GENETIC AND GENOMIC STUDIES ON WHEAT PRE-HARVEST SPROUTING RESISTANCE

by

MENG LIN

B.S., Huazhong Agricultural University, 2010

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Agronomy
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2017

Abstract

Wheat pre-harvest sprouting (PHS), germination of physiologically matured grains in a wheat spike before harvesting, can cause significant reduction in grain yield and end-use quality. Many quantitative trait loci (QTL) for PHS resistance have been reported in different sources. To determine the genetic architecture of PHS resistance and its relationship with grain color (GC) in US hard winter wheat, a genome-wide association study (GWAS) on both PHS resistance and GC was conducted using in a panel of 185 U.S. elite breeding lines and cultivars and 90K wheat SNP arrrays. PHS resistance was assessed by evaluating sprouting rates in wheat spikes harvested from both greenhouse and field experiments. Thirteen QTLs for PHS resistance were identified on 11 chromosomes in at least two experiments, and the effects of these QTLs varied among different environments. The common QTLs for PHS resistance and GC were identified on the long arms of the chromosome 3A and 3D, indicating pleiotropic effect of the two QTLs. Significant QTLs were also detected on chromosome arms 3AS and 4AL, which were not related to GC, suggesting that it is possible to improve PHS resistance in white wheat.

To identify markers closely linked to the 4AL QTL, genotyping-by-sequencing (GBS) technology was used to analyze a population of recombinant inbred lines (RILs) developed from a cross between two parents, “Tutoumai A” and “Siyang 936”, contrasting in 4AL QTL. Several closely linked GBS SNP markers to the 4AL QTL were identified and some of them were coverted to KASP for marker-assisted breeding.

To investigate effects of the two non-GC related QTLs on 3AS and 4AL, both QTLs were transferered from “Tutoumai A” and “AUS1408” into a susceptible US hard winter wheat breeding line, NW97S186, through marker-assisted backcrossing using the gene marker *TaPHS1* for 3AS QTL and a tightly linked KASP marker we developed for 4AL QTL. The 3AS QTL

(*TaPHS1*) significantly interacted with environments and genetic backgrounds, whereas 4AL QTL (*TaMKK3-A*) interacted with environments only. The two QTLs showed additive effects on PHS resistance, indicating pyramiding these two QTLs can increase PHS resistance.

To improve breeding selection efficiency, genomic prediction using genome-wide markers and marker-based prediction (MBP) using selected trait-linked markers were conducted in the association panel. Among the four genomic prediction methods evaluated, the ridge regression best linear unbiased prediction (rrBLUP) provides the best prediction among the tested methods (rrBLUP, BayesB, BayesC and BayesC0). However, MBP using 11 significant SNPs identified in the association study provides a better prediction than genomic prediction. Therefore, for traits that are controlled by a few major QTLs, MBP may be more effective than genomic selection.

GENETIC AND GENOMIC STUDIES ON WHEAT PRE-HARVEST SPROUTING RESISTANCE

by

MENG LIN

B.S., Huazhong Agricultural University, 2010

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Agronomy
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2017

Approved by:

Co-major Professor
Dr. Guihua Bai

Approved by:

Co-major Professor
Dr. Allan K Fritz

Copyright

Abstract

Wheat pre-harvest sprouting (PHS), germination of physiologically matured grains in a wheat spike before harvesting, can cause significant reduction in grain yield and end-use quality. Many quantitative trait loci (QTL) for PHS resistance have been reported in different sources. To determine the genetic architecture of PHS resistance and its relationship with grain color (GC) in US hard winter wheat, a genome-wide association study (GWAS) on both PHS resistance and GC was conducted using in a panel of 185 U.S. elite breeding lines and cultivars and 90K wheat SNP arrrays. PHS resistance was assessed by evaluating sprouting rates in wheat spikes harvested from both greenhouse and field experiments. Thirteen QTLs for PHS resistance were identified on 11 chromosomes in at least two experiments, and the effects of these QTLs varied among different environments. The common QTLs for PHS resistance and GC were identified on the long arms of the chromosome 3A and 3D, indicating pleiotropic effect of the two QTLs. Significant QTLs were also detected on chromosome arms 3AS and 4AL, which were not related to GC, suggesting that it is possible to improve PHS resistance in white wheat.

To identify markers closely linked to the 4AL QTL, genotyping-by-sequencing (GBS) technology was used to analyze a population of recombinant inbred lines (RILs) developed from a cross between two parents, “Tutoumai A” and “Siyang 936”, contrasting in 4AL QTL. Several closely linked GBS SNP markers to the 4AL QTL were identified and some of them were coverted to KASP for marker-assisted breeding.

To investigate effects of the two non-GC related QTLs on 3AS and 4AL, both QTLs were transferered from “Tutoumai A” and “AUS1408” into a susceptible US hard winter wheat breeding line, NW97S186, through marker-assisted backcrossing using the gene marker *TaPHS1* for 3AS QTL and a tightly linked KASP marker we developed for 4AL QTL. The 3AS QTL

(*TaPHS1*) significantly interacted with environments and genetic backgrounds, whereas 4AL QTL (*TaMKK3-A*) interacted with environments only. The two QTLs showed additive effects on PHS resistance, indicating pyramiding these two QTLs can increase PHS resistance.

To improve breeding selection efficiency, genomic prediction using genome-wide markers and marker-based prediction (MBP) using selected trait-linked markers were conducted in the association panel. Among the four genomic prediction methods evaluated, the ridge regression best linear unbiased prediction (rrBLUP) provides the best prediction among the tested methods (rrBLUP, BayesB, BayesC and BayesC0). However, MBP using 11 significant SNPs identified in the association study provides a better prediction than genomic prediction. Therefore, for traits that are controlled by a few major QTLs, MBP may be more effective than genomic selection.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to express my deepest appreciations to my major advisor, Dr. Guihua Bai for providing professional guidance and enthusiastic help throughout my graduate study. I appreciate all his contributions for time, ideas and funding to make my Ph.D. experience productive and stimulating. I am also thankful for the excellent example he has provided as a successful scientist and mentor. I am grateful to my co-advisor, Dr. Allan K Fritz, and the members of my advisory committee, Dr. Jianming Yu, Dr. Guorong Zhang, and Dr. Christopher Toomajian, for devoting their time to serve on my supervisory committee and providing great guidance during my graduate program. I also would like to thank Dr. Mark C Ungerer for serving as the outside chair of my final examination committee.

I would like to thank Dr. Shubing Liu for helping me with pre-harvest sprouting phenotyping, Dr. Dadong Zhang for teaching me the statistical software for genome-wide association analysis, and Dr. Shan Wang for helping with GBS data analysis. I am also very thankful to Drs. Paul St. Amand and Amy Bernardo in our lab for their advice and technical support throughout my research, and my lab members Drs. Zhenqi Su, Hui Chen, Yue Lu, Jin Cai, Dongtao Liu, Jixin Zhao, Yi Feng, Shoucai Ma, Feng Jin, Xinzhong Zhang, Na Liu and my friends Mingqin Shao, Yaoguang Li, Hadi Alipour, Abdulrahman Hashimi, Nosheen Fatima and John Curato for their help and friendship.

I would like to acknowledge the Triticeae Coordinated Agriculture Project (T-CAP) from the USDA National Institute of Food and Agriculture (NIFA) and USDA Hard Winter Wheat Genetics Research Unit for funding to support my Doctoral research project.

Last but not least, I would like to thank my family for their love and encouragement. I thank my parents for supporting me in all my pursuits, my husband, Zheng Hao, for being loving, patient, supportive and encouraging, and my son, Issac Hao, for bringing me a lot of joy and happiness.

Chapter 1 - Literature Review

Origin and agronomic importance of wheat

Wheat (*Triticum aestivum* L.) is one of the 'top three' cereal crops, with the total production right after maize and rice (http://faostat.fao.org/). It is grown widely from 67˚ N in southern Russia to 45˚ S in southwestern Australia and Argentina, including elevated regions in the tropical and sub-tropical areas (Feldman 1995). Wheat accounts for more than 20% food calories of the world population by providing large amount of starch and considerable levels of protein (Nevo et al. 2013). The 'gluten' proteins in wheat endosperm provide unique properties of dough made from wheat flour, which cannot be replaced by other grain crops. Besides human consumption, wheat production also provides a large portion of animal feed production with some protein rich supplements, such as soybeans and oilseed residues (Shewry 2009).

Common wheat is an allohexaploid species with 2n = 6x = 42 chromosomes consisting of three genomes A, B and D, and the evolution process is shown in Fig 1.1. The A genome donor is *Triticum urartu* (AuAu genome) (Chapman et al. 1976), the B genome donor is possibly *Aegilops speltoides* (SS genome) (Feldman 1976), and the D genome donor is *Aegilops tauschii* (DD genome) (Kihara 1944). Common wheat is derived through the hybridization between a domesticated tetroploid, *Triticum turgidum ssp. dicoccoides* (AABB genome) and the diploid *Aegilops tauschii* (DD genome) about 7,000-12,000 years ago, and *Triticum turgidum* (AABB genome) is derived through the hybridization between *Triticum urartu* (AA genome) and the B genome donor about 580-820 thousand years ago (Petersen et al. 2006; Salse et al. 2008; Marcussen et al. 2014). Unlike the A genome and D genome donors, the origin of the B genome remains controversial because it is still debatable whether the B genome originated from a single Aegilops species or an introgression of several Aegilops species (Sales et al. 2008). The genome

of Aegilop speltoides was designated as the S genome (Cox 1998; Wang et al. 1996), which is present in the Sitopsis section of Aegilops and is shared by a group of species (Slageren 1994), among which *Aegilops speltoides* is the closest extant species to the B genome donor (Feldman 2001).

The ancient cultivated diploid wheat, *T. monococcum*, also known as einkorn, was grown in the southern Levant of the Middle East and is still cultivated in limited areas there. However, the wild diploid wheat, *T. aegilopoides* and *T. uratu* can be widely found in the Middle East (Gill and Friebe 2002). The tetroploid wheat species, *T. turgidum* and *T. timopheevii*, have both cultivated and wild forms. The emmer wheat, also known as *T. turgidum ssp. dicoccum*, was an ancient cultivated wheat in southeastern Turkey (Gill and Friebe 2002; Heun et al. 1997). About 9,000 years ago, durum wheat (AABB), was selected and domesticated from a free-threshing mutant of emmer wheat, and has become the most widely cultivated tetroploid wheat since (Landi 1995). In the United States, durum wheat is grown in limited areas of North Dakota and surrounding states, and its common food products include spaghetti and macaroni (Gibson and Benson 2002). Common wheat is the major type of commercial wheat nowadays. It originated in Iran 6,700 years ago (Marcussen et al. 2014), was introduced to the U.S. in 1602, and has become one of the major cereal crops produced in the country (Gibson and Benson 2002).

Wheat pre-harvest sprouting

Pre-harvest sprouting (PHS) of wheat (*Triticum aestivum* L.) refers to the germination of wheat grains in matured spikes before harvest due to continuous wet weather during harvest seasons. PHS happens mainly due to reduced seed dormancy during domestication (Harlan 1992). Human selection artificially removed seeds that had prolonged dormancy to allow seeds to germinate uniformly after sowing (Lunn et al. 2002). Consequently, domesticated winter

wheats usually have the dormancy period of 0 to 12 weeks, which is shorter than their wild relatives (Mackey 1989). PHS is not only problematic in wheat production, but also impact other crops, such as rice and sorghum, by greatly reducing yield and grain quality (Dong et al. 2003; Steinbach et al. 1995).

Impact of PHS on wheat production

PHS can result in a significant reduction in wheat grain yield and grain end-use quality (Groos *et al*., 2002, Mares *et al*., 2005), which can cause economic losses for both grower and end user. Seed germination starts with water imbibition, which invokes the activation of alpha-amylase and other enzymes in the aleurone layer and embryo. Increased activity of alpha-amylase digests starch in the endosperm, thus reducing grain yield and nutritional quality (Imtiaz et al. 2008). Alpha-amylase activity can be evaluated by the Hagberg falling number test (Hagberg 1960). Furthermore, flour made from sprouted wheat grain contains hydrolyzed carbohydrate that usually results in sticky crumb and collapsed loaves (Derera et al. 1980). In durum wheat, sprouting not only reduces yield and test weight, but also causes higher cooking loss, poor color, reduced firmness and low stickiness of spaghetti (Grant et al. 1993; Manthey 2000).

PHS has been a major concern in wheat growing areas with maritime climate, such as northwest America, northwest Europe and north Japan (Lunn et al. 2002), or where high humidity occurs before harvesting, as in western Australia (Derera et al. 1980). PHS occurs about every four years in the western Australian wheat belt, because commercially grown high yielding varieties lack PHS resistance (Biddulph et al. 2005). In the U.S., PHS occurs frequently in the white wheat growing regions, such as northwestern states Washington, Oregon and Idaho and eastern states Michigan and New York (Briggle 1980). In the plain states where most U.S.

wheat is grown, North Dakota was attacked by PHS and lost about 12% of the hard red spring wheat and 19% of the durum in 1977 (Anonymous 1977), and north central Kansas and south central Nebraska experienced significant damage in hard red winter wheat when encountered continuous rainfall (Briggle 1980).

PHS and seed dormancy

PHS is mainly controlled by seed dormancy (SD), and it is also influenced by other factors, including red seed color (Gfeller and Svejda 1960; Groos et al. 2002), spike morphology (King and Richard 1984), physical barriers to water penetration (Gale 1989) and environmental factors such as temperature and moisture (Argel and Humphreys 1983; Ceccato et al. 2011).

It has long been considered that red-grained wheat tends to be more tolerant than white-grained wheat. Besides grain color, wheat with awns can absorb up to 30% more water than its near-isogenic lines without awns, thus increase sprouting by 40% ((King and Richard 1984)). Similarly, the club wheat heads can increase ear water absorption by 25% (King and Richard 1984). Also, PHS has been considered partially controlled by seed coat permeability to water. Recently, it has been shown that water imbibition rates are not significantly different between dormant and non-dormant genotypes until 18 h, and there is no evidence that water moves across the seed coat directly and into the endosperm (Judith et al. 2009).

Dormancy is usually defined as the failure of an intact viable seed to germinate under favorable conditions, including appropriate supply of oxygen, water and temperature (Gosling et al. 1983; King 1993; Bewley 1997). Abscisic acid (ABA) and gibberellic acid (GA) are important regulators of SD in species (Bewley 1997); besides, temperature and humidity during seed development also have effects on the length of SD (Argel et al. 1983, Ceccato et al. 2011).

ABA is important in regulating seed embryonic development, maturation and

germination (King 1982; Quatrano 1987). It performs as an inhibitor of embryonic germination in immature wheat grain (Quatrano et al. 1983), seeds of rape (Finkelstein et al. 1985) and soybean (Eisenberg et al. 1985). The levels of ABA in embryos and sensitivities of embryos to ABA can make differences in dormancy expression. ABA-deficient mutants of Arabidopsis, tomato and corn produce seeds that show reduced dormancy (Karssen et al. 1983; Quarrie 1987). It has also been found that ABA levels are similar in the whole seeds and the embryos of dormant grain and non-dormant grain (Walker-Simmons 1988). Therefore, it is likely that the embryo sensitivity to ABA is more important in dormancy regulation. ABA-insensitive mutants, such as maize *vp1* and Arabidopsis *abi1*, *abi2* and *abi3*, demonstrate reduced dormancy and viviparous germination (Koornneef et al. 1989; Koornneef et al. 1984; Le Page-Degivry et al. 1990). Germination of isolated embryos from dormant grain can be blocked by low concentrations (0.05 to 0.5 μM) of ABA, whereas germination of non-dormant grains can only be inhibited by 100 to 1000-fold greater ABA concentration (Walker-Simmons 1987 & 1988). ABA regulates SD from two aspects: on one hand, many of the ABA-responsive proteins can protect cells survive through the environmental stress, such as heat and drought; on the other hand, ABA can suppress the biosynthesis of proteins required in germination (Ried and Walker-Simmons 1990). As dormancy releases, ABA catabolism can be triggered, which results in a decrease in ABA content and an increase in its catabolic products (phaseic acid and dihydrophaseic acid) in the embryos (Jacobsen et al. 2002; Kushiro et al. 2004).

Unlike ABA, GA appears to promote the growth of the embryo during germination, rather than break the seed dormancy. Although some studies have shown that high GA concentrations (4-10 M) can overcome dormancy in some species (Bewley 1997), there is no solid evidence to show that GA is important for breaking dormancy. The activities of GA and

ABA may be linked, because GA only accumulates after ABA concentration reduces (Jacobsen et al. 2002; Ogawa et al. 2003). It is more likely that ABA might inhibit seed germination by repressing GA biosynthesis (Perez-Flores et al. 2003) and block GA signaling pathway in aleurone layer and embryos (Gomez-Cadenas et al. 2001; Gubler et al. 2002). It has also been shown that the GA-deficient mutants of Arabidopsis (*ga1-3*) and tomato (*gib1*) require exogenous GA to germinate (Groot and Karssen 1987; Koornneef and Van der Veen 1980), indicating that GA plays an essential role in germination process.

Temperature and humidity are two environmental factors that have a large effect on expression of SD. The effect of temperature on SD depends on the stage of grain development. In general, a high temperature during grain filling stage can induce short dormancy, while low temperature results in long dormancy (Lunn et al. 2002; Biddulph et al. 2005). However, high temperature (30 °C) results in high sprouting rates during germination compared to low temperature (10 °C) (Nyachiro et al. 2002), and seed dormancy can be released under 4 °C which corresponds to the decline in ABA content (Ali-Rachedi et al. 2004). Rainfall and high humidity during grain filling decrease the seed drying rate and dormancy, and drought stress and low humidity increase the grain drying rate and dormancy. It was shown that rainfall during the 20 days before harvesting had large influence on wheat grain germination rates (Mares 1993).

PHS resistance evaluation

Several methods have been developed to evaluate PHS resistance, including germination experiment with intact spikes or hand-threshed seeds, the Hagberg Falling Number method and the immunological test to measure alpha-amylase activity. Among these methods, intact spike sprouting test directly measures wheat PHS resistance, while the other three methods measure seed dormancy or alpha-amylase activity during seed germination.

The most straightforward method for PHS resistance evaluation is the intact spike sprouting test. In this method, physiologically matured spikes, characterized by the loss of green color of wheat spikes (Trethowan 1995), are harvested, air-dried and immersed in water for 4-6 h. Then the spikes are incubated in a moist chamber or on wet sand for 7 to 14 days. At the end of the experiment, PHS is scored as either sprouting rate or on a 1-9 scale with 1 for no visible sprouting and 9 for completely sprouted spikes (Baier 1987). The time periods to dry and incubate spikes may vary from study to study and heavily depend on types of experiments, materials used and grown environments of plants, but appropriate drying time and incubating time are usually determined to maximize the sprouting variance of extreme genotypes in the population under test.

Germination index (GI), another commonly used method for PHS resistance, measures seed dormancy, a major component of PHS resistance. To measure GI, spikes are harvested at physiological maturity, dried for a defined time period and then hand threshed. Fifty kernels from each accession are placed on a wet filter paper in a Petri dish. The Petri dish is incubated at room temperature for seven days, and germinated kernels are counted and removed daily. A weighted GI (modified after Walker-Simmons 1987) is calculated as

$$GI = \frac{7 \times N1 + 6 \times N2 + \cdots + 1 \times N7}{n\ days\ of\ test \times total\ number\ of\ kernels}$$

where *N1*, *N2*, … *N7* are the numbers of kernels germinated on day 1, day 2, till day 7.

Alpha-amylase activity is an indicator of PHS damage, and it can be measured by falling number (Hagberg 1960). Flour made from sprouted grains contains more degraded starch that results in low viscosity of the flour slurry. Falling number measures the time in seconds required for a stirrer-viscometer to fall a given distance through a heated, well mixed flour/water

suspension. The low falling number indicates low viscosity of the flour slurry, thus severe damage from PHS.

QTLs and candidate genes for PHS resistance

PHS resistance is a complex trait controlled by several major QTLs and minor QTLs. PHS resistance QTLs have been reported on almost all wheat chromosomes, among which the QTLs on chromosome 3AS, 4AL and 2BL have been studied mostly (Mori et al. 2005, Liu et al. 2008, Nakamura et al. 2011, Liu et al. 2013, Kato et al. 2001, Torada et al. 2005, Chen et al. 2008, Cabral et al. 2014, Torada et al. 2016, Kulwal et al. 2004 & 2012, Zhang et al. 2014). The 3AS QTL, designated as *TaPHS1*, has been cloned (Nakamura et al. 2011, Liu et al. 2013) which is a MOTHER OF FLOWERING TIME (*TaMFT*)-like gene, and positively regulates wheat PHS resistance. This gene explained 11.6% to 74.3% phenotypic variance in different mapping studies (Mori et al. 2005, Liu et al. 2008, Liu et al. 2010). Three single nucleotide polymorphisms (SNPs) in the gene have been associated with PHS resistance with one mutation in the promoter region (-222) (Nakamura et al. 2011), and two others in the gene-coding region (+646, +666) (Liu et al. 2013). The mutations in the coding region generate a mis-splicing site and a premature stop codon, resulting in a truncated nonfunctional transcript. Also the mis-splicing mutations associated with PHS susceptibility, and were involved in wheat domestication (Liu et al. 2015). *Phs1* that has been consistently mapped on chromosome 4AL is another major gene for both PHS resistance and seed dormancy, and explained 7.0% to 77.2% phenotypic variance (Kato et al. 2001, Mares et al. 2001, Mares et al. 2005, Torada et al. 2005, Chen et al. 2008, Ogbnnaya et al. 2008, Singh et al. 2010, Liu et al. 2011, Cabral et al. 2014). Two candidate genes were proposed for *Phs1*, *PM19-A1* and *PM19-A2*, by Barrero et al. (2015). However, the function of *PM19-A1* cannot be validated in the transgenic studies, and *PM19-A2*

falls out of the 4A QTL region, which makes the results unconvincing. Another candidate gene, *TaMKK3-A* (a mitogen-activated protein kinase kinase 3 (MKK3) gene) was identified by Torada et al. (2016) using a map-based cloning. A single SNP in *TaMKK3-A* causes a nonsynonymous amino acid substitution in the kinase domain, and is proposed as the causal SNP for seed dormancy. Another major QTL for PHS resistance is on chromosome 2BL, which has been identified in both bi-parental mapping and association mapping studies (Kulwal et al. 2004 & 2012, Liu et al. 2008, Munkvold et al. 2009, Singh et al. 2010, Rehman Arif et al. 2012). Zhang et al. (2013) found *TaSdr-B1* gene, an ortholog of the rice seed dormancy gene *OsSdr4*, to be associated with PHS resistance and located the gene on the 2BL chromosome.

QTLs for PHS resistance have also been identified on the long arms of group 3 chromosomes where GC QTLs are co-localized and chromosome 5A (Groos et al. 2002). In that study, 'Renan', a red PHS resistant line, was crossed with 'Récital', a white PHS susceptible line, and GC and PHS resistance were mapped simultaneously in the same population. The GC QTLs and PHS resistance QTLs on chromosome 3AL and 3DL were almost co-localized, whereas the QTLs for these two traits on 3BL were apart from each other for about 20 cM. Later, the 3BL and 3DL QTLs for PHS resistance were also identified by Kuwal et al. (2004), and the 3AL and 3BL QTLs by Mohan et al. (2009) and Fofana et al. (2009). Other important QTLs have been mapped on chromosomes 1A (Mares et al. 2005, Kumar et al. 2009, Lohwasser et al. 2013), 2D (Mares et al. 2002, Kulwal et al. 2004, Tan et al. 2006, Munkvold et al. 2009), 4B (Zanetti et al. 2000, Kato et al. 2001, Mori et al. 2005, Liu et al. 2011), 6B and 7D (Roy et al. 1999).

Wheat grain color and its impact on PHS resistance

Wheat varieties can be classified as red wheat and white wheat. Red wheat contains more phenolic acid in wheat bran than white wheat (Kim et al. 2006), and the more phenolic acid

results in bitter taste in red wheat flour. The degree of wheat kernel color can be related to the amount of catechin and catechin tannin in the seed coat of immature grain kernel (Miyamoto & Everson 1958). Proanthocyanidin (PA), the phenolic oligomers or polymers, is also important components of grain color pigment precursors in the pericarp of immature seeds (McCallum and Walker 1990). Both catechin and PAs are colorless, and can be converted to colored phlobaphene and anthocyanidins, respectively, when kernels get mature.

Wheat grain color (GC) has been associated with PHS, meaning that red-grained wheat is usually more resistant to PHS than white-grained wheat (Flintham 2000; Warner et al. 2000; Himi et al. 2002). Seed dormancy levels were increased in white wheat NS-67 after a single GC gene was added to the group 3 chromosomes (Flintham 2000). It has also been shown that the white-grained mutants of 'Chinese Spring' and 'AUS1490' had increased sprouting rates compared to the original lines, indicating that GC genes improved PHS resistance (Warner et al. 2000; Himi et al. 2002). Common QTLs for GC and PHS resistance have been identified on chromosome 2B, 3AL, 3BL, 3DL, 5A and 6B (Groos et al. 2002; Kumar et al. 2009). Although it is not clear how GC genes regulate grain germination at molecular levels, many studies have shown that GC genes can enhance PHS resistance either by accumulating catechin and PAs, the germination inhibitors (Miyamoto and Everson, 1958; Stoy and Sundin, 1976; McCallum and Walker 1990), or increasing the sensitivity of embryos to ABA (Himi et al. 2002).

GC evaluation

Wheat GC is mostly evaluated by a chromameter or by visual scoring. A chromameter decomposes color in the L*a*b color space, where 'L' evaluates black (0) to white (100), 'a' evaluates green (negative) to red (positive) and 'b' evaluates blue (negative) to yellow (positive). Multiple measures are supposed to be done on a sample of 20 g grains to determine GC for

different genotypes (Groos et al. 2002). Another method to evaluate wheat GC is to soak wheat kernels in 1 M sodium hydroxide (NaOH) solution to increase the color contrast, and visually score the color intensity using a scale of 1 (white) to 5 (dark red) (Torada et al. 2002; Bassoi et al. 2005; Kumar et al. 2009).

QTLs and candidate genes for GC

In early 1920s, Nilsson-Ehle (1914) found that GC was controlled by three genes, *R-A1*, *R-B1* and *R-D1*, on chromosomes group 3 (Sears 1944; Allan and Vogel 1965; Metzger and Silbaugh 1970). Mapping studies using bi-parental mapping populations verified the location of the *R* genes (Groos et al. 2002; Fofana et al. 2009), and identified novel QTLs for GC on chromosomes 2B, 2D, 5A, 5D, 6B, 7B and 7D (Groos et al. 2002; Kumar et al. 2009), indicating that GC is a complex trait controlled by more than three genes. Recently, Himi et al. (2011) identified the *Tamyb10* genes as the candidate genes for the GC trait. These genes are transcription factors of the flavonoid biosynthetic pathway and encode MYB domain proteins that are similar to the regulatory proteins for PAs synthesis in Arabidopsis.

Genetic markers used in plant mapping studies

Genetic markers are distinguishable characters in morphological traits or protein/DNA molecules that can be landmarks for agronomic traits or QTLs under selection in plant breeding. Genetic markers can be classified into classical markers and DNA markers (Xu 2010). Classical markers include morphological markers, cytological markers and protein markers (Jiang 2013). Morphological traits, such as seed color, plant height, leaf shape and flower color, can be used as indicators when they are linked to other agronomic traits of interest (Kadervel et al. 2015). Although morphological markers are still useful in modern plant breeding, they are limited in number and may have undesirable effects on plant development (Jiang 2013). Cytological

markers are the banding patterns of the chromosome, which can be used for chromosome identification, chromosome mutation detection (Santos et al. 2006) and physical mapping (Orellana et al. 1993). However, the application of cytological markers is limited in mapping and breeding due to the low resolution and technical demand. Protein markers were mainly used in the 1980s, and they are enzyme variants that are different in size and molecular weight and can be distinguished using electrophoresis.

DNA markers can be classified into low-throughput, medium-throughput, high-throughput and ultra high-throughput according to the throughput that data are generated (Mir et al. 2013). Restriction fragment length polymorphism (RFLP) (Grodziker et al. 1975), referred to as 'First generation molecular markers' (Jones et al. 2009), initiated the era of DNA markers, despite the low-throughput nature of this technology. Medium-throughput DNA markers include random amplified polymorphic DNA (RAPD) (Williams et al. 1990), sequence-tagged site (STS) (Olsen et al. 1989), expressed sequence tags (EST) (Adams et al. 1991), simple sequence repeats (SSRs) (Akkaya et al. 1992) and amplified fragment length polymorphisms (AFLPs) (Vos et al. 1995). Single nucleotide polymorphism (SNP) has become broadly used and can be easily genotyped with high-throughput technologies, due to its abundance and even distribution across the genome. As the next-generation sequencing (NGS) technology develops, the costs of sequencing have reduced from $60 to $1 per mega base (Thudi et al. 2012), which allows genomic/transcriptomic re-sequencing at affordable prices, and thus facilitates SNP discovery. Recently, SNP arrays, genotyping-by-sequencing (GBS) and kompetitive allele-specific PCR (KASP) are broadly used methods for SNP detection.

High-density SNP arrays have become an important tool in genetic studies and plant breeding (Xing 2014). Such arrays have been developed for major crops, such as the 44K rice

SNP array (Zhao et al. 2011), the 50K maize SNP array (Cook et al. 2012), and the 9K and 90K wheat SNP arrays (Cavanagh et al. 2013; Wang et al. 2014). However, the lack of flexibility and relatively high cost of these SNP arrays limit their application in plant breeding (Lateef 2015). By taking the advantages of NGS, GBS has been developed as a robust and cost-efficient sequencing approach that can identify and genotype SNPs simultaneously (Mir 2013). GBS uses restriction enzymes to mask the repetitive regions and reduce the complexity of genomes, thus allows reaching important genomic regions that are unreachable to sequence capture approaches (van Oeveren et al. 2011) and increasing the chance of sampling markers from gene rich regions. GBS can provide adequate SNPs for high-resolution mapping, genomic selection, germplasm characterization and other breeding applications (Huang et al. 2010; Elshire et al. 2011; Poland et al. 2012). Barley, wheat and maize are early successful examples of applying GBS in plant genetic studies (Poland et al. 2012; Mascher et al. 2013; Romay et al. 2013). It has been shown to be an effective tool for genetic studies in rice (Bandillo et al. 2013), sorghum (Morris et al. 2013) and soybean (Sonah et al. 2015). The major challenge for GBS is a high rate of missing data due to low sequencing depth, which can be partially solved by imputation and high coverage sequencing. KASP assay is a simple and flexible genotyping system, which is commonly used when a large number of samples need be genotyped with a small number of SNPs (Mir et al. 2013). Chen et al. (2010) developed new SNP genotyping assays that combined competitive allele-specific PCR and Fluorescence Resonance Energy Transfer (FRET), and Kbioscience UK later developed this technology into KASP assays. Although KASP assays came to the market very recently, they have been applied successfully in plant genetic studies (Allen et al. 2011; Mammadov et al. 2012).

Linkage mapping and Genome-Wide Association Study (GWAS)

With the rapid development of genetic markers, dissecting and mapping QTLs for complex traits receives great attentions in plant genetics study and plant breeding. Linkage mapping and Genome-wide Association Study (GWAS) are two prevailing methods in QTL mapping studies, and each of them has its own advantages.

Linkage mapping usually starts with a segregating mapping population. The population size of more than 150 lines is preferred (Collard et al. 2005) because a large population size can provide an observable number of recombinants and allow an accurate evaluation of the target trait (Doerge 2002). Recombinant inbred lines (RILs) and double haploid (DH) populations are most commonly used in linkage mapping because the genotypes can be maintained and evaluated in multiple years and locations, and F_2 and backcross (BC) populations are also used (Würschum 2012). Several approaches have been used for linkage mapping, including single marker analysis (SMA), interval mapping (IM), composite interval mapping (CIM) and multiple interval mapping (MIM) (Tanksley 1993). SMA uses t-test, analysis of variance (ANOVA) or simple linear regression to screen markers that potentially related to the trait under investigation (Young 1996). However, SMA cannot provide recombination frequency between the marker and the QTL because the QTL effect and location are confounded, thus unable to be estimated separately. To address this problem, genetic maps are introduced where genetic markers are linearly ordered. With such information, the likelihood of a QTL is tested throughout a linkage map, and the logarithm of the odds (LOD) scores are used to estimate the location of a QTL (Soller et al. 1979; Lander and Botstein 1989). CIM is introduced to remove the variation caused by other, especially linked QTL by including additional markers outside a defined window as cofactors (Zeng 1993&1994). Therefore, CIM can reduce the chance of discovering ‘ghost

QTLs'. MIM, proposed by (Jansen 1993; Jansen and Stam 1994; Jansen 1995), focuses on detecting epistatic effects among QTLs. However, both CIM and MIM are restricted to one-dimensional search along the genetic map, thus are challenged by the multiplicity of epistatic QTL effects (Doerge 2002). Permutation and bootstrap resampling are two methods to determine the threshold of a significant QTL, and an empirical threshold of LOD at 3.0 is often used in linkage mapping studies (Collard et al. 2005).

GWAS uses a diverse population that consists of accessions collected from different geographic origins or with complex relatedness to make associations between genetic loci and trait under investigation. Diverse populations, very different from populations used in linkage mapping, capture all the historical recombinations occurred in the sampled accessions (Myles et al. 2009). Therefore, high-density genetic markers are needed to cover the linkage disequilibrium (LD) structure across the genome (Lipka et al. 2015), in order to detect the genetic variants associated with phenotypic variance. Particularly, the nested association mapping (NAM) population, created by crossing diverse inbred lines to a common parent, combines both historical and recent recombination events. NAM populations have been proven to be successful in dissecting complex traits in maize and barley (McMullen et al. 2009; Yu et al. 2008; Poland et al. 2011; Maurer et al. 2015), due to the advantages of high genetic power, high allele richness, low sensitivity to genetic heterogeneity and high efficiency in using genome sequence (Yu et al. 2008).

In a diverse population, population structure and relationships among individuals can be non-negligible sources of false positive associations. Therefore, covariates for population structure and kinship are introduced into the statistic model for GWAS (Zhu et al. 2008). STRUCTURE (Pritchard et al. 2000) and principal component analysis (PCA) (Price et al. 2006)

are the mostly used method to describe the population structure. Kinship matrix represents the relatedness among individuals by using identity-by-state to estimate identity-by-descent (Loiselle et al. 1995). However, if the trait under investigation is correlated with a population structure, introducing structure as covariates can cause the loss of statistical power (Lipka et al. 2015). To solve the structure issue, Yu *et al*. (2006) proposed a mixed linear model for GWAS; based on that, many approaches have been developed to increase computational efficiency, such as efficient mixed-model association (EMMA) (Kang et al. 2008), EMMA eXpedited (EMMAX) (Kang et al. 2010), the compressed mixed linear model (Zhang et al. 2010; Li et al. 2014) and population parameters previously determined (P3D) (Zhang et al. 2010). Currently, these approaches are available in the user-friendly software packages TASSEL (Trait Analysis by aSSociation, Evolution and Linkage, Bradbury et al. 2007) and Genome Association and Prediction Integrated Tool (GAPIT) (Lipka et al. 2012). Although statistic models and software packages have been well developed, there are still some concerns in GWAS. One of them is called 'synthetic association' (Dickson et al. 2010; Orozco et al. 2010), when several low-frequency causal variants are in strong LD with a common variant. Under such circumstances, genetic variance cannot be properly estimated and it takes extra effort to identify the casual variants. Another concern is that SNPs cannot represent all possible genetic variations associated to a trait, therefore, it is important to include other sources of genetic variation in GWAS, such as epigenetic variation, transposons and copy number variation (Lipka et al. 2015).

Breeding strategies for PHS resistance

Although wheat PHS resistance is greatly influenced by morphological and environmental factors, breeding for genetically improved wheat with proper seed dormancy is the most effective method to protect wheat cultivars from PHS damages (Liu et al. 2008). Red-

grained wheat usually shows more resistance to PHS than white-grained wheat (Seshu and Sorrells 1986), as the red color genes on the long arms of group 3 chromosomes can have pleiotropic effects on PHS resistance (Nelson et al. 1995; Groos et al. 2002). White-grained wheat is popular due to the users' preference in Asian market (Amano and Torada 2002; Tan et al. 2006) and its economic benefits like high flour extraction rate (McCaig and Depauw 1992). However, white-grained wheat is usually vulnerable to PHS, thus breeding for PHS-resistant white wheat is extremely important in PHS-favorable environments, such as Australia and the USA (Morris and Paulsen 1989; Imtiaz et al. 2008).

Phenotypic selection and marker-assisted selection (MAS) are the most commonly used methods to breed for PHS-resistant lines. Artificial mist is widely used to create wetting treatment and induce germination on harvested wheat spikes (Hucl 1994; Groos et al. 2002; Imtiaz et al. 2008). Phenotypic selection is straightforward in PHS resistance improvement, but is time and labor consuming. Therefore, MAS has been applied in many breeding programs, as PHS resistance genes were cloned on chromosomes 3A (Nakamura *et al*. 2011; Liu *et al*. 2013) and 4A (Torada et al. 2016) and QTLs for PHS resistance identified across the genome. Molecular markers have been developed for the causal SNPs in both the promoter and coding region of the *TaPHS1* gene (Liu et al. 2015), and the causal SNP for the *TaMKK* gene (Torada et al. 2016). These markers can be used either to identify germplasm carrying these two genes, or to select resistant lines in breeding materials. The *TaPHS1* gene has been successfully integrated using MAS to increase PHS resistance in several studies (Kottearachchi et al. 2006; Gupta et al. 2008). However, not all PHS resistance QTLs are suitable for breeding, because many of them show significant genetic-by-environment interactions and do not have consistent effects across different environments (MAS in wheat, http://maswheat.ucdavis.edu/protocols/PHS/).

AUS1408 is an important source for PHS resistance in white-grained wheat, and it has been widely used in breeding projects (Amano and Torada 2002; Hucl and Matus-Cádiz 2002). Besides AUS1408, Zenkoujikomugi, 8019R1 and RyuuMai7 are critical PHS-resistant germplasm adapted to various environments in Japan (Kottearachchi et al. 2006; Amano and Torada 2002). In U.S., Clark's Cream and its derivative line Cayuga contain PHS resistance QTLs on chromosomes 1AS and 2B and show a level of tolerance similar to many red cultivars (MAS in wheat, http://maswheat.ucdavis.edu/protocols/PHS/). And in China, PHS resistance in most cultivars can be traced back to Wanxian White, Fulingxuxu White, Suiningtuotuo and Yongchuan White (Xiao et al. 2002). QTLs for PHS resistance have also been identified in *Aegilops tauschii* (Lan et al. 1997; Imtiaz et al. 2008) and *Triticum spelta* (Zanetti et al. 2000). However, due to some undesirable traits from the linkage drag, these QTLs have not been widely used in wheat breeding.

Breeding for PHS resistance for various environments is challenging in wheat breeding. In order to improve PHS resistance, more information on PHS resistance genetic architecture, PHS resistance pathways and gene regulations, genetic-by-environment interactions, and user-friendly markers and efficient selection method is required.

Table 1.1 Summary of previously reported QTL for wheat preharvest sprouting (PHS) resistance and grain color (GC) traits

Trait[a]	Chromosome	Parental Lines	Population Type[b]	References
GR	4A (*TaMKK3-A*)	Leader/Haruyokoi	BC3F2, BC4F2	Torada *et al*. (2016)
GI	4A (*PM19-A1&A2*)	-	MAGIC population	Barrero *et al*. (2015)
GI	2B (*TaSdr-B1*)	Yangmai/Zhongyou 9507	RIL	Zhang *et al*. (2014)
GR	4A	RL4452/AC Domain	DH	Cabral *et al*. (2014)
GI	3B, 4A, 7B			
FN	4A, 7D			
GI	3B (*TaDFR-B*)	-	AM panel	Bi *et al*. (2014)
GI	1D, 3A, 4A	Opata 85/W7984	RIL	Lohwasser *et al*. (2013)
GR	1A, 4A			
GI	6D	Chinese Spring/Synthetic 6x	BC2F2	
GI	1B, 1D, 2B, 3A, 3B, 4A, 4B, 4D, 5B, 6B, 7A	-	AM panel	
GR	1A, 1B, 1D, 2B, 3B, 4A, 5A, 6A, 7B			
GR	3A (*TaPHS1*)	Rio Blanco/NW97S186, Rio Blanco/NW97S078	RIL	Liu *et al*. (2013)
GI	1A, 1B, 2A, 2B, 2D, 3A, 3B, 4A, 4B, 5B, 6B, 7A	-	AM panel	Arif *et al*. (2012)
GR	1A, 1B, 1D, 2B, 4A, 5B			
GR	2B, 3D, 7B	-	AM panel	Jaiswal *et al*. (2012)
GR	1BS, 2BS, 2BL, 2DS, 4AL, 6DL, 7BS, 7DS	-	AM panel	Kulwal *et al*. (2012)
GR	3A (*TaMFT*)	Zen/Chinese Spring	CS(Zen3A)	Nakamura *et al*. (2011)
GI	4A, 4B, 5B	Tutoumai A/Siyang 936	RIL	Liu *et al*. (2011)
GR	4A, 4B, 5B			
GI, GR	3AL (*Vp-1A*)	Wanxianbaimaizi/Jing411	RIL	Chang *et al*. (2011)
GR	1A, 2B, 3A, 4A, 5B, 6B, 7A	W98616/Argent	DH	Singh *et al*. (2010)
GI	3BL (*Vp-1B*)	-	AM panel	Chang *et al*. (2010)
GI	3BL (*Vp-1B*)	Wanxianbaimaizi/Jing411	RIL	Chang *et al*. (2010)
GI	3BL (*Vp-1B*)	-	AM panel	Xia *et al*. (2009)
GI	3A, 3D, 4A, 4B, 7D	AC Domain/RL4452	DH	Rasul *et al*. (2009)

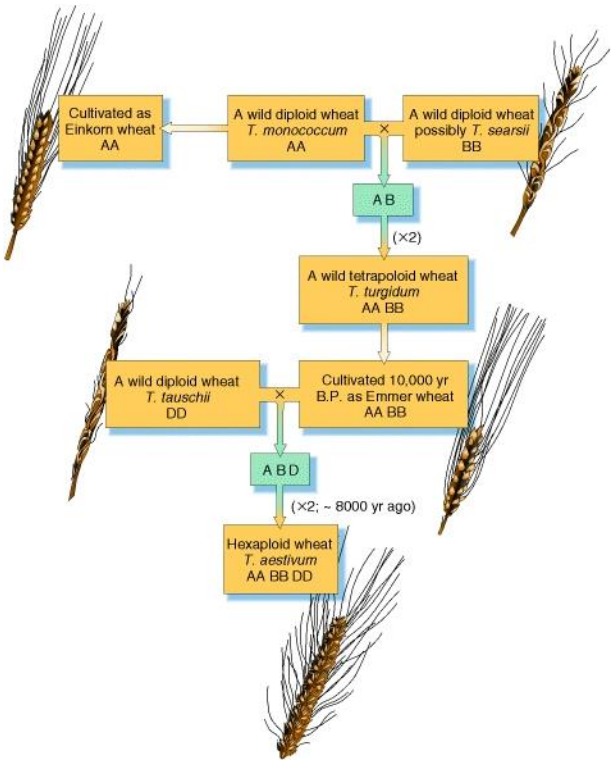
FN	4A, 4B			
GI	1A, 1B, 2B, 2D, 3A, 3D, 4A, 4D, 5B, 5D, 6D, 7D	Cayuga/Caledomia	DH	Munkvold *et al*. (2009)
GR	1A, 2A, 2B, 3B, 6A, 6B	PH132/WL711	RIL	Kumar *et al*. (2009)
GR	1AS, 2AL, 2DL, 3AL, 3BL	SPR8198/HD2329	RIL	Mohan *et al*. (2009)
GI	3BL, 3DL	SUN325B/QT7475	DH	Mares *et al*. (2009)
GI	3B, 3D	AC Domain/White-RL4137	DH	Fofana *et al*. (2009)
GR	3A, 3B, 3D, 5D			
FN	3B, 3D			
GI, GR	4AL	Tutoumai A/ Siyang 936	RIL	Chen *et al*. (2008)
GI, GR	4AL	Halberd/Cranbrook	DH	Zhang *et al*. (2008)
GI	4AL	OS21-5/Leader	BC5F2	Torada *et al*. (2008)
GI, GR	4A	CN19055/Auunello	RIL	Ogbonnaya *et al*. (2008)
GI, GR	3D, 4A	Syn37/Janz	BC1F7	Imtiaz *et al*. (2008)
GI	3BL (*Vp-1B*)	-	AM panel	Yang *et al*. (2007)
GI, GR	3D, 4A	-	AM panel	Ogbonnaya *et al*. (2007)
GI	3AmL, 4AmL, 5AmL	*T. boeoticum* L. Boiss(KT1-1)/*T. monococcum* L. (KT3-5)	RIL	Nakamura *et al*. (2007)
GI	2DS, 3AL, 4AL, 5BL	AUS1408/Cascadeds	DH	Tan *et al*. (2006)
GR	4AL	Kitamoe/Munstertaler	DH	Torada *et al*. (2005)
GR	3A, 4A, 4B	Zen/Chinese Spring	RIL	Mori *et al*. (2005)
GI	4AL	AUS1408/Janz, AUS1408/Cascades, SW95-50213/AUS1408, AUS1490/Janz, SW95-50213/Cunningham	DH	Mares *et al*. (2005)
GR	1AL, 1BL, 3DL, 4AL	ITMI*pop*	RIL	Lohwasser *et al*. (2005)
GR	3AL	SPR8198/HD2329	RIL	Kulwal *et al*. (2005)
GR	1A	Kyle/CI13102	RIL	Knox *et al*. (2005)
GR	2BL, 2DS, 3BL, 3DL	W7984/Opata85	RIL	Kulwal *et al*. (2004)
GR	3AS, 3AL	Zen/Chinese Spring	RIL	Osa *et al*. (2003)
GI	2D, 3A, 3D, 4A, 5A	AUS1408/Oxley	F2 (disomic)	Mares *et al*. (2002)
GR	3A, 3B, 3D, 5A	Renan/Recital	RIL	Groos *et al*. (2002)
GI	2AL, 2DL, 4AL	Halberd/Cranbrook	DH	Mares *et al*. (2001)

GR	4A, 4B, 4D	AC Domain/Haruyutaka	DH	Kato *et al*. (2001)
FN	1A, 1BS, 1DS, 2A, 2B, 3A, 3B, 4A, 4B, 4DL, 5A, 5B, 6A, 6D, 7B	Forno/Oberkulmer	RIL	Zanetti *et al*. (2000)
AA	1A, 1BS, 2A, 2B, 3A, 3B, 3DL, 4A, 4B, 4DL, 5A, 5B, 6D, 7B			
GR	3A, 3B	Langdon/ DIC	Langdon-DIC substituion lines	Watanabe *et al*. (2000)
GR	6B, 7D	SPR8198/HD2329	RIL	Roy *et al*. (1999)
GR	1AS, 2S, 2L	Clark's Cream/NY6432-18	RIL	Anderson *et al*. (1993)
GR	5DL, 6BL, 4AL, 3BL	NY6432-18/NY6432-10	RIL	
GC	7B, 7D	Purple Feed/Saratovskaya 29, Purple/Saratovskaya 29	F2, F3, NIL	Tereshchenko *et al*. (2012)
GC	3A, 3B, 3D (*Tamyb10* genes)	Zenkoji Komuji/Tamaizumi	DH	Himi *et al*. (2011)
GC	3A, 3B, 3D	AC Domain/White-RL4137	DH	Fofana *et al*. (2009)
GC	2B, 2D, 3B, 5D, 6B	PH132/WL711	RIL	Kumar *et al*. (2009)
GC	2A, 4B, 6B, 7B	Kofa/W9262-260D3	DH	Pozniak *et al*. (2007)
GC	3A, 3B, 3D, 5A	Renan/Recital	RIL	Groos *et al*. (2002)

[a]GR refers to germination rate, GI refers to germination index, FN refers to falling number, AA refers to α-amylase activity, and GC refers to grain color

[b]MAGIC population=multi-parent advanced generation inter-cross population, RIL=recombinant inbred lines, DH=double haploids, AM panel=association mapping panel, NIL=near isogenic lines

Figure 1.1 Hybridization events involved in the evolution of bread wheat



http://www.cerealsdb.uk.net/cerealgenomics/WheatBP/Documents/DOC_Evolution.php

Reference:

Adams MD, Kelley JM, Gocayne JD, Dubrick M, Polymeropoulos MH, Xiao H, Merril CR, Wu A, Olde B, Moreno RF, Kerlavage AR, McCombie WR, Venter JC (1991) Complementary DNA sequencing: expressed sequence tags and human genome project. Science 252:1651–1656

Akkaya MS, Bhagwat AA, Cregan PB (1992) Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132:1131–1139

Ali-Rachedi S, Bouinot D, WagnerM-H, BonnetM, Sotta B, Grappin P, Jullien M (2004) Changes in endogenous abscisic acid levels during dormancy release and maintenance of mature seeds: studies with the Cape Verde Islands ecotype, the dormant model of Arabidopsis thaliana. Planta 219:479-488

Allan RE, Vogel OA (1965) Monosomic analysis of red seed color in wheat. Crop Sci. 5:474–475

Allen AM, Barker GLA, Berry ST, Coghill JA, Gwilliam R, Kirby S, Robinson P, Brenchley RC, D`Amore R, McKenzie N, Waite D, Hall A, Bevan M, Edwards KJ (2011) Transcript-Specific, Single- Nucleotide Polymorphism Discovery and Linkage Analysis in Hexaploid Bread Wheat (Triticum aestivum L.). Plant Biotech. J. 9:1086-1099

Amano Y, Torada A (2002) Breeding of white-grained wheats for Japan. Euphytica 126:83–88

Anderson JA, Sorrells ME, Tanksley SD (1993) RFLP analysis of genomic regions associated with resistance to preharvest sprouting in wheat. Crop Sci. 33:453-459

Anonyous (1977) Crop-weather summary. North Dakota Crop and Livestock Reporting Service, report dated 4 October, 1977

Argel PJ, Humphreys LR (1983) Environmental effects on seed development and

hardseededness in Stylosanthes hamata cv. Verano. I. Temperature. Crop Pasture Sci. 34:261–270

Arif MR, Neumann K, Nagel M, Kobiljski B, Lohwasser U, Börner A (2012). An association mapping analysis of dormancy and pre-harvest sprouting in wheat. Euphytica 188:409-417

Baier AC (1987) Pre-harvest sprouting. Annu. Wheat Newsl 33:40

Bandillo N, Raghavan C, Muyco PA, Sevilla MAL, Lobina IT, Dilla-Ermita CJ, Tung CW, McCouch S, Thomson M, Mauleon R, Singh RK, Gregorio G, Redoña E, Leung H (2013). Multi-parent advanced generation inter-cross (MAGIC) populations in rice: progress and potential for genetics research and breeding. Rice 6:11

Barrero JM, Cavanagh C, Verbyla KL, Tibbits JF, Verbyla AP, Huang BE, Rosewarne GM, Stephen S, Wang P, Whan A, Rigault P, Hayden MJ, Gubler (2015) Transcriptomic analysis of wheat near-isogenic lines identifies PM19-A1 and A2 as candidates for a major dormancy QTL. Genome Biol. 16:1

Bassoi MC, Flintham J (2005) Relationship between grain colour and preharvest sprouting-resistance in wheat. Pesquisa Agropecuária Brasileira 40:981-988

Bi HH, Sun YW, Xiao YG, Xia LQ (2014) Characterization of DFR allelic variations and their associations with pre-harvest sprouting resistance in a set of red-grained Chinese wheat germplasm. Euphytica 195:197-207

Biddulph TB, Mares DJ, Plummer JA, Setter TL (2005) Drought and high temperature increases preharvest sprouting tolerance in a genotype without grain dormancy. Euphytica 143:277-283

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics

23:2633-2635

Briggle L W (1980) Pre-harvest sprout damage in wheat in the US. Cereal Res. Commun. 245-250

Cabral AL, Jordan MC, McCartney CA, You FM, Humphreys DG, MacLachlan R, Pozniak CJ (2014) Identification of candidate genes, regions and markers for pre-harvest sprouting resistance in wheat (Triticum aestivum L.). BMC Plant Biol. 14:1

Cavanagh CR, Chao S, Wang S, Huang BE, Stephen S, Kiani S, Forrest K, Saintenac C, Brown-Guedira GL, Akhunova A (2013) Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. Proc. Natl. Acad. Sci. U.S.A. 110:8057-8062

Ceccato DV, Daniel Bertero H, Batlla D (2011) Environmental control of dormancy in quinoa (Chenopodium quinoa) seeds: two potential genetic resources for preharvest sprouting tolerance. Seed Sci. Res. 21:133–141

Chang C, Feng JM, Si HQ, Yin B, Zhang HP, Ma CX (2010) Validating a novel allele of viviparous-1 (Vp-1Bf) associated with high seed dormancy of Chinese wheat landrace, Wanxianbaimaizi. Mol. Breed. 25:517-525

Chang C, Zhang HP, Feng JM, Yin B, Si HQ, Ma CX (2010) Identifying alleles of Viviparous-1B associated with pre-harvest sprouting in micro-core collections of Chinese wheat germplasm. Mol. Breed. 25:481-490

Chang C, Zhang HP, Zhao QX, Feng JM, Si HQ, Lu J, Ma CX (2011) Rich allelic variations of Viviparous-1A and their associations with seed dormancy/pre-harvest sprouting of common wheat. Euphytica 179:343-353

Chapman V, Miller TE, Riley R (1976) Equivalence of the A genome of bread wheat and that of

Triticum urartu. Genet. Res. 27:69-76

Chen W, Mingus J, Mammadov J, Backlund JE, Greene T, Thompson S, Kumpatla S (2010) KASPar: a simple and cost-effective system for SNP genotyping. In: Plant and Animal Genomes XVII conference, San Diego, USA, p 194

Chen CX, Cai SB, Bai GH (2008) A major QTL controlling seed dormancy and pre-harvest sprouting resistance on chromosome 4A in a Chinese wheat landrace. Mol. Breed. 21:351-358

Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica 142:169-196

Cook JP, McMullen MD, Holland JB, Tian F, Bradbury P, Ross-Ibarra J, Buckler ES, Flint-Garcia SA (2012) Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. Plant Physiol. 158:824-834

Cox TS (1998) Deepening the wheat gene pool. J. Crop Prod. 1:1–25

Derera N (1980) The audit of sprouting. Cereal Res. Commun. 8:15–22

Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB (2010) Rare variants create synthetic genome-wide associations. PLoS Biol 8:e1000294

Doerge RW (2002) Mapping and analysis of quantitative trait loci in experimental populations. Nat. Rev. Genet. 3:43-52

Dong Y, Tsuzuki E, Kamiunten H, Terao H, Lin D, Matsuo,M, Zheng Y (2003) Identification of quantitative trait loci associated with pre-harvest sprouting resistance in rice (*Oryza sativa* L.). Field Crop. Res. 81:133-139

Eisenberg AJ, Mascarenhas JP (1985) Abscisic acid and the regulation of synthesis of specific

seed proteins and their messenger RNAs during culture of soybean embryos. Planta 166:505-514

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6:e19379

Feldman M (1976) Wheats: Triticum spp. (Gramineae-Triticinae). Evolution of Crop Plants NW Simmonds, ed

Feldman M (1995) Wheats. 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

Finkelstein RR, Tenbarge KM, Shumway JE, Crouch ML (1985) Role of ABA in maturation of rapeseed embryos. Plant Physiol. 78:630-636

Flintham JE (2000) Different genetic components control coatimposed and embryo-imposed dormancy in wheat. Seed Sci. Res.10:43–50

Fofana B, Humphreys DG, Rasul G, Cloutier S, Brûlé-Babel A, Woods S, Lukow OM, Somers DJ (2009) Mapping quantitative trait loci controlling pre-harvest sprouting resistance in a red× white seeded spring wheat cross. Euphytica 165:509-521

Gale MD (1989) The genetics of preharvest sprouting in cereals, particularly in wheat. In:Derera NF (ed) Preharvest field sprouting in cereals. CRC Press, Boca Raton pp 85–110

Gfeller F, Svejda F (1960) Inheritance of post-harvest seed dormancy and kernel color in spring wheat lines. Can. J Plant Sci. 40:1–6

Gibson L, Benson G (2002) Origin, history, and uses of oat (Avena sativa) and wheat (Triticum

aestivum). Iowa State University, Department of Agronomy

Gill BS, Friebe B (2002) Cytogenetics, phylogeny and evolution of cultivated wheats. Bread Wheat-Improvement and Production 567

Gomez-Cadenas A, Zentella R, Walker-Simmons M, Ho T-HD (2001) Gibberellin/abscisic acid antagonism in barley aleurone cells: site of action of the protein kinase PKABA1 in relation to gibberellin signaling molecules. Plant Cell 13:667-679

Grant LA, Dick JW, Shelton DR (1993) Effects of drying temperature, starch damage, sprouting, and additives on spaghetti quality characteristics. Cereal Chem. 70:676–684

Grodzicker T, Williams J, Sharp P, Sambrook J (1975) Physical mapping of temperature sensitive mutants of adenovirus. Cold Spring Harb Symp Quant Biol 39:439–446

Groos C, Gay G, Perretant MR, Gervais L, Bernard M, Dedryver F, Charmet G. (2002) Study of the relationship between pre-harvest sprouting and grain color by quantitative trait loci analysis in a white × red grain bread-wheat cross. Theor. Appl. Genet. 104:39–47

Groot SPC, Karssen CM (1987) Gibberellins regulate seed germination in tomato by endosperm weakening: A study with gibberellin-deficient mutants. Planta 171:525-531

Gubler F, Chandler PM, White RG, Llewellyn DJ, Jacobsen JV (2002) Gibberellin signaling in barley aleurone cells. Control of SLN1 and GAMYB expression. Plant Physiol. 129:191-200

Gupta PK, Balyan HS, Kumar J, Kulwal PK, Kumar N, Mir RR, Kumar A, Prabhu KV (2008) QTL analysis and marker assisted selection for improvement in grain protein content and pre-harvest sprouting tolerance in bread wheat. p. 1–3. In: Appels R, Eastwood R, Lagudah E, Langridge P, Mackay M, McIntyre L, Sharp P (eds) Proceedings of 11th International Wheat Genet Symposium, Brisbane Australia, August 24–29, 2008. Sydney University

Press. http://hdl.handle.net/2123/3350

Hagberg S (1960) A rapid method for determining alpha-amylase activity. Cereal Chem. 37:218-222

Harlan JR (1992) Crops and man. Madison: American Society of Agronomy

Heun M, Schäfer-Pregl R, Klawan D, Castagna R, Accerbi M, Borghi B, Salamini F (1997) Site of einkorn wheat domestication identified by DNA fingerprinting. Science 278:1312-1314

Himi E, Mares DJ, Yanagisawa A, Noda K (2002) Effect of grain colour gene (R) on grain dormancy and sensitivity of the embryo to abscisic acid (ABA) in wheat. J Exp Bot. 53:1569–1574

Himi E, Maekawa M, Miura H, Noda K (2011) Development of PCR markers for Tamyb10 related to R-1, red grain color gene in wheat. Theor. Appl. Genet. 122:1561-1576

Huang X, Wei X, Tap S, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T, Zhang Z, Li M, Fan D, Guo Y, Wang A, Wang L, Deng L, Li W, Lu Y,Weng O, Liu K, Huang T, Zhou T, Jing Y, Li W, Lin Z (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. Nat. Genet. 42:961–967

Hucl P (1994) Repeatability of a simplified method for determining sprouting resistance in wheat. Plant Var. Seeds 7: 79–84

Hucl P, Matus-Cádiz M (2002) W98616, a white-seeded spring wheat with increased preharvest sprouting. Can. J. Plant Sci. 82:129-131

Imtiaz M, Ogbonnaya FC, Oman J, Ginkel MV (2008) Characterization of quantitative trait loci controlling genetic variation for preharvest sprouting in synthetic backcross derived wheat lines. Genetics 178:1725–1736

Jacobsen JV, Pearce DW, Poole AT, Pharis RP, Mander LN (2002) Abscisic acid, phaseic acid

and gibberellin contents associated with dormancy and germination in barley. Physiol. Plant 115:428-441

Jaiswal V, Mir RR, Mohan A, Balyan HS, Gupta PK (2012) Association mapping for pre-harvest sprouting tolerance in common wheat (Triticum aestivum L.). Euphytica 188:89-102

Jansen RC, Stam P (1994) High resolution of quantitative traits into multiple loci via interval mapping. Genetics 136:1447–1455

Jansen RC (1995) Genetic Mapping of Quantitative Trait Loci in Plants — a Novel Statistical Approach. Ph.D. thesis, CIP–data Koninklijke Biblotheek, Den Haag, The Netherlands

Jansen RC (1993) Interval mapping of multiple quantitative trait loci. Genetics 135:205–211

Jiang G (2013) Molecular markers and marker-assisted breeding in plants. Plant Breeding from Laboratories to Fields

Jones N, Ougham H, Thomas H, Pasakinskiene I (2009) Markers and mapping revisited: finding your gene. New Phytol. 183:935–966

Kadirvel P, Senthilvel S, Geethanjali S, Sujatha M, Varaprasad KS (2015) Genetic Markers, Trait Mapping and Marker-Assisted Selection in Plant Breeding. In Plant Biology and Biotechnology pp 65-88

Kang H, Zaitlen N, Wade C, Kirby A, Heckerman S, Daly M, Eskin E (2008) Efficient control of population structure in model organism association mapping. Genetics 178:1709-1723

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E (2010) Variance component model to account for sample structure in genome-wide association studies. Nat. Genet. 42:348 U110

Karssen CM, Brinkhorst-Van der Swan DLC, Breekland AE, Koorneef M (1983) Induction of dormancy during seed development by endogenous abscisic acid: studies on abscisic acid

deficient genotypes of Arabidopsis thaliana (L.) Heynh. Planta 157:158-165

Kato K, Nakamura W, Tabiki T, Miura H (2001) Detection of loci controlling seed dormancy on group 4 chromosomes of wheat and comparative mapping with rice and barley genomes. Theor. Appl. Genet. 102:980–985

Kihara H (1944) Discovery of the DD-analyser, one of the ancestors of Triticum vulgare. Agric. Hortic. 19:13-14

Kim KH, Tsao R, Yang R, Cui SW (2006) Phenolic acid profiles and antioxidant activities of wheat bran extracts and the effect of hydrolysis conditions. Food Chem. 95: 466-473

King RW (1982) Abscisic acid in seed development. In AA Khan, ed, The Physiology and Biochemistry of Seed Development, Dormancy and Germination. Elsevier Biomedical Press, Amsterdam, pp 157-181

King RW, Richards RA (1984) Water-uptake in relation to preharvest sprouting damage in wheat—ear characteristics. Aust. J. Agric. Res. 35:327–336

Knox RE, Clarke FR, Clarke JM, Fox SL (2005) Genetic analysis of pre-harvest sprouting in a durum wheat cross. Euphytica 143:261-264

Koornneef M, Hanhart CJ, Hilhorst HWM, Karssen CM (1989) In vivo inhibition of seed development and reserve protein accumulation in recombinants of abscisic acid biosynthesis and responsiveness mutants in Arabidopsis thaliana. Plant Physiol. 90: 463–469

Koornneef M, Reuling G, Karssen CM (1984) The isolation and characterization of abscisic acid-insensitive mutants of *Arabidopsis thaliana*. Physiol. Plant 61:377–383

Koornneef M, Van der Veen JH (1980) lnduction and analysis of gibberellin sensitive mutants in Arabidopsis thaliana (L.) Heynh. Theor. Appl. Genet. 58:257-263

Kottearachchi NS, Uchino N, Kato K, Miura H (2006) Increased grain dormancy in white-

grained wheat by introgression of pre-harvest sprouting tolerance QTL. Euphytica 152:421–428

Kulwal P, Ishikawa G, Benscher D, Feng ZY, Yu LX, Jadhav A, Mehetre S, Sorrells ME (2012) Association mapping for preharvest sprouting resistance in white winter wheat. Theor. Appl. Genet. 125:793–805

Kulwal PL, Singh R, Balyan HS, Gupta PK (2004) Genetic basis of pre-harvest sprouting tolerance using single-locus and two-locus QTL analyses in bread wheat. Funct. Integr. Genomics 4:94–101

Kulwal PL, Kumar N, Gaur A, Khurana P, Khurana JP, Tyagi AK, Balyan HS, Gupta PK (2005) Mapping of a major QTL for pre-harvest sprouting tolerance on chromosome 3A in bread wheat. Theor. Appl. Genet. 111:1052-1059

Kumar A, Kumar J, Singh R, Garg T, Chhuneja P, Balyan HS, Gupta PK (2009) QTL analysis for grain colour and pre-harvest sprouting in bread wheat. Plant Sci. 177:114-122

Kushiro T, Okamoto M, Nakabayashi K, Yamagishi K, Kitamura S, Asami T, Hirai N, Koshiba T, Kamiya Y, Nambara E (2004) The Arabidopsis cytochrome P450 CYP707A encodes ABA 8(-hydroxylase: key enzymes in ABA catabolism. EMBO J. 23:1647-1656

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199

Lan X, Liu D, Wang Z (1997) Inheritance in synthetic hexaploid wheat ‘RSP’of sprouting tolerance derived from Aegilops tauschii Cosson. Euphytica 95:321-323

Landi A (1995) Durum wheat, semolina and pasta quality characteristics for an Italian food company. Durum wheat quality in the Mediterranean region: 33-42

Lateef DD (2015) DNA Marker Technologies in Plants and Applications for Crop

Improvements. J. Biosci. Med 3:7-18

Le Page-Degivry MT, Brthe P, Garello G (1990) Involvement of endogenous abscisic acid in onset and release of Helianthus annuus embryo dormancy. Plant Physiol. 92:1164– 1168

Li M, Liu X, Bradbury P, Yu J, Zhang YM, Todhunter RJ, Buckler ES, Zhang Z (2014) Enrichment of statistical power for genome-wide association studies. BMC Biol. 12:73

Lipka AE, Tian F, Wang QS, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang ZW (2012) GAPIT: genome association and prediction integrated tool. Bioinformatics 28:2397-2399

Lipka AE, Kandianis CB, Hudson ME, Yu J, Drnevich J, Bradbury PJ, Gore MA (2015) From association to prediction: statistical methods for the dissection and selection of complex traits in plants. Curr. Opin Plant Biol. 24:110-118

Liu S, Sehgal SK, Li J, Lin M, Trick HN, Yu J, Gill BS, Bai G (2013) Cloning and characterization of a critical regulator for preharvest sprouting in wheat. Genetics 195:263–273

Liu S, Bai G (2010) Dissection and fine mapping of a major QTL for preharvest sprouting resistance in white wheat Rio Blanco. Theor. Appl. Genet. 121:1395-1404

Liu S, Bai G, Cai S, Chen C (2011) Dissection of genetic components of preharvest sprouting resistance in white wheat. Mol. Breed. 27:511–523

Liu S, Cai S, Graybosch R, Chen C, Bai G (2008) Quantitative trait loci for resistance to pre-harvest sprouting in US hard white winter wheat Rio Blanco. Theor. Appl. Genet. 117:691-699

Liu S, Sehgal SK, Li J, Lin M, Trick HN, Yu J, Gill BS, Bai G (2013) Cloning and characterization of a critical regulator for preharvest sprouting in wheat. Genetics 195:263-

273

Liu S, Sehgal SK, Lin M, Li J, Trick HN, Gill BS, Bai G (2015) Independent mis-splicing mutations in TaPHS1 causing loss of preharvest sprouting (PHS) resistance during wheat domestication. New Phytol. 208:928-935

Lohwasser U, Arif MR, Börner A (2013) Discovery of loci determining pre-harvest sprouting and dormancy in wheat and barley applying segregation and association mapping. Biol. Plantarum 57:663-674

Lohwasser U, Röder MS, Börner A (2005) QTL mapping of the domestication traits pre-harvest sprouting and dormancy in wheat (Triticum aestivum L.). Euphytica 143:247-249

Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic- structure of a tropical understory shrub, Psychotria officinalis (rubiaceae). Am. J. Bot. 82:1420-1425

Lunn GD, Kettlewell PS, Major BJ, Scott RK (2002) Variation in dormancy duration of the UK wheat cultivar Hornet due to environmental conditions during grain development. Euphytica 126: 89-97

MacKey J (1989) Seed dormancy in wild and weedy relatives of cereals. Chapter 2 in 'Preharvest Field Sprouting in Cereals', ed. D.F. Derera, CRC Press, USA, pp 15–25

Mammadov J. Chen W, Mingus J, Thompson S, Kumpatla S (2012) Development of Versatile Gene-Based SNP Assays in Maize (Zea mays L.). Mol. Breed. 29:779-790

Manthey F (2000) Effect of pre-harvest sprouting on durum wheat and pasta quality. In: Proceedings of the Third National Wheat Industry Research Forum, Las Vegas, NV, pp 63

Marcussen T, Sandve SR, Heier L, Spannagl M, Pfeifer M, Consortium TIWGS, Jakobsen KS, Wulff BBH, Steuernagel B, Mayer KFX, Olsen O-A (2014) Ancient hybridizations among the ancestral genomes of wheat. Science 345:288-230

Mares DJ, Mrva K (2001) Mapping quantitative trait loci associated with variation in grain dormancy in Australian wheat. Crop Pasture Sci. 52:1257-1265

Mares D, Mrva K, Cheong J, Williams K, Watson B, Storlie E, Sutherland M, Zou Y (2005) A QTL located on chromosome 4A associated with dormancy in white- and red-grained wheats of diverse origin. Theor. Appl. Genet. 111:1357–1364

Mares D, Mrva K, Tan MK, Sharp P (2002) Dormancy in white-grained wheat: progress towards identification of genes and molecular markers. Euphytica 126:47-53

Mares D, Rathjen J, Mrva K, Cheong J (2009) Genetic and environmental control of dormancy in white-grained wheat (Triticum aestivum L.). Euphytica 168:311-318

Mascher M, Wu S, Amand PS, Stein N, Poland J (2013) Application of genotyping-by-sequencing on semiconductor sequencing platforms: a comparison of genetic and reference-based marker ordering in barley. PLoS One 8:e76925

Maurer A, Draba V, Jiang Y, Schnaithmann F, Sharma R, Schumann E, Kilian B, Reif JC, Pillen K (2015) Modeling the genetic architecture of flowering time control in barley through nested association mapping. BMC genomics 16:1

McCaig TN, Depauw RM (1992) Breeding for preharvest sprouting tolerance in white-seed-coat spring wheat. Crop Sci. 32:19–23

McCallum JA, Walker JRL (1990) Proanthocyanidins in wheat bran. Cereal Chem. 67:282-285

McMullen MD, Kresovich S, Villeda HS, Bradbury P, Li HH, Sun Q, Flint-Garcia S, Thornsberry J, Acharya C, Bottoms C et al. (2009) Genetic properties of the maize nested association mapping population. Science 325:737-740

Metzger RJ, Silbaugh BA (1970) Location of genes for seed coat color in hexaploid wheat, Triticum aestivum L. Crop Sci. 10:495–496

Mir RR, Hiremath PJ, Riera-Lizarazu O, Varshney RK (2013) Evolving Molecular Marker Technologies in Plants: From RFLPs to GBS. In: Lübberstedt, T. and Varshney, R.K., Eds. Diagnostics in Plant Breeding, Springer, Berlin, 229-247

Miyamoto T, Everson EH (1958) Biochemical and physiological studies of wheat seed pigmentation. Agron. J. 50:733-734

Mohan M, Nair S, Bhagwat A, Krishna TG, Yano M, Bhatia CR, Sasaki T (1997) Genome mapping, molecular markers and marker-assisted selection in crop plants. Mol. Breed. 3:87-103

Mohan A, Kulwal P, Singh R, Kumar V, Mir RR, Kumar J, Prasad M, Balyan HS, Gupta PK (2009) Genome-wide QTL analysis for pre-harvest sprouting tolerance in bread wheat. Euphytica 168:319-329

Mori M, Uchino N, Chono M, Kato K, Miura H (2005) Mapping QTLs for grain dormancy on wheat chromosome 3A and the group 4 chromosomes, and their combined effect. Theor. Appl. Genet. 110:1315–1323

Morris CF, Paulsen GM (1989) Registration of 5 preharvest sprouting-resistant hard white winter wheat germplasm lines. Crop Sci. 29:246–247

Morris GP, Ramu P, Deshpande SP, Hash CT, Shah T, Upadhyaya HD, Riera-Lizarazu O, Brown PJ, Acharya CB, Mitchell SE, Harriman J, Glaubitz JC, Buckler ES, Kresovich S (2013) Population genomic and genome-wide association studies of agroclimatic traits in sorghum. Proc. Nati. Acad Sci. U.S.A. 110:453-458

Munkvold JD, Tanaka J, Benscher D, Sorrells ME (2009) Mapping quantitative trait loci for preharvest sprouting resistance in white wheat. Theor. Appl. Genet. 119:1223–1235

Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang ZW, Costich DE, Buckler ES (2009) Association

mapping: critical considerations shift from genotyping to experimental design. Plant Cell 21:2194-2202

Nakamura S, Abe F, Kawahigashi H, Nakazono K, Tagiri A, Matsumoto T, Utsugi S, Ogawa T, Handa H, Ishida H, Mori M, Kawaura K, Ogihara Y, Miura H (2011) A wheat homolog of MOTHER OF FT AND TFL1 acts in the regulation of germination. The Plant Cell 23:3215–3229

Nakamura S, Komatsuda T, Miura H (2007) Mapping diploid wheat homologues of Arabidopsis seed ABA signaling genes and QTLs for seed dormancy. Theor. Appl. Genet. 114:1129-1139

Nelson JC, Deynze AEV, Autrique E, Sorrells ME, Autrique E, Yun HL, Negre S, Bernard M, Leroy P (1995) Molecular mapping of wheat: homoeologous group 3. Genome 38:525–533

Nevo E, Korol AB, Beiles A, Fahima T (2013) Domestication of Wheats. Evolution of wild emmer and wheat improvement: population genetics, genetic resources, and genome organization of wheat's progenitor, *Triticum dicoccoides*. Springer Science & Business Media. pp4-7

Nilsson-Ehle H. (1911) Kreuzungsuntersuchungen an hafer und weizen. Lunds Universitets Arsskrift, N.F. Afd 2, Bd 7, No 6. pp 3-84

Nyachiro JM, Clarke FR, DePauw RM, Knox RE, Armstrong KC (2002) Temperature effects on seed germination and expression of seed dormancy in wheat. Euphytica 126:123-127

Ogawa M, Hanada A, Yamauchi Y, Kuwahara A, Kamiya Y, Yamaguchi S (2003) Gibberellin biosynthesis and response during Arabidopsis seed germination. Plant Cell 15:1591-1604

Ogbonnaya FC, Imtiaz M, DePauw RM (2007) Haplotype diversity of preharvest sprouting QTLs in wheat. Genome 50:107-118

Ogbonnaya FC, Imtiaz M, Ye G, Hearnden PR, Hernandez E, Eastwood RF, van Ginkel M, Shorter SC, Winchester JM (2008) Genetic and QTL analyses of seed dormancy and preharvest sprouting resistance in the wheat germplasm CN10955. Theor. Appl. Genet. 116:891–902

Olsen M, Hood L, Cantor C, Botstein D (1989) A common language for physical mapping of the human genome. Science 245:1434–1435

Orellana J, Fern.ndez-Calv.n B, V.zquez JF, Carrillo JM (1993) Mapping of genes controlling seed storageproteins and cytological markers on chromosome 1R of rye. Theor. Appl. Genet. 85:639–643

Orozco G, Barrett JC, Zeggini E (2010) Synthetic associations in the context of genome-wide association scan signals. Hum. Mol. Genet. 19:R137-R144

Osa M, Kato K, Mori M, Shindo C, Torada A, Miura H (2003) Mapping QTLs for seed dormancy and the Vp1 homologue on chromosome 3A in wheat. Theor. Appl. Genet. 106:1491-1496

Perez-Flores L, Carrari F, Osuna-Fernandez R, Rodrıguez MV, Encisco S, Stanelloni R, Sanchez RA, Bottini R, Iusem ND, Benech-Arnold R (2003) Expression analysis of a GA 20-oxidase in embryos from two sorghum lines with contrasting dormancy: possible participation of this gene in the hormonal control of germination. J. Exp. Bot. 54:2071-2079

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

Poland JA, Bradbury PJ, Buckler ES, Nelson RJ (2011) Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. Proc. Nati. Acad Sci.

U.S.A. 108:6893-6898

Poland JA, Brown PJ, Sorrells ME, Jannink JL (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PloS One, 7:e32253

Pozniak CJ, Knox RE, Clarke FR, Clarke JM (2007) Identification of QTL and association of a phytoene synthase gene with endosperm colour in durum wheat. Theor. Appl. Genet. 114:525-537

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 38:904-909

Pritchard JK, Stephens M, Rosenberg NA, Donnelly P (2000) Association mapping in structured populations. Am. J. Hum. Genet. 67:170-181

Quarrie SA (1987) Use of genotypes differing in endogenous abscisic acid levels in studies of physiology and development. In GV Hoad, JR Lenton, MB Jackson, RK Atkin, eds, Hormone Action in Plant Development. Butterworths, London, pp 89-105

Quatrano RS (1987) Regulation of gene expression by abscisic acid during angiosperm embryo development. In Oxford Surveys of Plant Molecular and Cell Biology, Vol 3. Oxford Univ. Press, Oxford, UK. In press

Quatrano RS, Ballo BL, Williamson JD, Hamblin MT, Mansfield M (1983) ABA controlled expression ofembryo-specific genes during wheat grain development. In R Goldberg, ed, Plant Molecular Biology, Liss Inc., New York, pp 343-353

Rasul G, Humphreys DG, Brule-Babel A, McCartney CA, Knox RE, DePauw RM, Somers DJ (2009) Mapping QTLs for pre-harvest sprouting traits in the spring wheat cross

‘RL4452/AC Domain’. Euphytica 168:363-378

Rathjen JR, Strounina EV, Mares DJ (2009) Water movement into dormant and non-dormant wheat (Triticum aestivum L.) grains. J. Exp. Bot. 60:1619-1631

Rehman Arif MA, Neumann K, Nagel M, Kobiljski B, Lohwasser U, Börner A (2012) An association mapping analysis of dormancy and pre-harvest sprouting in wheat. Euphytica 188:409–417

Ried JL, Walker-Simmons MK (1990) Synthesis of abscisic acid-responsive, heat-stable proteins in embryonic axes of dormant wheat grain. Plant Physiol. 93:662-667

Romay MC, Millard MJ, Glaubitz JC, Peiffer JA, Swarts KL, Casstevens TM, Elshire RJ, Acharya CB, Mitchell SE, Flint-Garcia SA (2013) Comprehensive genotyping of the USA national maize inbred seed bank. Genome Biol. 14:R55

Roy JK, Prasad M, Varshney RK, Balyan HS, Blake TK, Dhaliwal HS, Singh H, Edwards KJ, Gupta PK (1999) Identification of a microsatellite on chromosomes 6B and a STS on 7D of bread wheat showing an association with preharvest sprouting tolerance. Theor. Appl. Genet. 99:336-340

Salse J, Chagué V, Bolot S, Magdelenat G, Huneau C, Pont C, Belcram H, Couloux Gardais A, Evrard A, Segurens B, Charles M, Ravel C, Samain S, Charmet G, Boudet N, Chalhou B (2008) New insights into the origin of the B genome of hexaploid wheat: Evolutionary relationships at the SPA genomic region with the S genome of the diploid relative Aegilops speltoides. BMC Genomics 9:555

Santos AG, Livingston DP, Jellen EN, Wooten DR, Murphy JP (2006) A cytological marker associated with winter hardiness in oat. Crop Sci. 46:203–208

Sears ER (1994) Cytogenetic studies with polyploid species of wheat. II. Additional

chromosomal aberrations in Triticum vulgare. Genetics 29:232–246

Seshu DV, Sorrells ME (1986) Genetic studies on seed dormancy in rice. In: RI IR (ed) Rice genetics. IRRI, Manila, pp 369–382

Shewry PR (2009). Wheat. J. Exp. Bot. 60:1537-1553

Singh R, Matus-Cádiz M, Båga M, Hucl P, Chibbar RN (2010) Identification of genomic regions associated with seed dormancy in white grained wheat. Euphytica 174:391–408

Slageren MW van (1994) Wild wheats: a monograph of Aegilops L. and Amblyopyrum (Jaub. & Spach) Eig (Poaceae). Wageningen Agric. Univ. Press, Wageningen

Soller M, Brody T, Genizi A(1979) On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. Theor. Appl. Genet. 47:35–39

Sonah H, O'Donoughue L, Cober E, Rajcan I, Belzile F (2015) Identification of loci governing eight agronomic traits using a GBS-GWAS approach and validation by QTL mapping in soya bean. Plant Biotech. J. 13:211-221

Steinbach HS, Benech–Arnold RL, Kristof G, Sánchez RA, Marcucci-Poltri S (1995) Physiological basis of pre-harvest sprouting resistance in Sorghum bicolor (L.) Moench. ABA levels and sensitivity in developing embryos of sprouting-resistant and -susceptible varieties. J. Exp. Bot. 46:701-709

Stoy V, Sundin K (1976) Effects of growth regulating substances in cereal grain germination. Cereal Res. Commun. 4:157-163

Suzuki T, Matsuura T, Kawakami N, Noda K (2000) Accumulation and leakage of abscisic acid during embryo development and seed dormancy in wheat. Plant Growth Regul. 30:253-260

Tan MK, Sharp PJ, Lu MQ, Howes N (2006) Genetics of grain dormancy in a white wheat. Crop

Pasture Sci. 57:1157–1165

Tanksley SD (1993) Mapping polygenes. Annu. Rev. Genet. 27:205–233

Tereshchenko O, Gordeeva E, Arbuzova V, Börner A, Khlestkina E (2012) The D genome carries a gene determining purple grain colour in wheat. Cereal Res. Commun. 40:334-341

Thudi M, Li Y, Jackson SA, May GD, Varshney RK (2012) Current State-of-Art of Sequencing Technologies for Plant Genomics Research. Brief. Funct. Genomics 11:3-11

Torada A, Amano Y (2002) Effect of seed coat color on seed dormancy in different environments. Euphytica 126:99-105

Torada A, Ikeguchi S, Koike M (2005) Mapping and validation of PCR-based markers associated with a major QTL for seed dormancy in wheat. Euphytica 143:251–255

Torada A, Koike M, Ikeguchi S, Tsutsui I (2008) Mapping of a major locus controlling seed dormancy using backcrossed progenies in wheat (Triticum aestivum L.). Genome 51:426-432

Torada A, Koike M, Ogawa T, Takenouchi Y, Tadamura K, Wu J, Matsumoto T, Kawaura K, Ogihara Y (2016) A causal gene for seed dormancy on wheat chromosome 4A encodes a MAP kinase kinase. Curr. Biol. 26:782-787

Trethowan RM (1995) Evaluation and selection of bread wheat (Triticum aestivum L.) for preharvest sprouting tolerance. Crop Pasture Sci. 46:463-474

van Oeveren J, de Ruiter M, Jesse T, van der Poel H, Tang J, Yalcin F, Janssen A, Volpin H, Stormo KE, Bogden R, van Eijk MJT, Prins M (2011) Sequence-Based Physical Mapping of Complex Genomes by Whole Genome Profiling. Genome Res. 21:618-625

Vos P, Hogers R, Bleeker M, Reijans M, Van De Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids

Res. 23:4407–4414

Walker-Simmons M (1988) Enhancement of ABA responsiveness in wheat embryos by high temperature. Plant Cell Environ. 11:769-775

Walker-Simmons M (1987) Embryonic ABA levels and sensitivity in developing wheat embryos of sprouting resistant and susceptible cultivars. Plant Physiol. 84:61-66

Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE, Maccaferri M, Salvi S, Milner SG, Cattivelli L (2014) Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. Plant Biotech. J. 12:787-796

Wang RRC, Bothmer R von, Dvorak J, Fedak G, Linde-Laursen I, Muramatsu M (1996) Genome symbols in the Triticeae (Poaceae). In: Wang, R.R.-C., Jensen, K.B., Jaussi, C. (Eds.), Proc. 2nd Int. Triticeae Symp.. Utah State Univ., Logan, pp 29–34

Warner RL, Kudrna DA, Spaeth SC, Jones SS (2000) Dormancy in white-grain mutants of Chinese Spring wheat (Triticum aestivum L.). Seed Sci. Res. 10:51–60

Watanabe N, Ikebata N (2000) The effects of homoeologous group 3 chromosomes on grain colour dependent seed dormancy and brittle rachis in tetraploid wheat. Euphytica 115:215-220

Williams JGK, Kubelik AR, Livak KJ, Rafalsky JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18:6531–6535

Würschum T (2012) Mapping QTL for agronomic traits in breeding populations. Theor. Appl. Genet. 125:201-210

Xia LQ, Yang Y, Ma YZ, Chen XM, He ZH, Röder MS, Jhnes HD, Shewry PR (2009) What can the Viviparous-1 gene tell us about wheat pre-harvest sprouting? Euphytica 168:385-394

Xiao SH, Zhang XY, Yan CS, Lin H (2002) Germplasm improvement for preharvest sprouting resistance in Chinese white-grained wheat: an overview of the current strategy. Euphytica 126:35-38

Xing Y (2014) SNP Array: A Powerful Platform to Accelerate Genetic Studies and Breeding. J. Plant Bioch. Physiol.

Xu Y (2010) Molecular plant breeding. CAB International

Yang Y, Zhao XL, Xia LQ, Chen XM, Xia XC, Yu Z, He ZH, Röder M (2007) Development and validation of a Viviparous-1 STS marker for pre-harvest sprouting tolerance in Chinese wheats. Theor. Appl. Genet. 115:971-980

Young ND (1996) QTL mapping and quantitative disease resistance in plants. Annu. Rev. Phytopathol. 34:479-501

Yu JM, Holland JB, McMullen MD, Buckler ES (2008) Genetic design and statistical power of nested association mapping in maize. Genetics 178:539-551

Yu JM, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler E (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38:203-208

Zanetti S, Winzeler M, Keller M, Keller B, Messmer M (2000) Genetic analysis of pre-harvest sprouting resistance in a wheat x spelt cross. Crop Sci. Madison 40:1406-1417

Zeng ZB (1994) Precision mapping of quantitative trait loci. Genetics 136:1457–1468

Zeng ZB (1993) Theoretical basis of precision mapping of quantitative trait loci. Proc. Natl. Acad. Sci. U.S.A. 90:10972–10976

Zhang ZW, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu JM, Arnett DK, Ordovas JM, Buckler ES (2010) Mixed linear model approach adapted for genome-

wide association studies. Nat. Genet. 42:355-360

Zhang XQ, Li C, Tay A, Lance R, Mares D, Cheong J, Cakir M, Ma J, Appels R (2008) A new PCR-based marker on chromosome 4AL for resistance to pre-harvest sprouting in wheat (Triticum aestivum L.). Mol. Breed. 22:227-236

Zhang Y, Miao X, Xia X, He Z (2014) Cloning of seed dormancy genes (TaSdr) associated with tolerance to pre-harvest sprouting in common wheat and development of a functional marker. Theor. Appl. Genet. 127:855-866

Zhao K, Tung C-W, Eizenga GC, Wright MH, Ali ML, Price AH, Norton GJ, Islam MR, Reynolds A, Mezey J (2011) Genome-wide association mapping reveals a rich genetic architecture of complex traits in Oryza sativa. Nat. Commun. 2:467

Zhu CS, Gore MA, Buckler ES, Yu J (2008) Status and prospects of association mapping in plants. Plant Genome 1:5-20

Chapter 2 - Genotyping-by-Sequencing (GBS) Identified SNPs Tightly Linked to QTLs for Pre-harvest Sprouting Resistance

Abstract

Pre-harvest sprouting (PHS) is a major constraint to wheat production in many growing area worldwide. It reduces not only the end-use quality of wheat flour, but also grain yield. To identify markers tightly linked to the quantitative trait loci (QTLS) for PHS resistance and seed dormancy (SD), we evaluated 155 recombinant inbred lines (RILs) derived from the cross between a PHS-resistant parent TutoumaiA and a PHS-susceptible parent Siyang936 for single nucleotide polymorphisms (SNPs) using genotyping-by-sequencing (GBS), and for PHS resistance and SD using both field and greenhouse grown plants. Two SNPs, *GBS109947* and *GBS212432*, were mapped to a major QTL region for PHS and SD on chromosome 4AL, and delimited the QTL to a 2.9 cM interval. Two and nine additional SNPs were mapped to minor QTL regions for SD on chromosome 5B and 5A, respectively. Selected SNPs in these QTL regions were converted into kompetitive allele specific PCR (KASP) assays that can be easily used for marker-assisted selection to improve PHS resistance.

Introduction

Pre-harvest sprouting (PHS) in wheat (*Triticum aestivum L.*) can cause significant reduction in grain yield and grain end–use quality, thus a substantial reduction in grain price (Groos *et al*. 2002; Mares *et al*. 2005) due to germination of grain in a matured wheat spike before harvesting. It usually occurs when continuous wet weather is available before harvest. Growing PHS resistant cultivars is the most effective way to minimize the PHS damage, especially in wheat growing areas where wet weather occurs frequently during harvest seasons.

Seed dormancy (SD) has been considered the major factor that determines PHS resistance in wheat and other cereal crops (Bewley and Black 1982; Anderson *et al*. 1993; Mares and Mrva 2001; Ogbonnaya *et al*. 2008), although several other factors have also been considered to contribute to overall PHS resistance, including physical barriers to water penetration (Gale *et al*. 1989), spike morphology (King and Richardd 1984), red seed color (Gfeller and Svejda 1960; Groos *et al*. 2002) and environment factors such as temperature and moisture (Argel *et al*. 1983; Ceccato *et al*. 2011). Both PHS and SD are complex traits controlled by several quantitative genetic loci (QTLs). For PHS resistance, one QTL on chromosome 3AS showed a major effect on PHS resistance (Osa *et al*. 2003; Mori *et al*. 2005; Liu *et al*. 2008), and the casual gene of this QTL for both SD and PHS resistance has been cloned (Nakamura *et al*. 2011; Liu *et al*. 2013). Another major QTL has been identified on chromosome 4AL in different genetic backgrounds (Kato *et al*. 2001; Mares and Mrva 2005; Torada *et al*. 2005; Chen *et al*. 2008; Ogbonnaya *et al*. 2008). In addition, QTLs with minor effects have been reported on 2B (Kulwal *et al*. 2004; Munkvold *et al*. 2009), 3D (Imtiaz *et al*. 2008), 4B and 4D (Kato *et al*. 2001), 6B and 7D (Roy *et al*. 1999), and several other chromosomes (Anderson *et al*. 1993). For SD, major QTLs were reported on 3A (Osa *et al*. 2003; Mori *et al*. 2005) and 4A (Kato *et al*. 2001; Noda *et al*. 2002;

Mares and Mrva *et al*. 2005). However, how much SD contributes to PHS resistance still remains unknown. Therefore, simultaneously mapping QTLs for both PHS resistance and SD may reveal the genetic relationship between the two traits.

High-density genetic maps are essential for QTL fine mapping and delimiting the casual genes to very narrow genetic intervals (Liu *et al*. 2014). More recently, next generation sequencing (NGS) technology has been used for QTL mapping in many crops (Wicker *et al*. 2008; Kobayashi *et al*. 2014; Chen *et al*. 2014). Wheat is polyploid, thus has a large genome (~17 GB) and abundant repetitive DNA sequences, which complicates analysis of genetic variations and development of high-resolution genetic maps. Recently, a genotyping-by-sequencing (GBS) protocol has been adapted in wheat by using restriction digestion to reduce the complexity of the genome (Poland *et al*. 2012). GBS takes the advantages of NGS, and keeps the sequencing costs down by multiplexing samples using barcodes. Although complete reference genome sequences can increase the efficiency of SNP identification in different species (Poland *et al*. 2012; Spindel *et al*. 2013) and it is not available in wheat. Fortunately, analytical pipeline is now available for species with incomplete or no reference genome sequences (Mascher *et al*. 2013). The objectives of this study were to (1) fine map QTLs for both PHS resistance and SD in a Chinese landrace using GBS-SNPs, (2) develop closely linked DNA markers to the QTLs for marker-assisted selection in wheat breeding programs, and (3) elucidate the genetic relationship between SD and PHS resistance.

Materials and Methods

Plant materials and experimental design

A mapping population of 155 RILs derived from the cross TotoumaiA x Siyang936 was developed by single-seed decent. TotoumaiA is a white PHS-resistant Chinese landrace, while

Siyang936 is a white PHS-susceptible cultivar from China. Both parents and the RILs were evaluated for PHS resistance using plants collected from two field experiments (2005 and 2006) at Jiangsu Academy of Agriculture Sciences (JAAS), Nanjing, China, and from three greenhouse experiments (2005 to 2007) at Kansas State University (KSU), Manhattan, KS. Seed dormancy was evaluated using plants grown in the five experiments from 2004 to 2006 in both locations. Each experiment was arranged in a randomized complete block design with two replicates. Also, a natural population of 380 accessions from the USA and China was used to test allelic diversity of SNPs closely linked to the 4A QTL for PHS resistance and to evaluate the potential use of these SNPs in marker-assisted selection.

Evaluation of SD and PHS

In the greenhouse experiments, plants were grown at 22 ± 5 day/15 ± 2 night temperature with supplemental daylight of 12 h. Pre-harvest sprouting was evaluated in the laboratory using intact spikes. When wheat spikes reached physiological maturity, five spikes per RIL were harvested from each replicate and air-dried for 5 d in a greenhouse. Harvested spikes were stored at -20℃ to maintain dormancy. After all RILs were collected, spikes were air-dried again for 2 d and immersed in de-ionized water for 5 h. The wet spikes were incubated in a moist chamber set up in the laboratory at 22 ± 1℃ with 100% humidity maintained by running a humidifier for 30 min twice a day. At 7^{th} d of incubation, the numbers of germinated and non-germinated seeds in each spike were counted, and PHS resistance was measured as percentage of visible sprouted kernels (PVSK) in a spike. For SD test, 50 hand-threshed kernels from the remaining spikes in each RIL were evaluated for seed germination rate in the laboratory, and a weighted germination index (GI) was calculated to reflect SD as previously described (Chen *et al*. 2008).

In the field experiments, each RIL and their parents were sowed in a two-row plot with 4-

m-long at 0.25 m apart. At physiological maturity, when the spike and peduncle turned yellow, 20 spikes per plot (10 spikes per row) were harvested. Harvested spikes were stored and evaluated for both PHS and SD as previously described for the greenhouse experiment, with the exception that 10 spikes per RIL were used for field experiments instead of five for greenhouse experiments.

GBS library construction and SNP calling

Genomic DNA of parents and their RILs was extracted using a modified cetyltrimethyl ammonium bromide (CTAB) method (Saghai-Maroof *et al*. 1984). DNA concentration was quantified using Quant-iT™ PicoGreen® dsDNA Assay (Lifetechnologies) and normalized to 20ng per ul. The GBS library was constructed as previously described (Poland *et al*. 2012). In brief, DNA samples were digested with *HF- PstI* and *MspI* (New England BioLabs Inc., Ipswich, MA) and then ligated to barcoded adaptors and a Y common adaptor using T4 ligase (New England BioLabs Inc.). Ligation products were pooled and cleaned up using QIAquick PCR Purification Kit (Qiagen). Primers complementary to both adaptors were used for PCR. The PCR product was then cleaned up again using QIAquick PCR Purification Kit, size-selected with a range of 250 - 300 bp in an E-gel system (Life Technologies Inc., NY 14072) and concentration estimated by the Qubit 2.0 fluorometer using Qubit dsDNA HS Assay Kit (Life Technologies Inc., NY 14072). The size-selected library was sequenced on an Ion Proton system (Life Technologies Inc., NY 14072).

SNP calling used the pipeline developed by Saintenac *et al*. (2013). Reads generated by Ion Proton were trimmed by removing bases with phred33 quality score <15 from both sides. Reads were also removed if more than 20% of bases having quality score <15. Sequences from each parent were clustered, and the clusters that differed from each other by no more than three

mismatches were used as reference sequences. Reads were aligned to the reference using *bowtie* (Langmead *et al*. 2009) with parameter set at -v 3 –k 1. Since RILs were used in library construction, SNPs with heterozygotes >10% of total RILs were discarded to reduce the false positive results. SNPs with missing data <50% were used for mapping.

Genetic map construction and QTL analysis

A linkage map was constructed using SNP data from GBS (GBS-SNP) and previously reported SSR data (Liu *et al*. 2011) using Regression function in JoinMap version 4.0 (Van Ooijen and Voorrips 2006). Recombination fractions were converted into centiMorgans (cM) using the Kosambi function (Kosambi 1944). Composite interval mapping (CIM) was performed for each experiment and lines means across environments using WinQTLCart 2.5 (Wang *et al*. 2005). LOD threshold of 2.24 was determined from 1000 permutation tests (Doerge and Churchill 1996) to claim significant QTLs.

Results

GBS-SNP calling

A combination of *PstI* and *MspI* restriction enzymes was used to reduce the wheat genome complexity. GBS generated a total of 87 million reads in one run of Ion Proton. After initial filtering, 82 million reads met the quality score. A total of 3180 GBS-SNPs were called at <20 % missing data, and 8623 GBS-SNPs were called at <50% missing data in the population.

Map construction

Totally, all GBS-SNPs with <50% missing data and 93 SSRs were used to construct the linkage map, and 2029 GBS-SNPs and 43 SSRs were mapped into 63 linkage groups. The linkage map covered a total length of 2646.82 cM in genetic distance with an average interval length of 1.28 cM. The number of markers per linkage group ranged from 5 to 175.

Seed dormancy and PHS resistance in parents and RIL

The PVSK ranged from 6.8 to 48.4 % for Tutumai A and from 43.9 to 90.8 % for Siyang 936, and the GI ranged from 18.2 to 62.3 % for Tutoumai A and from 61.2 to 92.7 % for Siyang 936 in the five experiments conducted at JAAS and KSU. Tutoumai A had about 35 and 40 % lower PVSK and GI ratings than these for Siyang 936 in an average, although large variations in each trait were observed for each parents among experiments. Both traits showed continuous distributions in the RIL population, and transgressive segregation was observed for both traits, indicating that both parents might contribute favorable alleles.

QTL mapping

CIM detected four QTLs on different linkage groups. A major QTL was detected on chromosome 4A for both PHS resistance and SD with two SSRs and two SNPs mapped in the QTL region (Fig. 2.1A). One QTL each for PHS resistance were detected on chromosome 5B and 5A (Fig. 2.1B, 2.1C), and one QTL for both PHS resistance and SD was detected on chromosome 4B (Fig. 2.1D). Two GBS-SNPs were mapped to the 4A QTL region, two were mapped to the 5B QTL region, nine mapped to the 5A QTL region, and none were mapped to 4B QTL region.

To verify the genotypic data generated by GBS and to eliminate missing data for markers in QTL regions, 26 KASP assays were designed from the corresponding GBS sequences harboring SNPs that were mapped within or around these QTL regions. Eleven KASP-SNP markers amplified well and showed polymorphism between parents and among the RILs, and seven of them were remapped to three of the QTL regions (Table 2.1). The other four SNPs shifted position and moved outside the QTL regions after all missing data at these loci were filled by KASP-SNP and errors were corrected. Comparison between GBS-SNP and KASP-SNP

data found that seven SNPs showed exactly identical genotypes in the RILs between GBS and KASP assays, and four KASP-SNPs did not match with GBS-SNPs because two GBS-SNPs had a SNP calling error in one RIL, one had errors in five RILs, and one had errors in 16 RILs. Therefore, the average error rate for the eleven SNPs caused by either sequencing or SNP calling was 1.35%.

The QTL with the largest effect, *Qphs.pseru-4A,* was delimited to a 2.9 cM interval between *GBS212432* and *GBS109947* (Fig. 2.1A) and explained 8.3 to 17.2% phenotypic variances for PHS resistance and 9.4 to 26.5% for SD (Table 2.2). On one side of the QTL, both markers *Xbarc170* and *GBS109947* showed the largest effect on PHS resistance and SD among all markers tested in all the experiments (Table 2.3); on the other side of the QTL, however, *GBS212432* had much greater effects than *Xgwm397* on both traits measured (Table 2.3), thus *GBS212432* was more closely linked marker to the QTL than *Xgwm397,* and *GBS212432* and *GBS109947* flanked the QTL.

Qphs.pseru-5B was detected in two JAAS experiments and one KSU experiment that accounted for 5.5 -12.5% phenotypic variances on PHS. However, this QTL was not detected in any SD experiment (Fig. 2.1B; Table 2.2). Two SNPs mapped to this QTL region, and this QTL was linked closely to the SSR marker *Xbarc275* in these experiments (Fig. 2.1B).

Qphs.pseru-5A was another QTL identified for PHS resistance. It was detected in the two JAAS experiments and significant for the overall mean of germination rate, and explained 7.7% to 15.5% phenotypic variances (Fig. 2.1C; Table 2.2). Nine GBS-SNPs together with two SSRs were mapped to this QTL region, and the SSRs were the most closely linked markers to the QTL (Fig. 2.1C).

Qphs.pseru-4B was identified for both PHS resistance and SD in four experiments, and explained 6.3 to 8.7% phenotypic variances. However, GBS-SNPs were not mapped to the QTL region (Fig. 2.1D; Table 2.2).

Allele diversity of SNPs in 4A QTL region

The QTL on chromosome 4A was detected in three KSU experiments for PHS resistance and in all the experiments for SD. This QTL explained up to 22.3% of the phenotypic variance for PHS resistance and 28.7% of the phenotypic variance for longer SD over all the experiments (Table 2.2). It is more likely a stable QTL with a major effect on PHS resistance and SD. To evaluate the potential efficiency of marker-assisted selection using these markers, four markers tightly linked to the QTL were used to estimate the selection progress. Difference between mean sprouting rates of individuals carrying contrasting alleles of *GBS109947* was similar to that of *Xbarc170*. On the other side of the QTL, *GBS212432* showed a larger contrast in spouting rates between two alleles than that between two alleles of *Xgwm397* (Table 2.4), indicating *GBS212432* is the closer marker to the QTL than *Xgwm397*.

The four markers, *GBS109947*, *GBS212432*, *Xbarc170* and *Xgwm397*, were used to screen a natural population consisting of 205 U.S., 146 Chinese, 26 Japanese and 3 Korean wheat lines or cultivars. A total of 21 alleles of *Xbarc170* were identified with a low polymorphism information content (PIC) value of 0.11, and 14 alleles of *Xgwm397* were identified with a PIC value of 0.22. For *GBS212432*, 168 accessions had the same allele as Tutoumai A, 131 had the same allele as Siyang936, and 81 accessions showed neither parental genotypes. Surprisingly, at the locus of *GBS109947*, only 3 accessions carried the same allele as Siyang936, indicating Siyang936 had a rare allele at this locus (Table 2.5).

Discussion

Evaluation of PHS and SD

PHS is a complicated trait, and many factors may contribute to PHS resistance, including SD, seed color, and other morphological characteristics. In addition, environment factors, such as temperature and moisture during mature period, can also interfere the expression of PHS resistance. Therefore, repeated experiments are critical in providing increased accuracy in PHS resistance estimation. In this study, we conducted five experiments to estimate PHS resistance and SD. To exclude possible effects from morphological traits, spikes were harvested at physiological maturity, dried for a fixed period, and soaked in distilled water overnight. Therefore, environmental interference on phenotyping procedure was minimized. The sprouting index (SI) has been used as a standard method to measure the germination rate (Anderson *et al*. 1993; Kulwal *et al*. 2004). Chen *et al*. (2008) and Imtiaz *et al*. (2008) used percentage of visually sprouted seeds (VSS) to measure germination rate, and proved that VSS gave a more accurate PHS rating than SI. The current study used this same measurement to measure overall PHS resistance.

QTLs for PHS resistance and SD in wheat

In this study, four QTLs were detected for PHS resistance and two of them were detected for long seed dormancy. Many QTLs for PHS resistance have been reported on different chromosomes in previous studies. Anderson *et al*. (1993) detected several genetic regions on chromosomes 1AS, 3BL, 4AL, 5DL and 6BL associated with PHS resistance, whereas Zanetti *et al*. (2000) reported QTLs on chromosome 3B, 5A, 6A and 7B. QTLs for PHS resistance were detected on chromosome 5A and group 3 where the kernel color genes were previously reported (Groos *et al*. 2002), and also on chromosome 6B and 7B (Roy *et al*. 1999). For SD, major QTLs

were mainly reported on 3A (Osa *et al*. 2003; Mori *et al*. 2005) and 4A (Kato *et al*. 2001; Noda *et al*. 2002; Mares and Mrva *et al*. 2005). In this study, PHS resistance and SD were evaluated in the same experiments. Therefore, we were able to estimate QTL effects on both PHS resistance and seed dormancy.

The QTL on chromosome 4A was detected in three KSU greenhouse experiments for PHS resistance and all the experiments for SD, and explained up to 17.2% and 26.5% phenotypic variance for PHS resistance and SD, respectively. This indicated that *Qphs.pseru-4A* is a very stable QTL with a major effect on both PHS resistance and SD, and validated that SD was the most important factor for PHS resistance.

Another QTL on chromosome 5B was detected only for PHS resistance, not for SD, suggesting this QTL may contribute to PHS resistance due to factors other than SD. QTL for PHS resistance on chromosome 5B have been reported in previous studies (Groos *et al*. 2002; Tan *et al*. 2006), but we were unable to determine whether they are the same QTL due to lacking of common markers among these QTLs. Similarly, the QTL detected on 5A was also only for PHS resistance. Groos *et al*. (2002) and Nakamura *et al*. (2007) reported a QTL on chromosome 5AS for PHS resistance, but common markers were not found between those and our studies. One QTL was detected on chromosome 4B, and showed minor effects on PHS resistance and SD. QTL for PHS resistance and SD was also reported on chromosome 4B previously (Kato *et al*. 2001; Mori *et al*. 2005; Mohan *et al*. 2009; Rasul *et al*. 2009), but common markers among these QTLs are lacking to determine if they are the same QTL.

We were not able to detect the QTL for PHS resistance on chromosome 3A, *TaPHS1*, in this study. The functional SNP of *TaPHS1* is not polymorphic between TutoumaiA and Siyang936. Two SSRs closely linked to the 3A QTL, *Xbarc57* and *Xbarc321*, did not show

polymorphism in the population either. *Xwmc11* was the closest polymorphic marker to this QTL in this study (data not shown), but it was at least 30 cM away from the QTL (Song et al, 2005, Liu et al, 2008). Therefore, it is more likely tha both parents carry the same allele at the 3A QTL.

Efficiency of GBS and KASP

The application of GBS facilitates generation of high-density genetic maps at a low cost (Poland *et al*. 2012). High-resolution maps have been created with GBS-SNPs in sorghum, wheat, rice and barley, and maps saturated with GBS-SNPs have proven to be very useful for fine mapping of QTLs for different traits and identification of candidate genes for gene cloning (Poland *et al*. 2012; Saintenac *et al*. 2013; Liu *et al*. 2014; Spindel *et al*. 2013). One disadvantage of GBS-SNPs for mapping is a large number of missing data for some markers due to limitation in sequencing depth; therefore, imputation method is recommended to predict genotypes with missing data (Poland *et al*. 2012; Spindel *et al*. 2013; Sonah *et al*. 2013). Another way to increase data quality is to use high quality SNPs with missing data <20% without imputation (Liu *et al*. 2014), but such implement would probably result in loss of some important SNPs that have more than 20% missing data. In this study, we used a different strategy. At first, we used GBS-SNPs with <50% missing data to construct initial map to scan QTLs; and then convert GBS-SNPs from the QTL regions to KASP-SNPs to confirm GBS-SNPs in the QTL regions. Using this method, initially, more than 8,000 SNPs were scored from one Ion Proton run, together with SSR anchoring markers, a high-density genetic map was generated with 2029 SNPs and 43 SSRs. Missing data and sequencing errors may cause an expansion of genetic distance between markers in the initial genetic map, but it includes many more SNPs than the map developed using SNPs with <20% missing data. We validated GBS-SNPs with KASP-SNP assays, which minimized negative effect of missing data and corrected the sequencing errors in

the QTL regions, thus improved accuracy of in the QTL regions. Among 26 KASP assays designed, 11 worked very well in the RILs. Among these working KASP-SNPs, seven agreed with GBS-SNP calls among RILs. However, four had SNP call errors with one having wrong SNP calls in 16 RILs. These errors could be due to error from either sequencing or SNP calling pipeline. Thus, reducing sequence error and improve SNP call quality will minimize genotyping error. Conversion of GBS-SNP to KASP-SNP improves QTL mapping quality. Other KASP assays did not amplify well mainly because of short sequence reads that result in difficulty in primer design that cannot generate optimal primers for SNP amplification.

With new GBS-SNP map developed from the same population reported in the previous study (Liu *et al*. 2011), we not only identified the same QTL on chromosome 4A, 5B and 4B reported previously, but also a new QTL on 5A. The new QTL on 5A detected in this study, not in the previous study, is because the QTL was mapped in a large linkage group of GBS-SNPs and two SSRs; whereas in the previous study, the two SSRs did not form a linkage group thus were not used in QTL analysis. Therefore, GBS is an effective marker system for SNP discovery, and useful for new QTL identification and QTL fine mapping.

Mapping resolution was significantly increased in the 4A and the 5B QTL regions by adding GBS-SNPs in these regions. In our previous study, QTL in 4A was mapped in a 9.1 cM genetic interval (Chen *et al*. 2008), using GBS-SNPs in this study, it was mapped to a 2.9 cM interval between two SNPs, *GBS212432* and *GBS109947*. The 4A QTL showing major effect is a good candidate for map-based cloning of PHS resistance gene and the SNPs identified in this study laid a solid foundation for such work.

Application of SNPs in MAS

Since PHS is easily affected by environmental factors and phenotyping of PHS is time-consuming and labor intensive, marker-assisted selection provides a desirable approach to quickly deployment of PHS-resistant QTLs in breeding programs. *GBS212432* and *GBS109947* are the closest markers associated with QTL on chromosome 4A in the population used in this study. However, marker analysis in a natural population indicated that the susceptible allele of *GBS109947* is a rare allele, and it may provide false positive results when it is used as a diagnostic marker to screen a natural population. Since *Xbarc170* showed similar effect as *GBS109947*, it still is valuable marker for MAS. *GBS212432* showed good polymorphism in the natural population (Table 2.4), thus can be used together with *Xbarc170* to increase selection accuracy. In addition, SNPs and SSRs in *Qphs.pseru-5B*, *Qphs.pseru-5A* and *Qphs.pseru-4B* regions can also be valuable in pyramiding PHS resistance QTLs to achieve an increased level of PHS resistance.

References:

Anderson JA, Sorrells ME, Tanksley SD (1993) RFLP analysis of genomic regions associated with resistance to preharvest sprouting in wheat. Crop Sci 33:453–459

Argel PJ, Humphreys LR (1983) Environmental effects on seed development and hardseededness in Stylosanthes hamata cv. Verano. I. Temperature. Crop and Pasture Science 34: 261-270

Bewley JD, Black M (1982) Physiological and biochemistry of seeds in relation to germination, vol 2. Springer, Heidelberg, pp 61–81

Ceccato DV, Daniel Bertero H, Batlla D (2011) Environmental control of dormancy in quinoa (Chenopodium quinoa) seeds: two potential genetic resources for preharvest sprouting tolerance. Seed Science Research 21: 133-141

Chen CX, Cai SB, Bai GH (2008). A major QTL controlling seed dormancy and preharvest sprouting resistance on chromosome 4A in a Chinese wheat landrace. Molecular Breeding 21: 351-358.

Chen Z, Wang B, Dong X, Liu H, Ren L, Chen J, Hauck A, Song W, Lai J (2014) An ultra-high density bin-map for rapid QTL mapping for tassel and ear architecture in a large F-2 maize population. BMC Genomics 15 (433)

Doerge RW, Churchill GA (1996) Permutation tests for multiple loci affecting a quantitative character. Genetics 142:285–294

Gale MD (1989) The genetics of preharvest sprouting in cereals, particularly in wheat. In: Derera NF (ed) Preharvest field sprouting in cereals. CRC Press, Boca Raton, pp 85–110

Gfeller F, Svejda F (1960) Inheritance of post-harvest seed dormancy and kernel color in spring wheat lines. Can J Plant Sci 40:1–6

Gore MA, Fang DD, Poland JA, Zhang J, Percy RG, Cantrell RG, Thyssen G, Lipka AE (2014) Linkage map construction and quantitative trait locus analysis of agronomic and fiber quality traits in cotton. The Plant Genome 7(1)

Groos C, Gay G, Perretant MR, Gervais L, Bernard M, Dedryver F, Charmet G (2002) Study of the relationship between pre-harvest sprouting and grain color by quantitative trait loci analysis in a whitexred grain bread-wheat cross. Theor Appl Genet 104:39–47

Imtiaz M, Ogbonnaya FC, Oman J, Ginkel MV (2008) Characterization of quantitative trait loci controlling genetic variation for preharvest sprouting in synthetic backcrossderived wheat lines. Genetics 178:1725–1736

Kato K, Nakamura W, Tabiki T, Miura H, Sawada S (2001) Detection of loci controlling seed dormancy on group 4 chromosomes of wheat and comparative mapping with rice and barley genomes. Theor Appl Genet 102:980–985

King RW, Richards RA (1984) Water-uptake in relation to preharvest sprouting damage in wheat—ear characteristics. Aust J Agric Res 35:327–336

Kobayashi T, Yamamoto K, Suetsugu Y, Kuwazaki S, Hattori M, Jairin J, Matsumura M (2014) Genetic mapping of the rice resistance-breaking gene of the brown planthopper Nilaparvata lugens. Proceedings of the Royal Society B: Biological Sciences 281(1787), 20140726

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Kulwal PL, Singh R, Balyan HS, Gupta PK (2004) Genetic basis of pre-harvest sprouting tolerance using single-locus and two-locus QTL analyses in bread wheat. Funct Integr Genomics 4:94–101

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment

of short DNA sequences to the human genome. Genome Biol 10: R25

Liu H, Bayer M, Druka A, Russell JR, Hackett CA, Poland J, Ramsay L, Hedley P, Waugh R (2014) An evaluation of genotyping by sequencing (GBS) to map the Breviaristatum-e (ari-e) locus in cultivated barley. BMC genomics 15: 104-114

Liu S, Cai S, Graybosch R, Chen C, Bai G. (2008) Quantitative trait loci for resistance to pre-harvest sprouting in US hard white winter wheat Rio Blanco. Theor Appl Genet 117: 691-699

Liu S, Bai G, Cai S, Chen C (2011). Dissection of genetic components of preharvest sprouting resistance in white wheat. Molecular breeding, 27(4), 511-523.

Liu S, Sehgal SK, Li J, Lin M, Trick HN, Yu J, Gill BS, Bai G (2013) Cloning and characterization of a critical regulator for preharvest sprouting in wheat. Genetics 195: 263-273

Mares DJ, Mrva K (2001) Mapping quantitative trait loci associated with variation in grain dormancy in Australian wheat. Crop and Pasture Science 52: 1257-1265

Mares DJ, Mrva K, Cheong J, Williams K, Watson B, Storlie E, Sutherland M, Zou Y (2005) A QTL located on chromosome 4A associated with dormancy in white- and red-grained wheats of diverse origin. Theor Appl Genet 111:1357–1364

Mascher M, Wu S, Amand PS, Stein N, Poland J (2013). Application of genotyping-by-sequencing on semiconductor sequencing platforms: a comparison of genetic and reference-based marker ordering in barley. PloS one 8: e76925

Mohan A, Kulwal P, Singh R, Kumar V, Mir RR, Kumar J, Prasad M, Balyan HS, Gupta PK (2009) Genome-wide QTL analysis for pre-harvest sprouting tolerance in bread wheat. Euphytica 168: 319-329

Mori M, Uchino N, Chono M, Kato K, Miura H (2005) Mapping QTLs for grain dormancy on wheat chromosome 3A and group 4 chromosomes, and their combined eVect. Theor Appl Genet 110:1315–1323

Munkvold JD, Tanaka J, Benscher D, Sorrells ME (2009) Mapping quantitative trait loci for preharvest sprouting resistance in white wheat. Theor Appl Genet 119:1223–1235

Nakamura S, Komatsuda T, Miura H (2007) Mapping diploid wheat homologues of Arabidopsis seed ABA signaling genes and QTLs for seed dormancy. Theor Appl Genet 114: 1129-1139

Nakamura S, Abe F, Kawahigashi H, Nakazono K, Tagiri A, Matsumoto T, Utsugi S, Ogawa T, Handa H, Ishida H, Mori M, Kawaura K, Ogihara Y, Miura H (2011) A wheat homolog of MOTHER OF FT AND TFL1 acts in the regulation of germination. The Plant Cell Online 23: 3215-3229

Noda K, Matsuura T, Maekawa M, Taketa S (2002) Chromosomes responsible for sensitivity of embryo to abscisic acid and dormancy in wheat. Euphytica 123:203–209

Ogbonnaya FC, Imtiaz M, Ye G, Hearnden PR, Hernandez E, Eastwood RF, Ginkel MV, Shorter SC, Winchester JM (2008) Genetic and QTL analyses of seed dormancy and preharvest sprouting resistance in the wheat germplasm CN10955. Theor Appl Genet 116: 891–902

Osa M, Kato K, Mori M, Shindo C, Torada A, Miura H (2003) Mapping QTLs for seed dormancy and the Vp1 homologue on chromosome 3A in wheat. Theor Appl Genet 106:1491–1496

Poland JA, Brown PJ, Sorrells ME, Jannink JL (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PloS one 7: e32253

Rasul G, Humphreys DG, Brule-Babel A, McCartney CA, Knox RE, DePauw RM, Somers DJ (2009) Mapping QTLs for pre-harvest sprouting traits in the spring wheat cross 'RL4452/AC Domain'. Euphytica168: 363-378

Roy JK, Prasad M, Varshney RK, Balyan HS, Blake TK (1999) Identification of a microsatellite on chromosomes 6B and a STS on 7D of bread wheat showing an association with preharvest sprouting tolerance. Theor Appl Genet 99: 336–340

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA 81:8014–8018

Saintenac C, Jiang D, Wang S, Akhunov E (2013) Sequence-based mapping of the polyploid wheat genome. G3: Genes| Genomes| Genetics 3: 1105-1114

Sonah H, Bastien M, Iquira E, Tardivel A, Légaré G, Boyle B, Normandeau É, Laroche J, Larose S, Jean M, Belzile F (2013) An improved genotyping by sequencing (GBS) approach offering increased versatility and efficiency of SNP discovery and genotyping. PloS one 8(1) e54603

Song QJ, Shi JR, Singh S, Fickus EW, Costa JM, Lewis J, Gill BS, Ward R, Cregan PB. (2005). Development and mapping of microsatellite (SSR) markers in wheat. Theor Appl Genet 110: 550–560

Spindel J, Wright M, Chen C, Cobb J, Gage J, Harrington S, Lorieux M, Ahmadi N, Couch SM (2013) Bridging the genotyping gap: using genotyping by sequencing (GBS) to add high-density SNP markers and new value to traditional bi-parental mapping and breeding populations. Theor Appl Genet 126: 2699-2716

Tan MK, Sharp PJ, Lu MQ, Howes N (2006) Genetics of grain dormancy in a white wheat. Aust

J Agric Res 57: 1157–1165

Torada A, Ikegnchi S, Koike M (2005) Mapping and validation of PCR-based markers associated with a major QTL for seed dormancy in wheat. Euphytica 143:251–255

Van Ooijen JW (2006) JoinMap® 4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma BV, Wageningen, Netherlands

Wang S, Basten CJ, Zeng ZB (2005) Windows QTL Cartographer 2.5. Departmentof Statistics, North Carolina State University, Raliegh. http://statgen.ncsu.edu/qtlcart/WQTLCart.htm

Wicker T, Narechania A, Sabot F, Stein J, Vu GT, Graner A, Stein N (2008) Low-pass shotgun sequencing of the barley genome facilitates rapid identification of genes, conserved non-coding sequences and novel repeats. BMC genomics 9(1), 518

Zanetti S, Winzeler M, Keller M, Keller B, Messmer M (2000) Genetic analysis of pre-harvest sprouting resistance in a wheat x spelt cross. Crop Sci 40:1406–1417

Table 2.1 List of KASPar primers developed from GBS sequences

Primer name[a]	Position	Primer sequence (5'-3')
GBS_212432_T	*Qphs.pseru-4A*	TTCACAGCGCCTCGGCCGCCC
GBS_212432_S	*Qphs.pseru-4A*	TTCACAGCGCCTCGGCCGCCA
GBS_212432_R	*Qphs.pseru-4A*	GTACCACTCTGGTGCACTCC
GBS_109947_T	*Qphs.pseru-4A*	TTAGCCGTGTGACGCCGTGT
GBS_109947_S	*Qphs.pseru-4A*	TTAGCCGTGTGACGCCGTGC
GBS_109947_R	*Qphs.pseru-4A*	GCGTGAATTGCTGACCTCTC
GBS_963571_T	*Qphs.pseru-4A*	CGATCATAGCAGTGGAACGC
GBS_963571_S	*Qphs.pseru-4A*	CGATCATAGCAGTGGAACGT
GBS_963571_R	*Qphs.pseru-4A*	CTCGCACAGTGAAGGTCATT
GBS_T240557_T	*Qphs.pseru-5B*	CAGCTTCAGTGCCTTCCTCG
GBS_T240557_S	*Qphs.pseru-5B*	CAGCTTCAGTGCCTTCCTCA
GBS_T240557_R	*Qphs.pseru-5B*	GAGTGACGTCATCCACAAGG
GBS_T66183_T	*Qphs.pseru-5B*	GGTGGAGGGATTTGGATGATC
GBS_T66183_S	*Qphs.pseru-5B*	GGTGGAGGGATTTGGATGATA
GBS_T66183_R	*Qphs.pseru-5B*	CGTCCTCTTGCTTGATGGTA
GBS_T169803_T	*Qphs.pseru-5B*	GCAGTAATTTTAGTAGCATTC
GBS_T169803_S	*Qphs.pseru-5B*	GCAGTAATTTTAGTAGCATTT
GBS_T169803_R	*Qphs.pseru-5B*	TATTGCTTCATTAGAGGACA
GBS_T162884_T	*Qphs.pseru-4B*	CAAATGTCGCATGTGGCTGC
GBS_T162884_S	*Qphs.pseru-4B*	CAAATGTCGCATGTGGCTGA
GBS_T162884_R	*Qphs.pseru-4B*	CGCGTATGAGCATGATACCT

[a]T Forward primer with TutoumaiA allele, S Forward primer with Siyang936 allele, R Reverse primer

Table 2.2 Putative QTLs for preharvest sprouting resistance (PHS) and seed dormancy (SD) identified by composite interval mapping using spikes and seeds harvested from recombinant inbred lines grown in field trials of 2004, 2005 and 2006 (JAAS, Nanjing, China) and greenhouse trials of 2005, 2006 and 2007 (KSU, Manhattan, KS)

QTL and its location	Marker interval	2004 JAAS		2005 JAAS		2006JAAS		2005 KSU		2006 KSU		2007 KSU		Mean over all experiments	
		LOD[a]	R^2 (%)	LOD	R^2 (%)	LOD	R^2 (%)	LOD	R^2 (%)	LOD	R^2 (%)	LOD	R^2 (%)	LOD	R^2 (%)
PHS															
Qphs.pseru-4A	*GBS_212432/GBS_109947*	-[b]	-	0.448	0.9	0.506	1.0	3.810*	10.2	6.768*	16.6	4.009*	8.3	9.490*	17.2
Qphs.pseru-5B	*Xbarc346-2/TTM_62137_50*	-	-	2.473*	5.5	5.237*	12.5	0.247	0.6	4.150*	9.8	0.881	1.7	6.849*	12.7
Qphs.pseru-4B	*Xbarc20/Xwmc238*	-	-	0.465	0.9	0.282	0.6	0.523	1.2	3.084*	7.0	3.138*	6.3	0.290	0.4
Qphs.pseru-5A	*TTM_199619_7/TTM_12597_31*			7.171*	15.5	4.149*	8.9	0.160	0.4	0.336	0.7	0.779	1.4	4.680*	7.7
SD															
Qphs.pseru-4A	*GBS_212432/GBS_109947*	9.933*	21.6	4.927*	11.5	3.731*	9.4	9.234*	20.3	4.820*	13.3	-	-	11.029*	26.5
Qphs.pseru-5B	*Xbarc346-2/TTM_62137_50*	0.359	0.7	0.797	1.7	0.248	0.6	0.342	0.6	0.153	0.4	-	-	0.810	1.6
Qphs.pseru-4B	*Xbarc20/Xwmc238*	4.255*	8.4	1.499	3.3	0.523	1.2	4.281*	8.7	0.205	0.5	-	-	0.487	0.9
Qphs.pseru-5A	*TTM_199619/TTM_12597_31*	0.256	0.7	0.718	1.7	0.159	0.1	0.190	0.4	0.201	0.5	-	-	0.175	0.4

[a]LOD refers to logarithm of odds

[b]Trait was not evaluated in this location

* Significant quantitative trait locus (QTL) with a LOD value greater than the threshold of 2.24 determined by 1000 times of permutations

Table 2.3 Closely linked or flanking markers, LOD values, and coefficients of determination (R^2) of QTL for preharvest sprouting PHS) resistance and seed dormancy (SD) on chromosome 4AL estimated using the recombinant inbred lines (RILs) from TutoumaiA/Siyang 936 grown in JAAS Jiangsu Academy of Agricultural Sciences (JAAS), and Kansas State University (KSU), respectively.

Close or flanking markers of 4A QTL	position	2004 JAAS		2005 JAAS		2006JAAS		2005 KSU		2006 KSU		2007 KSU		Mean over experiments	
		LOD [a]	R^2 (%)	LOD	R^2 (%)	LOD	R^2 (%)	LOD	R^2 (%)	LOD	R^2 (%)	LOD	R^2 (%)	LOD	R^2 (%)
Preharvest sprouting															
Xgwm397	51.04	-[b]	-	0.094	0.2	0.144	0.3	1.619	4.2	2.872	7.0	0.078	0.1	4.352	8.0
GBS212432	60.52	-	-	0.337	0.7	0.000	0.0	3.245	8.3	5.011	11.9	3.010	6.2	8.417	14.7
GBS109947/GBS212432	62.53	-	-	0.321	0.6	0.028	0.1	3.810	10.2	6.768	16.6	3.107	6.7	9.490	17.2
GBS109947	63.43	-	-	0.232	0.4	0.172	0.3	3.365	8.5	6.577	15.2	2.607	5.4	7.854	13.5
Xbarc170	64.78	-	-	0.448	0.9	0.109	0.2	3.726	9.4	5.535	13.0	2.019	4.2	7.984	13.7
Seed dormancy															
Xgwm397	51.04	4.243	10.2	2.933	7.0	1.621	4.2	3.846	9.2	3.923	10.4	-	-	7.324	18.8
GBS212432	60.52	9.933	21.6	4.362	10.3	3.248	8.3	9.234	20.3	4.722	12.4	-	-	11.393	27.3
GBS109947/GBS212432	61.53	9.865	22.3	4.622	11.3	3.814	10.3	9.276	21.0	4.820	13.3	-	-	12.426	31.1
GBS109947	63.43	7.582	17.0	4.248	10.0	3.369	8.5	6.894	15.7	3.948	10.5	-	-	10.475	25.4
Xbarc170	64.78	7.604	17.1	4.927	11.5	3.730	9.4	6.928	15.8	3.934	10.4	-	-	11.029	26.5

[a.] LOD = logarithm of odds

[b.] Trait was not evaluated at this location

Table 2.4 Difference (Dif) in ratings of preharvest sprouting (PHS) and seed dormancy (SD) as reflected by a percentage of germinated seeds between resistance (R) and susceptible (S) alleles of two SNPs and two SSRs for the PHS resistance QTL on chromosome 4A

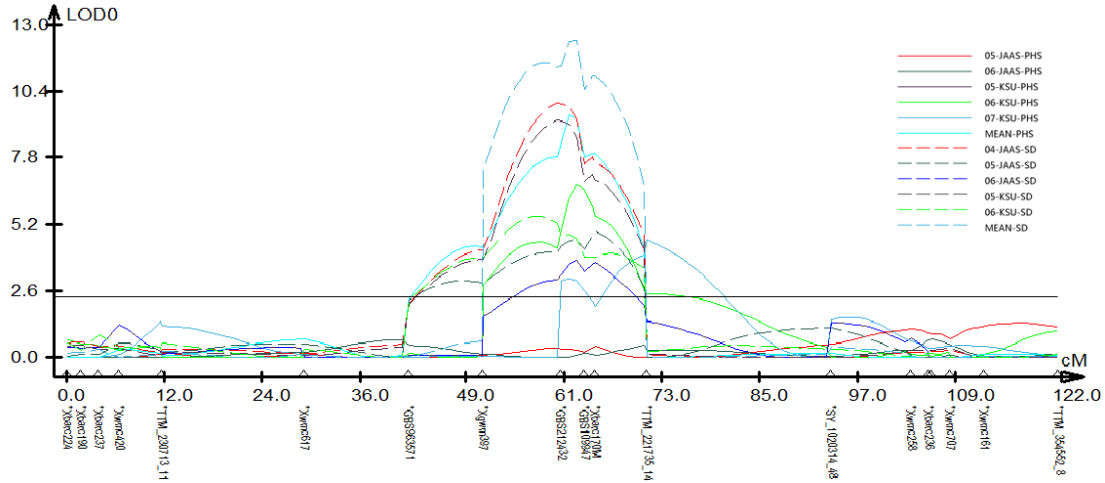
		PHS						SD					
Locus	genotype	2005 JAAS	2006 JAAS	2005 KSU	2006 KSU	2007 KSU	Mean over experiments	2004 JAAS	2005 JAAS	2006 JAAS	2005 KSU	2006 KSU	Mean over experiments
GBS109947	S	71.52	53.45	39.88	53.28	57.24	56.82	71.17	34.21	39.87	71.20	69.89	57.12
GBS109947	R	65.26	51.86	22.51	33.99	38.31	45.56	54.05	18.54	22.50	54.59	59.86	41.64
GBS109947	Dif	6.26	1.59	17.37	19.29	18.92	11.27	17.12	15.67	17.38	16.61	10.04	15.48
Xbarc170	S	72.44	54.41	39.88	52.87	57.80	57.29	71.64	34.15	39.88	71.70	69.93	57.33
Xbarc170	R	66.08	52.09	22.49	34.98	38.84	46.11	54.03	18.63	22.48	54.59	59.87	41.63
Xbarc170	Dif	6.36	2.32	17.39	17.89	18.96	11.18	17.61	15.52	17.40	17.12	10.06	15.70
GBS212432	S	72.50	54.98	38.51	52.55	58.03	57.13	71.55	33.73	38.50	71.61	70.10	56.96
GBS212432	R	64.83	50.52	23.10	34.19	36.29	45.04	53.00	18.32	23.09	53.57	58.89	41.10
GBS212432	Dif	7.67	4.46	15.40	18.36	21.74	12.09	18.55	15.42	15.41	18.04	11.21	15.86
Xgwm397	S	72.30	53.56	38.93	52.71	57.07	56.78	70.86	33.81	38.92	70.85	70.18	56.92
Xgwm397	R	66.23	51.89	25.19	36.73	40.97	47.02	57.00	20.00	25.18	56.89	60.13	43.72
Xgwm397	Dif	6.06	1.68	13.74	15.98	16.10	9.75	13.86	13.81	13.75	13.95	10.06	13.20

Table 2.5 Number of alleles and polymorphism information content (PIC) of SSRs and the allele frequency distributions of SNPs in the 4A QTL region in a natural population

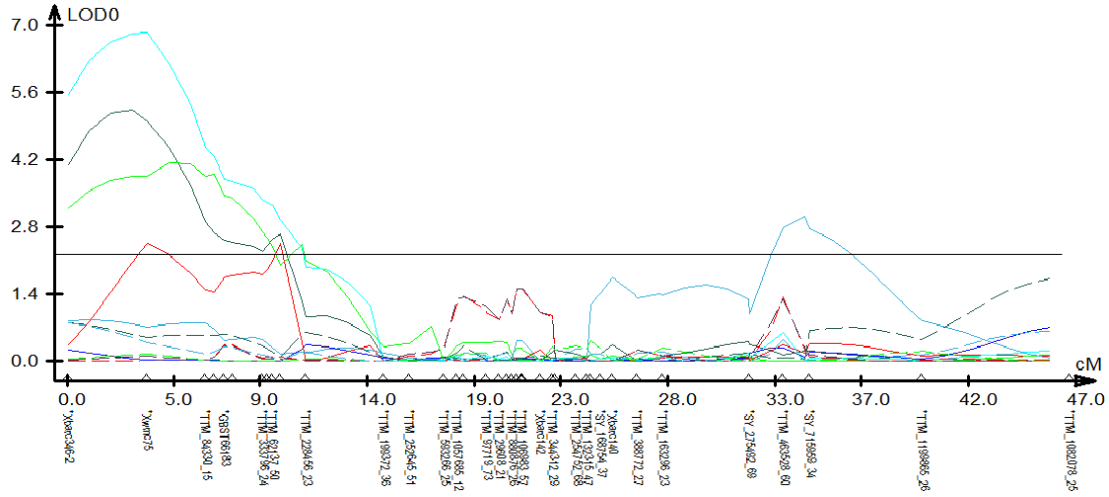
Markers	No. of alleles in the population				PIC	
Xgwm397	14				0.22	
Xbarc170	21				0.11	
	Same allele as Tutoumai A		Same allele as Siyang 936		Undetermined	
	No.	Freq.	No.	Freq.	No.	Freq.
GBS212432	168	0.44	131	0.35	81	0.21
GBS109947	356	0.94	3	0.01	21	0.05

Figure 2.1 Composite interval mapping (CIM) of QTLs for long seed dormancy (SD) and preharvest sprouting (PHS) resistance on chromosome 4A (A), 5B (B) 5A (C) and 4B (D) using SSR and SNP markers and phenotypic data from 10 experiments. The line parallel to the X-axis is the threshold line for the significant LOD value of 2.24 (*P < 0.05*). Genetic distances are shown in centiMorgans (cM).

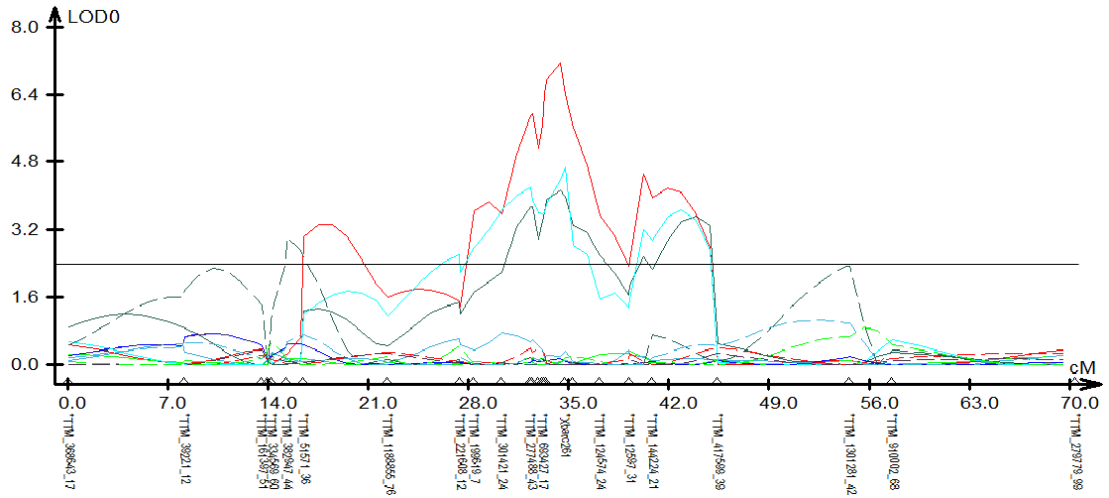
A.



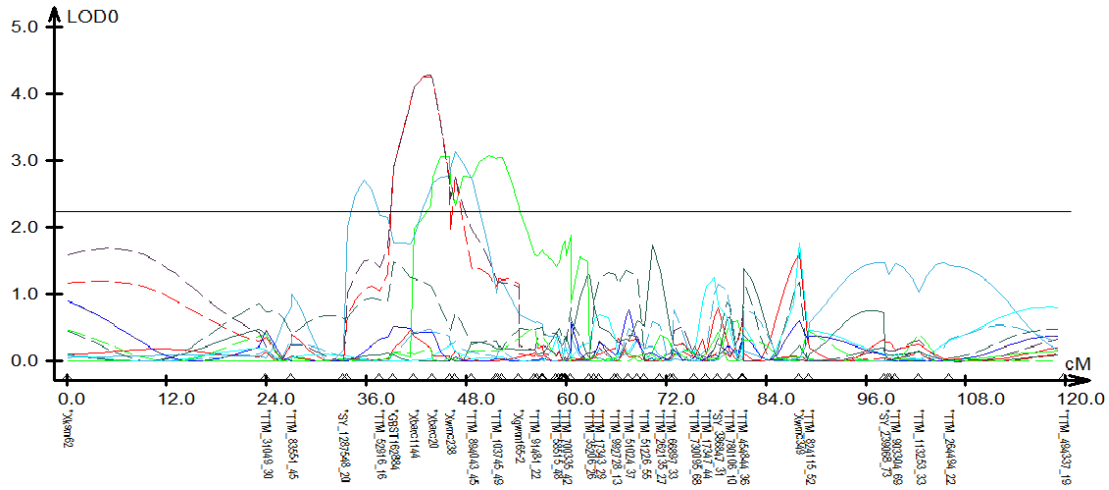
B.



C.


LOD0
8.0
6.4
4.8
3.2
1.6
0.0
cM
0.0
7.0
14.0
21.0
28.0
35.0
42.0
49.0
56.0
63.0
70.0
TTM_368643_17
TTM_39221_12
TTM_161397_51
TTM_334589_60
TTM_38294 7_44
TTM_5157 1_36
TTM_1168365_76
TTM_221608_12
TTM_199619_7
TTM_301421_24
TTM_227488_43
TTM_692437_17
Xbeec2B1
TTM_124574_24
TTM_12597_31
TTM_144224_21
TTM_417569_39
TTM_1301281_42
TTM_910092_68
TTM_279779_99

D.


LOD0
5.0
4.0
3.0
2.0
1.0
0.0
cM
0.0
12.0
24.0
36.0
48.0
60.0
72.0
84.0
96.0
108.0
120.0

Chapter 3 - Genome-wide Association Analysis on Pre-harvest Sprouting Resistance and Grain Color in U.S. Winter Wheat

Abstract

Pre-harvest sprouting (PHS) of wheat can cause substantial reduction in grain yield and end-use quality. Grain color (GC) together with other components affect PHS resistance. Several quantitative trait loci (QTLs) have been reported for PHS resistance, and two of them on chromosome 3AS (*TaPHS1*) and 4A have been cloned. To determine genetic architecture of PHS and GC and genetic relationships of the two traits, a genome-wide association study (GWAS) was conducted by evaluating a panel of 185 U.S. elite breeding lines and cultivars for sprouting rates of wheat spikes and GC in both greenhouse and field experiments. The panel was genotyped using the wheat 9K and 90K single nucleotide polymorphism (SNP) arrays. Four QTLs for GC on four chromosomes and 12 QTLs for PHS resistance on 10 chromosomes were identified in at least two experiments. QTLs for PHS resistance showed varied effects under different environments, and those on chromosomes 3AS, 3AL, 3B, 4AL and 7A were the more frequently identified QTLs. The common QTLs for GC and PHS resistance were identified on the long arms of the chromosome 3A and 3D. Wheat GC is regulated by the three known genes on group 3 chromosomes and additional genes from other chromosomes. These GC genes showed significant effects on PHS resistance in some environments. However, several other QTLs that did not affect grain color also played a significant role on PHS resistance. Therefore, it is possible to breed PHS-resistant white wheat by pyramiding these non-color related QTLs.

Introduction

Pre-harvest sprouting (PHS) of wheat (*Triticum aestivum* L.) refers to the germination of wheat grains in matured spikes before harvest due to continuous wet weather during harvest seasons. PHS can result in a significant reduction in wheat grain yield and grain end-use quality, thus a reduction in grain sale price (Groos *et al*., 2002, Mares *et al*., 2005). Growing PHS-resistant cultivars is the most effective way to minimize PHS damage. PHS resistance QTLs have been reported on almost all wheat chromosomes. One major QTL mapped on chromosome 3AS, designated as *TaPHS1*, has been cloned (Nakamura *et al*., 2011; Liu *et al*., 2013). Another major QTL on chromosome 4AL has been fine mapped with single nucleotide polymorphisms (SNPs) (Cabral *et al*., 2014; Barrero *et al*., 2015; Lin *et al*., 2015). Recently, several candidate genes have been reported for the 4A QTL in different studies (Barrero *et al*., 2015; Torada *et al*., 2016). In addition, several minor QTLs have also been reported on chromosomes 2B (Kulwal *et al*., 2004; Munkvold *et al*., 2009; Kulwal *et al*., 2012; Zhang *et al*., 2014), 3D (Imtiaz *et al*., 2008), 4B, 4D (Kato *et al*., 2001) and many others (Anderson *et al*., 1993).

Wheat grain color (GC) has long been associated with PHS, and red-grained wheats are usually more tolerant to PHS than the white-grained wheats (Flintham, 2000; Warner *et al*., 2000; Himi *et al*., 2002). The pigments, catechin and proanthocyanidins (PAs) synthesized through the flavonoid synthesis pathway, result in red GC (Miyamoto and Everson, 1958; McCallum and Wlker, 1990). Early cytogenetic studies suggested that three genes, *R-A1*, *R-B1* and *R-D1*, on homoeologous group 3 chromosomes control GC (Sears, 1944; Allan and Vogel, 1965; Metzger and Silbaugh, 1970), and show a pleiotropic effect on wheat PHS resistance by accumulating catechin, a precursor of the red pigment, that inhibits grain germination (Miyamoto and Everson, 1958; Stoy and Sundin, 1976). Flintham (2000) found that grain dormancy levels

were increased in white-grained wheat NS-67 after adding a single GC (*R*) gene to one of group 3 chromosomes. Groos *et al*. (2002) identified common QTLs for GC and PHS resistance on chromosomes 3AL, 3BL, 3DL and 5A in a white × red wheat cross. The white-grained mutants of 'Chinese Spring' and 'AUS1490' showed increased sprouting, indicating that *R* genes enhanced PHS tolerance (Warner *et al*., 2000; Himi *et al*., 2002). Recently, *Tamyb10* genes, the transcription factors of the flavonoid biosynthetic pathway, have been reported as candidate genes for the GC trait (Himi *et al*., 2011). However, how much these *R* genes contribute to PHS resistance remains unknown. Therefore, simultaneous genome-wide association studies (GWAS) on both traits may reveal the relationship between *R* genes and PHS resistance.

Genome-wide association studies have been conducted in many plant species to discover and validate QTLs and genes for various traits. By taking advantages of historical recombination events and linkage disequilibrium (LD) between causal genetic variants and nearby SNPs, GWAS detects statistical associations between genetic variations and phenotypic variations throughout the genome (Flint-Garcia *et al*., 2003; Nordborg and Weigel, 2008; Myles *et al*., 2009; Lipka *et al*., 2015). Therefore, GWAS can potentially increase mapping resolution by taking advantages of historical recombinations using highly diverse populations. To date, GWAS has not been reported for GC, and only several studies have been reported for wheat PHS resistance (Kulwal *et al*., 2012; Rehman Arif *et al*., 2012; Jaiswal *et al*., 2012; Albrecht *et al*., 2015). In the current study, we analyzed a panel of elite breeding lines and cultivars from major U.S. winter wheat breeding programs using the wheat 9K and 90K arrays to (1) study the phenotypic variance of PHS resistance in U.S. winter wheat, (2) identify genome-wide QTLs for GC and PHS resistance, and (3) determine the genetic relationship between GC and PHS resistance.

Materials and Methods

Plant materials

A set of 185 winter wheat accessions (Zhang *et al*., 2010) was assembled to include 130 hard winter wheat (HWW) and 55 soft winter wheat (SWW) accessions. A mapping population of 155 F_6 recombinant inbred lines (RILs) derived from the cross of Tutoumai A x Siyang 936 (Liu *et al*., 2008; Lin *et al*., 2015) was used to validate the SNPs that showed significant associations with the *Qphs.hwwgr-4A*.

Pre-harvest sprouting evaluation

In the greenhouse experiments, five plants per accession were grown in a 13 by 13 cm Dura-pot (Hummert Int. Topeka, KS) under the growth condition listed in Table 3.7 after vernalization for seven weeks at 4°C in a cold chamber. The GWAS panel was evaluated for PHS in the greenhouse experiments of fall (August-December) 2011, spring (January-May) and fall 2012, and spring 2013. All experiments were conducted in a randomized complete block design with two replications of five plants.

The GWAS panel was also planted for PHS resistance evaluation in the Kansas State University Rocky Ford Wheat Research Farm, Manhattan, KS and the Agricultural Research Center-Hays, Hays, KS, respectively, in the summers of 2013 and 2014. About 30 seeds per accession were planted in a 1.22-m-long single-row plot, and each experiment had two replications.

When wheat plants reached physiological maturity, similar to Zadoks scale 91 (Zadoks et al., 1974), spikes that lost green color (Trethowan 1995) were harvested from both greenhouse and field experiments, and evaluated for PHS in the lab. Five spikes per accession that were harvested from each replicate were air-dried for 5 d in a greenhouse, and then stored at -20°C to

maintain dormancy for PHS evaluation. After all accessions had been collected, the greenhouse-harvested spikes were air-dried 9 d and field-harvested spikes for 5 d at room temperature. The additional drying days were determined based on preliminary test results of randomly selected samples from field and greenhouse experiments that maximize phenotypic differences among genotypes. After the dried spikes had been immersed in de-ionized water for 12 h, they were enclosed in a moist chamber at 22±1°C with an attached humidifier that ran twice daily at 2 h each time to maintain high moisture in the chamber. After 7 d of incubation, the germinated and non-germinated kernels were hand-threshed and counted separately to calculate the percentage of germinated kernels from all five spikes of each replication.

Evaluation of grain color

Grain color was evaluated for grains harvested from one field experiment (2009-2010 Enid Oklahoma) and the fall 2011 greenhouse experiment at Manhattan KS. For each accession, ten seeds were soaked in 1 M sodium hydroxide (NaOH) for 1 h to increase the color contrast. Grain color intensity was determined visually using a scale of 1 to 4, where 1 represents white, 2 light red, 3 red and 4 dark red.

DNA isolation and genotyping

Leaf tissue was collected at the two-leaf stage, and genomic DNA was isolated using a modified cetyltrimethyl ammonium bromide (CTAB) method (Zhang *et al*., 2010). A total of 446 polymorphic SSR markers were selected to genotype the association panel based on PCR product quality, chromosome distribution in available genetic maps (http://wheat.pw.usda.gov/GG3/; verified 11 Aug. 2010), and previously reported associations with PHS resistance. One expressed sequence tag (EST), *ZXQ118* (Zhang *et al*., 2008), three gene markers of *PM19A1* and *PM19A2* (Barrero *et al*., 2015) and one gene marker of *TaMKK3*-

A were used to determine the association between PHS resistance and *Qphs.hwwgr-4A*. Five sequence-tagged sites (STS) from three *Tamyb10* genes (Himi *et al*., 2011) were analyzed to determine QTLs for GC. Amplification, separation and scoring of polyerase chain reaction (PCR) products followed Zhang *et al.* (2010).

The GWAS panel was also genotyped with the Wheat 9K and 90K SNP arrays (Cavanagh *et al*., 2013; Wang *et al*., 2014) at USDA-ARS Cereal Crops Research Unit (Fargo, ND). SNPs with less than 5% minor allele frequency (MAF) or with more than 15% missing data were removed. A total of 5,921 and 21,600 SNPs were scored from the 9K and 90K SNP arrays, respectively. Association analysis was initially conducted using the 9K genotypic data, and 28 non-redundant SNPs with $p < 0.001$ were then selected and pooled together with the 90K data. Totally, 21,628 SNPs were used for the final analysis. Also, one SNP in the promoter region and two SNPs in the coding region of the *TaPHS1* gene (Nakamura *et al*., 2011; Liu *et al*., 2013) were analyzed using three Kompetitive Allele Specific PCR (KASP) assays. Sequences that harbored significant SNPs and SSR markers were searched against the W7984 reference sequence to estimate their putative chromosome positions.

Population structure and kinship

Population structure was characterized by a set of 1500 SNPs that are evenly distributed on all the 21 wheat chromosomes using the admixture model in STRUCTURE 2.3.4 (Pritchard *et al*., 2000). *K*-values ran from 2 to 20 with 10 iterations set for each *k*-value. The burn-in time and replication number were set at 2×10^5 and 2×10^4, respectively. For each trait, Bayesian information criterion (BIC) (Schwarz, 1978) was applied to determine the optimum number of subpopulations. Marker-based kinship was estimated to approximate the probability of two individuals being identical by descent through adjusting the average probability of being

identical in state between random individuals (Yu *et al*., 2006). Kinship was calculated with the same set of 1,500 SNPs used for structure analysis using SPAGeDi package (Hardy & Vekemans, 2002).

Statistical analysis and genome-wide association analysis

Best linear unbiased predictions (BLUPs) were calculated for each accession evaluated in the greenhouse and field experiments using the 'lme4' package in R 3.2.2 (Bates *et al*., 2014) with year and location as random effects in the model. Genome-wide association analysis was conducted using two models: the generalized linear model (GLM) with the Q matrix as fixed effects, and the mixed linear model (MLM) with a Q matrix as fixed effects and a kinship matrix as random effects. These two models were applied to each experiment for GC and PHS resistance, and model fitness was determined based on the BIC values. Association analysis of SNP data was conducted using the genome association and prediction integrated tool (GAPIT) implemented in R (Lipka *et al*., 2012), and association analysis of SSR data was conducted using PROC MIXED procedure in SAS 9.3 (SAS Institute Inc., Cary, NC). A threshold of $p < 0.001$ was set to claim significant associations between SSR markers and the traits (GC and PHS resistance), and $p < 0.0001$ was set to claim significant associations between SNPs and the traits. Linkage disequilibrium and haplotype analyses of the significant SNPs were performed with HAPLOVIEW v.4.2 (http://www.broadinstitute.org/scientificcommunity/science/programs/medicalandpopulationgenetics/haploview/haploview). Color intensity of the LD plot was determined by the magnitude of pairwise D' value.

QTL analysis

A linkage map covering the 4A QTL region was constructed for the RIL population of Tutoumai A x Siyang 936 using KASP markers converted from significant SNPs from the association study, and previously mapped SSR markers in Liu *et al*. (2011) and GBS-SNPs in Lin *et al*. (2015) by JoinMap version 4.0 (Van Ooijen and Voorrips 2006). Recombination fractions were converted into centiMorgans (cM) using the Kosambi function (Kosambi, 1944). Interval mapping (IM) using sprouting data from the 2005 and 2006 greenhouse experiments and their combined mean was performed using WinQTLCart 2.5 (Wang *et al*., 2005). LOD thresholds to claim significant QTLs for each dataset were determined from 1000 permutations (Doerge and Churchill, 1996).

Results

Phenotypic variations in grain color and pre-harvest sprouting

Twenty-nine accessions were scored as white wheats, and 156 accessions as red wheats. The GC scores were highly consistent between greenhouse and field grown seeds (Fig. 3.1) with a high correlation coefficient of 0.87 ($P<0.0001$), indicating a low genotype-by-environment interaction for GC.

Significant correlations for sprouting rates were observed among most of the eight experiments (Table 3.1). Cluster analysis showed high similarities in sprouting rates of accessions among all the field experiments, but significant differences between the field and the greenhouse experiments (Fig. 3.2a). The broad sense heritability across all eight experiments was high (0.83), with 0.62 in the greenhouse experiments and 0.92 in the field experiments. The population could be roughly divided into three subgroups (Fig. 3.2b), with average sprouting rates of 13.9% in Group 1, 35.5% in Group 2 and 60.3% in Group 3, and average GC scores of

3.0 in Group 1, 2.6 in Group 2 and 1.7 in Group 3, indicating red wheats were more likely to have low sprouting rates. Most of the soft winter wheats were clustered to Group 1, as well as some hard white winter (HWW) wheat accessions from the *Regional Germplasm Observation Nursery* (RGON). The rest accessions from the RGON were mostly clustered to Group 2, whereas accessions from the Southern Regional Performance Nursery (SRPN) and the Northern Regional Performance Nursery (NRPN) were mainly clustered to Group 2 and Group 3.

Genome-wide association studies on grain color

According to the BIC values, the mixed model with population structure (K=3) and kinship fit the best for the GC trait, thus was applied in the following analysis. GWAS detected four significant QTLs on chromosomes 1B, 3A, 3B and 3D, which were represented by the gene markers for *Tamyb10* genes (5 STS markers) and closely linked SSRs (6) and SNPs (12) (Table 3.2). Three major QTLs for GC in the distal region of the long arms of group 3 chromosomes are significant for the data from both greenhouse and field experiments (Table 3.2). Among them, the QTL on 3DL, as indicated by significant markers *Tamyb10-D1* and 3 SNPs, showed the largest effect and explained up to 23.0% of the phenotypic variance for GC. The QTL on chromosome 3BL that was characterized by seven SNPs, two gene makers for *Tamyb10-B1*, and one SSR was significant in both greenhouse and field experiments, and explained up to 19.2% of the phenotypic variance. A QTL on 3AL showed a moderate effect on GC, and explained about 11.1% phenotypic variance. QTL on chromosomes 1B was identified in both field and greenhouse experiments and explained up to 11.7% of the phenotypic variances. Also, three SSRs, *Xwmc93*, *Xbarc145* and *Xbarc148*, were also significant for GC, but their positions cannot be determined because they were mapped to multiple chromosomes.

The association mapping population can be classified into eight genotypic groups based on the allele combinations of the gene markers of *Tamyb10* genes on group 3 chromosomes. The average GC scores in each group tended to increase as the number of red color alleles increases. However, red wheat accessions T154, LA02-923, MO040192 and NC04-15533 do not contain the red alleles (abd) at any of the three loci, whereas white accessions KS05HW15-2 and OK06848W carry the red allele of *Tamyb10-A1* (Abd), and white accessions KS05HW136-3 and CO03W139 carry the red allele of *Tamyb10-D1* (abD) (Fig. 3.3), suggesting that other genes besides *Tamyb10* may also contribute to GC, or the markers for *Tamyb10* genes may not be diagnostic in some genetic backgrounds.

Genome-wide association studies on pre-harvest sprouting resistance

Generalized linear model with Q matrix of k=3 was selected for GWAS on PHS resistance based on BIC values. Twelve QTLs on ten chromosomes were significant for PHS resistance in at least two experiments (Table 3.3). Among them, QTLs on chromosome 3AS, 3AL, 3B and 4AL were the most frequently identified QTLs for PHS resistance. The 3AS QTL was detected in the fall 2011 greenhouse experiment and all the four field experiments, and explained 9.5% to 15.8% of the phenotypic variances for PHS resistance. Significant markers included one SSR, *Xbarc57*, five SNPs from the SNP chips and one SNP developed from the *TaPHS1* gene sequence (Table 3.3), thus this QTL corresponds to *TaPHS1*. The 3AL QTL was identified by SNPs, SSR and the *Tamyb10-a1* gene marker in spring 2012 and all the field experiments, and explained 6.8% to 12.1% of the phenotypic variances. Thus this QTL corresponds to the *Tamyb10-a1* gene for GC on 3AL. The QTL on chromosome 4AL showed a wide range of effects among the experiments, and explained 9.9% to 47.6% phenotypic variance among the two greenhouse experiments (fall and spring 2012) and one field experiment (Manhattan, 2014).

The most significant SNPs for *Qphs.hwwgr-4A* were *Ex_c66324_1151*, *wsnp_Ex_c13031_20625900* and *wsnp_Ex_rep_c66324_64493429*. *ZXQ118*, an EST in the 4A QTL region (Zhang *et al*., 2008), and the gene markers for *PM19A1* were also significant in the fall 2012 greenhouse experiment and the mean sprouting rates over all the greenhouse experiments, but explained much lower phenotypic variation than the previous three markers (Table 3.3). The QTL on chromosome 3B was significant in two greenhouse experiments (fall 2011 and spring 2012) and all field experiments, and explained 7.0% to 12.3% of the phenotypic variances for PHS resistance. Two QTLs were identified on chromosome 3D with one at the distal end of the short arm (*Qphs.hwwgr-3DS*) and another at the distal end of the long arm (*Qphs.hwwgr-3DL*). *Qphs.hwwgr-3DS* was detected in the spring 2012 greenhouse, and two 2014 field experiments, and *Qphs.hwwgr-3DL* was significant in both field experiments in 2013. QTL identified on chromosome 7A was significant in the fall 2011 greenhouse and 2013 Manhattan field experiments and explained up to 13.5% of the phenotypic variance.

Some QTLs were only significant in a single environment. For example, the two QTLs identified on chromosomes 2B, *Qphs.hwwgr-2B.1* and *Qphs.hwwgr-2B.2*, were associated with PHS resistance each in one greenhouse experiment (spring 2013 and 2012, respectively), and the *Qgc.hwwgr-6B.1* was identified only in one field experiment (Manhattan, 2013) (Table 3.4). Therefore, these QTLs may be more sensitive to environmental conditions.

Relationships between grain color and pre-harvest sprouting resistance

Analysis of variance (ANOVA) was conducted by taking GC as the explanatory variable and PHS resistance as the response variable, and it showed that GC had significant effects on PHS resistance in all the field experiments ($P < 0.0001$), but not in any of the greenhouse experiments (Table 3.5). White wheat had significantly higher sprouting rates than red wheats (*P*

< 0.0001) in the field experiments, but the difference was not significant between red-grained accessions with different color scores (data not shown).

Common QTLs for GC and PHS resistance were identified on the long arms of chromosomes 3A and 3D (Table 3.6), but not on 3BL. The QTL on chromosome 3AL, identified by *Tamyb10-A1*, was significant for GC in both field and greenhouse experiments and for PHS resistance in all the field experiments. For the QTL on chromosome 3D as represented by *Tamyb10-D1*, one SNP was significant for GC in both experiments and also for PHS resistance in the 2013 field experiments at both Manhattan and Hays. Unlike the 3A and 3D QTLs, QTL on chromosome 3B, represented by the *Tamyb10-B1* as well as seven linked SNPs and one SSR, was significant for only GC, not PHS resistance in any experiments. Therefore, *Tamyb10-A1* and *Tamyb10-D1*, but not *Tamyb10-B1*, were very likely to have pleiotropic effects on PHS resistance under the field conditions.

Validation of the significant SNPs for the 4A QTL in a bi-parental population

Seventeen KASP assays were designed based on the sequences of the significant SNPs identified in the 4A QTL region for PHS resistance. Four of the KASP markers (Table 3.8) showed co-segregation among the F_6 RILs of "Tutoumai A" × "Siyang 936", and were mapped between the two previously reported flanking GBS SNPs (*GBS212432*, *GBS109947*) for the QTL (Lin *et al*., 2015) at 1.02 cM to *GBS212432* and 2.10 cM to *GBS109947* (Fig. 3.4). These four SNPs showed the highest LOD scores in all experiments, and explained up to 31.76% of the phenotypic variance in the population.

Linkage disequilibrium

Linkage disequilibrium (LD) parameter *D'* was calculated to determine the linkage relationship between SNPs from different QTLs and link the markers with unknown positions to

known QTLs. LD was calculated for the 125 SNPs that were significantly linked to nine PHS resistance QTLs in at least two experiments. Strong LD was detected for SNPs within each PHS resistance QTL region, but not between different QTLs (Fig. 3.5b), indicating that those QTLs for PHS resistance were independent. Pair-wise *D'* values were also estimated for the 17 SNPs that were tightly linked to the four GC QTLs. Similarly, strong LD was not detected among the SNPs linked to GC QTLs in group 3 chromosomes (Fig. 3.5a). Although SNPs in the 1B QTL showed high *D'* values (around 0.83) with SNPs in the 3D QTL, r^2 values that adjusts LD relationships by incorperating allele frequencies were low (around 0.08) between SNPs from the two QTLs.

Genetic positions of most significant SNPs for GC on chromosome 3D and *Tamyb10-D1* could not be determined using the W7984 reference sequence, and SNPs significantly related to GC on chromosome 3D, *D_GA8KES402JVT1Y_74* and *BS00067163_51*, were far apart from each other on the chromosome 3D. However, LD analysis suggested that these SNPs were tightly linked to *Tamyb10-D1*, and thus they linked to the same 3D QTL for GC (Fig. 3.5a).

Discussion

QTLs for grain color

Wheat GC has been a classic example for dissection of a quantitative trait (Nilsson-Ehle, 1909) and three genes on wheat chromosomes group 3 have long been proposed as the genes controlling wheat GC. Several previous studies have mapped the three genes as major QTLs as well as some minor QTLs on chromosomes 2B, 2D, 5A and 6B for GC (Groos *et al*., 2002; Kumar *et al*., 2009; Himi *et al*., 2011). Being the first association study for wheat GC, we not only validated the effects of these three GC genes, *Tamyb10-A1*, *Tamyb10-B1* and *Tamyb10-D1*, on the long arms of chromosomes group 3, but also identified a new QTL on the chromosome 1B

for GC, suggesting that QTLs on other chromosomes than these well-known QTLs on chromosomes group 3 may also play a role in regulating GC in some wheat germplasm lines.

Groos *et al*. (2002) mapped all group 3 QTLs in a bi-parental population, but they did not discuss their effects of each QTL. In this study, a diverse association panel makes it possible to compare the effects of all the three QTLs. Among the three genes on the chromosome group 3, *Tamyb10-D1* had the largest effect on GC (R^2=0.24) and *Tamyb10-A1* the smallest (R^2=0.11) in the association mapping panel whereas their minor allele frequencies (MAF) were similar (Table 3.2), indicating that the large effect of *Tamyb10-D1* was not due to a higher MAF than other two genes. On the other hand, one single gene changed GC from white to red, and adding one or two additional GC genes only slightly increased redness (Fig. 3.3). Besides, QTL on chromosome 1B also contribute to GC, which was not reported previously, thus it is likely a new QTL for GC. That the red allele of the 1B QTL presents in the four red wheat accessions that do not carry the red alleles (abd) at any of the three *Tamyb10* genes supports this assumption. Therefore, when breeding for white wheat cultivars, breeders not only need to remove the three *Tamyb10* genes, but also should watch for other genes that may contribute to GC.

In this study, wheat GC was visually scored after increasing color intensities using sodium hydroxide solution. High repeatability in GC between the greenhouse and field experiments (Fig. 3.1) indicates that the GC scoring method used in the experiments is highly repeatable. All of the four QTLs identified for GC were detected in both experiments, which provided genetic evidence that QTLs for GC are relatively stable across environments.

QTLs for pre-harvest sprouting resistance

QTLs for PHS resistance have been mapped on almost all wheat chromosomes in previous bi-parental mapping studies. Although association studies on PHS resistance have been

conducted using several types of markers (Kulwal *et al*., 2012; Rehman Arif *et al*., 2012), the current study is the first report to use high density SNPs for GWAS on PHS resistance. We identified 12 QTLs that were significant in at least two experiments.

For the QTL on 3AS, the causal gene (*TaPHS1*) has been cloned (Nakamura *et al*., 2011; Liu *et al*., 2013). One of the reported functional SNPs in the coding region (Liu *et al*., 2013) was significant in one greenhouse (fall 2011), whereas the functional SNP in the promoter region (Nakamura *et al*., 2011) was not significant in any of the experiments (Table 3.3). However, the most significant markers linked to the 3AS PHS resistance QTL were not the functional SNPs, which was probably due to environmental effects on phenotyping (Nakamura *et al*., 2011). Among the gene markers for *PM19A1* and *PM19A2,* only one of the candidate gene markers of *PM19A1* was significantly associated with PHS resistance in the fall 2012 experiment, although the 4A QTL showed an extremely large effect on PHS in that experiment (Table 3.3). This was probably due to the fact that the gene expression was affected by environments or the gene markers are not diagnostic. However, the gene marker for *TaMKK3-A* was in strong LD with the most significant SNPs for the 4A QTL, indicating that the *TaMKK3-A* is more likely to be the candidate gene for the 4A QTL.

The QTL identified at the distal end of chromosome 3DS was not reported previously. LD analysis indicated that *Qgc.hwwgr-3DS* is a different QTL from *Qgc.hwwgr-3AS* (Fig. 3.5b). For the QTL on chromosome 3B, the sequences of the linked SSR markers are not found in the W7984 reference sequence, thus we cannot determine whether or not the significant SSR markers and SNPs on 3B linked to the same QTL. Similarly, we cannot determine the QTL positions on chromosome 7A.

QTL identified on chromosome 1A could be the same QTL reported by Knox *et al*. (2005) in durum because *Xwmc183* was located near the QTL region mapped in our study based on the W7984 reference sequence. The QTL on chromosome 2D is the same QTL as *QPhs.ccsu-2D.4* (Mohan *et al*., 2009) because of the common SSR *Xgwm539*. However, we cannot determine whether the QTLs that were identified on chromosomes 1D, 5A, 5B, 6A and 6B were the same QTLs reported in previous studies (Kumar *et al*., 2009; Groos *et al*., 2002; Arif *et al*., 2012; Kulwal *et al*., 2004; Roy *et al*., 1999) due to the lack of common markers.

Variation of PHS resistance across environments

PHS is a complicated trait affected by many factors, including seed dormancy (SD) (Bewley and Black, 1982; Anderson *et al*., 1993; Mares and Mrva, 2001; Ogbonnaya *et al*., 2008), GC (Gfeller and Svejda, 1960; Groos *et al*., 2002), spike morphology, as well as environmental factors such as temperature, moisture and photoperiod after flowering (Argel *et al*., 1983; Ceccato *et al*., 2011). In the current study, PHS resistance of the tested accessions and QTL effects varied across environments with more variation observed among the greenhouse experiments than that among the field experiments (Fig. 3.2a). A total of four greenhouse experiments were conducted in the fall greenhouse cycles of 2011 and 2012 with the harvest time in winter, and the spring cycles of 2012 and 2013 with the harvest time in summer. The two seasons were highly different in growing and post-harvesting temperatures, which has been shown to influence PHS resistance (Nakamura *et al*., 2011； Barrero *et al*., 2015）. Meanwhile, in the field experiments at Manhattan and Hays, dry hot winds shortened maturity period, which greatly reduced environment effects on wheat PHS resistance. Therefore, PHS resistance was similar in the four field experiments.

Qphs.hwwgr-3AS and *Qphs.hwwgr-4A* were the major QTLs for PHS resistance, and most frequently identified in all experiments. However, *Qphs.hwwgr-3AS* was detected more frequently in the field experiments, while *Qphs.hwwgr-4A* was detected more frequently in the greenhouse conditions (Table 3.3), which might be due to high temperatures in field conditions during late grain maturation that suppressed the expression of *Qphs.hwwgr-4A* (Barrero et al., 2015).

According to the heat map derived from individual PHS ratings across all the experiments, the population can be roughly divided into three clusters (Fig. 3.2b). Most of the soft winter wheats had low germination rates, and were clustered to Group 1. Wheat cultivars from RGON were mostly clustered to Group 1 and Group 2, whereas accessions in SRPN and NRPN showed higher germination rates, and were mainly clustered to Group 2 and Group 3. These results indicated that the soft winter wheat accessions grown in the humid climate during harvest season had a higher selective pressure on PHS resistance than the hard winter wheat accessions from the Great Plains that are grown under relatively drier climate.

Validation of the markers for the QTL on 4A

In this study, a RIL population from "Tutoumai A" x "Siyang 936" was used to validate the position of significant SNPs for 4A PHS resistance QTL. Four polymorphic SNPs from GWAS were successfully mapped to the QTL region, and they are more closely linked to PHS resistance than previously reported flanking markers, *GBS212432* and *GBS109947*, for this QTL (Lin *et al*., 2015). This result indicates that GWAS provides more power to increase marker density and mapping resolution, whereas bi-parental populations can further validate the positions of new markers. Barrero *et al*. (2015) proposed *PM19A1* and *PM19A2* as the candidate genes for the 4A QTL and identified causal deletions in *PM19A1* and *PM19A2*. We analyzed the

markers developed based on the causal variation in Tutoumai A and Siyang 936, but did not find any polymorphism between the two parents. Therefore, a different gene or different causal SNP in the gene may control the PHS resistance of 4A QTL in this population, which was also supported by the results from the GWAS that the candidate gene markers contributed much lower phenotypic variation for PHS resistance than three other SNP markers (*Ex_c66324_1151*, *wsnp_Ex_c13031_20625900*, *wsnp_Ex_rep_c66324_64493429*) (Table 3.3).

Effect of grain color QTLs on pre-harvest sprouting resistance

GC has been considered as an important factor for PHS resistance, and previous studies showed that seed dormancy level of a white-grained wheat line was improved by the introgression of an *R* gene (Flinthman *et al*., 2000). In the current study, GC explained 26% to 44% of the phenotypic variance for PHS resistance, and *Tamyb10-A1* and *Tamyb10-D1* showed significant effects on both GC and PHS resistance, which agree with a previous study (Groos *et al*., 2002). *Tamyb10* genes encode R2R3-type MYB transcription factors, which regulate the accumulation of PA in the biosynthesis pathways (Himi *et al*., 2011). Therefore, it is possible that these transcription factors showed pleiotropic effects by regulating more than one metabolism pathway, and had effects on improving wheat PHS resistance. However, the GC gene on 3BL, *Tamyb10-B1*, did not show any effect on PHS resistance in this study (Table 3.2; Table 3.6).

In this study, GC was significantly related to PHS resistance in field experiments, but had barely any effect in the greenhouse experiments (Table 3.5). Also, the *Tamyb10-A1* gene affected PHS resistance in all of the four field experiments, and the *Tamyb10-D1* gene only affected PHS resistance in the 2013 experiments. Such results suggested that environmental factors could be important triggers of pleiotropic effects of the GC genes on PHS resistance. That *Tamyb10-B1*

did not show any effect on PHS resistance might be due to the field environments of this study that could not trigger the expression of pleiotropic effect of the gene.

Although some GC genes contributed to wheat PHS resistance, many QTLs for PHS resistance did not affect GC. Therefore, some red wheats can be highly susceptible to PHS, while some white wheats can be highly resistant (Torada *et al*., 2002; Bi *et al*., 2014). Breeding for PHS resistance, attention should be paid to these QTLs with a major effect on PHS in most environments without a pleiotropic effect on GC, such as these on 3AS and 4AL. Pyramiding several of these genes in one cultivar should be able to avoid PHS damage in U.S. HWW.

References

Albrecht T, Oberforster M, Kempf H, Ramgraber L, Schacht J, Kazman E, Zechner E, Neumayer A, Hartl L, Mohler V. Genome-wide association mapping of preharvest sprouting resistance in a diversity panel of European winter wheats. J Appl Genet. 2015;56(3):277-85. doi: 10.1007/s13353-015-0286-5

Allan RE, Vogel OA. Monosomic analysis of red seed color in wheat. Crop Sci. 1965;5:474–475.

Anderson JA, Sorrells ME, Tanksley SD. RFLP analysis of genomic regions associated with resistance to preharvest sprouting in wheat. Crop Sci. 1993;33:453–459.

Argel PJ, Humphreys LR. Environmental effects on seed development and hardseededness in Stylosanthes hamata cv. Verano. I. Temperature. Crop Pasture Sci. 1983;34:261-270.

Barrero JM, Cavanagh C, Verbyla KL, Tibbits JF, Verbyla AP, Huang BE, Rosewarne MG, Stephen S, Wang P, Whan A, Rigault P, Hayden JM, Gubler F. Transcriptomic analysis of wheat near-isogenic lines identifies *PM19-A1* and *A2* as candidates for a major dormancy QTL. Genome Biol. 2015;16:93.

Bates D, Maechler M, Bolker B, Walker S. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7. 2014;http://CRAN.R-project.org/package=lme4.

Bewley JD, Black M. Physiological and biochemistry of seeds in relation to germination, vol 2. Springer, Heidelberg. 1982;pp 61–81.

Bi HH, Sun YW, Xiao YG, Xia LQ. Characterization of *DFR* allelic variations and their associations with pre-harvest sprouting resistance in a set of red-grained Chinese wheat germplasm. Euphytica. 2014;195(2):197-207.

Cabral AL, Jordan MC, McCartney CA, You FM, Humphreys DG, MacLachlan R, Pozniak CJ. Identification of candidate genes, regions and markers for pre-harvest sprouting resistance in wheat (*Triticum aestivum L.*). BMC Plant Biol. 2014;14:340.

Cavanagh CR, Chao S, Wang S, Huang BE, Stephen S, Kiani S, Forrest K, Saintenac C, Brown-Guedira GL, Akhunova A *et al*. Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. P Natl Acad Sci USA. 2013;110:8057-8062.

Ceccato DV, Daniel Bertero H, Batlla D. Environmental control of dormancy in quinoa (*Chenopodium quinoa*) seeds: two potential genetic resources for preharvest sprouting tolerance. Seed Sci Res. 2011;21:133–141.

Chen CX, Cai SB, Bai G. A major QTL controlling seed dormancy and preharvest sprouting resistance on chromosome 4A in a Chinese wheat landrace. Mol Breed. 2008;21:351–358.

Doerge RW, Churchill GA. Permutation tests for multiple loci affecting a quantitative character. Genetics. 1996;142:285–294.

Flint-Garcia SA, Thornsberry JM, Buckler ES IV. Structure of linkage disequilibrium in plants. Annu Rev Plant Biol. 2003;54:357–374.

Flintham JE. Different genetic components control coatimposed and embryo-imposed dormancy in wheat. Seed Sci Res. 2000;10:43–50.

Gfeller F, Svejda F. Inheritance of post-harvest seed dormancy and kernel color in spring wheat lines. Can J Plant Sci. 1960;40:1–6.

Groos C, Gay G, Perretant MR, Gervais L, Bernard M, Dedryver F, Charmet G. Study of the relationship between pre-harvest sprouting and grain color by quantitative trait loci analysis in a white × red grain bread-wheat cross. Theor Appl Genet. 2002;104:39–47.

Hardy OJ, Vekemans X. spagedi: a versatile computer program to analyze spatial genetic structure at the individual or population levels. Mol Ecol Notes. 2002;2:618–620. doi: 10.1046/j.1471-8286.2002.00305.x.

Himi E, Maekawa M, Miura H, Noda K. Development of PCR markers for *Tamyb10* related to *R-1*, red grain color gene in wheat. Theor Appl Genet. 2011;122:1561-1576.

Himi E, Mares DJ, Yanagisawa A, Noda K. Effect of grain colour gene (*R*) on grain dormancy and sensitivity of the embryo to abscisic acid (ABA) in wheat. J Exp Bot. 2002;53(374):1569–1574.

Imtiaz M, Ogbonnaya FC, Oman J, Ginkel MV. Characterization of quantitative trait loci controlling genetic variation for preharvest sprouting in synthetic backcross derived wheat lines. Genetics. 2008;178:1725–1736.

Jaiswal V, Mir RR, Mohan A, Balyan HS, Gupta PK. Association mapping for pre-harvest sprouting tolerance in common wheat (*Triticum aestivum* L.). Euphytica. 2012;188:89–102.

Kato K, Nakamura W, Tabiki T, Miura H, Sawada S. Detection of loci controlling seed dormancy on group 4 chromosomes of wheat and comparative mapping with rice and barley genomes. Theor Appl Genet. 2001;102:980–985.

Knox RE, Clarke FR, Clarke JM, Fox SL. Genetic analysis of pre-harvest sprouting in a durum wheat cross. Euphytica. 2005;143:261-264.

Kosambi DD. The estimation of map distances from recombination values. Ann Eugen. 1944;12:172–175.

Kulwal PL, Singh R, Balyan HS, Gupta PK. Genetic basis of pre-harvest sprouting tolerance using single-locus and two-locus QTL analyses in bread wheat. Funct Integr Genomics. 2004;4:94–101.

Kulwal P, Ishikawa G, Benscher D, Feng Z, Yu LX, Jadhav A, Mehetre S, Sorrells ME. Association mapping for pre-harvest sprouting resistance in white winter wheat. Theor Appl Genet. 2012;125:793-805.

Kumar A, Kumar J, Singh R, Garg T, Chhuneja P, Balyan HS, Gupta PK. QTL analysis for grain colour and pre-harvest sprouting in bread wheat. Plant Sci. 2009;177:114-122.

Lin M, Cai S, Wang S, Liu S, Zhang G, Bai G. Genotyping-by-sequencing (GBS) identified SNP tightly linked to QTL for pre-harvest sprouting resistance. Theor Appl Genet. 2015;1-11.

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z. GAPIT: genome association and prediction integrated tool. Bioinformatics. 2012;28(18):2397-2399.

Lipka AE, Kandianis CB, Hudson ME, Yu J, Drnevich J, Bradbury PJ, Gore MA. From association to prediction: statistical methods for the dissection and selection of complex traits in plants. Curr Opin Plant Biol. 2015;24:110-118.

Liu S, Bai G, Cai S, Chen C. Dissection of genetic components of preharvest sprouting resistance in white wheat. Mol Breed. 2011;27:511–523.

Liu S, Cai S, Graybosch R, Chen C, Bai G. Quantitative trait loci for resistance to pre-harvest sprouting in US hard white winter wheat *Rio Blanco*. Theor Appl Genet. 2008;117:691–699.

Liu S, Sehgal SK, Li J, Lin M, Trick HN, Yu J, Gill BS, Bai G. Cloning and characterization of a critical regulator for preharvest sprouting in wheat. Genetics. 2013;195:263–273.

Mares DJ, Mrva K. Mapping quantitative trait loci associated with variation in grain dormancy in Australian wheat. Crop Pasture Sci. 2001;52:1257–1265.

Mares DJ, Mrva K, Cheong J, Williams K, Watson B, Storlie E, Sutherland M, Zou Y. A QTL located on chromosome 4A associated with dormancy in white- and red-grained wheats of diverse origin. Theor Appl Genet. 2005;111:1357–1364.

Matthews DE, Carollo VL, Lazo GR, Anderson OD. GrainGenes, the genome database for small-grain crops. Nucleic Acids Res. 2003;31:183–186.

McCallum JA, Walker JRL. Proanthocyanidins in wheat bran. Cereal Chem. 1990;67(3):282–285.

Metzger RJ, Silbaugh BA. Location of genes for seed coat color in hexaploid wheat, *Triticum aestivum L*. Crop Sci. 1970;10:495–496.

Miyamoto T, Everson EH. Biochemical and physiological studies of wheat seed pigmentation. Agron J. 1958;50:733–734.

Mohan A, Kulwal P, Singh R, Kumar V, Mir RR, Kumar J, Prasad M, Balyan HS, Gupta PK. Genome-wide QTL analysis for pre-harvest sprouting tolerance in bread wheat. Euphytica. 2009;168:319-329.

Mori M, Uchino N, Chono M, Kato K, Miura H. Mapping QTL for grain dormancy on wheat chromosome 3A and group 4 chromosomes, and their combined effect. Theor Appl Genet. 2005;110:1315–1323.

Munkvold JD, Tanaka J, Benscher D, Sorrells ME. Mapping quantitative trait loci for preharvest sprouting resistance in white wheat. Theor Appl Genet. 2009;119:1223–1235.

Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, Buckler ES. Association mapping: critical considerations shift from genotyping to experimental design. Plant Cell. 2009;21:2194–2202.

Nakamura S, Abe F, Kawahigashi H, Nakazono K, Tagiri A, Matsumoto T, Utsugi S, Ogawa T, Handa H, Ishida H, Mori M, Kawaura K, Ogihara Y, Miura H. A wheat homolog of MOTHER OF FT AND TFL1 acts in the regulation of germination. The Plant Cell Online. 2011;23:3215–3229.

Nilsson-Ehle H. Kreuzungsuntersuchungen an hafer und weizen. Lunds Universitets Arsskrift, N.F. Afd 2, Bd 7, No 6. 1911;pp 3-84.

Nordborg M, Weigel D. Next-generation genetics in plants. Nature. 2008;456:720–723.

Ogbonnaya FC, Imtiaz M, Ye G, Hearnden PR, Hernandez E, Eastwood RF, Ginkel MV, Shorter SC, Winchester JM. Genetic and QTL analyses of seed dormancy and preharvest sprouting resistance in the wheat germplasm CN10955. Theor Appl Genet. 2008;116:891–902.

Osa M, Kato K, Mori M, Shindo C, Torada A, Miura H. Mapping QTL for seed dormancy and the Vp1 homologue on chromosome 3A in wheat. Theor Appl Genet. 2003;106:1491–1496.

Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155:945–959.

Rehman Arif MA, Neumann K, Nagel M, Kobiljski B, Lohwasser U, Börner A. An association mapping analysis of dormancy and pre-harvest sprouting in wheat. Euphytica. 2012;188:409-417.

Roy JK, Prasad M, Varshney RK, Balyan HS, Blake TK. Identification of a microsatellite on chromosomes 6B and a STS on 7D of bread wheat showing an association with preharvest sprouting tolerance. Theor Appl Genet. 1999;99:336–340.

Schwarz G. Estimating the dimension of a model. Ann Stat. 1978;6:461-464.

Sears ER. Cytogenetic studies with polyploid species of wheat. II. Additional chromosomal aberrations in *Triticum vulgare*. Genetics. 1944;29:232–246.

Stoy V, Sundin K. Effects of growth regulating substances in cereal grain germination. Cereal Res Commun. 1976;4:157-163.

Torada A, Amano Y. Effect of seed coat color on seed dormancy in different environments. Euphytica. 2002;126(1):99-105.

Torada A, Ikegnchi S, Koike M. Mapping and validation of PCR-based markers associated with a major QTL for seed dormancy in wheat. Euphytica. 2005;143:251–255.

Torada A, Koike M, Ogawa T, Takenouchi Y, Tadamura K, Wu J, Matsumoto T, Kawaura K, Ogihara Y. A Causal Gene for Seed Dormancy on Wheat Chromosome 4A Encodes a MAP Kinase Kinase. Curr Biol. 2016;http://dx.doi.org/10.1016/j.cub.2016.01.063.

Trethowan RM. Evaluation and selection of bread wheat (*Triticum aestivum* L) for preharvest sprouting tolerance. Aust J Agric Res. 1995;46:463–474.

Van Ooijen JW. JoinMap® 4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma BV, Wageningen, Netherlands. 2006.

Wang S, Basten CJ, Zeng ZB. Windows QTL Cartographer 2.5. Departmentof Statistics, North Carolina State University, Raliegh. 2005;http://statgen.ncsu.edu/qtlcart/WQTLCart.htm.

Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE, Maccaferri M, Salvi S, Milner SG, Cattivelli L, Mastrangelo AM, Whan A, Stephen S, Barker G, Wieseke R, Plieske J, International Wheat Genome Sequencing Consortium, Lillemo M, Mather D, Appels R, Dolferus R, Brown-Guedira G, Korol A, Akhunova AR, Feuillet C, Salse J, Morgante M, Pozniak C, Luo M-C, Dvorak J, *et al*. Characterization of polyploid wheat genomic

diversity using a high-density 90 000 single nucleotide polymorphism array. Plant Biotech J. 2014;12:787-796.

Warner RL, Kudrna DA, Spaeth SC, Jones SS. Dormancy in white-grain mutants of Chinese Spring wheat (*Triticum aestivum L.*). Seed Sci Res. 2000;10:51–60.

Yu J, Pressoir G, Briggs WH, Bi IV, Yamasak M, Doeble JF, McMullen MD, Gaut BS, Nielson DM, Holland JB, Kresovich S, Kresovich S. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet. 2006;38(2):203-208.

ZADOKS JC, CHANG TT, KONZAK CF. A decimal code for the growth stages of cereals. Weed Res. 1974;14(6): 415-421.

Zhang XQ, Li C, Tay A, Lance R, Mares D, Cheong J, Cakir M, Ma J, Appels R. A new PCR-based marker on chromosome 4AL for resistance to pre-harvest sprouting in wheat (Triticum aestivum L.). Mol Breed. 2008;22:227-236.

Zhang D, Bai G, Zhu C, Yu J, Carver BF. Genetic diversity, population structure, and linkage disequilibrium in U.S. elite winter wheat. Plant Genome. 2010;3:117-127.

Zhang Y, Miao X, Xia X, He Z. Cloning of seed dormancy genes (*TaSdr*) associated with tolerance to pre-harvest sprouting in common wheat and development of a functional marker. Theor Appl Genet. 2014;127:855-866.

Table 3.1 Pairwise correlation coefficients among germination rates from all eight experiments and best linear unbiased predictions (BLUP) of the all greenhouse experiments and all field experiments.

Corr Coeff	2011F	2012S	2012F	2013S	2013_MH	2013_Hays	2014_MH	2014_Hays	GH_BLUP
2012S	0.238***								
2012F	0.468***	0.313***							
2013S	0.167*	0.543***	0.337***						
2013_MH	0.276***	0.171*	0.212**	0.153*					
2013_Hays	0.407***	0.227**	0.402***	0.237***	0.741***				
2014_MH	0.384***	0.312***	0.576***	0.242***	0.686***	0.721***			
2014_Hays	0.358***	0.228**	0.431***	0.266***	0.747***	0.755***	0.821***		
GH_BLUP	0.559***	0.718***	0.821***	0.710***	0.268***	0.437***	0.551***	0.448***	
Field_BLUP	0.399***	0.265***	0.463***	0.253***	0.869***	0.890***	0.909***	0.929***	0.484***

*$p < 0.05$; **$p < 0.01$, ***$p < 0.001$

Table 3.2 Quantitative trait loci (QTL) identified for wheat grain color (GC) evaluated for the seeds harvested from the field experiment of Enid, OK, in,2010 (Enid2010) and from the greenhouse (GH) experiment conducted in Manhattan KS, 2011 (GH2011).

Chromosome	Marker name	Marker type	Chromosome Position (cM)[a]	Positive allele	Enid2010		GH2011		Mean	
					p	R^2 (%)[b]	*p*	R^2 (%)	*p*	R^2 (%)
1B	*Ra_c35710_395*	90K	58.08	0.92	7.31E-06	9.9	2.19E-05	9.3	3.16E-06	10.8
1B	*RAC875_c1188_531*	90K	58.08	0.92	4.62E-06	10.4	8.06E-06	10.3	1.33E-06	11.7
3A	*Xwmc559-1*	SSR	107.20	0.94	1.25E-04	9.8	5.00E-04	10.6	1.57E-04	10.8
3A	*Tamyb10-A1-66*	STS	114.02	0.63	8.75E-06	7.5	1.86E-04	10.2	2.25E-05	9.4
3A	*Tamyb10-A1-74*	STS	114.02	0.61	3.12E-06	8.5	6.50E-05	11.1	7.32E-06	10.4
3B	*BS00040742_51*	90K	68.26	0.36	9.42E-06	9.7	-	-	2.82E-05	8.6
3B	*Tdurum_contig100004_204*	90K	-	0.38	4.81E-06	10.3	1.00E-05	10.1	7.54E-06	9.9
3B	*BS00025679_51*	90K	76.22	0.58	2.45E-05	8.7	1.88E-06	11.8	6.30E-06	10.1
3B	*Kukri_c60633_121*	90K	76.22	0.35	8.53E-06	9.8	2.35E-06	11.6	4.23E-06	10.5
3B	*Kukri_c60633_257*	90K	76.22	0.33	3.63E-05	8.3	5.35E-06	10.7	9.98E-06	9.6
3B	*Excalibur_rep_c97324_623*	90K	76.22	0.35	5.23E-06	10.3	1.64E-06	12.0	2.50E-06	11.0
3B	*Tamyb10-B1-1*	STS	77.36	0.26	2.26E-05	11.1	1.06E-05	11.0	7.46E-06	11.8
3B	*Tamyb10-B1-2*	STS	77.36	0.26	7.32E-07	11.1	8.21E-07	11.0	3.22E-07	11.8
3B	*Xbarc84*	SSR	80.77	0.31	1.89E-05	4.1	-	-	1.16E-04	6.4
3D	*GENE-1785_118*	90K	-	0.42	4.74E-06	10.4	2.42E-09	19.2	6.49E-08	14.8
3D	*D_GA8KES402JVT1Y_74*	90K	11.37	0.54	1.32E-07	14.0	3.31E-10	21.5	1.79E-09	18.7
3D	*BS00067163_51*	90K	92.34	0.52	5.36E-08	15.0	8.39E-11	23.2	7.49E-10	19.7
3D	*BS00063075_51*	90K	-	0.72	8.85E-05	7.5	5.03E-06	10.8	9.74E-06	9.7
3D	*Tamyb10-D1-93*	STS	-	0.56	4.31E-11	21.9	3.66E-13	17.5	6.51E-13	21.0
3D	*Xbarc376*	SSR	-	0.94	-	-	5.00E-04	24.6	7.00E-04	23.3
1A/1D	*Xwmc93*	SSR	-	0.37	-	-	6.10E-05	5.3	3.00E-04	7.2
1A/2D/3B	*Xbarc145*	SSR	-	0.11	-	-	6.62E-05	5.0	1.65E-04	6.7
1A/1D/3A/5B	*Xbarc148*	SSR	-	0.70	2.50E-07	16.6	1.03E-06	18.7	2.27E-07	18.4

[a] The marker positions in the chromosome based on W7984 reference map

[b] Phenotypic variance explained by a significant marker significantly related to grain color

Table 3.3 Quantitative trait loci (QTLs) of wheat pre-harvest sprouting resistance identified in at least two of the experiments using sprouting rates (%) evaluated in the fall 2011 (2011F), spring 2012 (2012S), fall 2012 (2012F) and spring 2013 (2013S) greenhouse experiments, the 2013 and 2014 Manhattan (2013MH and 2014MH) and 2013 and 2014 Hays (2013Hays and 2014Hays) field experiments, and using the best linear unbiased predictions (BLUP) of each accession from all the greenhouse (GH_BLUP) and field (Field_BLUP) experiments

Chromosome	SNP	Type	Position	Resistance allele freq.	2011F		2012S		2012F		2013S		2013MH		2013Hays		2014MH		2014Hays		GH_BLUP		Field_BLUP	
					p-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2
1A	BS00011787_51	90K	34.28	0.49	-	-	-	-	-	-	-	-	3.30E-05	0.092	-	-	-	-	-	-	-	-	-	-
1A	Kukri_c22508_119	90K	-	0.51	-	-	-	-	-	-	-	-	5.43E-05	0.086	-	-	-	-	-	-	-	-	-	-
1A	Kukri_c60564_136	90K	50.20	0.93	6.54E-05	0.091	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1A	IACX742	90K	51.33	0.93	7.44E-05	0.089	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1A	BS00094925_51	90K	55.73	0.93	7.44E-05	0.089	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1D	Ex_c6765_2118	90K	48.90	0.83	-	-	-	-	-	-	5.74E-04	0.067	-	-	-	-	-	-	-	-	-	-	-	-
1D	wsnp_Ku_c19622_29138795	90K	-	0.67	-	-	-	-	-	-	3.38E-04	0.073	-	-	-	-	-	-	-	-	-	-	-	-
1D	GWM337	SSR	-	0.18	-	-	4.98E-04	0.178	-	-	1.30E-04	0.133	-	-	-	-	-	-	-	-	4.70E-04	0.135	-	-
2D	BobWhite_c1477_315	90K	-	0.33	-	-	-	-	-	-	9.48E-04	0.062	-	-	-	-	-	-	-	-	-	-	-	-
2D	GWM539	SSR	-	0.11	-	-	-	-	-	-	-	-	5.00E-04	0.117	-	-	-	-	6.00E-04	0.151	-	-	3.38E-04	0.167
3A	wsnp_Ex_rep_c67702_66370241	9K	9.12	0.78	4.42E-05	0.095	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3A	wsnp_Ra_c2339_4506620	9K	9.12	0.38	-	-	-	-	-	-	-	-	5.28E-05	0.087	-	-	-	-	-	-	-	-	5.66E-05	0.085
3A	BS00094057_51	90K	9.12	0.76	9.06E-05	0.087	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3A	wsnp_Ex_c10014_16477392	90K	9.12	0.65	-	-	-	-	-	-	-	-	6.91E-05	0.084	-	-	-	-	-	-	-	-	7.22E-05	0.083
3A	RAC875_c76948_970	90K	-	0.74	8.80E-05	0.088	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3A	TaPHS1.1	STS	-	0.78	5.55E-05	0.093	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3A	BARC57.1	SSR	-	0.79	-	-	-	-	-	-	-	-	5.13E-06	0.140	8.52E-06	0.135	3.38E-05	0.158	3.00E-04	0.118	-	-	5.90E-06	0.154
3A	BARC57.2	SSR	-	0.79	-	-	-	-	-	-	-	-	6.74E-06	0.140	6.56E-06	0.142	9.46E-05	0.153	2.00E-04	0.137	-	-	3.78E-06	0.159
3AL	wsnp_Ex_c24085_33332723	90K	121.90	0.89	-	-	9.46E-04	0.062	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3AL	wsnp_BM137927A_Ta_2_1	90K	121.90	0.89	-	-	7.41E-04	0.064	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3AL	GENE.1464_73	90K	164.20	0.88	-	-	7.41E-04	0.064	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3AL	wsnp_Ku_c5359_9531713	9K	L201.88	0.89	-	-	2.50E-04	0.076	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3AL	Tamyb10-A1-66	STS	114.02	0.61	-	-	-	-	-	-	-	-	4.67E-07	0.121	1.82E-05	0.092	3.00E-04	0.068	2.83E-05	0.086	-	-	2.01E-06	0.109
3AL	Tamyb10-A1-74	STS	114.02	0.59	-	-	-	-	-	-	-	-	8.07E-07	0.116	3.47E-05	0.086	-	-	1.16E-04	0.073	-	-	1.12E-05	0.094
3AL	WMC559	SSR	107.20	0.89	-	-	-	-	-	-	-	-	-	-	9.00E-04	0.104	-	-	-	-	-	-	-	-
3B	wsnp_BE446087B_Ta_2_1	9K	46.66	0.91	-	-	-	-	-	-	-	-	-	-	3.22E-05	0.093	-	-	-	-	-	-	-	-
3B	RAC875_c15722_1081	90K	50.26	0.90	-	-	-	-	-	-	-	-	-	-	6.75E-05	0.085	-	-	-	-	-	-	-	-
3BL	BARC77	SSR	-	0.94	-	-	5.33E-04	0.123	-	-	-	-	4.00E-04	0.076	-	-	-	-	2.00E-04	0.095	-	-	4.82E-04	0.081
3BL	GWM108	SSR	-	0.81	-	-	-	-	-	-	-	-	9.00E-04	0.070	-	-	2.00E-04	0.075	-	-	-	-	5.68E-04	0.073
3BL	GWM181	SSR	-	0.92	3.00E-04	0.122	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3BL	GWM247	SSR	-	0.92	9.00E-04	0.098	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3D	IAAV1578	90K	0.00	0.44	-	-	9.77E-04	0.061	-	-	-	-	-	-	-	-	2.63E-05	0.098	-	-	-	-	-	-
3D	BS00021687_51	90K	0.00	0.44	-	-	-	-	-	-	-	-	-	-	-	-	3.70E-05	0.095	-	-	-	-	-	-
3D	BS00080151_51	90K	0.00	0.44	-	-	-	-	-	-	-	-	-	-	-	-	3.70E-05	0.095	-	-	-	-	-	-

Chromosome	SNP	Type	Position	Resistance allele freq.	2011F		2012S		2012F		2013S		2013MH		2013Hays		2014MH		2014Hays		GH_BLUP		Field_BLUP	
					p-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2
3D	BS00085464_51	90K	0.00	0.46	-	-	-	-	-	-	-	-	-	-	-	-	9.86E-06	0.109	9.58E-05	0.073	-	-	-	-
3D	BS00108577_51	90K	0.00	0.47	-	-	-	-	-	-	-	-	-	-	-	-	1.87E-05	0.102	-	-	-	-	-	-
3D	Excalibur_c20559_98	90K	0.00	0.47	-	-	-	-	-	-	-	-	-	-	-	-	1.87E-05	0.102	-	-	-	-	-	-
3D	IAAV2980	90K	0.00	0.47	-	-	-	-	-	-	-	-	-	-	-	-	1.10E-05	0.108	-	-	-	-	-	-
3D	D_contig59199_227	90K	0.00	0.47	-	-	-	-	-	-	-	-	-	-	-	-	1.16E-05	0.107	-	-	-	-	-	-
3D	BS00022669_51	90K	-	0.45	-	-	-	-	-	-	-	-	-	-	-	-	4.43E-05	0.093	-	-	-	-	-	-
3D	BS00076298_51	90K	-	0.48	-	-	-	-	-	-	-	-	-	-	-	-	4.94E-05	0.091	-	-	-	-	-	-
3D	BobWhite_c621_1218	90K	-	0.31	-	-	-	-	-	-	-	-	-	-	-	-	5.28E-05	0.091	-	-	-	-	-	-
3D	BS00067117_51	90K	-	0.47	-	-	-	-	-	-	-	-	-	-	-	-	8.26E-06	0.111	8.56E-05	0.074	-	-	-	-
3D	CAP8_c5043_351	90K	-	0.47	-	-	-	-	-	-	-	-	-	-	-	-	1.87E-05	0.102	-	-	-	-	-	-
3D	Excalibur_c9485_351	90K	-	0.47	-	-	-	-	-	-	-	-	-	-	-	-	1.87E-05	0.102	-	-	-	-	-	-
3D	Kukri_c50527_241	90K	-	0.35	-	-	8.46E-04	0.063	-	-	-	-	-	-	-	-	2.97E-05	0.097	-	-	-	-	-	-
3D	tplb0062k24_584	90K	-	0.45	-	-	-	-	-	-	-	-	-	-	-	-	6.08E-05	0.089	-	-	-	-	-	-
3D	BobWhite_c3111_636	90K	-	0.32	-	-	-	-	-	-	-	-	-	-	-	-	1.15E-05	0.108	8.88E-05	0.074	-	-	7.09E-05	0.083
3DL	BS00067163_51	90K	92.34	0.52	-	-	-	-	-	-	-	-	7.79E-05	0.083	-	-	-	-	-	-	-	-	-	-
3DL	GENE-1785_118	90K	-	0.42	-	-	-	-	-	-	-	-	8.22E-05	0.082	-	-	-	-	-	-	-	-	-	-
3D	Tamyb10-D1-93	STS	-	0.56	-	-	-	-	-	-	-	-	6.00E-04	0.057	3.00E-04	0.067	-	-	-	-	-	-	-	-
4A	BS00037019_51	90K	76.40	0.19	-	-	-	-	2.22E-09	0.217	-	-	-	-	-	-	-	-	-	-	1.67E-06	0.133	-	-
4A	Ex_c66324_1151	90K	76.97	0.42	-	-	-	-	3.67E-17	0.476	-	-	-	-	-	-	2.37E-05	0.099	-	-	1.34E-12	0.315	-	-
4A	wsnp_Ex_c13031_20625900	90K	76.97	0.42	-	-	-	-	1.79E-16	0.451	-	-	-	-	-	-	2.93E-05	0.097	-	-	2.54E-12	0.306	-	-
4A	wsnp_Ex_rep_c66324_64493429	90K	76.97	0.43	-	-	-	-	1.51E-16	0.453	-	-	-	-	-	-	2.60E-05	0.098	-	-	2.99E-12	0.304	-	-
4A	BS00072025_51	90K	76.97	0.32	-	-	-	-	1.20E-09	0.225	-	-	-	-	-	-	-	-	-	-	1.18E-06	0.137	-	-
4A	IAAV615	90K	76.97	0.19	-	-	-	-	1.74E-09	0.220	-	-	-	-	-	-	-	-	-	-	1.40E-06	0.135	-	-
4A	IACX2890	90K	76.97	0.22	-	-	-	-	3.84E-05	0.097	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4A	RAC875_c21369_425	90K	78.11	0.61	-	-	-	-	7.53E-05	0.090	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4A	wsnp_Ku_c4342_7887834	90K	78.11	0.61	-	-	-	-	7.53E-05	0.090	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4A	BS00023151_51	90K	78.11	0.18	-	-	-	-	4.40E-07	0.150	-	-	-	-	-	-	-	-	-	-	2.31E-05	0.103	-	-
4A	wsnp_Ex_c19207_28125389	9K	82.65	0.11	-	-	-	-	2.53E-05	0.102	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4A	Excalibur_c30378_673	90K	90.65	0.47	-	-	-	-	5.28E-06	0.120	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4A	RAC875_c11524_553	90K	90.65	0.62	-	-	-	-	2.65E-06	0.129	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4A	wsnp_Ex_c11619_18714738	90K	90.65	0.59	-	-	-	-	6.61E-06	0.118	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4A	PM19A1K1	STS	-	0.68	-	-	-	-	3.22E-05	0.099	-	-	-	-	-	-	-	-	-	-	2.99E-05	0.100	-	-
4A	wsnp_Ex_rep_c104448_89161562	9K	-	0.19	-	-	-	-	1.21E-08	0.195	-	-	-	-	-	-	-	-	-	-	1.79E-05	0.106	-	-
4A	wsnp_Ex_c612_1213451	9K	-	0.45	-	-	-	-	2.61E-06	0.129	-	-	-	-	-	-	-	-	-	-	4.12E-05	0.096	-	-
4A	wsnp_JD_c38619_27992279	90K	-	0.45	-	-	-	-	4.46E-06	0.122	-	-	-	-	-	-	-	-	-	-	4.43E-05	0.095	-	-
4A	ZXQ118	STS	-	0.17	-	-	-	-	3.39E-05	0.144	-	-	-	-	-	-	-	-	-	-	2.78E-05	0.109	-	-
4A	BARC236	SSR	92.92	0.23	-		-		-		-		7.41E-05		0.141		-		-		-		-	
4A	WMC757	SSR	-	0.11	-		-		9.84E-04		0.119		-		-		-		-		-		-	
5A	Excalibur_c34426_723	90K	35.36	0.84	-	-	-	-	-	-	4.20E-04	0.070	-	-	-	-	-	-	-	-	-	-	-	-
5A	BobWhite_c4004_61	90K	35.36	0.83	-	-	-	-	-	-	2.62E-04	0.076	-	-	-	-	-	-	-	-	-	-	-	-
5A	BS00021873_51	90K	35.36	0.84	-	-	-	-	-	-	3.26E-04	0.073	-	-	-	-	-	-	-	-	-	-	-	-

Chromosome	SNP	Type	Position	Resistance allele freq.	2011F		2012S		2012F		2013S		2013MH		2013Hays		2014MH		2014Hays		GH_BLUP		Field_BLUP	
					p-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2	*p*-value	R^2
5A	Excalibur_c54774_408	90K	47.99	0.77	-	-	5.60E-04	0.067	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5A	Excalibur_c24051_1028	90K	-	0.66	-	-	7.88E-04	0.063	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6B.2	wsnp_Ex_c19525_28494827	90K	94.46	0.91	-	-	-	-	-	-	-	-	1.21E-05	0.102	8.19E-06	0.108	-	-	-	-	-	-	2.81E-05	0.093
6B.2	Excalibur_c15109_942	90K	95.60	0.64	-	-	-	-	-	-	4.29E-04	0.070	-	-	-	-	-	-	-	-	-	-	-	-
6B.2	RAC875_c7332_955	90K	-	0.91	-	-	-	-	-	-	-	-	-	-	4.51E-06	0.115	-	-	-	-	-	-	5.90E-05	0.085
6B	GWM88	SSR	-	0.91	-	-	9.55E-04	0.089	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7A	Excalibur_c53632_204	90K	68.49	0.88	4.04E-05	0.096	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7A	wsnp_Ex_c8614_14453388	9K	69.63	0.95	4.59E-06	0.121	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7A	wsnp_Ex_c26509_35755018	9K	69.63	0.95	1.41E-06	0.135	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7A	wsnp_Ex_c38981_46383475	9K	69.63	0.95	1.41E-06	0.135	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7A	wsnp_Ex_rep_c68405_67220388	9K	-	0.95	2.50E-05	0.102	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7A	WMC603-1	SSR	-	0.92	1.00E-04	0.254	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7A	GWM130	SSR	-	0.94	-	-	-	-	-	-	-	-	5.00E-04	0.106	-	-	-	-	-	-	-	-	-	-
1A/1D/3A/5B	BARC148	SSR	-	0.83	-	-	-	-	-	-	-	-	1.38E-08	0.228	3.38E-05	0.119	-	-	2.22E-05	0.145	-	-	3.43E-06	0.147
5ABD	BARC232-1	SSR	-	0.92	-	-	-	-	-	-	-	-	6.00E-04	0.103	-	-	4.40E-05	0.126	3.97E-05	0.136	-	-	3.80E-04	0.109

[a] The positions of markers on W7984 reference sequence

[b] Phenotypic variance explained by a significant marker significantly related to pre-harvest sprouting resistance

Table 3.4 Quantitative trait loci (QTL) for pre-harvest sprouting resistance identified in only one of the experiments conducted in the spring 2012 (2012S) and spring 2013 (2013S) greenhouse experiments and the 2013 Manhattan field experiment (2013MH)

Chromosome	Marker name	Marker type	Chromosome position (cM)[a]	Positive allele frequency	*p*	R^2 (%)[b]	Experiment
2B.1	*Excalibur_c1787_1199*	90K	7.97	0.81	2.97E-04	7.4	2013S
2B.1	*BS00044806_51*	90K	10.24	0.72	4.98E-04	6.9	2013S
2B.1	*Tdurum_contig51145_476*	90K	10.24	0.76	3.65E-04	7.2	2013S
2B.1	*BS00022203_51*	90K	-	0.17	7.08E-04	6.5	2013S
2B.1	*Excalibur_c3524_318*	90K	-	0.81	7.05E-04	6.5	2013S
2B.1	*Kukri_c16758_443*	90K	10.24	0.73	1.31E-04	8.3	2013S
2B.1	*wsnp_JD_c3288_4296662*	9K	10.24	0.77	7.57E-04	6.4	2013S
2B.1	*BS00065556_51*	90K	-	0.77	2.56E-04	7.6	2013S
2B.2	*wsnp_Ex_c13865_21720466*	9K	83.07	0.42	6.42E-04	6.6	2012S
2B.2	*wsnp_RFL_Contig3273_3319580*	90K	83.07	0.41	7.30E-04	6.4	2012S
2B.2	*RAC875_c26697_589*	90K	83.07	0.35	4.27E-04	7.0	2012S
2B.2	*Tdurum_contig28795_322*	90K	-	0.41	7.63E-04	6.4	2012S
6B.1	*RAC875_c23251_624*	90K	43.28	0.89	2.59E-05	9.4	2013MH
6B.1	*BS00066799_51*	90K	43.28	0.92	7.13E-05	8.3	2013MH
6B.1	*CAP8_c1361_367*	90K	43.28	0.89	2.59E-05	9.4	2013MH

[a] The positions of markers on W7984 reference sequence

[b] Phenotypic variance explained by a marker that was significantly associated with pre-harvest sprouting resistance

Table 3.5 Effect of grain color (GC) that was evaluated in the field at Enid, OK in 2010 (2010Enid) and the greenhouse at Manhattan KS 2011 (2011F_GH) on pre-harvest sprouting (PHS) resistance evaluated in four greenhouse experiments (GH_experiments) conducted in Manhattan, KS, and four field experiments conducted in Manhattan (MH) and Hays, KS in 2013 and 2014, respectively.

Experiments	2010Enid (GC)		2011F_GH (GC)	
	p	R^2 **(%)**[a]	***p***	R^2 **(%)**[a]
GH_experiments (PHS)	NS	-	NS	-
2013MH (PHS)	<2e-16	43.7	<2e-16	44.5
2013Hays (PHS)	<2e-16	42.9	<2e-16	43.6
2014MH (PHS)	1.27E-10	26.3	3.13E-11	27.5
2014Hays (PHS)	1.13E-15	35.6	3.41E-15	34.8
Field_BLUP [b](PHS)	<2e-16	43.9	<2e-16	44.1

[a] Phenotypic variance explained by grain color in each PHS experiment, which is derived from the analysis of variance (ANOVA) where grain color (GC) was used as the explanatory variable and pre-harvest sprouting (PHS) resistance as the response variable.

[b] Field_BLUP=Best Linear Unbiased Predictions calculated from all four field experiment.

Table 3.6 Common Quantitative trait loci (QTL) identified for grain color evaluated in 2010 field (2010Enid) and 2011 greenhouse (GH) experiments and pre-harvest sprouting resistance evaluated in Manhattan (MH) and Hays in 2013 and 2014 experiments, respectively

Chromosome	Marker name	Chromosome position (cM)[a]	Grain color						PHS resistance									
			2010Enid		2011GH		Mean		2013MH		2013Hays		2014MH		2014Hays		Field_BLUP	
			p	R^2 (%)[b]	*p*	R^2 (%)[b]	*p*	R^2 (%)[b]	*p*	R^2 (%)[c]	*p*	R^2 (%)[c]	*p*	R^2 (%)[c]	*p*	R^2 (%)[c]	*p*	R^2 (%)[c]
3AL	*Xwmc559-1*	107.20	1.25E-04	9.8	5.00E-04	10.6	1.57E-04	10.8	-	-	9.00E-04	10.4	-	-	-	-	-	-
3AL	*Tamyb10-A1-66*	114.02	8.75E-06	7.5	1.86E-04	10.2	2.25E-05	9.4	4.67E-07	12.1	1.82E-05	9.2	3.00E-04	6.8	2.83E-05	8.6	2.01E-06	10.9
3AL	*Tamyb10-A1-74*	114.02	3.12E-06	8.5	6.50E-05	11.1	7.32E-06	10.4	8.07E-07	11.6	3.47E-05	8.6	-	-	1.16E-04	7.3	1.12E-05	9.4
3DL	*BS00067163_51*	92.34	5.36E-08	15.0	8.39E-11	23.2	7.49E-10	19.7	7.79E-05	8.3	-	-	-	-	-	-	-	-
3DL	*Tamyb10-D1-93*	-	4.31E-11	21.9	3.66E-13	17.5	6.51E-13	21.0	6.00E-04	5.7	3.00E-04	6.7	-	-	-	-	-	-
1A/1D/3A/5B	*Xbarc148*	-	2.50E-07	16.6	1.03E-06	18.7	2.27E-07	18.4	1.38E-08	22.8	3.38E-05	11.9	-	-	2.22E-05	14.5	3.43E-06	14.7

[a] The marker positions in a chromosome based on W7984 reference map

[b] Phenotypic variance explained by a marker that is significantly associated with grain color

[c] Phenotypic variance explained by a marker that is significantly associated with pre-harvest sprouting resistance

Table 3.7 Environmental statistics of greenhouse and fields in Manhattan and Hays

	Day Temp.[a]	Night Temp.	Day length (h)
GH(May-June)	25±5	20±2	12
GH(Dec-Jan)	22±3	17±2	12
	Max Temp.[b]	**Min Temp.**	**Precip. (cm)**
2013MH	36.1	7.8	12.3
2014MH	33.9	9.4	27.6
2013Hays	41.1	5.3	6.6
2014Hays	40.0	5.9	19.4

[a] Greenhouse Day/Night temperature (°C) is expressed as Mean ± Standard Deviation

[b] Field temperature range (°C) and precipitation are calculated from May 1st to June 15th in 2013 and 2014. Data is from "www.usclimatedata.com"

Table 3.8 Kompetitive Allele Specific PCR assays developed from significant SNPs for the 4A pre-harvest sprouting resistance quantitative trait locus

KASP		Primer Sequence (5' to 3')	90K SNP	Seqence
KASP3743	Forward[T]*	GAAGGTGACCAAGTTCATGCTTCAGT TTGGCCAACCATGT[T]	*wsnp_Ex_rep_c6632 4_64493429*	TTAGAGAAGTCATGTTGCCAAGTACAACAGGTATTGTACCGACAAGGTCGTTATCA TTGAGGAATAGGAAGCTGAGTTGAGTCAGTTTGGCCAACCATGT[T/C]GGAACAAC ACCTTCAAATGAGTTCTCGCCAAGGGAAAGAGTTTGGAGGTATGGACAAGATGCA AAGCCCAATGGAATCTGACCTGTGAAACTATTACCTT
	Forward[S]	GAAGGTCGGAGTCAACGGATTTCAG TTTGGCCAACCATGT[C]		
	Reverse	TCTTGTCCATACCTCCAAAC		
KASP8081	Forward[T]	GAAGGTGACCAAGTTCATGCTGGTC CATCGTACTCGCAAAA[T]	*BS00037019_51*	AATCAGAACCCATCGCCCAATGTCCAGAACGGTCCATCGTACTCGCAAAA[T/C]CAT AACCCTTCTCCTGTTGCCCAGAACAGTCCATTGTTTTTGCAACACCA
	Forward[S]	GAAGGTCGGAGTCAACGGATTGGTC CATCGTACTCGCAAAA[C]		
	Reverse	AATGGACTGTTCTGGGCAAC		
KASP34562	Forward[T]	GAAGGTGACCAAGTTCATGCTTGGA GTCTGAAAGCATTCG[A]	*IAAV615*	ATGCACTCTGTTTGACTGCTTCTGTCCCTTACTTTGAGGATTCAGAATTAAGCTCTG TTTTTGCCTCCGTCTGCCAGAACTTGGAGTCTGAAAGCATTCG[A/G]CTCTATTAAAT TCAGGGTATTTTTATTGTCTGAATATTTGATTTGTGTTTTCCTATGATGCATGGAAAT TTGTAATCTCTGTCGGATTAAGGATTATTTA
	Forward[S]	GAAGGTCGGAGTCAACGGATTTGGA GTCTGAAAGCATTCG[G]		
	Reverse	TCCATGCATCATAGGAAAACA		
KASP34586	Forward[T]	GAAGGTGACCAAGTTCATGCTAAGG GGAGGTGATCGTGGA[T]	*IACX2890*	AGTGGCACCCGCATCGTTGATCGCGCACAATGCCGGAGTCGAAGGGGAGGTGATC GTGGA[T/G]AAAATCAAGGACAGCGAGTGGGAATTCGGCTACAACGCGATGACCGA CAAGCACGAGAA
	Forward[S]	GAAGGTCGGAGTCAACGGATTAAGG GGAGGTGATCGTGGA[G]		
	Reverse	TTGTAGCCGAATTCCCACTC		

*[T] is the SNP allele detected in Tutoumai A, [S] is the SNP allele detected in Siyang 936

Figure 3.1 Frequency distribution of grain color (GC) scores evaluated using a 1 to 4 scale (white, light red, red and dark red) in the association mapping population. The seeds were harvested from the Manhattan 2011 greenhouse (2011MH) experiment and the Enid 2010 field (2010 ENID) experiment.

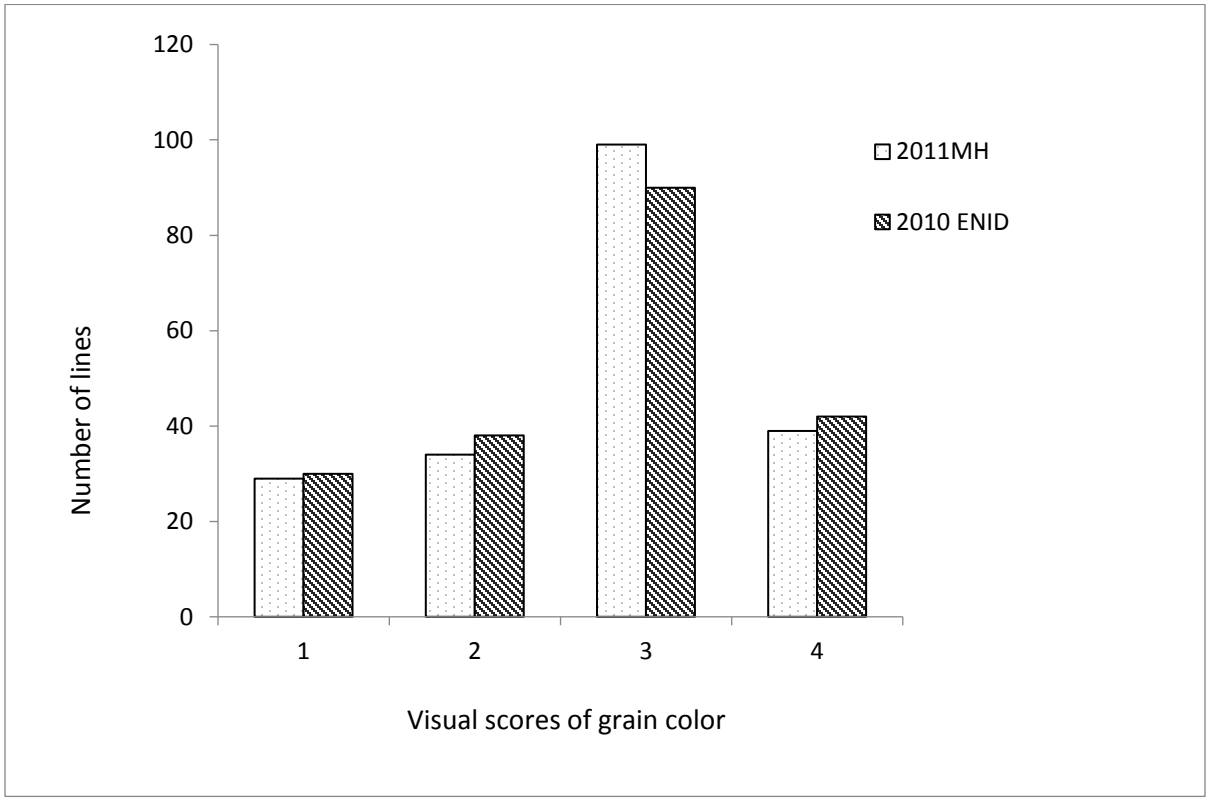


Figure 3.2 **Heatmaps showing (a) the relationships of pre-harvest sprouting data among four greenhouse (GH) experiments conducted at Manhattan, KS in fall 2011(11F_GH), fall 2012(12F_GH), spring 2012(12S_GH), spring 2013(13S_GH) and four field experiments conducted at Manhattan in 2013 (13MH_FD) and 2014 (14MH_FD), and Hays in 2013 (14Hays_FD) and 2014 (14Hays_FD), and (b) the relationships and grouping of wheat accessions that were determined using the mean pre-harvest sprouting data collected from all four greenhouse and four field experiments. Similarity levels increase from light yellow (the lowest similarity) to dark red (the highest similarity).**

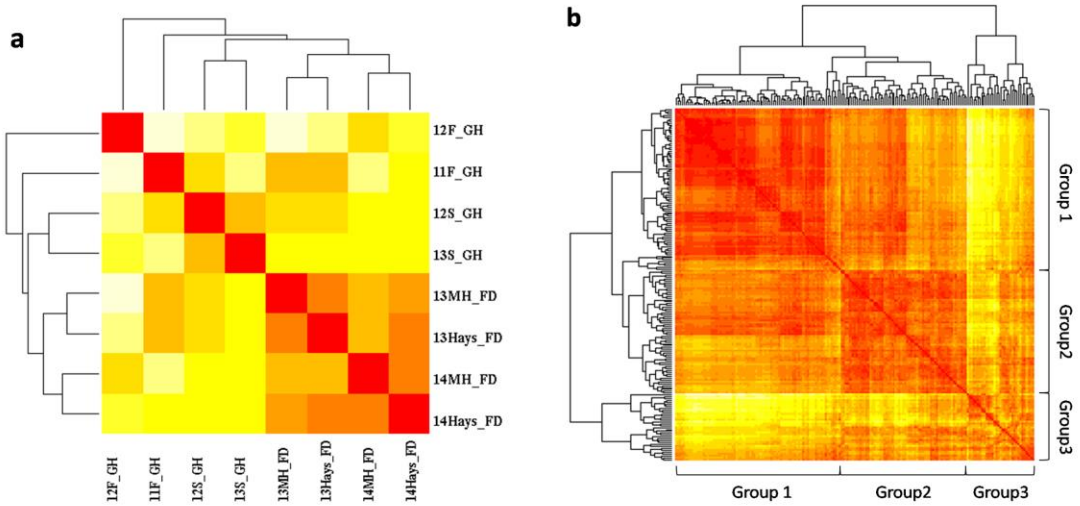


Figure 3.3 Distribution of grain color (GC) scores in the association mapping population predicted by *Tamyb10* gene markers. Six allele combinations of three GC genes on chromosomes A, B and D separated 185 accessions into eight genotypes. Lower case represents a white grain allele and upper case represents a red grain allele in each locus. The three letters in each genotype represent three gene loci in the chromosomes A, B and D, respectively, e.g. Abc indicates red allele on 3A and white alleles on 3B and 3D. GC scores used a 1-4 scale with 1 for white grain and 4 for red grain.

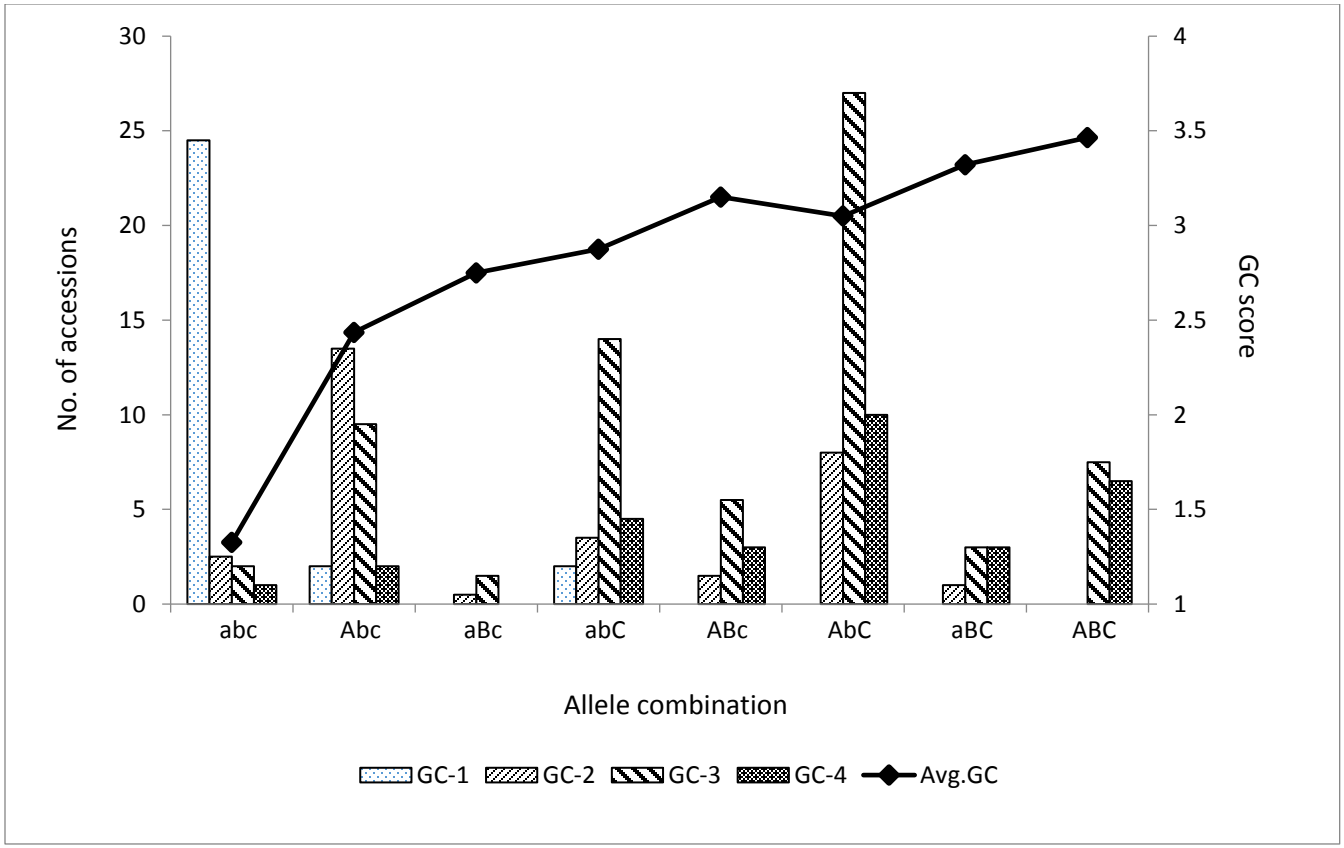


Figure 3.4 Interval mapping (IM) of a quantitative trait locus (QTL) for pre-harvest sprouting (PHS) resistance on chromosome 4A using SSRs, GBS-SNPs and SNPs identified from genome-wide association study (GWAS). The line parallel to the X-axis is the threshold line for the significant LOD value of 2.42 ($P < 0.05$). Genetic distances are in centiMorgans (cM).

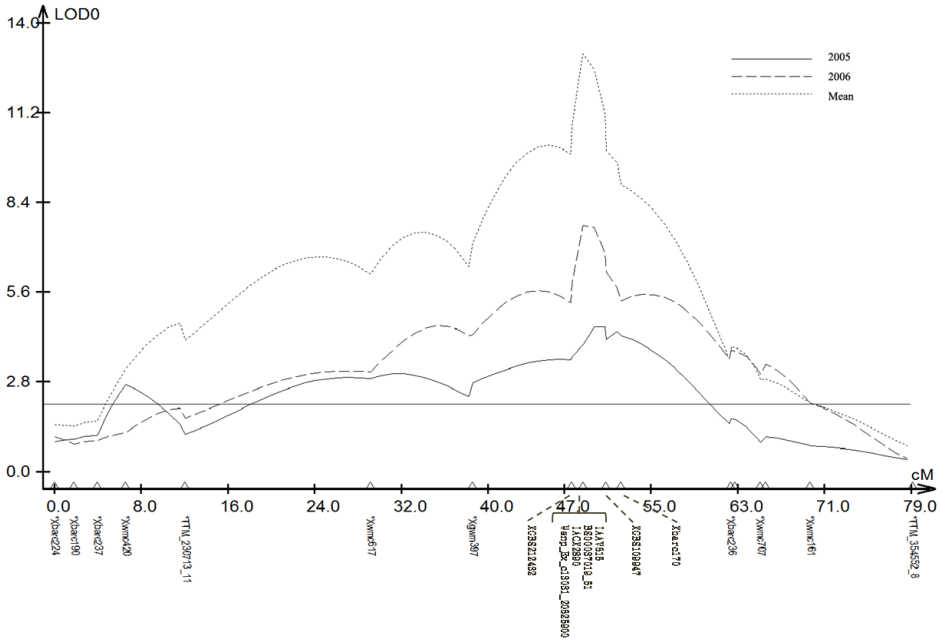
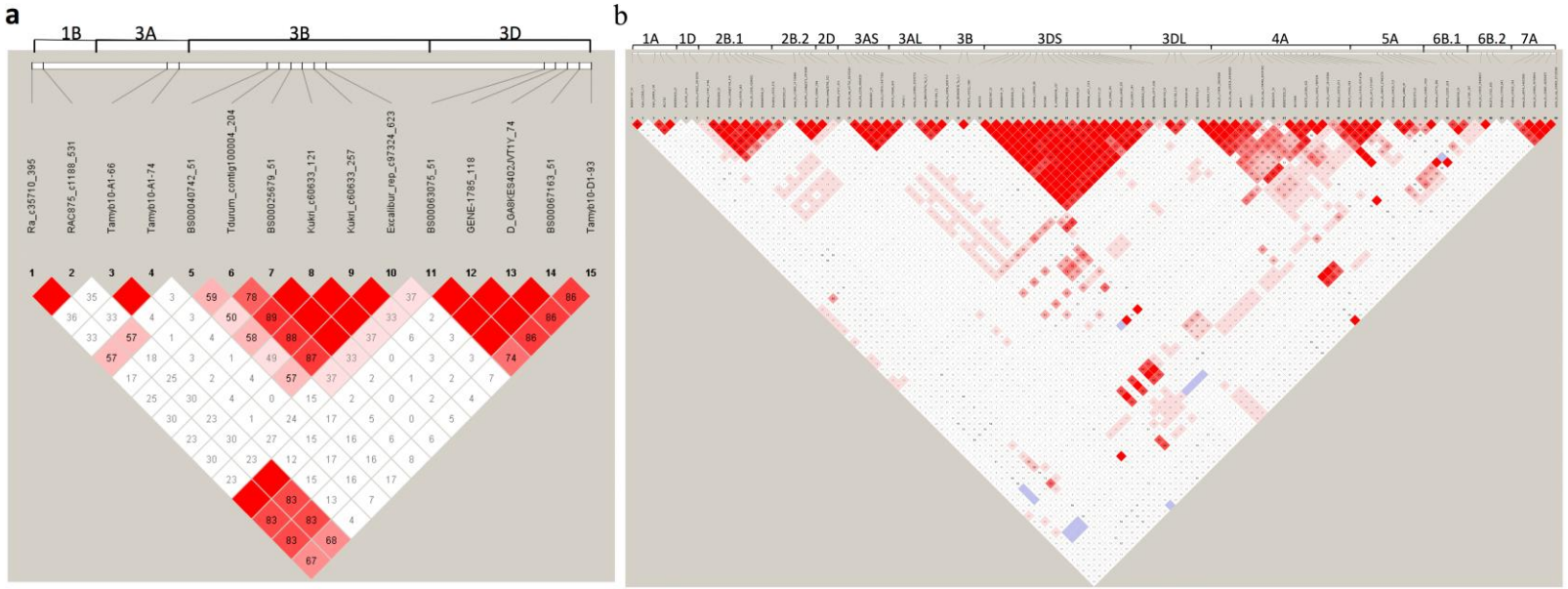


Figure 3.5 LD plots of SNP markers that showed significantly association with GC (a) and pre-harvest sprouting (PHS) resistance (b). The chromosome numbers are labeled above the chromosome maps (the long white bar) and marker names are labeled between the LD plot and chromosome maps.



Chapter 4 - Effects of *TaPHS1* and *TaMKK3-A* Genes on Wheat PHS Resistance

Abstract

TaPHS1 on wheat chromosome 3AS and *TaMKK3-A* on chromosome 4AL are two cloned genes that show a major effect on pre-harvest sprouting (PHS) resistance, and are independent from grain color (GC). In this study, we used marker-assisted backcrossing (MAB) to introgress *TaPHS1* and *TaMKK3-A* from two PHS resistant sources, "Tutoumai A" and AUS1408, to a susceptible wheat line, NW97S186, to investigate individual effects of the two genes and their combined effects in different environments. *TaPHS1* showed a significant main effect and interactions with environments and genetic backgrounds (GBG), whereas the *TaMKK3-A* gene had a significant main effect and only interacted with environments. The two genes showed additive effects on PHS resistance and the combined effects of *TaPHS1* and *TaMKK3-A* was larger in the greenhouse than that in the field, indicating pyramiding these two QTLs can increase PHS resistance, and such effect is more obvious when wheat plants are grown in a mild environment such as in the greenhouse than in a dry and hot environment during maturation.

Introduction

PHS resistance is a complex trait controlled by several major QTLs and many minor QTLs. PHS resistance QTLs have been reported on almost all wheat chromosomes, among which causal genes for the non-GC related QTLs on chromosome 3AS and 4AL have been cloned and designated as *TaPHS1* and *TaMKK3-A*, respectively (Nakamura et al. 2011, Liu et al. 2013, Torada et al. 2016). *TaPHS1*, annotated as a MOTHER OF FLOWERING TIME (*TaMFT*)-like gene, positively regulates wheat PHS resistance. Three single nucleotide polymorphisms (SNPs) have been identified to associate with PHS resistance. One SNP in the promoter region (-222) increases seed dormancy at low temperatures during seed development (Nakamura et al. 2011), and two other SNPs in the gene-coding region (+646, +666) decrease seed dormancy by generating a mis-splicing site and a premature stop codon, respectively, to form a truncated nonfunctional transcript and thus increase PHS susceptiblity (Liu et al. 2013). These mis-splicing *TaPHS1* mutation were involved in wheat domestication (Liu et al. 2015). Another major gene, *Phs1*, for both PHS resistance and seed dormancy was mapped on chromosome arm 4AL in both white and red wheat (Kato et al. 2001, Mares et al. 2001, Mares et al. 2005, Torada et al. 2005, Chen et al. 2008, Ogbnnaya et al. 2008, Singh et al. 2010, Liu et al. 2011, Cabral et al. 2014). *TaMKK3-A*, a mitogen-activated protein kinase kinase 3 (MKK3) was cloned by map-based cloning as the candidate gene of *Phs1* (Torada et al. 2016). A single SNP that causes a nonsynonymous amino acid substitution in the kinase domain was reported to be the functional SNP in the gene (Torada et al. 2016).

MFT has been considered a negative regulator of ABA sensitivity for seed germination in Arabidopsis (Xi et al. 2010), and *TaPHS1* is proposed as a messenger that coordinates performance between tissues in seed germination (Nakamura et al. 2011). Similarly, protein

kinases play critical roles in signal transduction pathways, and *MKK* genes are important in protein phosphorylation in ABA signaling (Torada et al. 2016). However, how *TaPHS1* and *TaMKK3-A* interact with each other to regulate seed dormancy and PHS resistance is still unknown. The current study was to investigate the individual and combined genetic effects of the two genes in different environments by transferring both *TaPHS1* or/and *TaMKK3-A* into a susceptible wheat line using marker-assisted backcross (MAB).

Materials and methods

Plant materials and PHS evaluation

"Tutoumai A" is a highly PHS resistant Chinese landrace (Chen et al. 2008), and AUS1408 is a spring wheat line from the Transvaal region of South Africa. They are both white-grained wheat, and have been used as resistant parents in the 4AL QTL mapping studies (Chen et al. 2008, Zhang et al. 2008). Although they were not reported to carry the 3A QTL, both accessions carry the *TaPHS1* resistance allele when they were assayed with the *TaPHS1* function marker. Therefore, "Tutoumai A" and AUS1408 were used as the donors for both *TaPHS1* and *TaMKK3-A*. NW97S186, a PHS susceptible hard white winter wheat cultivar developed by USDA-ARS at the University of Nebraska-Lincoln, was used as the common recurrent parent. The backcross procedure is described in Fig. 4.1. In brief, "Tutoumai A" and AUS1408 were crossed to NW97S186, respectively, to obtain "Tutoumai A"/NW97S186 F_1 and AUS1408/NW97S186 F_1. Their F_1 plants were backcrossed to NW97S186 twice to develop BC_2F_1 plants. The BC_1F_1 plants and BC_2F_1 plants were genotyped with the two gene markers in the *TaPHS1* coding region and one SNP tightly linked to *TaMKK3-A* to select the heterozygous plants for both genes to be used for further backcrossing or generation advancement. At least 10 heterozygous plants with both genes were identified among the BC_2F_1's in each cross. The

selected BC_2F_1 plants were selfed and the double homozygous BC_2F_2 were selected with the same markers (Fig. 4.1). The selected double homozygous BC_2F_2 and later generations were used to evaluate germination rate in the greenhouse experiments conducted at Manhattan in fall of 2015 and spring of 2016, as well as in the field of Manhattan and Hays in 2016 as described in Chapter 3. The physiologically matured spikes were dried for 10 d before germination in all the four experiments.

Statistical analysis

Four-way analysis of variance (ANOVA) was conducted using PROC GLM procedure in SAS 9.3 (SAS institute Inc., Cary, NC) with environment, genetic background and genotypes as fixed effects. Environments referred to the four experiments and genetic backgrounds referred to the two donors, "Tutoumai A" and AUS1408. Only homozygous genotypes of the *TaPHS1* and *TaMKK3-A* genes were studied, with lower case letters for susceptible genotypes and upper case letters for resistant genotypes. Least-squared means were compared under the protection of overall *F*-test at a significant level of 0.05.

Results

Selection of backcrossing progenies

Among the 42 double homozygotes selected from BC_2F_2 lines of the cross of NW97S186/"Tutoumai A" (N/T), seven were the AABB genotype, where 'A' represented the resistance allele of *TaPHS1* and 'B' represented the resistance allele of *TaMKK3-A*, 11 lines were AAbb genotype, 15 lines were aaBB genotype and nine lines were aabb genotype. Among the selected 44 BC_2F_2 progenies of the NW97S186/AUS1408 (N/A) cross, 18 lines were AABB genotype, nine lines were AAbb genotype, 11 lines were aaBB genotype and six lines were aabb genotype.

In each backcross population, the parents showed germination rates that were close to the extreme germination rates in the selected progenies (Table 4.1), indicating no obvious transgressive segregation and that "Tutoumai A" or AUS1408 carry the resistance alleles for both QTLs. In each experiment, the mean germination rates were similar between the two backcrossing populations. The spring greenhouse experiment showed the highest mean germination rates of 68.7% and 58.1% in the N/T and N/A populations, respectively, while the Manhattan field experiment showed the lowest mean germination rates of 34.2% and 34.8% in the N/T and N/A populations, respectively (Table 4.1). Generally, larger variances in germination rates were observed in the greenhouse experiments than in the field experiments, indicating that the growing environments greatly influence the expression of PHS resistance genes in wheat (Table 4.1).

Effects of *TaPHS1* and *TaMKK3-A* genes on PHS resistance in the greenhouse and field experiments

Overall ANOVA revealed that environments, genetic backgrounds (GBG), and genotypes (*TaPHS1* and *TaMKK3-A*) could explain 56.4% of the phenotypic variance for PHS resistance. Significant main effects were identified for environment and genotypes, and interactions were significant for environment by *TaPHS1*, environment by *TaMKK3-A*, and environment by GBG by *TaPHS1* (Table 4.3). Therefore, the main effect of *TaPHS1* needs to be investigated in each donor background under different environments, whereas the effect of *TaMKK3-A* could be estimated in the four environments without considering the donor background effect.

Effects of *TaPHS1* from "Tutoumai A" were significant on PHS resistance in the spring and fall greenhouse experiments with 29.4% and 22.5% reduction in germination rates,

respectively (Fig. 4.2). However, the effects of *TaPHS1* from AUS1408 were significant in the spring greenhouse experiment, and both Manhattan and Hays field experiments with 26.5%, 14.1% and 18.7% reduction in germination rates, respectively (Fig. 4.2). *TaMKK3-A* showed significant effects on PHS resistance in the spring and fall greenhouse experiments, and the Manhattan field experiments with 18.8%, 22.8% and 9.6% reduction in germination rates, respectively (Fig. 4.3).

Combined genetic effects of *TaPHS1* and *TaMKK3-A*

The combined effects of *TaPHS1* and *TaMKK3-A* varied with different genetic backgrounds across environments. In the N/T population, the combined effect was significant in the greenhouse experiments, but not in the field experiments. In the greenhouse experiments, adding either of the resistance genes (AA or BB) significantly reduced germination rates, and a more reduction in germination rate was observed when a wheat line carried both resistance genes compared to a line with a single gene (Table 4.4). In the N/A population, introgression of a single resistance gene did not increase PHS resistance significantly, whereas adding both genes significantly reduced germination rate (Table 4.4). The effect of combining *TaPHS1* with *TaMKK3-A* on PHS resistance was larger in the greenhouse experiments than in the field experiments, suggesting that the greenhouse conditions were more favorable to the expression of both genes in this study.

Discussion

PHS resistance is a complex trait that is not only controlled by seed dormancy (SD) (Bewley and Black 1982, Anderson *et al*. 1993), but also affected by GC (Gfeller and Svejda 1960, Groos *et al*. 2002), spike morphology, as well as environmental factors such as temperature, moisture and photoperiod after flowering (Argel *et al*. 1983, Ceccato *et al*. 2011).

In this study, we demonstrated that both cloned genes, *TaPHS1* and *TaMKK3-A*, for PHS resistance showed significant interactions with the environments (Table 4.3). On average, larger individual and combined effects of the two genes were detected in the greenhouse experiments than in the field experiments. This observation was possibly due to the fact that the plants had extended maturation period under greenhouse conditions that favors the gene expression. In addition, *TaMKK3-A* showed a larger effect in the fall greenhouse experiment than in other experiments (Fig. 4.3), suggesting that low temperature might up-regulate the expression of 4A QTL (Barrero et al., 2015). However, *TaPHS1* was more effective on reducing germination rate for plants grown in the spring greenhouse where temperature fore wheat seed development was higher than the fall greenhouse (Fig. 4.2), which was contradictory to the previous result that low temperature increased *TaPHS1* expression level in developing seeds (Nakamura et al. 2011). Other environmental factors such as humidity, photoperiod or light quality might also contributed to such discrepancy, because the *TaPHS1* gene might show a similar response to those environmental factors as *FT*-like and *TFL1*-like genes did in other species (Rohde and Bhalerao 2007, Shalit et al. 2009, Nakamura et al. 2011). *TaPHS1* and *TaMKK3-A* demonstrated various effects on germination rates (Fig. 4.2 & 4.3) in the two field experiments where they had similar temperature but different precipitations, indicating that humidity might also play an important role in affecting those gene expressions.

Significant environment by GBG by *TaPHS1* was observed in this study. In the fall greenhouse experiment, *TaPHS1* from "Tutoumai A" significantly reduced reduced germination rates, whereas *TaPHS1* from AUS1408 did not (Fig. 4.2). However, the result was opposite in the two field experiments (Fig. 4.2). Considering we did not conduct background marker-assisted

selection, it is likely that *TaPHS1* interacted with other QTLs in both "Tutoumai A" and AUS1408 backgrounds.

TaPHS1 and *TaMKK3-A* are the two major genes cloned for PHS resistance. In the current study, one of the two genes may not provide adequate protection from PHS resistance in some experiments, but pyramiding both genes can significantly reduce germination rates in most experiments (Table 4.4). Significant genetic-by-environment interactions of the two genes indicate that *TaPHS1* and *TaMKK3-A* can be more effective in wheat planting areas with mild climate during maturation. Gene markers for *TaPHS1* and SNPs closely link to *TaMKK3-A* have been shown to be useful in MAB, thus they can be applied in breeding for selecting the two genes to improve PHS resistance. However, in this study, the selected backcrossing progenies with both the resistance genes still showed higher average germination rates than their resistant donors in most experiments (Table 4.1 & 4.4), suggesting that other minor genes may be present in both donor parents and identifying and pyramiding these minor resistance QTLs with *TaPHS1* and *TaMKK3-A* can enhance levels of wheat PHS resistance.

References

Anderson JA, Sorrells ME, Tanksley SD (1993) RFLP analysis of genomic regions associated with resistance to preharvest sprouting in wheat. Crop Sci 33:453–459

Argel PJ, Humphreys LR (1983) Environmental effects on seed development and hardseededness in Stylosanthes hamata cv. Verano. I. Temperature. Crop Pasture Sci 34:261-270

Barrero JM, Cavanagh C, Verbyla KL, Tibbits JF, Verbyla AP, Huang BE, Rosewarne MG, Stephen S, Wang P, Whan A, Rigault P, Hayden JM, Gubler F (2015) Transcriptomic analysis of wheat near-isogenic lines identifies *PM19-A1* and *A2* as candidates for a major dormancy QTL. Genome Biol 16:93

Bewley JD, Black M (1982) Physiological and biochemistry of seeds in relation to germination, vol 2. Springer, Heidelberg pp 61–81

Cabral AL, Jordan MC, McCartney CA, You FM, Humphreys DG, MacLachlan R, Pozniak CJ (2014) Identification of candidate genes, regions and markers for pre-harvest sprouting resistance in wheat (*Triticum aestivum* L.). BMC Plant Biol 14:1

Ceccato DV, Daniel Bertero H, Batlla D (2011) Environmental control of dormancy in quinoa (*Chenopodium quinoa*) seeds: two potential genetic resources for preharvest sprouting tolerance. Seed Sci Res 21:133–141

Chen CX, Cai SB, Bai GH (2008) A major QTL controlling seed dormancy and pre-harvest sprouting resistance on chromosome 4A in a Chinese wheat landrace. Mol Breed 21:351-358

Gfeller F, Svejda F (1960) Inheritance of post-harvest seed dormancy and kernel color in spring wheat lines. Can J Plant Sci 40:1–6

Groos C, Gay G, Perretant MR, Gervais L, Bernard M, Dedryver F, Charmet G (2002) Study of the relationship between pre-harvest sprouting and grain color by quantitative trait loci analysis in a white × red grain bread-wheat cross. Theor Appl Genet 104:39–47

Kato K, Nakamura W, Tabiki T, Miura H (2001) Detection of loci controlling seed dormancy on group 4 chromosomes of wheat and comparative mapping with rice and barley genomes. Theor Appl Genet 102:980–985

Liu S, Bai G, Cai S, Chen C (2011) Dissection of genetic components of preharvest sprouting resistance in white wheat. Mol Breed 27:511–523

Liu S, Sehgal SK, Li J, Lin M, Trick HN, Yu J, Gill BS, Bai G (2013) Cloning and characterization of a critical regulator for preharvest sprouting in wheat. Genetics 195:263-273

Liu S, Sehgal SK, Lin M, Li J, Trick HN, Gill BS, Bai G (2015) Independent mis-splicing mutations in *TaPHS1* causing loss of preharvest sprouting (PHS) resistance during wheat domestication. New Phytol 208:928-935

Mares DJ (1987) Pre-harvest sprouting tolerance in white-grained wheat. In '4th International Symposium on Preharvest Sprouting in Cereals'. (Ed. DJ Mares) pp 64–74. (Westview Press: Boulder, CO)

Mares D, Mrva K, Cheong J, Williams K, Watson B, Storlie E, Sutherland M, Zou Y (2005) A QTL located on chromosome 4A associated with dormancy in white- and red-grained wheats of diverse origin. Theor Appl Genet 111:1357–1364

Mares DJ, Mrva K (2001) Mapping quantitative trait loci associated with variation in grain dormancy in Australian wheat. Crop Pasture Sci 52:1257-1265

Nakamura S, Abe F, Kawahigashi H, Nakazono K, Tagiri A, Matsumoto T, Utsugi S, Ogawa T,

Handa H, Ishida H, Mori M, Kawaura K, Ogihara Y, Miura H (2011) A wheat homolog of MOTHER OF FT AND TFL1 acts in the regulation of germination. The Plant Cell 23:3215–3229

Ogbonnaya FC, Imtiaz M, Ye G, Hearnden PR, Hernandez E, Eastwood RF, van Ginkel M, Shorter SC, Winchester JM (2008) Genetic and QTL analyses of seed dormancy and preharvest sprouting resistance in the wheat germplasm CN10955. Theor Appl Genet 116:891–902

Rohde A and Bhalerao RP (2007) Plant dormancy in the perennial context. Trends Plant Sci 12: 217–223

Shalit A, Rozman A, Goldshmidt A, Alvarez JP, Bowman JL, Eshed Y, Lifschitz E (2009) The flowering hormone florigen functions as a general systemic regulator of growth and termination. Proc Natl Acad Sci USA 106: 8392–8397

Singh R, Matus-Cádiz M, Båga M, Hucl P, Chibbar RN (2010) Identification of genomic regions associated with seed dormancy in white grained wheat. Euphytica 174:391–408

Torada A, Ikeguchi S, Koike M (2005) Mapping and validation of PCR-based markers associated with a major QTL for seed dormancy in wheat. Euphytica 143:251–255

Torada A, Koike M, Ogawa T, Takenouchi Y, Tadamura K, Wu J, Matsumoto T, Kawaura K, Ogihara Y (2016) A causal gene for seed dormancy on wheat chromosome 4A encodes a *MAP* kinase kinase. Curr Biol 26:782-787

Xi W, Liu C, Hou X, Yu H (2010) MOTHER OF FT AND TFL1 regulates seed germination through a negative feedback loop modulating ABA signaling in Arabidopsis. Plant Cell 22: 1733–1748

Table 4.1 Summary of germination rates of NW97S186, "Tutoumai A", AUS1408 and their selected backcross progenies in the 2015 fall & 2016 spring greenhouse experiments (GH_Fall and GH_Spring) and in the 2016 Manhattan & Hays field experiments

Population	Statistics/Parents	GH_Fall	GH_Spring	Field_MH	Field_Hays
Selected BC_2F_n* of N/T cross	**Mean**	0.44	0.69	0.34	0.51
	Variance	0.05	0.05	0.03	0.02
	Range	0.07-0.92	0.24-0.99	0.08-0.82	0.23-0.81
	NW97S186	0.82	0.98	0.67	0.91
	Tutoumai A	0.03	0.15	0.21	0.10
Selected BC_2F_n of N/A cross	**Mean**	0.43	0.58	0.35	0.50
	Variance	0.04	0.04	0.02	0.03
	Range	0.08-0.80	0.29-0.96	0.05-0.69	0.10-0.77
	NW97S186	0.74	0.86	0.67	0.66
	AUS1408	0.10	0.17	0.00	0.07

*Selected double homozygous BC_2F_2 were used to evaluate germination rate in the 2015 fall greenhouse experiment (GH_Fall), BC_2F_3 in the 2016 spring greenhouse experiment (GH_Spring), and BC_2F_4 in the 2016 Manhattan (MH) & Hays field experiments.

Table 4.2 Environmental statistics of greenhouse and field experiments conducted in Manhattan and Hays, KS

Experiment	Day Temp.*	Night Temp.	Day length (h)
GH (May-June)	25±5	20±2	12
GH (Dec.-Jan.)	22±3	17±2	12
	Max. Temp.†	**Min. Temp.**	**Precip. (cm)**
2016MH	35.0	5.6	15.8
2016Hays	35.0	3.2	7.9

* Greenhouse day/night temperature (°C) is expressed as mean ± standard deviation

† Field temperature range (°C) and precipitation are calculated from May 1st to June 15th in 2016. Max. Temp.=maximum temperature, Min. Temp.=minimum temperature, Precip.=precipitation. Data is from www.usclimatedata.com.

Table 4.3 Overall analysis of variance (ANOVA) of germination rates of the selected backcross progenies of NW97S186/"Tutoumai A" and NW97S186/AUS1408 in the 2015 fall & 2016 spring greenhouse experiments and in the 2016 Manhattan & Hays field experiments

Source*	DF	Type III SS	Mean Square	F Value	Pr > F
Env	3	0.972	0.972	52.72	<.0001†
GeneticBG	1	0.021	0.021	0.94	0.3324
Env*GeneticBG	3	0.017	0.017	0.94	0.4224
TaPHS1	1	1.839	1.839	84.01	<.0001*
Env* *TaPHS1*	3	0.238	0.238	6.40	0.0003*
GeneticBG* *TaPHS1*	1	0.006	0.006	0.27	0.6039
Env*GeneticBG* *TaPHS1*	3	0.252	0.252	4.99	0.0021*
TaMKK3-A	1	1.591	1.591	72.66	<.0001*
Env* *TaMKK3-A*	3	0.315	0.315	5.19	0.0016
GeneticBG* *TaMKK3-A*	1	0.072	0.072	3.31	0.0699
Env*GeneticBG* *TaMKK3-A*	3	0.000	0.000	0.02	0.9953
TaPHS1* TaMKK3-A	1	0.079	0.079	3.63	0.0578
Env* *TaPHS1* TaMKK3-A*	3	0.002	0.002	0.56	0.644
GeneticBG* *TaPHS1* *TaMKK3-A*	1	0.011	0.011	0.51	0.476
Error	312	6.83	0.02	-	-

*Env=environment, GeneticBG=genetic background

†Significant effects at the level of 0.05

Table 4.4 Combined genetic effects of TaPHS1 and TaMKK3-A genes from "Tutoumai A" and AUS1408 in both greenhouse and field experiments in Manhattan (MH) and Hays

Population	Genotype	GH_Spring	GH_Fall	Field_MH	Field_Hays
Selected BC_2F_n* of N/T cross	**AABB**	0.399[a†]	0.192[a]	0.253[a]	0.421[a]
	AAbb	0.642[b]	0.450[b]	0.363[a]	0.589[a]
	aaBB	0.722[b]	0.408[b]	0.306[a]	0.494[a]
	aabb	0.907[c]	0.684[c]	0.446[a]	0.511[a]
Selected BC_2F_n of N/A cross	**AABB**	0.435[a]	0.288[a]	0.243[a]	0.404[a]
	Aabb	0.571[ab]	0.591[b]	0.367[ab]	0.472[ab]
	aaBB	0.674[b]	0.460[b]	0.444[b]	0.623[b]
	aabb	0.861[b]	0.538[b]	0.454[b]	0.626[b]

*Selected double homozygous BC_2F_2 were used to evaluate germination rate in the 2015 fall greenhouse experiment (GH_Fall), BC_2F_3 in the 2016 spring greenhouse experiment (GH_Spring), and BC_2F_4 in the 2016 Manhattan (MH) & Hays field experiments

†Comparisons were made between genotypes within each genetic background and each experiment, and different letters indicate statistical difference at the significant level of 0.05

Figure 4.1 A workflow diagram of the backcrossing project to transfer QTLs on 3AS and 4AL from "Tutoumai A" and AUS1408 to NW97S186

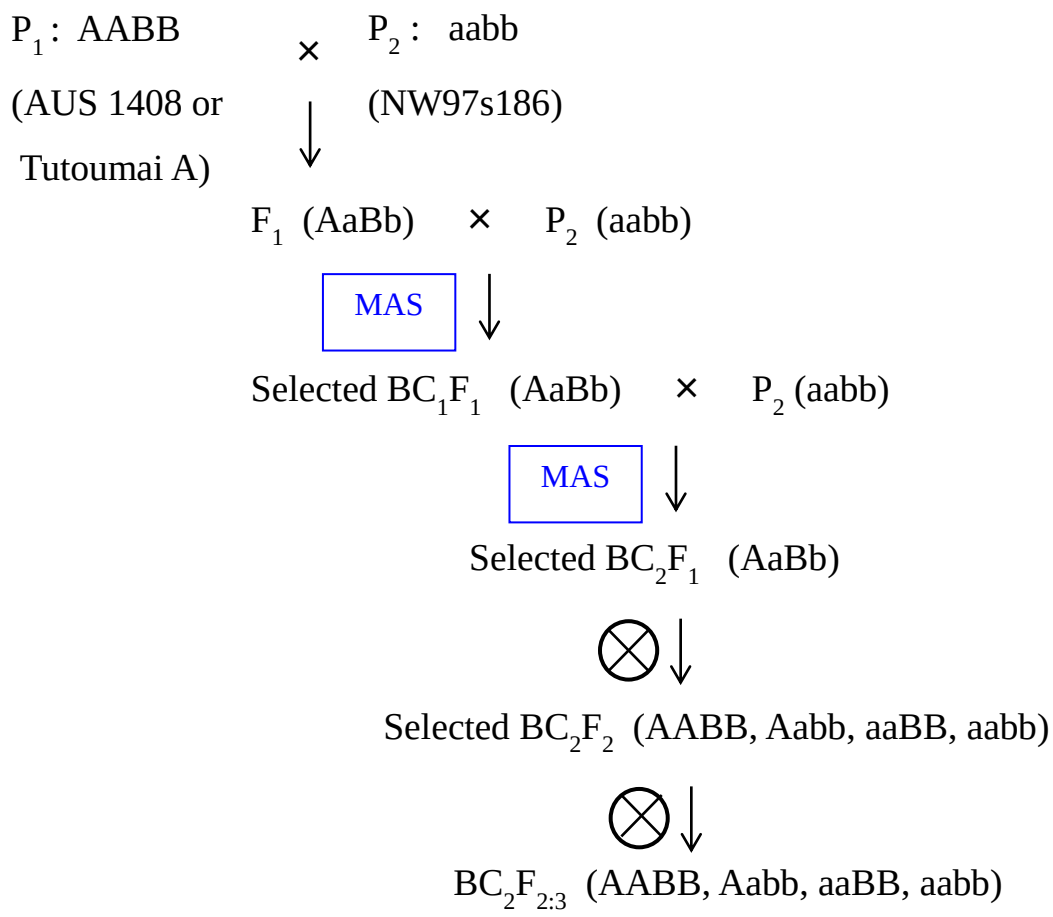


Figure 4.2 Effects of *TaPHS1* gene from AUS1408 and "Tutoumai A" backgrounds on germination rates evaluated in the 2015 fall & 2016 spring greenhouse experiments (GH_Fall and GH_Spring) and in the 2016 Manhattan & Hays field experiments

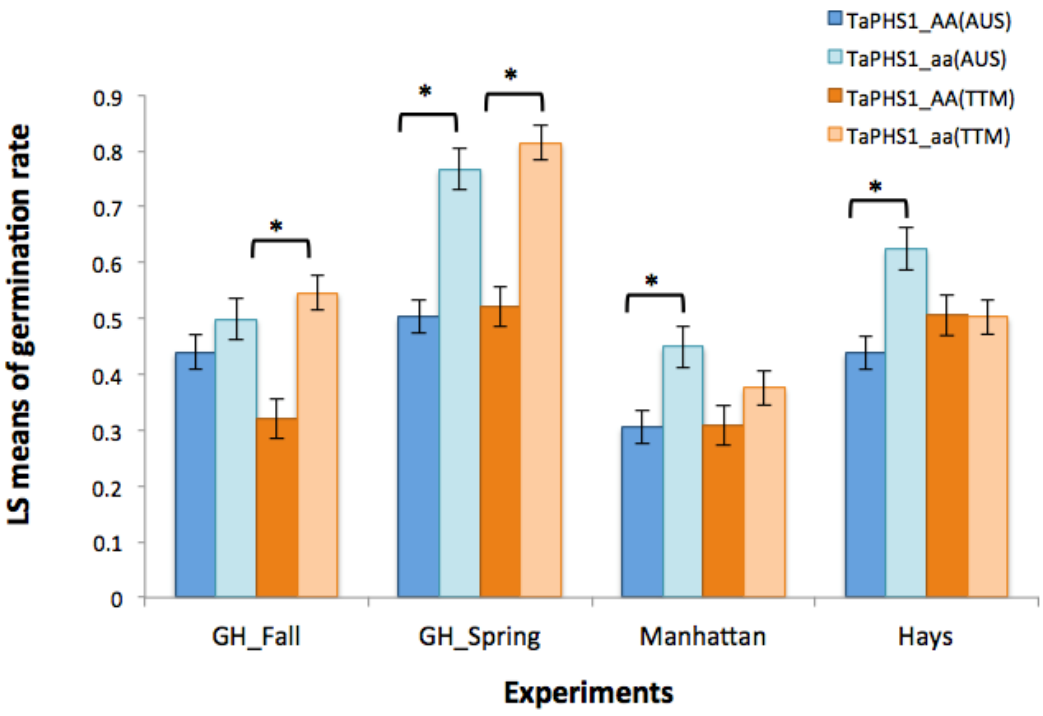
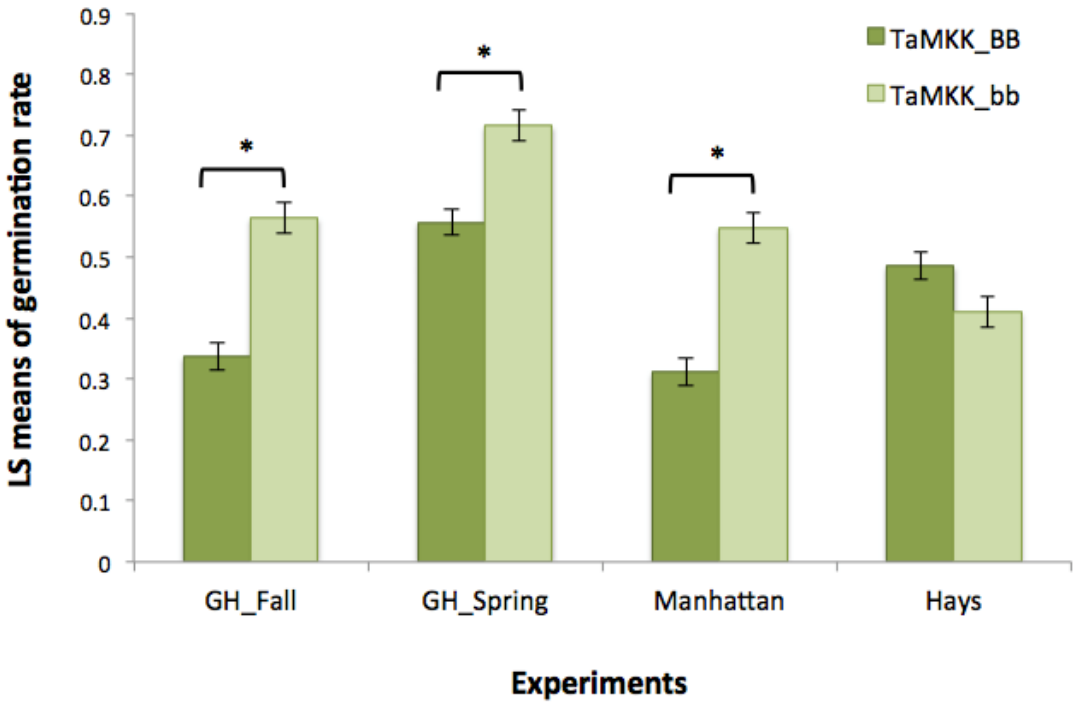


Figure 4.3 **Effects of *TaMKK3-A* gene on germination rates evaluated in the 2015 fall & 2016 spring greenhouse experiments (GH_Fall and GH_Spring) and in the 2016 Manhattan & Hays field experiments**



Chapter 5 - Genomic Prediction and Marker-Based Prediction on Wheat Pre-harvest Sprouting Resistance

Abstract

Wheat pre-harvest sprouting (PHS) can cause significant reduction in wheat end-use quality, and thus in grain sale price. Evaluation of a large number of wheat lines for PHS resistance in wheat breeding is a laborious and effort consuming tesk, and genetic markers can predict the PHS resistance by taking advantages of linkage disequilibrium. In this study, a panel of 185 U.S. cultivars and elite lines was used to compare prediction accuracy between genomic prediction and marker-based prediction (MBP). This panel was genotyped using the 9K iSelect SNP assays and evaluated for PHS resistance in Manhattan and Hays in both 2013 and 2014, respectively. Genome-wide association study (GWAS) identified 11 SNPs in three QTLs on chromosomes 3A, 4A and 2B with the best linear unbiased prediction (BLUP) of each accession. Four methods (ridge regression BLUP, BayesC0, BayesB and BayesC) were compared for genomic prediction accuracy, and rrBLUP provided better prediction accuracy than the three Bayesian methods on average. However, MBP using significant SNPs identified in the association study provided better prediction than the genomic prediction, therefore, can be more effective method to predict quantitative traits that are mainly controlled by a few major quantitative trait loci (QTLs).

Introduction

Molecular markers have been used in breeding for quantitative traits since the 1980s, and markers linked to quantitative trait loci (QTLs) can be used to predict performance of traits of interest in plant breeding (Bernardo 2008). Besides marker-assisted selection, markers linked to QTLs can also be applied to predict performance of the target traits by estimating their effects using a multiple-regression model (Edwards and Johnson 1994, Koebner 2003). However, marker-based prediction (MBP) can only be used when closely linked markers to these QTLs are available. Unfortunately, markers for most QTLs with minor effects remain to be identified, and thus prediction accuracy of MBP is low when a trait is controlled by many minor QTLs (Bernardo 2001).

Development of next generation sequencing (NGS) rapidly reduces the cost per sample for high-throughput SNP genotyping, genomic prediction can be implemented to accelerate breeding process when phenotyping is difficult or time consuming (Poland et al. 2012a). Genomic prediction, also referred as genomic selection, is to estimate all marker effects across the genome simultaneously in a training population without testing their significance, and to calculate the genomic estimated breeding values (GEBVs) in a testing population based on their molecular marker data (Meuwissen et al. 2001). The genomic prediction model, therefore, is able to include small effect QTL that are usually hard to be identified through QTL mapping (Meuwissen et al. 2001). Ridge regression and Bayesian approaches have been proposed to model the additive genetic effects and predict GEBVs by Meuwissen et al. (2001). Ridge regression assumes equal variance of all markers and penalizes the size of the regression coefficients, which results in an equal shrinkage of marker effect (Whittaker et al. 2000). BayesB

assumes non-zero loci-specific variances for each marker, BayesC assumes loci-specific variances that can be zero with probability π for each marker, and BayesC0 assumes equal genetic variance at each locus (https://github.com/reworkhow/JWAS.jl). Cross-validation has been widely applied to estimate genomic prediction accuracy. Among the frequently used cross-validation schemes, *k*-fold cross-validation provides "stable" estimates of model predictability, repeated random sub-sampling cross-validation can better assess the sample size effect, and leave-one-out cross-validation can be used to identify "outlier" observations in the training population (Yu et al. 2015). Recently, genomic prediction has been evaluated for various traits in several crops (Heffner et al. 2009 & 2011, Poland et al. 2012b, Würschum et al. 2013, Rutkoski et al. 2014, Arruda et al. 2015, Spindel et al. 2015), but not for wheat pre-harvest sprouting (PHS) resistance yet.

PHS resistance is a complex trait controlled by several major QTLs and many minor QTLs. PHS resistance QTLs have been reported on almost all wheat chromosomes, among which the QTLs on chromosome 3AS, 4AL and 2BL have been reported as main effect QTLs in many studies (Mori et al. 2005, Liu et al. 2008, Nakamura et al. 2011, Liu et al. 2013, Kato et al. 2001, Torada et al. 2005, Chen et al. 2008, Cabral et al. 2014, Torada et al. 2016, Kulwal et al. 2004 & 2012, Zhang et al. 2014). PHS can occur unexpectedly in most wheat planting areas in the U.S., which significantly reduces grain end-use quality and grain sales price. Therefore, breeding for PHS resistance is critical to reduce economic losses due to PHS. However, evaluation of a large number of wheat lines for PHS resistance is time and effort consuming, and genetic markers can predict phenotypic performance by taking advantages of linkage disequilibrium. In this study, both MBP and genomic prediction were applied in a wheat diversity panel to (1) compare genomic prediction accuracy among different prediction methodes

and (2) compare prediction accuracy between MBP and genomic prediction to identify an efficient selection method.

Materials and methods

Plant materials and pre-harvest sprouting evaluation

A diversity panel of 185 U.S. wheat elite lines and cultivars (See Chapter 3 for detail) was used in this study. PHS resistance was evaluated in four field experiments. The experiments were conducted in both Manhattan and Hays, KS in 2013 and 2014, and designated as Hays13, Manhattan13, Hays14, and Manhattan14, respectively. Experimental design for sprouting experiments was described in Chapter 3.

SNP genotyping

The wheat diversity panel was genotyped with the Wheat 9K SNP Arrays (Cavanagh et al. 2013) at USDA-ARS Cereal Crops Research Unit (Fargo, ND). SNPs with less than 5% minor allele frequency (MAF) or with more than 15% missing data were removed. A total of 5,921 from the 9K SNP array were used for genomic prediction.

Genome-wide association analysis

Best linear unbiased predictions (BLUPs) were calculated for each accession evaluated in the field experiments using the 'lme4' package in R 3.2.2 (Bates *et al*., 2014) with year and location as random effects in the model. Genome-wide association analysis (GWAS) was described in Chapter 3.

Genomic prediction and marker-based prediction

Ridge regression and three Bayes methods were applied in genomic prediction for PHS resistance. Ridge regression best linear unbiased predictor (rrBLUP) was implemented using r package rrBLUP (Endelman 2011). BayesB, BayesC and BayesC0 were applied using Julia

package JWAS (https://github.com/reworkhow/JWAS.jl). Markers that were identified to be significantly related to PHS resistance in GWAS were applied in marker-based prediction (MBP) using multiple-linear regression.

Leave-one-out cross-validation was performed to assess prediction accuracy, which was measured as the Pearson correlation between the observed germination rates and the predicted germination rates. Cross-validation within a single experiment was conducted using models from all the four genomic prediction methods to compare prediction accuracy among predictive methods. Cross-validation across experiments was conducted using models from rrBLUP and MBP.

Results

Phenotypic data

The distributions of germination rates in the four field experiments followed the similar trend, with 43% to 69% of the accessions showing a germination rate less than 20% (Fig. 5.1). The broad sense heritability was high (0.92) across environments, indicating high repeatability of the experiments. The means and standard deviations of germination rates in each experiment are shown in Table 5.1. The mean germination rates varied among the four experiments. Generally, 2014 experiments had higher mean germination rates than that of 2013's.

Prediction model accuracies

Four methods (rrBLUP, BayesB, BayesC and BayesC0) were applied to genomic prediction on PHS resistance, and leave-one-out cross-validations were conducted for each experiment using each method. The rrBLUP method provided the best predictions in most experiments, as well as using the BLUPs of germination rates (Table 5.2). BayesB and BayesC

showed very similar prediction accuracies, and performed better than BayesC0 in all the experiments (Table 5.2).

Prediction models constructed from each experiment using rrBLUP were validated in the remaining three experiments, and the prediction accuracies ranged from 0.200 to 0.439 (Table 5.3). The prediction model built from Manhattan13 experiment provided best predictions in all the experiments than the models from the rest of other experiments, whereas the model from Hays13 showed the lowest prediction accuracies in most cases (Table 5.3). In addition, all of Manhattan13, Manhattan14 and Hays14 experiments had poor predictions for the Hays13 experiment compared with their predictions on the rest experiments (Table 5.3).

Marker-based prediction accuracies

GWAS conducted with the BLUPs of germination rates from the four experiments identified 11 SNPs related to PHS resistance at a significant level of 10^{-4} (Table 5.4). These SNPs were in three QTL regions on chromosomes 3A, 4A and 2B, and they explained 31% of the totally phenotypic variance. Marker-based prediction (MBP) was conducted with these 11 SNPs using the least-squared regression, and models constructed from each experiment were used to predict germination rates in the rest of the experiments. The prediction accuracies ranged from 0.340 to 0.517 (Table 5.5). The model from Manhattan14 experiment made the best prediction on average, whereas the model from Manhattan13 experiment was the worst (Table 5.5). The mean prediction accuracy of MBP was 0.43, which was much higher than that of genomic prediction (0.27) (Fig. 5.2).

Discussion

PHS resistance is a complex trait that can be greatly influenced by environmental factors, such as temperature, moisture and photoperiod after flowering (Argel *et al*. 1983, Ceccato *et al*.

2011), and thus demonstrates significant genetic-by-environment interactions (Nakamura et al. 2011, Barrero et al. 2015, Lin et al. 2016). In the current study, different temperature ranges and precipitations in 2013 and 2014 (Table 5.1) could be the major sources of phenotypic differences between years, whereas variation in precipitation in Manhattan and Hays (Table 5.1) could be the main cause of phenotypic differences between locations. In the wheat diversity panel used, 157 are red winter wheat accessions, and most of them tended to have low germination rates in all the four experiments (Fig. 5.1), suggesting grain color showed significant effect on PHS resistance in the diversity panel.

Among the four methods of genomic prediction, rrBLUP showed the best prediction accuracy and was the most computationally efficient method (Table 5.2). BayesB and BayesC provided more accurate prediction on PHS resistance than BayesC0 (Table 5.2). This might be due to the fact that BayesB and BayesC allow unequal variance for each marker in the assumptions that is more reasonable than the assumption of BayesC0. In the Hays13 experiment, genetic variance likely explained only a small portion of the phenotypic variance, thus the model constructed from Hays13 experiment might not be able to provide accurate estimation of breeding values. Therefore, not only the model from the Hays13 experiment cannot accurately predict other experiments, but also it was poorly predicted by the models developed from the rest of other experiments (Table 5.3). For the traits that show large genetic-by-environment interactions, BLUPs can reduce environmental variances to increase genomic prediction accuracy.

The absolute value of regression coefficients ($|\beta|$) of all SNPs identified in GWAS were greater or equal to 0.002 in the genomic prediction model (Table 5.4), meaning that they were among the most important markers in the model given the fact that only 96 out of 5,921 SNPs

had $|\beta|$ greater than 0.002. Such result agrees with the statement that markers showing large effects in the genomic prediction model may be linked to major QTLs (Bernardo and Yu 2007). However, among the 9 SNPs that had $|\beta|$ greater than 0.003, four of them hardly showed any effect in GWAS (data not shown), indicating that although genomic prediction model is able to captures genetic variance, it is not effective for QTL identification.

In the current study, MBP using 11 SNPs that were significantly related to PHS resistance in GWAS provided more accurate predictions (0.340 to 0.517) than genomic prediction (0.200 to 0.439). The descrepancy in prediction accuracy could be due to the fact that PHS resistance is mainly controlled by several major QTLs, and GWAS was able to identify most of these QTLs; therefore, SNPs significantly related to the trait captured genetic variance very well in this study. However, genetic effects of trait-related SNPs might be underestimated using rrBLUP, and thus prediction accuracy in genomic prediction was reduced.
Genomic prediction is described as a black-box procedure (Haley et al. 2006), which does not require dissection of molecular mechanisms underlying the regulation of quantitative traits (Bernardo and Yu 2007). Prediction accuracy is determined by the size of training population, heritability, the number of QTL and the genetic architecture of the trait and the number of markers available (Daetwyler et al. 2008, Daetwyler et al. 2010, Combs and Bernardo 2013). In this study, although the heritability was high (0.92), genomic prediction accuracies were lower than expected (0.200 to 0.439), which could result from a small sample size (185) and limited number of markers (5,921) besides the influence from genetic architecture of PHS resistance. It has also been shown that population structure has great influence on prediction accuracy (Windhausen et al. 2012). Making predictions in the diversity panel with relatively loose genetic relationship could be another reason of low prediction accuracy. However, we can expect

improved accuracy when prediction is conducted in a breeding population derived from lines in the training population. Although QTLs have been identified on almost all the wheat chromosomes, several major PHS resistance QTLs take account of most of the genetic variance in germination rate. Therefore, it is very promising to use MBP in wheat breeding to improve PHS resistance.

Reference:

Aisawi, KAB, Reynolds MP, Singh RP, Foulkes MJ (2015) The physiological basis of the genetic progress in yield potential of CIMMYT spring wheat cultivars from 1966 to 2009. Crop Sci 55:1749

Argel PJ, Humphreys LR (1983) Environmental effects on seed development and hardseededness in Stylosanthes hamata cv. Verano. I. Temperature. Crop Pasture Sci 34:261-270

Barrero JM, Cavanagh C, Verbyla KL, Tibbits JF, Verbyla AP, Huang BE, Rosewarne MG, Stephen S, Wang P, Whan A, Rigault P, Hayden JM, Gubler F (2015) Transcriptomic analysis of wheat near-isogenic lines identifies PM19-A1 and A2 as candidates for a major dormancy QTL. Genome Biol 16:93

Bernardo R (2001) What if we knew all the genes for a quantitative trait in hybrid crops? Crop Sci 41:1–4

Bernardo R (2008) Molecular markers and selection for complex traits in plants: learning from the last 20 years. Crop Sci 48:1649–1664

Bernardo R, Yu J (2007) Prospects for genomewide selection for quantitative traits in maize. Crop Science 47:1082-1090

Cabral AL, Jordan MC, McCartney CA, You FM, Humphreys DG, MacLachlan R, Pozniak CJ (2014) Identification of candidate genes, regions and markers for pre-harvest sprouting resistance in wheat (*Triticum aestivum* L.). BMC Plant Biol 14:1

Cavanagh CR, Chao S, Wang S, Huang BE, Stephen S, Kiani S, Forrest K, Saintenac C, Brown-Guedira GL, Akhunova A et al. (2013) Genome-wide comparative diversity uncovers

multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. P Natl Acad Sci USA 110:8057-8062

Ceccato DV, Daniel Bertero H, Batlla D (2011) Environmental control of dormancy in quinoa (*Chenopodium quinoa*) seeds: two potential genetic resources for preharvest sprouting tolerance. Seed Sci Res 21:133–141

Chen CX, Cai SB, Bai GH (2008) A major QTL controlling seed dormancy and pre-harvest sprouting resistance on chromosome 4A in a Chinese wheat landrace. Mol Breed 21:351-358

Combs E, Bernardo R (2013) Accuracy of Genome-wide Selection for Different Traits with Constant Population Size, Heritability, and Number of Markers. The Plant Genome 6 (1)

Daetwyler HD, Villanueva B, Woolliams JA (2008) Accuracy of Predicting the Genetic Risk of Disease Using a Genome-Wide Approach. PLoS ONE 3:e3395

Daetwyler HD, Pong-Wong R, Villanueva B, Woolliams JA (2010) The Impact of Genetic Architecture on Genome-Wide Evaluation Methods. Genetics 185:1021 1031

Edwards M, Johnson L (1994) RFLPs for rapid recurrent selection. p. 33–40. In Analysis of molecular marker data. Joint Plant Breeding Symposia Series. ASA, Madison, WI

Endelman JB (2011) Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. Plant Genome J 4:250

Haley, CS, Koning DJ, Elsen JM, Veerkamp RF (2006) Research needs in numerical genomics and quantitative genetics of livestock. SABRE Workshop, 13 June 2006, Edinburgh. Available at http://www.sabre-eu.eu/Portals/0/Research%20Needs%20in%20Numerical%20Genomics%20(SABRE%20Workshop%20report).pdf (verifi ed 15 Feb. 2007). SABRE, European Union

Heffner EL, Jannink JL, Sorrells ME (2011) Genomic Selection Accuracy using Multifamily Prediction Models in a Wheat Breeding Program. Plant Gen 4:65-75

Heffner EL, Sorrells ME, Jannink JL (2009) Genomic selection for crop improvement. Crop Sci. 49:1

Kato K, Nakamura W, Tabiki T, Miura H (2001) Detection of loci controlling seed dormancy on group 4 chromosomes of wheat and comparative mapping with rice and barley genomes. Theor Appl Genet 102:980–985

Koebner R (2003) MAS in cereals: Green for maize, amber for rice, still red for wheat and barley. In Marker assisted selection: A fast track to increase genetic gain in plant and animal breeding? Turin, Italy. 17–18 Oct. 2003. Available at http://www.fao.org/biotech/docs/Koebner.pdf (verifi ed 15 Feb. 2007) FAO, Rome

Kulwal P, Ishikawa G, Benscher D, Feng ZY, Yu LX, Jadhav A, Mehetre S, Sorrells ME (2012) Association mapping for preharvest sprouting resistance in white winter wheat. Theor Appl Genet 125:793–805

Kulwal PL, Singh R, Balyan HS, Gupta PK (2004) Genetic basis of pre-harvest sprouting tolerance using single-locus and two-locus QTL analyses in bread wheat. Funct Integr Genomics 4:94–101

Lin M, Zhang D, Liu S, Zhang G, Yu J, Fritz AK, Bai G (2016) Genome-wide association analysis on pre-harvest sprouting resistance and grain color in US winter wheat. BMC genomics 17:794

Liu S, Cai S, Graybosch R, Chen C, Bai G (2008) Quantitative trait loci for resistance to pre-harvest sprouting in US hard white winter wheat Rio Blanco. Theor Appl Genet 117:691-699

Liu S, Sehgal SK, Li J, Lin M, Trick HN, Yu J, Gill BS, Bai G (2013) Cloning and characterization of a critical regulator for preharvest sprouting in wheat. Genetics 195:263-273

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819-1829

Mori M, Uchino N, Chono M, Kato K, Miura H (2005) Mapping QTLs for grain dormancy on wheat chromosome 3A and the group 4 chromosomes, and their combined effect. Theor Appl Genet 110:1315–1323

Nakamura S, Abe F, Kawahigashi H, Nakazono K, Tagiri A, Matsumoto T, Utsugi S, Ogawa T, Handa H, Ishida H, Mori M, Kawaura K, Ogihara Y, Miura H (2011) A wheat homolog of MOTHER OF FT AND TFL1 acts in the regulation of germination. The Plant Cell 23:3215–3229

Poland J, Endelman J, Dawson J, Rutkoski J, Wu S, Manes Y, et al. (2012b) Genomic selection in wheat breeding using genotypingby-sequencing. Plant Genome 5:103

Poland JA, Brown PJ, Sorrells ME, Jannink JL (2012a) Development of High-Density Genetic Maps for Barley and Wheat Using a Novel Two-Enzyme Genotyping-by-Sequencing Approach. PLoS ONE 7:e32253

Rutkoski JE, Poland JA, Singh RP, Huerta-Espino J, Bhavani S et al. (2014) Genomic Selection for Quantitative Adult Plant Stem Rust Resistance in Wheat. The Plant Genome 7(3)

Spindel J, Begum H, Akdemir D, Virk P, Collard B et al. (2015) Genomic Selection and Association Mapping in Rice (*Oryza sativa*): Effect of Trait Genetic Architecture, Training Population Composition, Marker Number and Statistical Model on Accuracy of Rice Genomic Selection in Elite, Tropical Rice Breeding Lines. PLoS Genet 11:e1004982

Torada A, Ikeguchi S, Koike M (2005) Mapping and validation of PCR-based markers associated with a major QTL for seed dormancy in wheat. Euphytica 143:251–255

Torada A, Koike M, Ogawa T, Takenouchi Y, Tadamura K, Wu J, Matsumoto T, Kawaura K, Ogihara Y (2016) A causal gene for seed dormancy on wheat chromosome 4A encodes a MAP kinase kinase. Curr Biol 26:782-787

Whittaker JC, Thompson R, Denham MC (2000) Marker-assisted selection using ridge regression. Genet Res 75:249–252

Windhausen VS, Atlin GN, Hickey JM, Crossa J, Jannink JL et al. (2012) Effectiveness of Genomic Prediction of Maize Hybrid Performance in Different Breeding Populations and Environments. G3: Genes|Genomes|Genetics 2:1427-1436

Würschum T, Reif J, Kraft T, Janssen G, Zhao Y (2013) Genomic selection in sugar beet breeding populations. BMC Genet 14:1-8

Yu X, Guo T, Li X, Yu J (2015) Empirical evaluation for models, cross-validation schemes, and population structure in genomic selection. Annul meeting of ASA, CSSA, SSSA, Minneapolis, MN

Zhang Y, Miao X, Xia X, He Z (2014) Cloning of seed dormancy genes (*TaSdr*) associated with tolerance to pre-harvest sprouting in common wheat and development of a functional marker. Theor Appl Genet 127:855-866

Table 5.1 Means and standard errors of germination rates and environmental statistics of the four field experiments conducted at Manhattan (MH) and Hays in 2013 and 2014, respectively

Statistics	MH_13	Hays_13	MH_14	Hays_14
Mean	0.181	0.224	0.359	0.279
SE	0.221	0.250	0.296	0.275
Max. Temp.[a]	36.1	33.9	41.1	40.0
Min. Temp.[a]	7.8	9.4	5.3	5.9
Precip. (cm)[a]	12.3	27.6	6.6	19.4

[a] Max./Min. Temp.=field maximum and minimum temperatures (°C) and precipitation in cm from May 1st to June 15th in 2013 (MH_13 and Hays_13) and 2014 (MH_14 and Hays_14) field experiments. Data is adapted from "www.usclimatedata.com"

Table 5.2 Genomic prediction accuracy estimated by leave-one-out cross-validation in each of the four field experiments conducted at Manhattan and Hays in 2013 and 2014, respectively, using ridge-regression best linear unbiased prediction (rrBLUP) and three Bayesian methods

Method	MH_13	Hays_13	MH_14	Hays_14	BLUP	Mean accuracy
rrBLUP	0.439	0.262	0.246	0.305	0.337	0.318
Bayes C0	0.298	0.139	0.178	0.242	0.235	0.218
Bayes B	0.341	0.155	0.376	0.299	0.28	0.29
Bayes C	0.334	0.149	0.365	0.293	0.286	0.285

Table 5.3 Genomic prediction accuracy estimated by leave-one-out cross-validation in different experiments conducted at Manhattan (MH) and Hays in 2013 and 2014, respectively, using ridge-regression best linear unbiased prediction (rrBLUP) method

	MH_13_Train	Hays_13_Train	MH_14_Train	Hays_14_Train	BLUP_Train
MH_13_Validation	0.439	0.27	0.347	0.358	0.386
Hays_13__Validation	0.264	0.262	0.203	0.200	0.251
MH_14__Validation	0.323	0.208	0.246	0.252	0.281
Hays_14__Validation	0.363	0.218	0.274	0.305	0.315
BLUP_Validation	-	-	-	-	0.337

Table 5.4 Significant SNPs identified in GWAS using best linear unbiased predictions (BLUPs) for each accession from the four field experiments conducted at Manhattan and Hays in 2013 and 2014, respectively, and their coefficients estimated in genomic prediction using ridge-regression best linear unbiased prediction (rrBLUP) method

SNP	Chr.[a]	Position (cM)	*p*-value	MAF[b]	R^{2}[c]	β[d]
wsnp_Ra_c2339_4506620	3A	6.19	2.72E-05	0.497	0.080	0.0031
wsnp_Ex_c10014_16477392	3A	4.08	7.57E-05	0.284	0.071	-0.0024
wsnp_Ex_c9485_15724984	3A	4.08	2.80E-04	0.197	0.059	0.0020
TaPHS1.2	3A	5.82	2.98E-04	0.243	0.059	-0.0030
TaPHS1.hap	3A	5.8	2.98E-04	0.243	0.059	0.0030
wsnp_Ex_rep_c67702_66370241	3A	4.08	3.29E-04	0.216	0.058	0.0024
wsnp_Ex_rep_c67635_66291944	3A	6.46	6.12E-04	0.451	0.052	-0.0026
wsnp_Ex_c10014_16476905	3A	4.08	8.48E-04	0.341	0.050	-0.0026
wsnp_Ex_c13031_20625900	4A	129.34	6.57E-04	0.430	0.052	0.0039
wsnp_Ex_rep_c66324_64493429	4A	129.34	7.82E-04	0.435	0.050	0.0038
wsnp_CAP8_c4576_2228073	2B	36.88	8.05E-04	0.495	0.050	-0.0022

[a] Chr.=chromosome

[b]Minor allele frequency

[c]Phenotypic variance explained by SNPs significantly related to pre-harvest sprouting resistance in the genome-wide association study

[d] Coefficients estimated for these SNPs in genomic prediction using ridge regression method

Table 5.5 Marker-based prediction accuracy estimated by leave-one-out cross-validation in different experiments conducted at Manhattan (MH) and Hays in 2013 and 2014, respectively, using ridge-regression best linear unbiased prediction (rrBLUP) method

	MH_13_Train	Hays_13_Train	MH_14_Train	Hays_14_Train	BLUP_Train
MH_13_Validation	0.421	0.372	0.378	0.386	0.401
Hays_13_Validation	0.340	0.374	0.369	0.366	0.377
MH_14_Validation	0.424	0.454	0.517	0.494	0.495
Hays_14_Validation	0.410	0.426	0.466	0.466	0.461
BLUP_Validation	-	-	-	-	0.483

Figure 5.1 Distributions of germination rates in (a) 2013 Manhattan, (b) 2013 Hays, (c) 2014 Manhattan, and (d) 2014 Hays field experiments

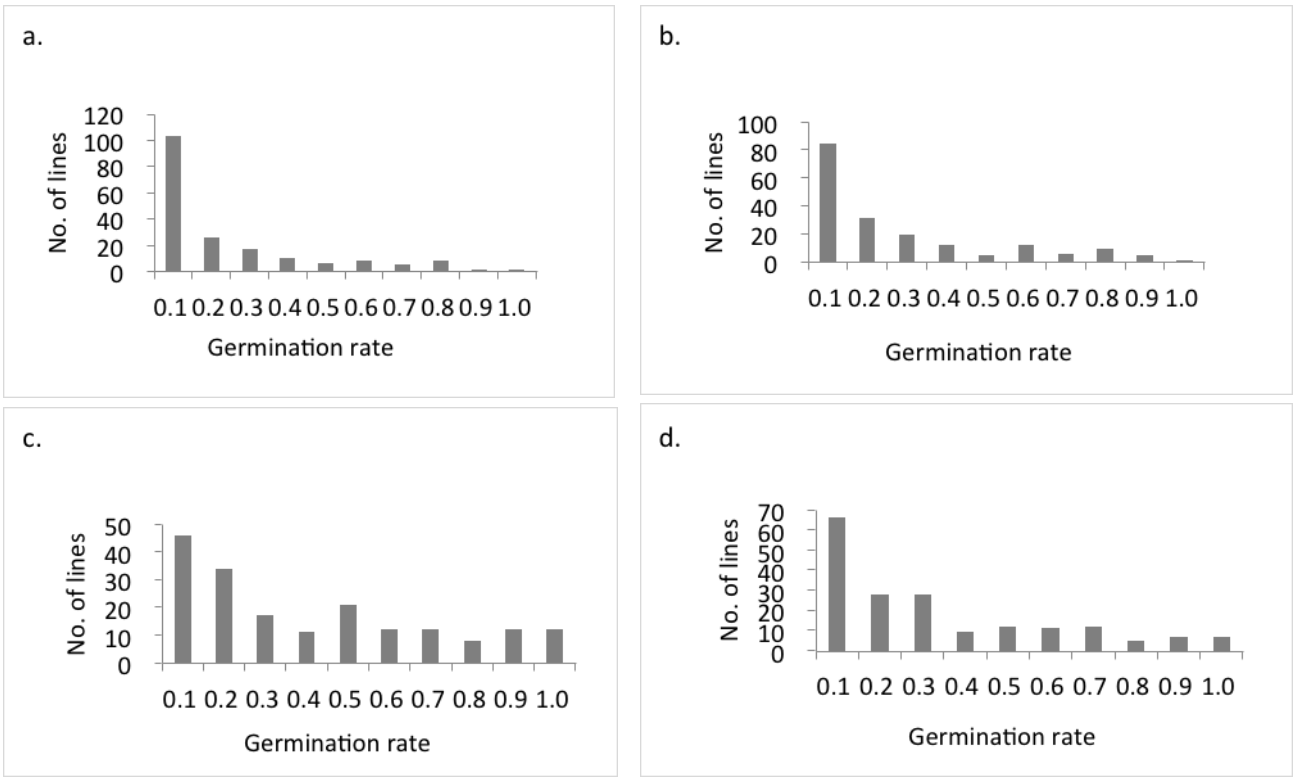
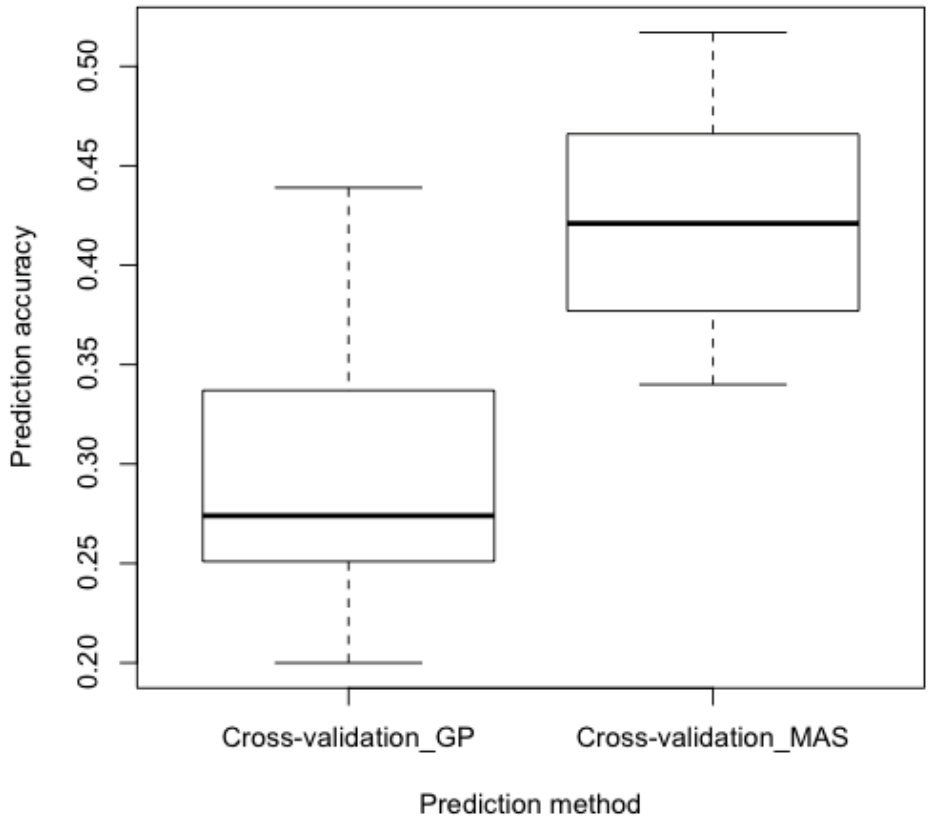


Figure 5.2 Box-plot to compare prediction accuracies estimated by genomic prediction (GP) and marker-based prediction (MBP)



Appendix A – A list for wheat grain color, *Tamyb10* alleles and germination rates of the association panel evaluated in the greenhouse and field experiments

No.	Accession	Grain Color		*Tamyb10* Genotype			Germination Rate							
		2011F_GH	2010Enid_Field	*Tamyb10-A1*	*Tamyb10-B1*	*Tamyb10-D1*	2011F	2012S	2012F	2013S	2013_MH	2013_Hays	2014_MH	2014_Hays
A001	Atlas66	4	4	R	R	r	0.006	0.016	0.000	0.154	0.124	0.109	0.524	0.202
A002	OK04505	4	2.5	R	r	R	0.014	0.098	0.627	0.075	0.022	0.148	0.260	0.085
A003	KS05HW136-3	1	1	r	r	R	0.070	0.159	0.559	0.462	0.173	0.486	0.928	0.967
A004	T158	3	2	r	R	R	0.028	0.013	0.013	0.004	0.342	0.064	0.525	0.559
A005	KS980554-12-~9	3	3	R	r	r	0.682	0.727	0.900	0.528	0.416	0.776	0.252	0.059
A006	KS980512-2-2	1	1	r	r	r	0.047	0.325	0.431	0.290	0.277	0.541	0.993	0.959
A007	TX04M410211	2	3	R	R	r	0.158	0.421	0.760	0.509	0.054	0.066	0.038	0.005
A008	N98L20040-44	3	3	R	R	R	0.004	0.526	0.195	0.849	0.718	0.562	0.153	0.075
A009	NI04420	3	3	R	r	R	0.179	0.393	0.904	0.362	0.894	0.884	0.150	0.116
A010	Duster	4	4	R	R	R	0.029	0.130	0.725	0.414	0.140	0.048	0.458	0.614
A011	OK02522W	1	1	r	r	r	0.610	0.196	0.968	0.493	0.076	0.247	0.509	0.648
A012	Scout 66	3	3	R	R	r	0.022	0.840	0.944	0.131	0.079	0.330	0.156	0.231
A013	AP04T8211	4	4	r	R	R	0.000	0.108	0.596	0.104	0.036	0.000	0.057	0.022
A014	HV9W96-1271R-1	3	4	r	R	R	0.402	0.209	0.986	0.373	0.381	0.147	0.759	0.738
A015	NE04424	2	3	R	r	R	0.076	0.650	0.920	0.756	0.056	0.108	0.094	0.036
A016	CO02W237	1	1	r	r	r	0.101	0.698	0.958	0.645	0.225	0.110	0.968	0.987
A017	OK03825-5403-6	3	3	R	R	R	0.015	0.327	0.122	0.795	0.080	0.132	0.436	0.545
A018	TX04V075080	2	2	R	r	r	0.021	0.095	0.513	0.378	0.072	0.017	0.000	0.017
A019	SD06165	2	2	R	r	r	0.009	0.026	0.027	0.268	0.007	0.012	0.061	0.350
A020	NX03Y2489	1	1	r	r	r	0.260	0.306	0.896	0.223	0.109	0.273	0.840	0.634
A021	NI04427	3	3	R	R	r	0.010	0.010	0.259	0.145	0.015	0.032	0.130	0.101
A022	Endurance	3	3	R	R	r	0.000	0.533	0.250	0.429	0.068	0.000	0.675	0.215
A023	TAM-107	3	2	r	R	R	0.127	0.034	0.348	0.081	0.768	0.782	0.599	0.278
A024	AP05T2413	3	2	r	r	R	0.002	0.227	0.021	0.643	0.251	0.599	0.623	0.257
A025	HV9W03-539R	3	2	R	R	R	0.017	0.962	0.044	0.522	0.734	0.588	0.009	0.018
A026	CO03064	3	3	r	R	R	0.007	0.194	0.014	0.449	0.047	0.134	0.301	0.116

A027	TX02A0252	3	3	R	R	R	0.015	0.479	0.000	0.167	0.000	0.007	0.650	0.479
A028	Kharkof	3	3	R	R	r	0.010	0.010	0.104	0.476	0.266	0.582	0.210	0.166
A029	SD06173	3	3	R	R	R	0.009	0.035	0.453	0.140	0.020	0.056	0.261	0.251
A030	NX04Y2107	3	3	r	R	r	0.003	0.027	0.012	0.129	0.083	0.058	0.566	0.368
A031	NE05548	3	2	R	R	R	0.255	0.073	0.493	0.432	0.038	0.055	0.114	0.102
A032	Deliver	2	1.5	R	r	R	0.092	0.460	0.832	0.666	0.097	0.247	0.419	0.207
A033	Trego	1	1	R	r	r	0.031	0.251	0.238	0.523	0.095	0.231	0.417	0.677
A034	HV9W03-696R-1	3	3	R	r	r	0.206	0.656	0.688	0.711	0.906	0.718	0.600	0.437
A035	NE05426	3	2	r	R	r	0.016	0.288	0.726	0.374	0.046	0.356	0.085	0.064
A036	CO03W054	1	1	r	r	r	0.000	0.005	0.000	0.006	0.020	0.046	0.405	0.404
A037	TX03A0148	3	3	R	r	R	0.108	0.431	0.981	0.175	0.307	0.566	0.601	0.424
A038	Antelope	1	1	r	r	r	0.120	0.210	0.112	0.371	0.135	0.140	0.920	0.828
A039	SD03164-1	2	2	R	r	R	0.023	0.268	0.960	0.827	0.287	0.094	0.461	0.165
A040	NW04Y2188	3	3	R	r	r	0.000	0.000	0.004	0.179	0.505	0.253	0.475	0.208
A041	NE05549	2	2	R	r	r	0.007	0.016	0.059	0.266	0.520	0.621	0.877	0.381
A042	OK Bullet	3	3	R	r	R	0.015	0.219	.	0.700	0.028	0.063	0.672	0.351
A043	OK03716W	3	3	R	r	R	0.020	0.122	0.900	0.513	0.215	0.043	0.720	0.879
A044	OK00514-05806	3	3	R	r	R	0.000	0.561	0.081	0.421	0.046	0.030	0.603	0.261
A045	AP06T3832	4	4	r	r	R	0.018	0.625	0.000	0.643	0.024	0.026	0.148	0.104
A046	HV9W02-942R	3	3	R	r	r	0.003	0.033	0.184	0.076	0.004	0.012	0.145	0.036
A047	NE05430	3	3	R	R	R	0.028	0.017	0.948	0.066	0.375	0.352	0.013	0.004
A048	CO03W139	1	1	r	r	R	0.066	0.014	0.943	0.349	0.577	0.787	0.905	0.872
A049	TX03A0563	3	3	r	r	R	0.014	0.038	0.754	0.525	0.064	0.040	0.038	0.000
A050	Wesley	3	3	-9	r	R	0.054	0.614	0.041	0.483	0.191	0.007	0.012	0.077
A051	NE02533	2	2	R	r	r	0.016	0.141	0.119	0.124	0.088	0.051	0.024	0.000
A052	NE05569	2	2	R	r	r	0.019	0.009	0.372	0.350	0.000	0.014	0.296	0.126
A053	Overley	2	2	R	r	r	0.008	0.109	0.764	0.412	0.058	0.140	0.128	0.004
A054	OK05903C	1	1	r	r	r	0.043	0.103	0.005	0.414	0.055	0.133	0.895	0.663
A055	Century	3	3	R	r	R	0.026	0.111	0.476	0.214	0.061	0.188	0.444	0.158
A056	KS05HW15-2	1	1	R	r	r	0.000	0.112	0.129	0.101	0.934	0.709	0.265	0.477
A057	T151	3	3	r	r	R	0.007	0.290	0.521	0.603	0.000	0.000	0.805	0.573
A058	KS970093-8-9-#1	2	2	R	r	r	0.004	0.000	0.004	0.017	0.000	0.016	0.062	0.282
A059	CO03W239	1	1	r	r	r	0.000	0.071	0.069	0.309	0.180	0.081	0.966	0.752
A060	TX04A001246	4	3	R	r	R	0.019	0.241	0.000	0.346	0.049	0.019	0.989	0.921

A061	Jerry	4	4	R	R	R	0.019	0.144	0.457	0.152	0.415	0.644	0.224	0.156
A062	NE02558	3	3	R	R	r	0.021	0.008	0.003	0.243	0.028	0.019	0.068	0.081
A063	MT0495	3	4	R	r	R	0.022	0.026	0.427	0.004	0.608	0.214	0.644	0.377
A064	Fuller	3	2	R	r	r	0.011	0.085	0.055	0.130	0.000	0.014	0.037	0.008
A065	OK03522	4	3	r	r	R	0.009	0.264	0.052	0.710	0.032	0.000	0.871	0.661
A066	KS05HW121-2	1	1	r	r	r	0.025	0.497	0.772	0.441	0.094	0.213	0.805	0.654
A067	T153	3	4	R	R	R	0.007	0.551	0.010	0.364	0.081	0.068	0.131	0.886
A068	KS970187-1-10	3	3	r	R	R	0.019	0.073	0.825	0.004	0.040	0.197	0.441	0.513
A069	CO03W043	1	1	r	r	r	0.016	0.239	0.778	0.427	0.045	0.209	0.879	0.641
A070	TX01V5134RC-3	3	4	R	R	r	0.000	0.030	0.020	0.044	0.007	0.031	0.433	0.173
A071	SD06W117	1	1	r	r	r	0.026	0.849	0.764	0.489	0.212	0.244	0.703	0.508
A072	SD05210	4	4	R	r	R	0.000	0.079	0.000	0.397	0.020	0.015	0.140	0.088
A073	NW03666	3	3	R	r	R	0.039	0.009	0.074	0.024	0.074	0.048	0.022	0.037
A074	MTS0531	1	1	r	r	r	0.026	0.256	0.065	0.323	0.199	0.034	1.000	0.611
A075	Centerfield	4	4	R	r	R	0.007	0.208	0.000	0.267	0.000	0.015	0.016	0.017
A076	OK04525	2	3	R	r	R	0.065	0.000	0.163	0.024	0.145	0.033	0.088	0.008
A077	OK03305	3	2	R	r	R	0.000	0.116	0.044	0.275	0.000	0.000	0.928	0.946
A078	T154	4	4	r	r	r	0.019	0.085	0.158	0.307	0.014	0.038	0.757	0.534
A079	NE05496	3	3	r	r	R	0.013	0.072	0.142	0.632	0.123	0.048	0.020	0.000
A080	TX04M410164	4	4	R	R	r	0.000	0.034	0.012	0.111	0.119	0.211	0.873	0.944
A081	SD06069	2.5	4	R	r	r	0.009	0.210	0.908	0.705	0.175	0.139	0.431	0.549
A082	SD05W030	1	1	r	r	r	0.010	0.581	0.866	0.359	0.110	0.187	0.481	0.570
A083	chisholm	4	3	R	r	R	0.209	0.207	0.629	0.356	0.452	0.463	0.072	0.116
A084	Guymon	1	1	r	r	r	0.013	0.373	0.312	0.234	0.079	0.384	0.910	0.940
A085	OK05830	4	4	R	r	R	0.016	0.489	0.370	0.518	0.097	0.078	0.185	0.135
A086	OK02405	4	4	R	R	R	0.000	0.015	0.068	0.000	0.005	0.013	0.087	0.000
A087	KS010957K~4	3	3	R	r	R	0.049	0.214	0.641	0.719	0.536	0.546	0.035	0.036
A088	NE06619	2	2	R	R	r	0.008	0.076	0.895	0.387	0.041	0.349	0.603	0.410
A089	MTS04120	4	4	r	r	R	0.007	0.283	0.880	0.155	0.008	0.083	0.104	0.070
A090	TX06A001239	4	4	R	r	r	0.039	0.016	0.316	0.420	0.129	0.009	0.225	0.103
A091	TXHT006F8-CS06/472-STA34	3	3	r	r	r	0.033	0.943	0.687	0.583	0.089	0.097	0.403	0.214
A092	MO011126	3	3	r	r	R	0.098	0.072	0.992	0.374	0.226	0.378	0.195	0.156
A093	OH02-7217	2	2	R	r	r	0.018	0.441	0.556	0.676	0.084	0.319	0.127	0.016

A094	MD99W483-06-9	4	3	R	r	R	0.080	0.407	0.858	0.717	0.600	0.336	0.049	0.000
A095	OK04507	3	3	R	r	R	0.000	0.299	0.691	0.700	0.038	0.105	0.021	0.013
A096	KS020304K~3	3	3	R	R	r	0.070	0.921	0.587	0.441	0.390	0.545	0.183	0.008
A097	KS010143K-11	2	2.5	R	r	R	0.328	0.190	1.000	0.585	0.699	0.732	0.307	0.020
A098	TX05A001334	4	4	R	r	R	0.360	0.274	0.959	0.333	0.760	0.964	0.057	0.018
A099	TX06A001376	4	4	r	R	R	0.170	0.018	0.627	0.240	0.127	0.007	0.032	0.018
A100	VA03W-412	2	3	R	r	R	0.033	0.567	0.995	0.839	0.202	0.305	0.092	0.014
A101	OH03-41-45	3	2	r	R	R	0.006	0.230	0.316	0.576	0.655	0.493	0.879	0.853
A102	OK05312	3	4	R	R	R	0.000	0.477	0.143	0.765	0.000	0.018	0.422	0.086
A103	HV9W05-881R	4	4	R	R	R	0.278	0.012	0.987	0.183	0.180	0.266	0.725	0.117
A104	NE06436	4	4	R	R	R	0.025	0.081	0.632	0.263	0.000	0.000	0.544	0.292
A105	NW05M6011-6-1	1	1	r	r	r	0.034	0.294	0.146	0.512	0.038	0.019	0.982	0.791
A106	TX06A001431	3	3	r	R	r	0.032	0.438	0.288	0.663	0.270	0.204	0.164	0.052
A107	TXHT023F7-CS06/607-STA07/40	4	4	R	R	r	0.042	0.207	0.004	0.650	0.111	0.273	0.063	0.059
A108	AR97044-10-2	3	3	R	r	r	0.012	0.149	0.024	0.386	0.160	0.000	0.044	0.017
A109	P02444A1-23-9	2	2	R	R	r	0.023	0.514	0.895	0.647	0.045	0.152	0.302	0.180
A110	VA05W-414	3	3	R	r	R	0.003	0.438	0.248	0.345	0.109	0.022	0.828	0.290
A111	OK05511	3	2	r	r	R	0.341	0.008	0.881	0.277	0.120	0.311	0.091	0.241
A112	SD07W041	2.5	1	r	r	r	0.191	0.694	0.892	0.423	0.070	0.015	0.293	0.655
A113	SD07204	3	3	R	R	R	0.092	0.613	0.202	0.560	0.475	0.708	0.150	0.055
A114	NW05M6015-25-4	1	1	r	r	r	0.163	0.335	0.386	0.235	0.082	0.137	0.200	0.243
A115	TXHT001F8-CS06/325-PRE07/75	2.5	2.5	r	R	r	0.298	0.790	0.992	0.777	0.148	0.703	0.426	0.391
A116	CO04W210	1	1	r	r	r	0.028	0.004	0.015	0.010	0.007	0.041	0.625	0.618
A117	KY96C-0769-7-3	3	2	r	r	R	0.085	0.528	0.213	0.343	0.411	0.073	0.245	0.404
A118	P03207A1-7	3	3	R	R	R	0.130	0.444	0.737	0.905	0.102	0.139	0.105	0.046
A119	KS07HW25	1	1	r	r	r	0.279	0.138	0.763	0.523	0.045	0.300	0.726	0.463
A120	SD07220	4	4	r	R	R	0.004	0.257	0.000	0.445	0.022	0.042	0.548	0.214
A121	KS010379M-2	3	3	R	R	r	0.070	0.026	0.139	0.177	0.008	0.000	0.083	0.031
A122	NE06472	3	3	R	R	R	0.003	0.000	0.000	0.036	0.042	0.000	0.013	0.014
A123	Roane	4	4	R	R	r	0.000	0.000	0.600	0.548	0.328	0.685	0.523	0.333
A124	OH02-12678	3	3	R	R	r	0.052	0.084	0.320	0.513	0.386	0.398	0.372	0.407
A125	LA02-923	2	2	r	r	r	0.042	0.104	0.003	0.215	0.016	0.088	0.745	0.535

A126	SD05W148-1	1	1	r	r	r	0.027	0.124	0.219	0.476	0.010	0.188	0.589	0.706
A127	KS010514-9TM-10	2	2	r	r	R	0.000	0.044	0.000	0.496	0.036	0.200	0.584	0.801
A128	N02Y5117	2	2	R	r	r	0.108	0.064	0.301	0.170	0.152	0.158	0.229	0.069
A129	INW0411	3	3	R	r	r	0.005	0.211	0.811	0.245	0.021	0.242	0.108	0.029
A130	MO040192	2	2	r	r	r	0.000	0.023	0.003	0.024	0.043	0.034	0.100	0.028
A131	KS07HW81	1	1	r	r	r	0.006	0.099	0.016	0.095	0.054	0.167	0.190	0.207
A132	U07-698-9	1	1	r	r	r	0.167	0.191	0.506	0.471	0.788	0.861	0.769	0.493
A133	TX05V5614	3	3	R	r	R	0.177	0.152	0.949	0.552	0.045	0.514	0.018	0.078
A134	Branson	2	2	R	r	R	0.000	0.413	0.004	0.603	0.013	0.048	0.193	0.020
A135	IL00-8530	3	3	R	r	R	0.064	0.005	0.953	0.667	0.733	0.856	0.115	0.189
A136	IL02-18228	4	4	R	r	R	0.006	0.529	0.517	0.755	0.695	0.673	0.012	0.032
A137	KS07HW117	1	1	r	r	r	0.020	0.208	0.006	0.341	0.029	0.007	0.851	0.669
A138	NE06549	3	3	R	R	r	0.040	0.207	0.532	0.351	0.000	0.000	0.021	0.000
A139	TX06A001084	3	3	r	r	R	0.000	0.125	0.000	0.360	0.021	0.240	0.058	0.087
A140	Bess	2	2	r	r	R	0.023	0.029	0.000	0.231	0.000	0.000	0.023	0.003
A141	IL02-19463	2	2	R	r	R	0.014	0.058	0.183	0.337	0.578	0.008	0.847	0.757
A142	Mocha exp.	3	3	r	R	R	0.000	0.101	0.101	0.409	0.052	0.106	0.083	0.008
A143	Pioneer Brand 26R61	2	3	R	r	R	0.067	0.270	0.863	0.310	0.060	0.315	0.294	0.478
A144	NC04-15533	3	3	r	r	r	0.029	0.203	0.825	0.689	0.067	0.294	0.148	0.017
A145	M03-3616-C	3	2	R	R	r	0.017	0.243	0.216	0.171	0.025	0.061	0.127	0.029
A146	W98007V1	4	4	R	R	R	0.000	0.066	0.004	0.439	0.199	0.483	0.216	0.260
A147	Arena exp.	2	2	R	r	r	0.062	0.341	0.783	0.398	0.090	0.278	0.405	0.237
A148	Coker 9553	3	3	R	r	r	0.000	0.000	0.000	0.015	0.041	0.017	0.020	0.053
A149	VA05W-258	2	2	R	r	r	0.077	0.351	0.580	0.511	0.170	0.102	0.341	0.303
A150	B030543	4	3	R	R	R	0.370	0.066	0.577	0.077	0.728	0.852	0.146	0.000
A151	W98008J1	3	3	R	r	R	0.012	0.193	0.013	0.596	0.268	0.511	0.637	0.251
A152	OK05122	3	3	r	r	R	0.039	0.074	0.384	0.618	0.016	0.000	0.379	0.432
A153	OK06210	4	4	R	R	R	0.000	0.372	0.575	0.821	0.019	0.007	0.374	0.249
A154	India exp.	3	3	R	r	R	0.025	0.025	0.321	0.525	0.403	0.147	0.088	0.064
A155	G69202	5	4	.	r	R	0.004	0.015	0.106	0.050	0.227	0.615	0.047	0.049
A156	USG 3555	4	4	.	u	r	0.037	0.295	0.590	0.132	0.006	0.000	0.384	0.237
A157	LA01138D-52	3	3	r	r	R	0.003	0.007	0.045	0.338	0.064	0.059	0.434	0.219
A158	VA05W-78	2.5	4	r	R	R	0.015	0.410	0.004	0.465	0.540	0.038	0.135	0.249
A1	OK05723	3	3	R	r	r	0.0	0.2	0.9	0.2	0.007	0.028	0.464	0.528

59	W						04	44	12	41				
A160	OK06319	3	3	R	r	R	0.003	0.561	0.000	0.579	0.010	0.000	0.697	0.272
A161	D04*5513	3	3	R	r	R	0.004	0.140	0.156	0.158	0.000	0.000	0.194	0.064
A162	M04-4566	3	3	r	r	R	0.016	0.008	0.150	0.102	0.000	0.000	0.209	0.125
A163	NC03-6228	3	3	r	r	R	0.000	0.000	0.000	0.069	0.023	0.000	0.095	0.035
A164	AR96077-7-2	4	4	r	r	R	0.074	0.597	0.842	0.678	0.336	0.890	0.188	0.252
A165	D04-5012	3	4	R	r	R	0.012	1.000	0.344	0.687	0.392	0.606	0.418	0.198
A166	G59160	3	4	R	r	r	0.023	0.191	0.692	0.442	0.071	0.193	0.183	0.093
A167	OK01420W	3	3	R	r	r	0.041	0.241	0.900	0.604	0.223	0.102	0.505	0.170
A168	OK06528	4	4	r	R	R	0.004	0.062	0.277	0.296	0.028	0.101	0.284	0.100
A169	OK06518	3	4	R	r	R	0.008	0.432	0.149	0.697	0.007	0.000	0.159	0.118
A170	KY97C-0321-02-01	3	3	R	R	R	0.227	0.463	0.923	0.457	0.083	0.155	0.615	0.158
A171	M04-4802	3	3	R	r	R	0.127	0.447	0.017	0.672	0.106	0.084	0.407	0.116
A172	AR97124-4-3	4	4	R	r	R	0.113	0.118	0.982	0.486	0.525	0.700	0.431	0.164
A173	GA991336-6E9	2	2	R	r	r	0.000	0.252	0.771	0.442	0.053	0.167	0.117	0.572
A174	G61505	3	3	r	r	R	0.118	0.503	0.842	0.371	0.025	0.073	0.355	0.259
A175	OK05134	4	3	R	r	R	0.015	0.173	0.090	0.713	0.025	0.027	0.305	0.184
A176	OK06313	3	3	R	r	R	0.002	0.019	0.032	0.032	0.761	0.241	0.029	0.004
A177	KY97C-0519-04-07	3	4	r	R	R	0.634	0.320	0.783	0.500	0.092	0.536	0.239	0.119
A178	M04*5109	3	3	r	r	R	0.008	0.351	0.931	0.615	0.017	0.079	0.050	0.062
A179	VA04W-259	3	3	R	r	R	0.004	0.007	0.003	0.203	0.000	0.000	0.338	0.209
A180	MD01W233-06-1	3	3	R	r	R	0.650	0.379	1.000	0.656	0.300	0.706	0.037	0.000
A181	GA991209-6E33	2	3	R	r	r	0.004	0.071	0.355	0.161	0.042	0.031	0.191	0.312
A182	G41732	2	2	R	r	r	0.006	0.120	0.121	0.192	0.012	0.034	0.179	0.112
A183	OK06848W	1	1	R	r	r	0.008	0.323	0.619	0.470	0.275	0.264	0.977	0.804
A184	W06-202B	4	4	r	r	R	0.040	0.044	0.628	0.440	0.112	0.555	0.085	0.018
A185	LA99005UC-31-3-C	3	3	r	r	R	0.781	0.916	0.981	0.654	0.691	0.709	0.217	0.019