes in chromosome 6DS of wheat. *Theor Appl Genet*, 111: 243–249.
- Liu ZJ, Zhu J, Cui Y, Liang Y, Wu HB, et al. 2012. Identification and comparative mapping of a powdery mildew resistance gene derived from wild emmer (*Triticum turgidum* var. *dicoccoides*) on chromosome 2BS. *Theor Appl Genet* 124: 1041–1049.
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using realtime quantitative PCR and the $2^{-\Delta\Delta C(T)}$ Method. *Methods* 25(4): 402–408.
- Lu HJ, Faris JD. 2006. Macro- and micro-colinearity between the genomic region of wheat chromosome 5B containing the *Tsn1* gene and the rice genome. *Funct Integr Genomics*, 6: 90–103.
- Luo et al., 2013. A 4-gigabase physical map unlocks the structure and evolution of the complex genome of *Aegilops tauschii*, the wheat D-genome progenitor. *Proc. Natl Acad. Sci.*, 110(19): 7940–7945.
- Luo et al., 2017. Genome sequence of the progenitor of the wheat D genome *Aegilops tauschii*. *Nature* 551: 498–502.

- Maas FB, Patterson FL, Foster JE, Ohm HW. 1989. Expression and inheritance of resistance of ELS6404-160 durum wheat to Hessian fly. *Crop Sci*, 29: 23–28.
- Mamaev BM. 1975. *Evolution of Gall Forming Insects: Gall Midges*. Wetherby, UK: Br. Libr. 317 pp
- Marchler-Bauer et al. 2015. CDD: NCBI's conserved domain database. *Nucleic Acids Res.*, 43(D): 222-6.
- Martin, T. J., T. L. Harvey, and J. H. Hatchett. 1982. Registration of greenbug and Hessian fly resistant wheat germplasm. *Crop Sci.*, 22: 1089.
- Martin-Sanchez et al., 2003. A new Hessian fly resistance gene (H30) transferred from the wild grass *Aegilops triuncialis* to hexaploid wheat. *Theor App Gen*, 106: 1248–1255.
- McIntosh et al., 2012. Catalogue of gene symbols. Mac-Gene 2012. Mishima, Shizuoka, Japan: National Institution of Genetics.
- Michelmore and Meyers, 1998. Clusters of resistance genes in plants evolve by divergent selection and a birth-and death process. *Genome Res*. 8: 1113–1130.
- Miranda et al., 2010. Genetic mapping of an *Aegilops tauschii*-derived Hessian Fly Resistance Gene in Common Wheat. *Crop science* 50: 612-616.
- Morse et al., 2005. Determining suitable internal standards for mRNA quantification of increasing cancer progression in human breast cells by real-time reverse transcriptase polymerase chain reaction. *Anal Biochem*, 342(1): 69–77.
- Naber N, El Bouhssini M, Lhaloui S. 2003. Biotypes of Hessian fly (Diptera; Cecidomyiidae) in Morocco. *J Appl Entomol*, 127: 174–176.
- Noble WB, Suneson JA: Differentiation of the two genetic factors for resistance to Hessian fly in Dawson wheat. *J Agri Res* 1943, 67:27–32.
- Nombela et al. 2003. The root-knot nematode resistance gene *Mi-1.2* of tomato is responsible for resistance against the whitefly *Bemisia tabaci*. *Mol. Plant Microbe Interact*. 16: 645–649.
- Obanni M, Ohm HW, Foster JE, Patterson FL. 1989. Genetics of resistance of PI 422297 durum wheat to the Hessian fly. *Crop Sci*, 29: 249–252.
- Obanni M, Patterson FL, Foster JE, Ohm HW. 1988. Genetic analyses of resistance of durum wheat PI 428435 to the Hessian fly. *Crop Sci*, 28: 223–226.
- Oellermann et al., 1983. Inheritance of resistance in Luso wheat to Hessian fly. *Crop Sci*, 23: 221–224.
- Patterson et al., 1994. Registration of eight Hessian fly resistant common winter wheat germplasm lines (Carol, Erin, Flynn, Iris, Joy, Karen, Lola, and Molly). *Crop Sci.*, 34: 315—316.

- Patterson FL, Foster JE, Ohm HW. 1988. Gene H16 in wheat for resistance to Hessian fly. *Crop Sci*, 28: 652–654.
- Pedigo, L.P. 2002. *Entomology and pest management*. 3rd ed. Prentice Hall, Englewood Cliffs, NJ.
- Periyannan et al., 2013. The gene Sr33, an ortholog of barley Mla genes, encodes resistance to wheat stem rust race Ug99. *Science*, 341(6147): 786-788.
- Porter et al., 2009. Insects which challenge global wheat production: 189–201. In Carver, B. (ed.) *Wheat: Science and trade*. Blackwell Publishing, Ames, IA.
- Price PW. 2005. Adaptive radiation of gall-inducing insects. *Basic Appl. Ecol.* 6:413–21
- Ratcliffe et al., 2000. Biotype composition of Hessian fly (Diptera: Cecidomyiidae) populations from the southeastern, midwestern, and northwestern United States and virulence to resistance genes in wheat. *J Econ Entomol*, 93: 1319–1328.
- Ratcliffe R, Hatchett J. 1997. Biology and genetics of the Hessian fly and resistance in wheat. In: Bondari K (ed) *New developments in entomology*. Res Signpost, Sci Inf Guild, Trivandrum: 47–56.
- Ratcliffe RH. 2012. Breeding for Hessian fly resistance in wheat. In: Radcliffe EB, Hutchison WD (eds), *Radcliffe's IPM world textbook*, University of Minnesota, St. Paul <http://ipmworld.umn.edu>. Accessed 6 Aug 2012
- Ren et al., 2016. *Bph32*, a novel gene encoding an unknown SCR domain containing protein, confers resistance against the brown planthopper in rice. *Sci Rep*, 6: 37645.
- Richter et al., 1995. R New rust resistance specificities associated with recombination in the Rp1 complex in maize. *Genetics* 141: 373–381.
- Roskam HC. 2005. Phylogeny of gall midges (Cecidomyiidae). See Reference 89, pp. 307–19
- Rossi et al. 1998. The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. *PNAS*. 95: 9750–9754.
- Saintenac et al., 2013. Identification of wheat gene Sr35 that confers resistance to Ug99 stem rust race group. *Science*, 341: 783–786.
- Saitou and Nei 1987. The Neighbor-joining Method: A new method for reconstructing phylogenetic trees. *Mol. Bio. Evol.* 4(4): 406-425.
- Sama et al. 2014. A putative candidate for the recessive gall midge resistance gene *gm3* in rice identified and validated. *Theor. Appl. Genet.* 127: 113–124.
- Sardesai et al., 2005. Identification and mapping of H32, a new wheat gene conferring resistance to Hessian fly. *Theor App Gen* 2005, 111: 1167–1173.

- Sears ER, Sears LMS. 1978. The telocentric chromosomes of common wheat. In: Ramanujam S (ed) Proceedings of the 5th International Wheat Genetics Symposium: 389-407.
- Sears ER. 1954. The aneuploids of common wheat. University of Missouri Agricultural Experiment Station, Research Bulletin, 572: 1-58.
- Sears ER. 1966. Chromosome mapping with the aid of telocentrics. Proceedings of the 2nd International Wheat Genetics Symposium: 370-381.
- Sehgal et al., 2012. P0027 Sequence-ready, Physical Maps of Chromosomes 1D, 4D, and 6D of Hexaploid Wheat, PAGXX
- Shands RG, Cartwright WB. 1953. A fifth gene conditioning Hessian fly response in common wheat. *Agro J*, 45: 302–307.
- Singh 2017. Genetic diversity of wheat wild relative, *Aegilops tauschii*, for wheat improvement. (Ph.D. thesis) Kansas State University, Manhattan, KS, USA
- Smit et al., 2013-2015. RepeatMasker Open-4.0
- Solovyev et al., 2006. Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biology*, 7 Suppl1:S10.1-12.
- Sorrells et al., 2003. Comparative DNA sequence analysis of wheat and rice genomes. *Genome Res*, 13: 1818–1827.
- Stebbins NB, Patterson FL, Gallun RL 1983. Inheritance of resistance of PI 94587 wheat to biotypes B and D of Hessian fly. *Crop Sci*, 23: 251–253.
- Stebbins NB, Patterson FL, Gallun RL. 1982. Interrelationships among wheat genes H3, H6, H9, and H10 for Hessian fly resistance. *Crop Sci*, 22: 1029–1032.
- Steuernagel et al., 2016. Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nature Biotechnology*, 34: 652-655.
- Stuart et al. 2012. Gall Midges (Hessian Flies) as Plant Pathogens. *Annual Review of Phytopathology*. 50(1): 339-357.
- Stuart JJ, Hatchett JH. 1987. Morphogenesis and cytology of the salivary gland of the Hessian fly, *Mayetiola destructor* (Diptera: Cecidomyiidae). *Ann. Entomol. Soc. Am.* 80:475–82
- Suneson JA, Noble WB. 1950.: Further differentiation of genetic factors in wheat for resistance to the Hessian fly. *US Dep Agric Tech Bull*, 1004: 8
- Takahashi et al., 2002. RCY1, an *Arabidopsis thaliana* RPP8/HRT family resistance gene, conferring resistance to cucumber mosaic virus requires salicylic acid, ethylene and a novel signal transduction mechanism. *The Plant Journal*, 32: 655–667.

- Tamura et al., 2014. Map-based cloning and characterization of a brown planthopper resistance gene *BPH26* from *Oryza sativa* L. ssp. indica cultivar ADR52. *Sci Rep*, 4: 5872.
- Tanaka M. 1961. New amphidiploids, synthesized 6X-wheats, derived from Emmer wheat · *Aegilops squarrosa*. *Wheat Info Serv*, 12: 11.
- Thind et al., 2017. Rapid cloning of genes in hexaploid wheat using cultivar-specific long-range chromosome assembly. *Nature Biotechnology*, 35: 793–796.
- Tikhonov et al., 1999. Colinearity and its exceptions in orthologous *adh* regions of maize and sorghum. *Proc. Natl. Acad. Sci.*, 96: 7409–7414.
- Tiwari et al., 2012. Endosperm tolerance of paternal aneuploidy allows radiation hybrid mapping of the wheat D-genome and a measure of c ray-induced chromosome breaks. *PLoS ONE*, 7: e48815.
- Tiwari et al., 2016. A whole-genome, radiation hybrid mapping resource of hexaploid wheat. *The Plant Journal*, 86: 195–207.
- Torto-Alalibo TA, Collmer CW, Lindeberg M, Bird D, Collmer A, Tyler BM. 2009. Common and contrasting themes in host cell–targeted effectors from bacterial, fungal, oomycete and nematode plant symbionts described using the gene ontology. *BMC Microbiol.* 9:S3
- Valarik M, Linkiewicz AM, Dubcovsky J. 2006. A microcolinearity study at the earliness per se gene *Eps-Am1* region reveals an ancient duplication that preceded the wheat-rice divergence. *Theor Appl Genet*, 112: 945–957.
- Wang et al., 2006. Genetic characterization and molecular mapping of Hessian fly resistance genes derived from *Aegilops tauschii* in synthetic wheat. *Theor App Gen*, 113: 611–618.
- Wang et al., 2015. Map-based cloning and characterization of *BPH29*, a B3 domain-containing recessive gene conferring brown planthopper resistance in rice. *J Exp Bot*, 66: 6035–6045.
- Williams et al., 2003. Phenotypic assessment and mapped markers for H31, a new wheat gene conferring resistance to Hessian fly (Diptera: Cecidomyiidae). *Theor App Gen*, 107: 1516–1523.
- Yahiaoui et al., 2004. Genome analysis at different ploidy levels allows cloning of the powdery mildew resistance gene *Pm3b* from hexaploid wheat. *Plant J.*, 37: 528–538.
- Zhao et al., 2016a. Allelic diversity in an NLR gene *BPH9* enables rice to combat planthopper variation. *Proc Natl Acad Sci*, 113: 12850–12855.
- Zhao H, Liu X, Chen M. 2006. *H22*, a major resistance gene to the Hessian fly (*Mayetiola destructor*), is mapped to the distal region of wheat chromosome 1DS. *Theor App Gen*, 113: 1491–1496.