ASSOCIATION MAPPING AND GENETIC DIVERSITY STUDIES OF AGRONOMIC AND QUALITY TRAITS IN DURUM WHEAT [*TRITICUM TURGIDUM* L. VAR. *DURUM* (DESF.)]

A Dissertation
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Marina Nikolayevna Johnson

In Partial Fulfillment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

Major Department:
Plant Sciences

November 2017

Fargo, North Dakota

North Dakota State University
Graduate School

Title

Association Mapping and Genetic Diversity Studies of Agronomic and Quality Traits in Durum Wheat [*Triticum turgidum* L. var. *durum* (Desf.)]

By

Marina Nikolayevna Johnson

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

DOCTOR OF PHILOSOPHY

SUPERVISORY COMMITTEE:

Dr. Elias M. Elias
Chair

Dr. Robert S. Brueggeman

Dr. Shiaoman Chao

Dr. Justin D. Faris

Dr. Frank A. Manthey

Dr. Phillip E. McClean

Approved:

November 17, 2017
Date

Richard D. Horsley
Department Chair

ABSTRACT

Genetic diversity studies in breeding programs are important to identify parental lines for hybridization and introgression of desirable alleles into elite germplasm. The genetic diversity analysis of 283 North Dakota State University (NDSU) advanced durum wheat breeding lines developed during the last 20 years indicated that the population was structured according to its breeding history. Total genetic diversity analysis ($H_T = 0.334$) showed adequate level of genetic variation. The results will help in breeding efforts to broaden the genetic base and select lines for crossing as well as for genetic and genomic studies to facilitate the combination of desirable alleles. The quantitative nature of important target traits, combined with environmental effects, makes it difficult to bring the desirable improvement in durum wheat to meet the expectations of all the stakeholders involved in the durum wheat industry. With an objective to identify molecular markers for marker-assisted breeding (MAB), the present study attempted to identify marker-trait associations for six agronomic and 29 quality traits using a genome-wide association study (GWAS) mapping approach. The study used two types of phenotypic datasets, a historic unbalanced dataset belonging to a total of 80 environments collected over a period of 16 years and a balanced dataset collected from two environments, to identify the applicability of historic unbalanced phenotypic data for GWAS analysis. A total of 292 QTL were identified for agronomic and quality traits, with 10 QTL showing major effects (R^2 >15%). Over 45% of QTL for agronomic and quality traits were present in both the unbalanced and balanced datasets, with about 50% of those present in both environments in the balanced dataset. Genome-wide association mapping studies identified several candidate markers for use in marker-assisted selection (MAS) for height, gluten strength, distribution of small kernels, polyphenol oxidase (PPO) activity, and yield.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Dr. Elias Elias, for his encouragement, guidance and support throughout my studies. His knowledge and wisdom were priceless. Thank you, Dr. Elias for all your hard work, which is very much appreciated! I am grateful for all my committee members, Dr. Brueggeman, Dr. Chao, Dr. Manthey, Dr. McClean, and Dr. Faris for their time, valuable council, and constructive feedback on my project. I am very grateful to Dr. Manthey and his team, especially Yu (Gaby) Liu and Hiroshi Ando, for helping me analyze "countless" number of grain and semolina samples. I am grateful for Dr. McClean and his team in analyzing my data. I appreciated the patient and clear way Dr. McClean explained GWAS beyond software usage. Many thank-yous go to Dr. Chao and her team for genotyping my samples! I am thankful for all the faculty of the Department of Plant Sciences, Plant Pathology, and USDA for being excellent examples of passionate and unbiased scientists that are always eager to learn. Special thank you goes to durum wheat breeding team and students in our project Stan Stansyk, Sally Mann, Megan Shwago, Wesam AbuHammed, Atena Oladzad Abbasabadi, Evan Salsman, Patricia Cabas Luhmann, Ali Hosseinirad for making harvest a fun and memorable time. Thank you for your friendship beyond work place! Very special thanks go to my parents, Nikolai and Nina, for putting up with their daughter's many years of schooling. I am so thankful for always being there for me while being thousand miles away. I am so thankful that Arlen and Ronna Gilbertson have taken the role of a second set of parents. Grateful for their love and care and a great support throughout my studies. Thank you for my wonderful friends, fellow students and mentors, Dr. and Mrs. Deckard, Dr. Ajay Kumar, Rahil Ashtari, and Naa Korkoi Ardayfio for assistance, support, snacks and good times. I am so grateful for the love and support of my best friend and husband, Ben Johnson. Thank you so much for being who you are and being there for me in every step of the way!

DEDICATION

To a very special person in my life, Arlen Gilbertson, for setting up an excellent example of integrity and always believing in me. I know you are very proud of my accomplishments looking down from Heaven!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFLP.....Amplified Fragment Length Polymorphism

AM.....Association Mapping

CLOSS.....Cooking Loss

CV.....Coefficient of Variance

CWT.....Pasta Cooked Weight

FIRM.....Pasta Firmness

FN.....Falling Number

GI.....Gluten Index

HT.....Plant Height

LK.....Percent Large Kernels

MAS.....Marker Assisted Selection

MIXO.....Mixograph

MK.....Percent Medium Kernels

PCA.....Principle Component Analysis

PCR.....Polymerase Chain Reaction

QTL.....Quantitative Trait Loci

RFLP.....Restriction Fragment Length Polymorphism

SASH.....Semolina Ash

SDS.....Sodium dodecyl sulphate–micro sedimentation

SEXT.....Semolina Extraction

SK.....Percent Small Kernels

SNP.....Single Nucleotide Polymorphism

SPROT .....Semolina Protein

SSR .....Simple Sequence Repeats

TEXT .....Total Extraction

TKW .....Thousand Kernel Weight

TWT.....Test Weight

TYP.....Total Yellow Pigment

WG.....Wet Gluten

WPROT.....Whole-Wheat Protein

WTS .....Work to Shear

YLD .....Yield

LIST OF APPENDIX TABLES

LIST OF APPENDIX FIGURES

GENERAL INTRODUCTION

Wheat (*Triticum* spp.) is one of the most widely cultivated crop of all cereals. It can be categorized by kernel texture into hard and soft and further subcategorized according to growing season and kernel color (Hoseney and Declour, 2010). Tetraploid durum possesses the hardest kernel of all and is more drought tolerant than hexaploid bread wheat (*T. aestivum*). It is cultivated in semi-arid regions of the world, such as Mediterranean Europe, the Middle East, North Africa, and the North American Great Plains (Elias and Manthey, 2005). In 2016, U.S. durum wheat acreage was an estimated 2.37 million (1.92 million in 2017), with approximately 62% of the total acreage in North Dakota (58% in 2017) (National Agricultural Statistics Service, 2017).

The agronomic performance of durum wheat has improved significantly worldwide as a result of public and private sector breeding efforts. Most of the cultivars grown in the Northern Plains were developed by the durum wheat breeding program established at North Dakota State University (NDSU) in 1929 (Joppa and Williams, 1988). The major focus of the durum wheat breeding program at NDSU is to develop durum wheat cultivars possessing characteristics that maximize economic return to producers as well as provide excellent quality durum wheat for the domestic pasta industry and the international export market.

A better understanding of the genetic composition of durum wheat's agronomic and quality traits is an essential precondition for more effective and targeted breeding activities (Araus et al., 2002; Salekdeh et al., 2009). Association mapping (AM) is one of the approaches in genomics that enables the identification and selection of chromosome regions harboring Quantitative Trait Loci (QTL) that control agronomic and quality traits (Tuberosa and Salvi, 2006; Collins et al., 2008). The basic objective of AM studies is to detect statistical associations

between genotypes and phenotypes in samples of unrelated individuals on the basis of linkage disequilibrium (LD), which is the nonrandom association of alleles at different loci (Zondervan and Cardon, 2004). Association mapping analysis offers greater precision in QTL location than linkage analysis performed with the population derived from a bi-parental cross due to its increased mapping resolution, reduced research time, and greater allele number (Yu and Buckler, 2006; Rafalski, 2010). The results of AM experiments are valuable for marker-assisted breeding programs as they allow for the identification of the desirable allelic variants at the major loci controlling the target traits (Heffner et al., 2009). Also, AM is a powerful approach for assessing genetic diversity in the breeding program. This information about the genetic diversity and population structure in elite breeding material is necessary for crop improvement (Inghelandt et al., 2010). Association mapping enables breeders to assess the allelic combinations selected over generations and thus provides genetic haplotypes for future crop improvement through marker-assisted selection (Kishore et al., 2012).

References

Araus, J.L., G.A. Slafer, M.P. Reynolds, C. Royo. 2002. Plant breeding and drought in C3 cereals: what should we breed for? Annals Bot. 89:925-940.

Collins, N.C., F. Tardieu, R. Tuberosa. 2008. Quantitative trait loci and crop performance under abiotic stress: where do we stand? Plant Physiol.147:469-486.

Elias, E.M., and F.A. Manthey. 2005. End products: present and future uses. Pages 63-86 in: Durum Wheat Breeding Current Approaches and Future Strategies. C. Royo, M.M. Nachit, N. DiFonzo, J.L. Araus, W.H. Pfeiffer, and G.A. Slafer, eds. Food Products Press: New York, NY.

Heffner, E.L., E. Mark, M.E. Sorrells, J.L. Jannink. 2009. Genomic selection for crop improvement. Crop Sci. 49:1-12.

Hoseney, C.R., and J.A. Delcour. 2010. Principles of cereal science and technology. St. Paul, MN: AACC International, 2010.

Inghelandt, D.V., A.E. Melchinger, C.L. Lebreton, B. Stich. 2010. Population structure and genetic diversity in a commercial maize breeding program assessed with SSR and SNP markers. Theor. Appl. Genet. 120:1289-1299.

Joppa, L.R., and N.D. Williams. 1988. Genetics and breeding of durum wheat in the United States. Pages 47-48 in: Durum Chemistry and Technology. F. Giuseppe, and L. Claudia. American Association of Cereal Chemist: St. Paul, MN.

Kishore, G., S. Gupta, A. Pandey. 2012. Assessment of population genetic diversity of Fagopyrum tataricum using SSR molecular marker. Biochem. Systematics Ecol. 43: 32-41.

NASS-National Agricultural Statistics Service. 2016. National statistics: wheat, durum, North Dakota. United States Department of Agriculture. http://usda.mannlib.cornell.edu/usda/current/CropValuSu/CropValuSu-02-24-2016.pdf. Accessed August, 2016.

Rafalski A.J. 2010. Association genetics in plant improvement. Curr. Opin. Plant Biol. 13:174-180.

Salekdeh, G.H., M. Reynolds, J. Bennett, J. Boyer. 2009. Conceptual framework for drought phenotyping during molecular breeding. Trends Plant Sci.14:488-496.

Tuberosa, R., and S. Salvi. 2006. Genomics-based approaches to improve drought tolerance of crops. Trends Plant Sci. 11:405-412.

Yu, J., and E.S. Buckler. 2006. Genetic association mapping and genome organization of maize. Curr. Opin. Biotechnol. 17:155-160.

Zondervan, K.T., and L.R. Cardon. 2004. The complex interplay among factors that influence allelic association. Nat. Rev. Genet. 5:89-100.

CHAPTER 1. LITERATURE REVIEW

Evolution of Durum Wheat

Durum wheat [*Triticum turgidum* L. var. *durum* (Desf.)] belongs to the grass family Poaceae, which evolved 50-70 million years ago (Mya); the sub-family Pooideae, which diverged from Poaceae around 20 Mya; and the tribe Triticeae, from which wheat takes its name (Huang et al., 2002; Levy and Feldman, 2002). Divergence from the tribe resulted in three other important cereal crops: barley (*Hordeum vulgare* L.), rye (*Secale cereal*), and triticale (× *Triticosecale*). The tribe Triticeae is presumably relatively young, with critical differentiation starting during the Pleistocene Epoch (Sakamoto, 1973). Comparative mapping has shown conserved syntenic relationships among genomes of different species in grass, including common wheat (*Triticum aestivum*), maize (*Zea mays* ssp. *mays* L), and sorghum [Sorghum bicolor L.) Moench], and rice (*Oryza Sativa* L), which suggests many of them originated from the same ancestor (Moore et al., 1995). Knowing the evolutionary relationship between these species helps in understanding the evolution of the grass genome and aids in the assembly of genome sequences from other pooid grasses (Luo et al., 2007).

The species included in the genus *Triticum* are *Triticum monococum* L. (einkorn wheat: genome A^mA^m), *Triticum urartu* (genome AA), *Triticum turgidum* L. (genome AABB), *Triticum aestivum* L. (genome AABBDD), *Triticum timopheevii* (genome AAGG), and *Triticum zhukovskyi* (genome AAAAGG) (http://www.ncbi.nlm.nih.gov). *Triticum urartu* exists only in the wild form, whereas *T. aestivum* and *T. zhukovskyi* exist only in cultivated forms. The other species, *T. monococcum*, *T. turgidum*, and *T. timopheevii*, have both a wild and a domesticated form.

About 300,000-500,000 years ago, the wild diploid wheat *T. urartu* hybridized with the B genome ancestor that is closely related to the goat grass *Aegilops speltoides*. This hybridization was followed by chromosome doubling to produce an allotetraploid wild emmer wheat, *T. turgidum ssp. dicoccoides* (2n = 4x = 28, AABB) (Huang et al., 2002; Levy and Feldman, 2002). About 12,000 years ago, hunter-gatherers began to cultivate wild emmer, subconsciously selecting for desired traits and gradually creating a cultivated emmer (*T. dicoccum*, 2n=4x=28, genome AABB). About 8,500 years ago, natural mutation changed the ears of emmer to a more easily threshed type that later evolved into the free-threshing ears of durum wheat (Dvorak et al., 2006).

Domestication and Dissemination of Durum Wheat

The domestication of plants and animals is the major factor that led to human civilization. Domestication was an outcome of the selection process, the purpose of which was to accelerate traits usable and desirable by humans (Gepts, 2004). Durum wheat originated in the Mediterranean Sea region and is the most commonly cultivated form of allotetraploid wheat. Wheat cultivation began about 12,000 years ago along the Fertile Crescent in what is known as the Levantine corridor (Israel, Jordan, Syria, Lebanon, and southern Turkey) (Zohary and Hopf, 2000). Genetic analysis of wheat domestication has been achieved using the mapping population developed from *T. dicoccoides* and the cultivated tetraploid wheat 'Langdon' (Peleg et al., 2011; Peng et al., 2003). Genetic improvements, such as thicker stems, tougher rachis, looser glumes, increased number of grains per spikelet, rapid and uniform germination, and particularly larger seeds are probably the most important signs of domestication, but are not necessarily reliable indicators (Ayal et al., 2005). *Triticum dicoccoides* possessed tough glumes and fragile rachis that made it susceptible to shattering, a mechanism of natural seed dispersal. Upon

domestication, each of the major cereals underwent genetic modifications that led to shatter-resistance phenotypes, allowing early farmers to harvest grain more efficiently (Faris and Gill, 2002; Ayal et al., 2005). Yield, grain size, plant height, and heading date are among other traits that were modified during the domestication and breeding process (Peng et al., 2003). Domestication and breeding for certain traits over time are likely to have narrowed the genetic diversity of durum wheat (Thuillet et al., 2005).

It is essential to trace the domestication process to find functional and regulatory genes that were eliminated from cultivars during domestication and breeding. Although *T. dicoccoides* possesses agriculturally deleterious features, it also possesses important beneficial traits, such as resistance to biotic and abiotic stresses, a high protein content, and alpha amylase inhibitors (Cakmak et al., 2004; Uauy et al., 2006). Other traits possessed are a high photosynthetic yield, salt and drought tolerance, herbicide resistance, earliness, high nitrogen content, high yield, short stature, and a high tillering capacity (Cakmak et al., 2004; Uauy et al., 2006). Among the 75 domestication QTL effects for 11 traits, wild QTL alleles of *T. dicoccoides* for 18 effects (24%) were agriculturally beneficial (Peng et al., 2003).

After domestication occurred, cultivated tetraploid emmer wheat spread east and south through the Mesopotamian plain to India and west through Anatolia to the Mediterranean coastal region (8,000 years ago), to the Balkans and Danube regions (7,000 years ago), and to Europe (7,000 years ago). Cultivated tetraploid wheat reached the United Kingdom and Scandinavia about 5000 years ago. It was introduced to Central Asia and China about 3,000 years ago from what is now Iran, and, later, to Africa from Egypt (Matsuoka, 2011; Ozkan et al., 2011).

In 1521, durum wheat seeds were introduced to Mexico by the Spaniards. At the beginning of the 1600s, farmers in California planted seeds of durum wheat transported by

explorers, traders, settlers, soldiers, and Spanish missionaries. On the East Coast, immigrants from Europe carried these seeds with them, while settlers from Southern Russia introduced durum wheat to the Great Plains in what is now Kansas in 1850s. (Damania, 2013).

U.S. Commercialization of Durum Wheat

Although durum wheat was introduced to U.S. growers at various times beginning around 1850, it failed to become a commercial crop because it was maladapted to the humid conditions of the Great Plains and possessed different milling properties (Ball, 1930); millers were not equipped to process the hard durum wheat grain, so refused to mill it (Paulsen and Shroyer, 2008).

Durum wheat was reintroduced to growers after 1900, when various cultivars of hard red spring wheat and durum wheat were collected from around the world and evaluated by the U.S. Department of Agriculture. The superior resistance of durum wheat to stem rust during a 1904 epidemic resulted in a rapid increase in durum production (Olmsted and Rhode, 2011). This increase in production coincided with the failure of the durum wheat crop in Europe in 1911, which allowed durum wheat growers in the Northern Plains to sell their durum at a higher price than their hard red spring wheat, making durum wheat production an appealing option (Isern, 2000).

Durum wheat is milled primarily to produce semolina. Pasta is made from semolina and water that is formed into dough and extruded under vacuum through a die. Although durum is also used to make bread in some parts of the world, this constitutes a relatively small proportion of worldwide durum usage (Troccolli et al., 2000)

A durum wheat breeding program was established at North Dakota State University (NDSU) in 1929, and it remains the only public durum wheat breeding program in the United

States. The objectives of the durum wheat breeding program include the improvement of agronomic performance, disease resistance, grain quality, semolina quality, and pasta cooking quality. Variation in the genetic traits of durum wheat must be fully surveyed to meet the broad requirements for durum grain. Domestication and selection improved the genetic uniformity of the crop, but decreased its genetic diversity. Most cultivars developed in the Northern Plains and Canada possess over 95% of North American germplasm in their background, although to increase the genetic diversity in the program at NDSU, germplasm has been introduced from Europe, the International Maize and Wheat Improvement Center (CIMMYT), and the International Center for Agricultural Research in the Dry Area (ICARDA) (Royo et al., 2009).

Diversity in Durum Wheat

Genetic diversity is the basis for genetic improvement and the foundation for survival, adaptation and evolution in time and space (Nevo and Beiles, 1989). In recent times, molecular markers have been widely used in wheat breeding programs to assess the genetic diversity available to the breeder. Molecular marker profiles account for the effects due to selection and genetic drift that occur over breeding cycles, making it possible to accurately portray the relationship among genotypes. As a result of modern breeding, it has been suggested that genetic diversity in wheat has been increasingly narrowed (Autrique et al., 1996), making it more difficult to breed for adaptation to biotic stresses, such as fungal pathogens, and abiotic stresses, such as drought or salt tolerance. Autrique et al. (1996) assessed 113 improved cultivars and landraces with restriction fragment length polymorphism (RFLP) markers. They observed a low genetic distance (many similar alleles) for improved cultivars and some landraces from Morocco and Jordan, while the other landraces from Greece, Italy, Turkey, Ethiopia, Syria, Algeria, and Cyprus showed a larger genetic distance. Other investigations of the genetic diversity of large

durum wheat germplasm collections, including both landraces and improved cultivars from different growing areas, highlighted the presence of broad genetic variation for morphological, agronomic, and quality traits (Martosa et al., 2005; Soleimani et al., 2002; Uddin and Boerner, 2008; Yildrim et al., 2011). A substantial level of genetic variation within modern cultivars of Canadian durum wheat detected by Amplified Fragment Length Polymorphism (AFLP) exists despite rigorous selection pressure in breeding programs. (Soleimani et al., 2002). Martosa et al. (2005) carried out a study to determine phylogenetic relationships and genetic diversity in durum wheat cultivars released in Italy and Spain throughout the 20th century using AFLP markers. Results indicated that the extent of genetic diversity in Italian and Spanish durum wheat remained constant over 100 years. High genetic variability of the Turkish landraces of durum wheat was reported by Yildirim et al. (2011) using microsatellite markers. Based on simple sequence repeats (SSR) markers, Abouzied et al. (2013) reported moderate genetic variability among and within populations of durum wheat cultivars and landraces.

Molecular Characterization of Durum Wheat

Tetraploid *T. turgidum* L. (2n=4x=28, genome AABB) represents a classical example of successful evolution through allopolyploidy. The genome size of tetraploid wheat is estimated to be 13,000 Mb (Arumuganathan and Earle, 1991). Durum wheat is a relatively young polyploid with extensive collinearity between the two genomes (Nachit et al., 2001).

The first genetic map for durum wheat was constructed in 1998 from a population of 65 recombinant inbred lines (RILs) using 198 RFLP, one polymerase chain reaction (PCR) marker, seven biochemical markers, and seven morphological markers. The total distance covered was 1,352 cM, with an average of 6.3 cM between two loci (Blanco et al., 1998). Since that time, several types of molecular markers have been used to construct durum wheat linkage maps,

including SSR (Roder et al., 1998; Korzun et al., 1999), AFLP (Lotti et al., 2000), and DarT (Manovani et al., 2008).

The abundance and genotyping cost effectiveness of single nucleotide polymorphism (SNP) markers led to the large-scale development and utilization of these markers in many crops, including wheat. Single nucleotide polymorphism markers allowed for the discovery of more than 2,500 SNP markers, saturating the durum wheat map with a marker density of 0.8 cM/marker (van Poecke et al., 2013). Wang et al. (2014) developed a high-density (90K) wheat SNP array that included 8,000 SNP markers from durum cultivars.

A high-density tetraploid wheat consensus map was created by merging genetic maps from 13 independent tetraploid wheat mapping populations. Ten of them were genotyped with the recently developed Illumina 9K and 90K wheat SNP arrays (Cavanagh et al., 2013; Wang et al., 2014). The consensus map ended up harboring 26,626 SNP and 791 SSR markers and spanning 2,631 cM of all 14 durum wheat chromosomes, with a marker density of 0.087 cM/marker, corresponding to 11 markers per 1 cM. Such high-resolution molecular maps can be applied in marker assisted selections (MAS) for desirable traits, positional cloning, and phylogenic studies.

Genetics of Durum Wheat Agronomic Traits

Yield, 1000-kernel weight, test weight, plant height, lodging, days to heading, disease resistance, and falling number are among agronomic traits that durum breeders evaluate while developing cultivars with improved performance. Most agronomic traits are quantitative in nature, controlled by many genes, and highly influenced by environment, genotype × environment interaction, and epistasis interaction (McCartney et al., 2005, Li et al., 2016). Previous genetic research has uncovered a number of QTL affecting plant height

(Worland et al., 1998; Li et al., 2016), heading date (Marza, et al., 2006; Reif et al., 2011; Li et al., 2016) and lodging (Keller et al., 1999; Marza et al., 2006; Li et al., 2016). Durum breeders must consider interactions between the above-mentioned and other agronomic and quality traits when selecting for superior genotypes.

Yield

Grain yield frequently is used as the main criterion for agronomic performance in crops. It is a complex, multicomponent trait greatly influenced by various plant morphological and physiological characteristics and growing conditions (Mohsin et al., 2009). Thus, success in breeding for improved yield requires combining the yield-related components (traits) associated with yield response (Li et al., 2011). Each yield component has its own genetic system and is influenced by environmental variations (Kahrizi et al., 2010). During the growing season, durum wheat may undergo severe weather conditions that can affect potential grain yield, so the ability of cultivars to adapt to different environments is an important characteristic (Ceccarelli et al., 1991). Yield stability refers to the ability of the plant genotype to express yield potential over a wide assortment of environments with minimal loss from the environmental stresses, diseases, and pests.

Days to heading

Heading date is one of the most important traits in cereal crops (Kitagawa et al., 2012). The appropriate transition from the vegetative to the reproductive stage is a critical adaptive trait. Crops with the appropriate heading time relative to the target environment can avoid pathogen interactions and negative environmental effects (such as hot temperatures), allowing them to maximize yield potential (Snape et al., 2001). Heading time in wheat is determined by three major genetic factors: the vernalization requirement, response to photoperiod, and earliness *per*

se (Borlaug, 1983; Kato et al., 1988; McIntosh et al., 2003). Photoperiod insensitivity (day length neutrality) and vernalization responses enable the cultivars to be grown in different latitudes and seasons. Wheat genotypes sensitive to photoperiod require long days to begin flowering, while genotypes insensitive to photoperiod flower independent of day length. Earliness, the time needed for a plant to reach a certain developmental stage, plays a major role in a plant's adaptation to its environment (Law and Worland, 1997).

Several studies in the northern Great Plains have found that photoperiod-insensitive lines are earlier to head and shorter than their photoperiod-sensitive counterparts; in these studies, the grain yield for the insensitive line was the same or higher than for the sensitive line, averaged over all environments (Busch et al., 1984; Marshall et al., 1989). The environmental conditions prevailing during the period from grain set to maturity can have a significant effect on grain yield. Photoperiod-insensitive cultivars tend to perform best in warmer, drier environments due to their avoidance of heat and moisture stress during grain filling (Fisher and Maurer, 1976; Musick and Dusek, 1980; Dyck et al., 2004).

Plant height

Plant height is one of the parameters associated with yield (Khush, 1999). Previous genetic research has uncovered a number of genes affecting plant height (*Rht* genes); however, only *Rht-B1b (Rht 1)*, *Rht-D1b (Rht 2)*, and *Rht8c* have been used extensively in agriculture (Borlaug 1968; Worland et al., 1998). Alleles for reduced height, found at the *Rht-B1* and *Rht-D1* loci on chromosomes 4B and 4D, respectively, were first identified in the 1930s on the Japanese cultivar Norin 10 (*Triticum aestvum* L.). The presence of mutant alleles at either *Rht-B1* or *Rht-D1*confers a semi-dwarf growth habit, while the presence of alleles at both loci confers a dwarf growth habit (Kalous et al., 2011).

Plant height is a major agronomic criterion in wheat breeding because it often affects grain yield by influencing resistance to lodging, number of tillers, and seeds per plant (Donald and Hamblin, 1976). According to several studies, semi-dwarf wheat has a higher grain yield potential and reduced lodging under high input and adequate moisture-growing conditions (Waddington et al., 1986; Knott, 1986; Worland and Snape, 2001; Hedden, 2003). McNeal et al. (1972) found that in low-yield environments, tall lines were superior to semi-dwarf ones. Kahrizi et al. (2010) reported a positive correlation between plant height with such morphological traits as peduncle length, flag length, leaf dry weight, stem dry weight, and spike dry weight and leaf area ratio. McCartney et al. (2005) demonstrated that reduced height at one of the major QTL, *QHt.crc-4D*, was significantly associated with decreased lodging, grain yield, test weight, and 1000-grain weight and increased time to maturity.

Lodging

Lodging in wheat occurs when the plant shoots are permanently displaced from an upright position (Pinthus, 1997). It may cause up to a 50% yield reduction and often results in reduced grain quality, greater drying costs, and a slower harvest (Stapper and Fischer, 1990). The main causes of lodging are strong winds, heavy rains, diseases, and pests that damage the root, stem, or crown. Lodging risk can be reduced by introducing the semi-dwarfing genes *Rht-B1b* and *Rht-D1b* into breeding materials (Wilhelm et al., 2013). However, stem lodging is also influenced by other factors, including main shoot weight, stem thickness, and second internode diameter (Sarker et al., 2007). Lodging represents a quantitative trait and is difficult to assess on a phenotypic basis. Few studies have identified QTL for stem diameter (Keller et al., 1999; Hai et al., 2005), stem wall width, and stem pushing resistance (Hai et al., 2005). Ma (2009) showed that a gene involved in the biosynthesis of lignin (COMT) was expressed more in a lodging-

resistant cultivar than a lodging-susceptible cultivar. Berry and Berry (2015) reported the number of QTL associated with lodging and noted that individual QTL with the largest estimated effects on lodging resistance were for height, stem diameter, stem strength, root spread, and root depth. Effectively improving the lodging resistance in wheat requires simultaneous selection for the traits that contribute to it.

Leaf disease

There are three very important leaf fungal diseases in North Dakota: tan spot [*Pyrenophora tritici-repentis* (PTR) (Died.) Drechs.], Septoria/Stagonspora nodorum blotch (SNB) [*Phaeosphaeria nodorum* (E. Müller) Hedjaroude], and Septoria tritici blotch (STB) [*Zymoseptoria tritici* (Desm.) Quaedvlieg & Crous, comb. nov.]. These diseases can cause severe yield losses under favorable to pathogenic conditions by reducing the photosynthetic area of leaves (Friskop and Liu, 2016). The need for simultaneous resistance to several diseases, change in the pathogen population, and complex genetic control of the trait are challenges when breeding for disease resistance. Several genes and QTL for resistance to tan spot have been reported (Faris et al., 1997; Faris and Friesen, 2005; Chu et al., 2008; Chu et al., 2010). Septoria/Stagonspora nodorum blotch flag leaf resistance has been reported on chromosome 2A, 2D, 3A, 3B, 5B, 6B, and 7B (Faris et al., 2009; Adhikari et al., 2011; Francki et al., 2011; Gurung et al., 2014). Resistance to STB was identified using bi-parental populations and reported on chromosomes 3A and 6B in spring wheat by Eriksen et al. (2003) and on chromosome 2B by Chartrain et al. (2009). Genome-wide association approaches have been used successfully for identifying SNP associated with these diseases, allowing for the elimination of long linkage blocks that can limit identification of closely-linked markers (Tommasini et al., 2007; Miedaner et al., 2013; Kollers et al., 2014).

Fusarium head blight (FHB), commonly known as scab, is presently one of the most detrimental fungal diseases for cultivated tetraploid durum wheat. *Fusarium graminearum* (sexual stage - *Gibberella zeae*) is a pathogenic, filamentous fungus that infects wheat, barley (*Hordeum vulgare* L.), oats (*Avena sativa*), and other small grain cereals, causing symptoms of FHB. Symptoms begin to appear at the point of infection in the form of water-soaked brown spots present on the glumes that eventually spread up and down the rachis. The production of sporodochia at the base of the infected glumes gives rise to a pinkish color on severely infected spikes. The lesions increase in size until the whole spikelet is covered and spread to the neighboring spikelets if the environment is conducive. Peduncles immediately below the inflorescence may become a discolored brown. With time, tissue of the inflorescence often becomes blighted (Gilbert and Tekauz, 1995). Studies of the genetic structure of populations of *G. zeae* collected from infected cereals in the United States have shown a large diversity among these populations (Zeller et al., 2004). The reason for the diversity may include extensive inter-population genetic exchange across large geographic regions, while control may include the use of resistant host genotypes and fungicides. While many QTL for FHB resistance have been found and are successfully used in hexaploid wheat (*Triticum aestvum*), most current wheat cultivars are highly susceptible, in part due to the narrow genetic diversity for FHB resistance in elite durum wheat germplasm (Elias et al., 2005; Gilbert and Tekauz, 1995). While many QTL for FHB resistance have been found and are successfully used in hexaploid wheat (*Triticum aestvum*), to date, a relatively small number of QTL for FHB resistance in durum wheat have been identified that provide satisfactory resistance (Prat et al., 2017). Therefore, ongoing research on new FHB virulence and regular surveillance to monitor possible changes in toxin chemotypes are important to prevent major FHB outbreaks in the future.

Falling number

Falling number (FN) is an important agronomic and end-use quality trait in durum wheat. It measures the rheological properties (changes in viscosity) of starch and the degree of starch damage. During pre-harvest sprouting caused by wet conditions before harvest, α-amylase enzyme activity is increased, resulting in starch damage. Thus, FN indirectly measures α-amylase enzyme activity. Falling number is defined as the time it takes a stirrer-viscometer to fall through a heated and ground whole wheat flour-water suspension (MacArthur et al., 1981). Low FN values generally indicate greater starch damage caused by higher levels of α-amylase enzyme activity. Values higher than 350 seconds are considered to represent sound wheat kernels, lacking α-amylase enzyme activity (Donnelly, 1980). In durum wheat, Dick et al. (1974) found no effect of extensive sprouting on semolina yield or spaghetti cooking quality. Donnelly et al. (1980) found that FN values negatively correlated with the percent damage values and that sprout damage of 10% or higher had a negative effect on spaghetti cooking quality and shelf-life stability.

Molecular markers associated with FN would be useful in selecting against durum wheat lines with low FN in early breeding generations. Six major QTL located on chromosomes 1B, 2A, 2B, 6B, and 7B explained 45% of FN variation in hard white spring wheat (Zhang et al., 2014). Several studies identified a number of QTL for FN using bi-parental and AM studies and reported the association of some QTL for FN with QTL for pre-harvest sprouting, late maturity, α-amylase, and seed dormancy (Kunert et al., 2007; Kulwal et al., 2012; Rasul et al., 2009; Mohler et al., 2014).

Genetics of Durum Wheat End Use Quality Traits

Although yield stability represents the major breeding goal in developing improved durum wheat cultivars, the stability of grain quality parameters is an important requirement for the milling and pasta industries (Rharrabti et al., 2003a). The grain quality of durum wheat is a complex characteristic that includes several components (Troccoli et al., 2000). Grain quality, which is controlled genetically as well as influenced by the growing environment, determines semolina quality, which in turn determines pasta processing and cooking quality parameters (Rharrabti et al., 2003b). To meet the broad quality requirements of the miller, end-product producer, and consumer, high quality durum wheat grain should possess a high protein content, a high percent of vitreous kernels, high FN, and a bright yellow color. Semolina protein quantity and quality are the most important components in pasta cooking quality (Dick and Matuso, 1988). Protein quality is associated with gluten strength. Semolina protein contains gliadin and glutenin proteins that upon hydration and mixing form a protein network often referred to as gluten. The firmness and resiliency of the cooked pasta products and stability during cooking is known to be associated with gluten strength (Sissons, 2008).

Grain Quality Parameters

Test weight, 1000-kernel weight, and kernel size

Test weight, 1000-kernel weight, and kernel size are interrelated. Test weight is a measure of the bulk density of wheat grain expressed in pounds per bushel (lb/bu), kilograms per hectoliter (kg/hl), or kilograms per cubic meter (kg m^{-3}). Grain volume weight as an indicator of grain quality is an important parameter for marketing purposes. For example, the U.S, grade for durum wheat is dependent on test weight. U.S, Nos.1, 2, 3, and 4 grades require a minimum test weight of 60.0, 58.0, 56.0, and 54.1 lb/bu, respectively (74.8, 72.3, 70.0, and 67.5 kg/hL,

respectively). Test weight has been reported to be significantly correlated with kernel weight, kernel diameter, and kernel size distribution (Barmore et al., 1965; Matsuo and Dexter, 1980). Kernel weight alone does not guarantee a high test weight. Test weight has been reported to be highly correlated with packing efficiency, which can be more related to kernel shape than kernel size (Yamazaki and Briggle, 1969).

Thousand-kernel weight (TKW) is a trait important from both an agronomic and quality standpoint. It is one of the most important yield components in wheat. Thousand-kernel weight is more heritable and stable under different environments than other yield components, such as spikes per plant and kernels per spike (Xiao and He, 2003). It is usually used in breeding programs to estimate the agronomic yield among various wheat genotypes (Baril, 1992; Troccoli et al., 2000). In addition, breeders and flour millers use1000-kernel weight as a complement to test weight (TWT).

Thousand-kernel weight is an indirect measure of average kernel size and shows a positive correlation with semolina yield (Dexter et al., 1987; Abaye et al., 1997). Like most agronomic and quality traits, TKW is a quantitative trait represented by continuous variation in a segregating population. Thousand-kernel weight is highly influenced by the environment during grain filling (number of spikes, number of fertile florets per spikelet, and kernel dimensions and volume) (Blumenthal et al., 1991; Schuler et al., 1994). During domestication and high selection intensity, genetic diversity for TKW has been reduced and allelic diversity at selected loci lowered (Wang et al., 2014). In durum wheat, Elouafi and Nachit (2004) reported and mapped five major QTL for TKW that explained 32% of the total variation, of which 25% was of a genetic nature. They also showed a major QTL for genotype by environment interaction around the centromere on chromosome 6B. Two major QTL for TKW were identified on chromosomes

3B and 4B by Russo et al (2014) using SNP markers that corresponded with those identified by Prashant et al. (2012) and Patil et al. (2013). Three QTL for kernel weight, kernel diameter, and kernel size distribution were detected on chromosomes 2A, 5B, and 7A (Tsilo et al., 2010). In the study Tsilo et al. (2010) reported a positive correlation for the percentage of large kernels with a single-kernel diameter ($r = 0.93$ at $P \leq 0.001$) and a negative correlation of the latter with the percentage of medium and small kernels, suggesting that kernel diameter and kernel size distribution have some genes in common.

Test weight is quantitative in nature and is highly influenced by the genotype by environment interaction (Rharrabti et al., 2003). Among the 11 putative QTL associated with TWT that were detected on chromosomes 1AL, 1BL, 2BL, 3BS, 3BL, 5AL, 6AL, 6BS, and 7BS, five were detected in one environment only, three in two environments, and three in five or more environments. Most of these QTL co-localized with those for TKW (Graziani et al., 2014). Patil et al. (2013) identified five QTL for TWT located over five chromosomes. The consistent QTL on chromosome 2A between the marker interval *Xgwm71.2–Xubc835.4* with a pleiotropic effect on TWT and TKW was suggested for use in early generation selection to improve TWT and TKW (Patil et al., 2013).

Grain protein concentration

Grain protein concentration is an important quality trait in durum wheat for pasta manufacturing due to its effect on the firmness of cooked pasta and tolerance to overcooking (Dexter and Matsuo, 1977). Protein quantity is a major determinant for the quality of pasta dried at high temperatures, while both protein quantity and its composition are important components for high-quality pasta dried at a low temperature (Dick and Matuso, 1988; Novaro et al., 1993). Generally, durum wheat grain protein ranges from 10 to 18% at a 12% moisture basis. Grain

protein content of 11% or lower results in a poor-quality pasta regardless of drying temperature regimes (D'Egidio et al., 1990).

Protein content is a quantitative trait with low heritability that is largely influenced by the environment, and, in some cases, has been found to be negatively correlated with grain yield (Johnson et al., 1985; Oury and Godin, 2007; Blanco et al., 2012). Elevated air temperature during grain filling results in an increase in grain protein percentage due to decreased starch synthesis. Although the increased protein percentage is a positive quality factor, the changes in the protein composition due to high air temperature can negate this positive effect. High air temperature during grain filling often results in an increase in the gliadin to glutenin ratio, which in turn results in reduced gluten strength and end-product quality (Moldestad et al., 2011).

Some major QTL for grain protein concentration were detected in a number of environments and populations on almost all chromosomes in both hexaploid and tetraploid wheat. In durum wheat, grain protein concentration QTL have been reported on chromosomes 2AS, 2 BL, 3AS, 4 AL, 7AL, 7AS, and 7BL (Zhang et al., 2008; Raman et al., 2009; Suprayogi et al., 2009; Sun et al., 2010; Blanco et al., 2012). A gene for high grain protein content located near the centromere of chromosome 6B of the 'Langdon' (DIC-6B) substitution accounted for 66% of the variation in a mapping population of 85 RILs (Joppa et al., 1997). Cloning of the gene *Gpc-B1* on 6BS revealed that its effect on grain protein concentration is due to improved N remobilization (Brevis and Dubcovsky, 2010). Some major QTL have been reported to show no reduction in grain yield in some environments, while the same QTL have been associated with a reduction in grain yield in others, indicating an environment and genotype by environment interaction (Joppa et al., 1997; Groos et al., 2003; Prasad et al., 2003; Borner et al., 2007; Suprayogi et al., 2009; Sun et al., 2010; Blanco et al., 2012; Terasawa et al., 2016). These

findings suggest that the use of these QTL is limited to certain environments and genetic backgrounds, with additional studies required to determine their utility.

Kernel vitreousness

Kernel vitreousness is an important quality factor worldwide. It is defined as the percentage of kernels, based on weight, having a translucent endosperm. Non-vitreous kernels with a starchy endosperm are opaque due to air spaces between starch granules (REF). The degree of kernel translucency, and thus the apparent degree of vitreousness, is related to the degree of kernel compactness. Generally, the more desirable coarse semolina is produced from highly vitreous wheats (Hoseney, 1987).

Durum wheat kernels tend to be harder than hexaploid wheat due to the absence of starch granule proteins called puroindolins (Baldwin, 2001). The *puroindoline* gene mapped on chromosome 5D in bread wheat is absent in tetraploid wheat (AABB) (Igrejas et al., 2002; Chantret et al., 2005). However, the *puroindoline* gene does not account for the full range of variation in kernel hardness, and several minor QTL have been mapped on different chromosomes in previous studies in hexaploid wheat (Sourdille et al., 1996). The *puroindoline b-2* gene was discovered in the durum wheat Langdon; it has at least two variants in one of the A and B genomes in some durum wheat, suggesting that the *puroindoline b-2* gene comprises a multigene family (Chen et al., 2011).

Flour- and Semolina-related Characteristics

Milling extraction

Wheat milling yield is an important criterion for the milling industry. Increases in semolina yield mean increases in profits for durum wheat millers. Several factors can affect the extraction rate: the amount of germ, thickness of the bran, size of the grain, kernel size

distribution, and depth of the crease (Marshall et al., 1984, 1986; Yoon et al., 2002). Grain hardness also plays an important role in semolina extraction, affecting the endosperm fracture pattern, starch damage, particle size, and the ease of separating bran from endosperm (Stenvert, 1972). Hard and soft wheat differ in the strength with which starch granules are attached to the protein matrix. Durum wheat has the hardest kernel of all due to the absence of the hardness (Ha) locus, which is present in hexaploid wheat on chromosome 5D (Igrejas et al., 2002; Chantret et al., 2005). In the preliminary analysis for hexaploid wheat, two regions were identified on chromosomes 3A and 7D that were significantly associated with milling yield, accounting for 22% and 19% of genetic variation, respectively. Another region identified with AFLP on chromosome 5A accounted for 19% of genetic variation (Parker et al., 1999).

Ash content

Ash content is an indicator of semolina/flour contamination with bran. It is related to kernel hardness, kernel size, kernel uniformity and bran thickness (Posner, 2009; Yoon et al., 2002). Mineral accumulation in the grain primarily depends on the translocation from the leaves to the developing kernels after the onset of senescence (Wardlaw, 1990). The heritability of ash content is greater than that for grain yield in durum wheat, indicating that the trait is under genetic control (Araus et al., 1998). Kernel ash content is highly influenced by environment and genotype by environment interaction (Fares et al., 1995). Favorable growing conditions result in increased ash content in the whole grain due to the increased uptake of minerals from the soil (Cubadda, 1988). High correlation has been observed between ash content in the whole grain and semolina. Under drought stress, the re-mobilization of minerals from the vegetative organs is much higher than that in well-watered conditions, which leads to a decrease in ash content in

the flag leaf at maturity and an increase in ash content in the grain at maturity (Merah et al., 1999).

Protein- and dough-related traits

There are two major wheat protein fractions, glutenins and gliadins that together represent about 80% of the total protein in typical wheat flour. The ratio of the two protein fractions determines the strength of gluten that is formed during hydration and mixing. The gliadin: glutenin ratio is affected by environment and predetermines gluten strength and, therefore, the quality of the end-product. Pasta made from semolina that forms a strong gluten is generally more resistant to overcooking. The glutenins are mostly responsible for the elasticity of dough, while extensibility is mainly conferred by the gliadins (Shewry et al., 1995). Glutenin consists of high molecular weight (HMW) and low molecular weight (LMW) subunits whose genes reside on chromosomes 1A and 1B (Payne and Lawrence, 1983; Payne et al., 1984; D'Ovidio and Masci, 2004). The HMW glutenin subunits (HMW-GS) are particularly important for determining dough elasticity (Anjum et al., 2007).

Different studies have identified the best glutenin subunit alleles, contributing to the increase of durum quality in modern cultivars through the breeding process (Peña et al., 1994; Ruiz and Carrillo, 1995). However, several other studies identified QTL associated with gluten strength on other chromosomes as well (Blanco et al., 1998; Elouafi et al., 2000; Patil et al., 2009; Kumar et al., 2013). Wheat breeding programs evaluate gluten strength predominantly by measuring the sodium dodecylsulphate (SDS)-sedimentation volume as a faster test with fewer required materials compared to the gluten index test and the alveograph and mixograph methods (Zhang et al., 2008); SDS volume has been shown to correlate with gluten strength and spaghetti cooking quality (Dexter et al., 1980; D'Egidio et al., 1990).

Yellow pigment content

Yellow pigment found in the endosperm of durum wheat is responsible for the yellow appearance of pasta. Consumers perceive pasta color as an important quality factor. Processors prefer clear, bright yellow semolina, which generally produces a superior end-product. Yellow pigment in durum wheat endosperm is caused primarily by carotenoids, whose content can be affected by genotype and growing environment. Taghouti et al. (2010) reported that variation due to genotype was greater than that due to the environment. These results are similar to other studies reporting the greater influence of genotype than environment on yellow pigment content. (Rharrabti et al., 2003; Pozniak et al., 2007).

The genetics of yellow pigment concentration in durum wheat has been studied extensively and is controlled largely by genetic factors with additive effects (Elouafi et al., 2001; Singh et al., 2009). The QTL that mapped to chromosomes 7A and 7B have been shown to explain in many cases over 50% of the observed phenotypic variability (Parker et al., 1998; Elouafiet al., 2001; Zhang et al., 2008). Other studies have reported minor QTL for color traits on both arms of these chromosomes (Blanco et al., 2011; Roncallo et al., 2012). Pozniak et al. (2007) mapped the genes coding for phytoene synthase (*Psy*) on the 7 and 5 chromosome groups. Phytoene synthase is a transferase enzyme involved in the biosynthesis of carotenoids in the biosynthesis pathway. In their study, the *Psy1–1* locus co-segregated with the 7B QTL. Minor QTL for yellow pigment content were essentially detected on all the durum wheat chromosomes, indicating the complexity of inheritance of this color trait (Parker et al., 1998; Pozniak et al., 2007; Zhang et al., 2008; Reimer et al., 2008; Colasuonno et al., 2014). A high level of carotenoid pigments in semolina does not, however, solely guarantee a high yellow color for the pasta itself. The degree of yellowness is additionally affected by the semolina extraction rate

(Dexter and Matsuo, 1978; Borrelli et al., 1999), oxidative degradation by the lipoxygenase (LOX) enzyme during processing, and processing conditions (Borrelli et al 2003). Lipoxygenase levels in durum wheat are cultivar-related and depend on the environment (Troccoli et al., 2000). Research by De Simone et al. (2010) showed that the activity of LOX in semolina was determined by the transcript level, presence of number of LOX isoforms, and the amount of LOX enzyme.

Polyphenol oxidase

Polyphenol oxidase (PPO) catalyzes the oxidation of phenolic compounds into o-quinones in the presence of molecular oxygen. The generated o-quinones further react with amines and thiol groups or undergo self-polymerization to produce dark/brown colored polyphenols (Anderson and Morris, 2003). Although high PPO activity can be beneficial in a variety of products, such as prunes, dark raisins, tea, and coffee, the brown color tends to mask the yellow color in semolina end-products during processing when it reaches substantial levels (Sissons, 2008). In mature grain, it was observed that wheat pericarp contained the highest PPO activity, with little activity found in wheat white flour, and none observed in the embryo (Marsh and Galliard, 1986). Milling wheat grain at a higher flour extraction rate raises the darkening effect in end products (Baik et al., 1994). Polyphenol oxidase activity varies with genotype, and durum wheat cultivars have lower PPO activity than common wheat cultivars (Lamkin et al., 1981).

Previous studies reported that PPO activity was mainly controlled by the genes located on homologous group 2 chromosomes (Si et al., 2012; Zhang et al., 2005). Major QTL for PPO activity were detected on chromosome arm 2AL between the markers *Ppo-A1* and

RAC875_c9845_762 1.4 cM apart. They explained 48.6–58.4 % of the phenotypic variance and were identified using the wheat 90 K iSelect array (Zhai et al., 2016).

Pasta Cooking-related Traits

Cooked quality for pasta is determined by its cooked firmness, cooking loss, and cooked weight. Cooked firmness determines the chewing characteristics of pasta, which should be firm to the bite (*al dente*). Cooking loss is the percent solids lost to the cooking water. Cooked weight measures the water-absorbing capacity of the pasta during cooking. Durum wheat is mainly used for pasta; the cultivars that meet the requirements of high-quality pasta products receive premium prices in the global market. Research has shown a significant positive correlation between cooking quality and protein quantity and quality (Dexter and Matsou, 1980; Sissons et al., 2005). Generally, high-protein semolina with strong gluten produces pasta with better cooking parameters and tolerance to overcooking as compared to low-protein and weak-gluten semolina. QTL affecting pasta firmness and cooking loss were detected on chromosomes 5A and 7B overlapping with QTL for grain protein content and wet gluten content (Zhang et al., 2008).

Genome-wide Association Study (GWAS) vs. Bi-parental QTL Mapping

Genome-wide association study (GWAS), or association mapping (AM), is an alternative approach to traditional bi-parental QTL mapping or linkage mapping (LM). It enables the identification and selection of chromosome regions harboring QTL responsible for phenotypic variation (Tuberosa and Salvi, 2006; Collins et al., 2008; Myles et al., 2009). Although both approaches use recombination's ability to break up the genome into fragments that can be correlated with phenotypic variation, AM has advantages over traditional bi-parental QTL mapping. These include: (1) increased mapping resolution, (2) reduced research time, and (3)

the ability to detect greater allele number (Yu and Buckler, 2006). Greater mapping resolution is achieved through many rounds of unrelated genotypes in a natural experiment. Reduced research time is achieved by eliminating the step of developing a bi-parental population. Association mapping studies detect statistical associations between genotypic and phenotypic data in samples of unrelated individuals (Zondervan and Cardon, 2004). The ability to detect greater allele numbers also comes from the nature of the population used in AM. Non-random mating has generated complex patterns of population structure and relatedness in crops and wild plants (Flint-Garcia et al., 2005; Nordborg et al., 2005). To avoid false-positive associations (Type I errors), it is important to account for population structure and relatedness with kinship and Q matrices in AM studies.

The results of AM experiments are valuable for marker-assisted breeding programs as they allow for identification of the desirable allelic variants at the major loci controlling the target traits (Heffner et al., 2009). Also, AM is a powerful approach for assessing genetic diversity in the breeding program. This information about the genetic diversity and population structure in elite breeding material is necessary for crop improvement (Inghelandt et al., 2010). Association mapping enables breeders to assess the allelic combinations selected over generations, providing genetic haplotypes for future crop improvement through MAS (Kishore et al., 2012).

Marker-assisted selection (MAS) in Durum Wheat

To increase breeding progress, some breeding programs are employing larger population sizes and using double haploids, which add to the cost of line development. To improve the selection for quantitative traits, breeders need to find efficient tools and strategies to reduce the cost. Marker-assisted selection, also known as marker-assisted breeding (MAB), provides an

opportunity for wheat breeders to introgress/pyramid genes of interest into breeding lines and to identify QTL in germplasm to be used as parents. Marker-assisted selection can be more cost-effective and precise than extensive phenotyping to incorporate donor genes for traits: (1) that are expensive or time-consuming to measure, (2) whose expression is environmental or developmental dependent, (3) require backcross breeding for recessive alleles, or (4) require pyramiding for desirable traits (Xu and Crouch, 2008). Marker-assisted selection also offers an opportunity for selection in early generations, thus improving the selection response. Previous and on-going work show that markers can be effectively used for genetic assessment and MAB in many crops, including durum wheat (Liu et al., 2008; Haeberle et al., 2009; Ruan et al., 2012; Kumar et al., 2013; AbuHammad et al., 2016;). Gene-based high-throughput genotyping results in more effective genetic mapping/genome analysis and opens new opportunities for its integration in wheat breeding programs worldwide. High-density SNP detection platforms have been developed for wheat with the ability to detect several thousand SNP and are showing promise as a tool for genome-wide association studies and genomic selection strategies (Chao et al., 2010; Paux et al., 2010).

Unbalanced Dataset for QTL Detection Using GWAS

The collection of phenotypic data requires extensive effort and resources. Breeding programs collect large amounts of phenotypic data for selection purposes from their advanced breeding lines each year. However, the number of such advanced breeding lines tested each year is relatively small, and some are replaced by other breeding lines, resulting in an unbalanced dataset for advanced breeding lines developed over time. If this historic unbalanced data set could be combined with the more affordable marker genotyping, the results could be used for genome-wide association studies (GWAS), saving significant resources and providing useful

information about marker trait association within individual breeding programs. Applying QTL mapping within breeding program activities would greatly advance the utility of QTL detection; the marker–QTL associations detected would be immediately useful for MAS.

Only a limited number of studies have been done comparing the efficacy of unbalanced datasets for QTL detection using GWAS. Simulation studies by Wang et al. (2012) estimated the optimum number of breeding lines to be greater than 384 for accurately predicting major and minor QTL using an unbalanced dataset.

References

Abaye, A.O., D.E. Brann, M.M. Alley, and C.A. Griffey. 1997. Winter durum wheat: Do we have all the answers? Publication 424-802. Crop and Soil Environmental Sci. Virginia Cooperative Extension.

Abouzied, H.M., S.M.M. Eldemery, K.F. Abdellatif. 2013. SSR based genetic diversity assessment in tetraploid and hexaploid wheat populations. Br. Biotechnol. J. 3: 390–404.

AbuHammad, W.A., S. Mamidi, A. Kumar et al. 2016. Identification and validation of a major cadmium accumulation locus and closely associated SNP markers in North Dakota durum wheat cultivars. Mol. Breeding. 36:112-120.

Adhikari, T.B., E.W. Jackson, S. Gurung, J.M. Hansen, and Bonman. 2011. Association mapping of quantitative resistance to Phaeosphaeria nodorum in spring wheat landraces from the USDA National Small Grains Collection. Phytopathol. 101:1301–1310.

Anderson, J.V., and C.F. Morris. 2003. Purification and analysis of wheat grain polyphenol oxidase (PPO) protein. Cereal Chem. 80:135-143.

Anjum, F.M., M.R. Khan, A. Din, M. Saeed, I. Pasha, M.U. Arshad. 2007. Wheat gluten: high molecular weight glutenin subunits-structure, genetics, and relation to dough elasticity. J. Food Sci. 72:56-63.

Araus, J.L., T. Amaro, J. Casadesus, A. Asbati, and M.M. Nachit. 1998. Relationship between ash content, carbon isotope discrimination and yield in durum wheat Australian J. Plant Physiol. 25:835-842.

Arumuganathan, E., and E.D. Earle. 1991. Nuclear DNA content of some important plant species. Plant Mol. Biol. 9:208–218.

Autrique, E., M.M. Nachit, P. Monneveux, S.D. Tanksley, and M.E. Sorrells.1996. Genetic diversity in durum wheat based on RFLPs, morphophysiological traits, and coefficient of parentage. Crop Sci. 36:735-742.

Ayal, S., R. Ophir, and A.A. Levy. 2005. Genomics of tetraploid wheat domestication. Wheat Inf. Ser. 100:185-204.

Baik, B.K., Z. Czuchajowska, and Y. Pomeranz, 1994. Comparison of polyphenol oxidase activities in wheats and flours from Australian and US cultivars. J. Cereal Sci. 19:291-296.

Baldwin, P.M. 2001. Starch granule-associated proteins and polypeptides: a review. Starch. 53:475-503.

Ball, C.R. 1930. The history of American wheat improvement. Agric. Hist. 4: 48–71.

Baril, C.P. 1992. Factor regression for interpreting genotype–environment interaction in bread-wheat trials. Theor. Appl. Genet. 83:1022–1026.

Barmore, M.A., and R.K. Golumbic. 1965. Wight per bushel and flour yield of Pacific Northwest white wheat. Cereal Sci. Today. 10:72-77.

Berry, P.M., and S.T. Berry. 2015. Understanding the genetic control of lodging-associated plant characters in winter wheat (Triticum aestvum L.). Euphytica. 205:671-689.

Blanco, A., G. Mangini, A. Giancaspro, S. Giove, P. Colasuonno, R. Simeone, A. Signorile, P. De Vita, A.M. Mastrangelo, L. Cattivelli, A. Gadaleta. 2012. Relationships between grain protein content and grain yield components through quantitative trait locus analyses in a recombinant inbred line population derived from two elite durum wheat cultivars. Mol. Breeding. 30:79-92.

Blanco, A., M.P. Bellomo, A. Cenci, C. De Giovanni, R. D'Ovidio, E. Iacono, B. Laddomada, M.A. Pagnotta, E. Porceddu, and A.A. Sciancalepore. 1998. Genetic linkage map of durum wheat. Theor. Appl. Genet. 97:721-728.

Blanco, A., P. Colasuonno, A. Gadaleta, G. Mangini, A. Schiavulli, R. Simeone, A.M. Digesù, P. De Vita, A.M. Mastrangelo, and L. Cattivelli. 2011. Quantitative trait loci for yellow pigment concentration and individual carotenoid compounds in durum wheat. J. Cereal Sci. 54:255-264.

Blumenthal, C.S., F. Bekes, I.L. Batey, C.W. Wrigley, H.J. Moss, D.J. Mares, E.W.R Barlow. 1991. Interpretation of grain quality results from wheat variety trials with reference to high temperature stress Aust. J. Agric. Res. 42:325-334.

Borlaug, N.E. 1968. Wheat breeding and its impact on world food supply. Pages 1-36 in: Proceedings of the 3rd International Wheat Genetics Symposium, Canberra, Australia, 5–9 August 1968. Edited by K.W. Finlay and K.W. Shepard. Australian Academy of Science, Canberra, Australia.

Borlaug, N.E. 1983. Contributions of conventional plant breeding to food production. Sci. 219:689–693.

Börner, A., E. Schumann, A. Fürste, H. Cöster, B. Leithold, M.S. Röder, W.E. Weber. 2002. Mapping of quantitative trait loci for agronomic important characters in hexaploid wheat (Triticum aestivum L.). Theor. Appl. Genet. 105:921–936.

Borrelli, G.M., A. Troccoli, N. Di Fonzo, C. Fares. 1999. Durum wheat lipoxygenase activity and other quality parameters that affect pasta color. Cereal Chem. 76:335–340.

Brevis, J.C., and J. Dubcovsky. 2010. Effects of the chromosome region including the grain protein content locus Gpc-B1 on wheat grain and protein yield. Crop. Sci. 50:93–104.

Busch, R.H., F.A. Elsayed, and R.E. Heiner. 1984. Effect of day length insensitivity on agronomic performance in spring wheat. Crop Sci. 21:57-60.

Cakmak, I.A., E. Torun, E. Millet, M. Feldman, T. Fahima, A.B. Korol, E. Nervo, H.J.O. Braun, and H. Zkan. 2004. Triticum dicoccoides: an important genetic resource for increasing zinc and iron concentration in modern cultivated wheat. Soil Sci. Plant. Nutr. 50:1047-1054.

Cavanagh, C.R., S.M. Chao, S.C. Wang, B.F. Huang, S. Stephen, S. Kiani, K. Forrest, C. Saintenac, G.L. Brown-Guedira, A. Akhunova, D. See, G.H. Bai, M. Pumphrey, L. Tomar, D.B. Wong, S. Kong, M. Reynolds, M.L. da Silva, H. Bockelman, L. Talbert, J.A. Anderson, J. Dreisigacker, S. Baenziger, A. Carter, V. Korzun, P.L. Morrell, J. Dubcovsky, M.K. Morell, M.E. Sorrells, M.J. Hayden, and E. Akhunov. 2013. Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. Proc. Natl. Acad. Sci. 110:8057-8062.

Ceccarelli, S., E. Acevedo, and S. Grando. 1991. Breeding for yield stability in unpredictable environments: single traits, interaction between traits, and architecture of genotypes. Euphytica. 56:169-185.

Chantret, N., J. Salse, F. Sabo, S. Rahman, A. Bellec, B. Laubin, I. Dubois, C. Dossat, P. Sourdille, P. Joudrier, M.F. Gautier, L. Cattolico, M. Beckert, S. Aubourg, J. Weissenbach, M. Caboche, M. Bernard, P. Leroy, and B. Chalhoub. 2005. Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (Triticum and Aegilops). Plant Cell. 17:1033-1045.

Chao, S., J. Dubcovsky, J. Dvorak, M.-C. Luo, S.P. Baenziger, R. Matnyazov, D.R. Clark, L.E. Talbert, J.A. Anderson, S. Dreisigacker, K. Glover, J. Chen, K. Campbell, P.L. Bruckner, J.C. Rudd, S. Haley, B.F. Carver, S. Perry, M.E. Sorrells, and E.D. Akhunov. 2010. Population- and genome-specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (Triticum aestivum L.). BMC Genomics. 11:727-735.

Chartrain, L., P. Sourdille, M. Bernard, J. Brown. 2009. Identification and location of Stb9, a gene for resistance to septoria tritici blotch in wheat cultivars Courtot and Tonic. Plant Pathol. 58:547-555.

Chen, F., H.X. Xu, F.Y. Zhang, X.C. Xia, Z.H. He, D.W. Wang, Z.D. Dong, K.H. Zhan, X.Y. Cheng, D.Q. Cui. 2011. Physical mapping of puroindoline b-2 genes and molecular characterization of a novel variant in durum wheat (Triticum turgidum L.). Mol. Biol. 28:153-161.

Chu, C., S. Chao, T.L. Friesen, J.D. Faris, S. Zhong, S.S. Xu. 2010. Identification of novel tan spot resistance QTLs using an SSR-based linkage map of tetraploid wheat. Mol. Breed. 25:327–338.

Chu, C., Z. Niu, S. Zhong, S. Chao, T. Friesen, and S. Halley. 2011. Identification and molecular mapping of two QTLs with major effects for resistance to Fusarium head blight in wheat. Theor. Appl. Genet. 123:1107–1119.

Chu, C.G, T. Friesen, S. Xu, and J. Faris. 2008. Identification of novel tan spot resistance loci beyond the known host-selective toxin insensitivity genes in wheat. Theor. Appl. Genet. 117: 873–881.

Clarke, J.M., W.A. Norvell, F.R. Clarke, and W.T. Buckley. 2002. Concentration of cadmium and other elements in the grain of near-isogenic durum lines. Can. J. Plant Sci. 82:27–33.

Colasuonno, P., A. Gadaleta, A. Giancaspro, D. Nigro, S. Giove, O. Incerti, G. Mangini, A. Signorile, R. Simeone, A. Blanco. 2014. Development of a high-density SNP-based linkage map and detection of yellow pigment content QTLs in durum wheat. Mol. Breeding. 34:1563–1578.

Collins, N.C., F. Tardieu, R. Tuberosa. 2008. Quantitative trait loci and crop performance under abiotic stress: where do we stand? Plant Physiol.147:469-486.

Cubadda, R. Evaluation of durum wheat, semolina and pasta in Europe. Pages 217-228 in: Durum Wheat: Chemistry and Technology. G. Fabriani and C. Lintas, eds. American Association of Cereal Chemistry: St. Paul, MN.

D'Egidio, M.G., B.M. Mariani, S. Nardi, P. Novaro, and R. Cubadda. 1990. Chemical and technological variables and their relationships: a predictive equation for pasta cooking quality. Cereal Chem. 67:275-281.

Damania, A.A.B. 2013. History of Introduction and Cultivation of Durum Wheat in North America. Poster presented at: Water, Food, Energy, and Innovation for a Sustainable World. ASA, CSSA and SSSA International Annual Meetings, Tampa, FL. 3-6 Nov. Poster 116–2.

De Simone, V., V. Menzo, A.M. De Leonardis, D.B. Ficco, D. Trono, L. Cattivelli, and P. De Vita. 2010. Different mechanisms control lipoxygenase activity in durum wheat kernels. J. Cereal Sci. 52:121-128.

Dexter, J.E., and R.R. Matsuo. 1977. Influence of protein content on some durum wheat quality parameters. Can. J. Plant Sci. 57:717-727.

Dexter, J.E., and R.R. Matsuo. 1980. Relationship between durum wheat protein properties and pasta dough rheology and spaghetti cooking quality. J. Agric. Food Chem. 28:899–902.

Dexter, J.E., R.R. Matsuo, D.G. Martin. 1987. The relationship of durum wheat test weight to milling performance and spaghetti quality. Cereal Food World. 32:772–777.

Dexter, J.E., R.R. Matsuo, F.G. Kosmolak, D. Leisle, and B.A. Marchylo. 1980. The suitability of the SDS-sedimentation test for assessing gluten strength in durum wheat. Can. J. Plant Sci. 60:25-29.

Dick, J.W., and R.R. Matsou. 1988. Durum wheat and pasta products. Pages 507-547 in: Wheat chemistry and technology, Volume 1. Y. Pomeranz, eds. American Association of Cereal Chemistry: St. Paul, MN.

Dick, J.W., D.E. Walsh, and K.A. Gilles. 1974. The effect of field sprouting on the quality of durum wheat. Cereal Chem. 51:180-188.

Donald, C.M., and J. Hamblin. 1976. The biological yield and harvest index of cereals as agronomic and plant breeding criteria. Adv. Agron. 28:361-405.

Donnelly, B. 1980. Effect of sprout damage on durum wheat quality. Macaroni J. 62:8-10.

D'Ovidio, R., and S. Masci. 2004. The low-molecular-weight glutenin subunits of wheat gluten. J. Cereal Sci. 39: 321–339.

Dvorak, J., E.D. Akhunov, A.R. Akhunov, K.R. Deal, and M.C. Luo. 2006. Molecular characterization of a diagnostic DNA marker for domesticated tetraploid wheat provides evidence for gene flow from wild tetraploid wheat to hexaploid wheat. Mol. Biol. Evol. 23:1386–1396.

Dyck, J.A., M.A. Matus-Cadiz, P. Hucl, L. Talbert, T. Hunt, J.P. Dubuc, H. Nass, G. Clayton, J. Dobb, and J. Quick. 2004. Agronomic performance of hard red spring wheat isolines sensitive and insensitive to photoperiod. Crop Sci. 44:1976– 1981.

Elias, E., F. Manthey, R. Stack, and S. Kianian. 2005. Breeding efforts to develop Fusarium head blight resistant durum wheat in North Dakota. Pages 25-26 in: Proceedings of the National Fusarium Head Blight Forum. S. Canty, J. Lewis, L. Siler, R. Ward, eds. East Lansing: Michigan State University, Milwaukee.

Elouafi, I., and M.M. Nachit. 2004. A genetic linkage map of the Durum × Triticum dicoccoides backcross population based on SSRs and AFLP markers, and QTL analysis for milling traits. Theor. Appl. Genet. 108:401–413.

Elouafi, I., M.M. Nachit, and L.M. Martin. 2001. Identification of a microsatellite onchromosome 7B showing a strong linkage with yellow pigment in durum wheat (Triticum turgidum L. var. durum). Hereditas. 135:255-261.

Eriksen, L., F. Borum, and A. Jahoor. 2003. Inheritance and localisation of resistance to Mycosphaerella graminicola causing septoria tritici blotch and plant height in the wheat (Triticum aestivum L.) genome with DNA markers. Theor. Appl. Genet. 107: 515–527.

Fares, C., A. Troccoli, and N. Di Fonzo. Use of friction debranning to evaluate ash distribution in Italian durum wheat cultivars. Cereal Chem. 73:232-234.

Faris, J., and B.S. Gill. 2002. Genomic targeting and high resolution mapping of the domestication gene Q in wheat. Genome. 45:706–718.

Faris, J., J. Anderson, L. Francl, J. Jordahl. 1997. RFLP mapping of resistance to chlorosis induction by Pyrenophora tritici-repentis in wheat. Theor. Appl. Genet. 94: 98–103.

Faris, J., T. Friesen. 2005. Identification of quantitative trait loci for race-nonspecific resistance to tan spot in wheat. Theor. Appl. Genet. 111: 386–392.

Faris, J.D., and T.L. Friesen. 2009. Reevaluation of a tetraploid wheat population indicates that the Tsn1-ToxA interaction is the only factor governing Stagonospora nodorum blotch susceptibility. Phytopathol. 99: 906–912.

Fischer, R.A., and R.O. Maurer. 1976. Crop temperature modification and yield potential in dwarf spring wheat. Ann. Appl. Biol. 80: 283–299.

Flint-Garcia, S.A., A.-C. Thuillet, J. Yu, G. Pressoir, S.M. Romero, S.E. Mitchell, J. Doebley, S. Kresovich, M.M. Goodman, and E.S. Buckler. 2005. Maize association population: a high-resolution platform for quantitative trait locus dissection. Plant J. 44:1054–1064.

Francki, M., M. Shankar, E. Walker, R. Loughman, H. Golzar et al. 2011. New quantitative trait loci in wheat for flag leaf resistance to Stagonospora nodorum blotch. Phytopathol. 101:1278–1284.

Friskop, A., and Z. Liu. 2016. Fungal Leaf Spot Diseases of Wheat: tan spot, septoria/stagonospora nodorum blotch and septoria tritici blotch. North Dakota State University Extension Service. PP–1249 (Revised).

Gepts, P. 2004. Crop domestication as a long-term selection experiment. Plant Breeding Rev. 24:1–44.

Gilbert, J., and A. Tekauz. 1995. Effects of Fusarium head blight and seed treatment on germination, emergence, and seedling vigour of spring wheat. Can. J. Plant Pathol. 17:252-259.

Graziani, M., M. Maccaferri, C. Royo, F. Salvatorelli and R. Tuberosa. 2014. QTL dissection of yield components and morpho-physiological traits in a durum wheat elite population tested in contrasting thermo-pluviometric conditions. Crop Pasture Sci. 65:80-95.

Groos, C, N. Robert, E. Bervas, G. Charmet. 2003. Genetic analysis of grain protein-content, grain yield and thousand-kernel weight in bread wheat. Theor. Appl. Genet. 106:1032–1040.

Gurung, S., S. Mamidi, J.M. Bonman, M. Xiong, G. Brown-Guedira, T.B. Adhikari. 2014. Genome-wide association study reveals novel quantitative trait loci associated with resistance to multiple leaf spot diseases of spring wheat. PLoS One. 108-179.

Haeberle, J., J. Holzapfel, G. Schweizer, and L. Hartl. 2009. A major QTL for resistance against Fusarium head blight in European winter wheat. Theor. Appl. Genet. 119:325-332.

Hai, L., H. Guo, S. Xiao, G. Jiang, X. Zhang, C. Yan, Z. Xin and J. Jia. 2005. Quantitative trait loci (QTL) of stem strength and related traits in a doubled-haploid population of wheat (Triticum aestivum L.). Euphytica. 141:1–9.

Hedden, P. 2003. Constructing dwarf rice. Nat. Biotechnol. 21:873–874.

Heffner, E.L., E. Mark, M.E. Sorrells, J.L. Jannink. 2009. Genomic selection for crop improvement. Crop Sci. 49:1–12.

Hoseney, R.C. 1987. Wheat hardness. Cereal Foods World. 32:320-322.

Huang, S., A. Sirikhachornkit, X. Su, J. Faris, B. Gill, R. Haselkorn, and P. Gornicki. 2002. Genes encoding plastid acetyl-CoA caroxylase and 3-phosphoglycerate kinase of the Triticum/Aegilops complex and the evolutionary history of polyploidy wheat. Proc. Nat. Academy Sci. 99:8133–8138.

Igrejas, G., P. Leroy, G. Charmet, T. Gaborit, D. Marion, and G. Branlard. 2002. Mapping QTLs for grain hardness and puroindoline content in wheat (Triticum aestivum L.). Theor. Appl. Genet. 106:19–27.

Inghelandt, D.V., A.E. Melchinger, C.L. Lebreton, B. Stich. 2010. Population structure and genetic diversity in a commercial maize breeding program assessed with SSR and SNP markers. Theor. Appl. Genet. 120:1289-1299.

Isern, T.D. 2000. Wheat explorer the world over: Mark Carleton of Kansas. Kans. Hist. 23:12–25.

Johnson, V.A., P.J. Mattern, C.J. Peterson, and S.L. Khur. 1985. Improvement of wheat protein by traditional breeding and genetic techniques. Cereal Chem. 62:350-355.

Joppa, L.R., C. Du, G.E. Hart, and G.A. Hareland. 1997. mapping gene(s) for grain protein in tetraploid wheat (triticum turgidum l.) using a population of recombinant inbred chromosome lines. Crop Sci. 37:1586-1589.

Kahrizi, D., M. Maniee, R. Mohammadi, and K. Gheghamirza. 2010. Estimation of genetic parameters related to morpho-agronomic traits of Durum Wheat (Triticum turgidum var. durum) Biharean Biologist. 4:93-97.

Kalous, J.R., J.M. Martin, J.D. Sherman, N.K. Blake, S.P. Lanning, and L.E. Talbert. 2011. Phenotypic variation and patterns of linkage disequilibrium associated with introduced genes in spring wheat. Crop Sci. 51:2466–2478.

Kato, K., H. Yamagata. 1988. Method for evaluation of chilling requirement and narrow-sense earliness of wheat cultivars. Japan J. Breed. 38:172–186.

Keller, M., C. Karutz, J.E. Schmid, P. Stamp, M. Winzeler, B. Keller, M.M. Messmer. 1999. Quantitative trait loci for lodging resistance in a segregating wheat/spelt population. Theor. Appl. Genet. 98:1171–1182.

Khush, G.S. 1999. Green revolution: preparing for the 21st century. Genome, 42:646–655.

Kishore, G., S. Gupta, A. Pandey. 2012. Assessment of population genetic diversity of Fagopyrum tataricum using SSR molecular marker. Biochem. Systematics Ecol. 43: 32-41.

Kitagawa, S., S. Shimada, K. Murai. 2012. Effect of Ppd-1 on the expression of flowering-time genes in vegetative and reproductive growth stages of wheat. Genes. Genet. Syst. 87:161–168.

Knott, D.R. 1986. Effect of genes for photoperiodism, semidwarfism, and awns on agronomic characters in a wheat cross. Crop Sci. 26:1158–1162.

Kollers, S., B. Rodemann, J. Ling, V. Korzun, E. Ebmeyer et al. 2014. Genome-wide association mapping of tan spot resistance (Pyrenophora tritici-repentis) in European winter wheat. Mol. Breeding. 34:363–371.

Korzun, V., M.S. Röder, K. Wendehake, A. Pasqualone, C. Lotti, M.W. Ganal and A. Blanco. 1999. Integration of dinucleotide microsatellites from hexaploid bread wheat into a genetic linkage map of durum wheat. Theor. Appl. Genet. 98:1202–1207.

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

Kumar, A., E.M. Elias, F. Ghavami, X. Xu, S. Jain, F. Manthey, M. Mergoum, M.S. Alamri, P. M.A. Kianian, S.F. Kianian. 2013. A major QTL for gluten strength in durum wheat (Triticum turgidum L. var. durum). J. Cereal Sci. 57:21-29.

Kunert, A., A.A. Naz, O. Dedeck, K. Pillen, and J. Léon. 2007. AB-QTL analysis in winter wheat: I. Synthetic hexaploid wheat (T. turgidum ssp. dicoccoides × T. tauschii) as a source of favourable alleles for milling and baking quality traits. Theor. Appl. Genet. 115:683–695.

Lamkin, W. M., B.S. Miller, S.W. Nelson, D.D.Taylor, and M.S. Lee. 1981. Polyphenol oxidase activities of hard red winter, soft red winter, hard red spring, white common, club, and durum wheat cultivars. Cereal Chem. 58:27-31.

Law, C.N., and A.J. Worland. 1997. Genetic analyses of some flowering time and adaptative traits in wheat. New Phytol. 137:19–28.

Levy, A.A., and M. Feldman. 2002. The impact of polyploidy on grass genome evolution. Plant Physiol. 130:1587–1593.

Li, C., G. Bai, B.F. Carver, S. Chao, Z. Wang. 2016. Mapping quantitative trait loci for plant adaptation and morphology traits in wheat using single nucleotide polymorphisms. Euphytica. 208:299-312.

Li, P., J. Chen, and P. Wu. 2011. Agronomic Characteristics and grain yield of 30 spring wheat genotypes under drought stress and nonstress conditions. Agron. J. 103:1619-1628.

Lin, F., S.L. Xue, Z.Z. Zhang, C.Q. Zhang, Z.X. Kong, G.Q. Yao, D.G. Tian, H.L. Zhu, C.J. Li, Y. Cao. 2006. Mapping QTL associated with resistance to Fusarium head blight in the Nanda2419 × Wangshuibai population. II: Type I resistance. Theor. Appl. Genet. 112:528–535.

Liu, S., M.O. Pumphrey, B.S. Gill, H.N. Trick, J.X. Zhang, J. Dolezel, B. Chalhoub, and J.A. Anderson. 2008. Toward positional cloning of Fhb1, a major QTL for Fusarium head blight resistance in wheat. Cer. Res. Comm. 36:195-201.

Lotti, C., S. Salvi, A. Pasqualone, R. Tuberosa, and A. Blanco. 2000. Integration of AFLP markers into an RFLP-based map of durum wheat. Plant Breed. 119:393–401.

Luo, M.-C., Z.-L.Yang, F.M. You, T. Kawahara, J.G.Waines, , Dvorak, J. 2007. The structure of wild and domesticated emmer wheat populations, gene flow between them, and the site of emmer domestication. Theor. Appl. Genet. 114:947–959.

Ma, Q. 2009. The expression of caffeic acid 3-0-methyltransferase in two wheat genotypes differing in lodging resistance. J. Exp. Bot. 60:2763–2771.

MacArthur, L.A., B.L. D'Appolonia, and O.J. Banasik. 1981. The falling number test – what is it and how does it work? North Dakota Farm Res. 38:15-22.

Marsh, D. R., and T. Galliard. 1986. Measurements of polyphenol oxidase activity in wheat-milling fractions. J. Cereal Sci. 4:241-248.

Marshall, D.R., D.J. Mares, H.J. Moss, and F.W. Ellison. 1986. Effect of grain shape and size on milling yields in wheat. II. Experimental studies. Aust. J. Agric. Res. 37:331-342.

Marshall, D.R., F.W. Ellison, D.J. Mares. 1984. Effects of grain shape and size on milling yield in wheat. I. Theoretical analysis based on simple geometrical models. Aust. J. Agric. Res. 35:619–630.

Marshall, L., R. Busch, F. Cholick, I. Edwards, and R. Frohberg.1989. Agronomic performance of spring wheat isolines differing for daylength response. Crop Sci. 29:752–757.

Martosa, V., C. Royo, Y. Rharrabtia, L.F. Garcia del Moral. 2005. Using AFLPs to determine phylogenetic relationships and genetic erosion in durum wheat cultivars released in Italy and Spain throughout the 20th century. Field Crops Res. 91:107–116.

Marza, F., G.H. Bai, B.F. Carver, W.C. Zhou. 2006. Quantitative trait loci for yield and related traits in the wheat population 'Ning7840' x 'Clark'. Theor. Appl. Genet. 112:688–698.

Matsuo, R.R., and J.E. Dexter. 1980. Relationship between some durum wheat characteristics and semolina milling properties. Can. J. Plant Sci. 60:49-53.

Matsuoka, F. 2011. Evolution of Polyploid Triticum wheats under cultivation: The role of domestication, natural hybridization and allopolyploid speciation in their diversification. Plant Cell Physiol. 52:750–764.

McCartney, C.A., D.J. Somers, D.G. Humphreys, O. Lukow, N. Ames, J. Noll, S. Cloutier, B.D. McCallum. 2005. Mapping quantitative trait loci controlling agronomic traits in the spring wheat cross RL4452 × 'AC Domain'. Genome. 48:870-883.

McIntosh, R.A., Y. Yamazaki, J. Dubcovsky, W.J. Rogers, R. Appels, and X.C. Xia. 2013. Catalogue of gene symbols for wheat. In Proceedings of the 12th International Wheat Genetics Symposium, Yokohama, Japan, 8–13 September 2013.

McNeal, F.H., M.A. Berg, V.R. Stewart, and D.E. Baldridge. 1972. Agronomic response of three height classes of spring wheat, Triticum aestivum L., compared at different yield levels. Agron. J. 64: 362–364.

Merah, O., Deleens, E., Monneveux, P. 1999. Grain yield, carbon isotope discrimination, mineral and silicon content in durum wheat under different precipitation regimes. Physiologia Plantarum. 107:387-394.

Miedaner, T., Y, Zhao, M. Gowda, C.F. Longin, V. Korzun et al. 2013. Genetic architecture of resistance to septoria tritici blotch in European wheat. BMC Genomics. 14:858-864.

Mohler, V., T. Albercht, K. Mrva, G. Schweizer, and L. Hartle. 2014. Genetic analysis of falling number in three bi-parental common winter wheat populations. Plant Breeding. 133:448-453.

Mohsin, T., N. Khan, and F.N. Naqvi. 2009. Heritability, phenotypic correlation and path coefficient studies for some agronomic characters in synthetic elite lines of wheat J. Food Agric. Environ. 7:278–282.

Moldestad, A., E.M. Fergestad, B. Hoel, A.O. Skjelvag, A.K. Uhlen. 2011. Effect of temperature variation during grain filling on wheat gluten resistance. J. Cereal Sci. 53:347-354.

Moore, G., K.M. Devos, Z. Wang, and M.D. Gale. 1995. Grasses, line up and form a circle. Current Biol. 5:737–739.

Musick, J.T., and D.A. Dusek. 1980. Planting date and water deficit effects on development and yield of irrigated winter wheat. Agron. J. 72:45–52.

Myles, S., J. Peiffer, P.J. Brown, E.S. Ersoz, Z. Zhang, D.E. Costich, and E.S. Buckler. 2009. Association Mapping: Critical Considerations Shift from Genotyping to Experimental Design. Plant Cell. 21:2194-2202.

Nachit, M., I. Elouafi, M.A. Pagnotta, A. El Saleh, E. Iacono, M. Labhilili, A. Asbati, M. Azrak, H. Hazzam, D. Benscher, M. Khairallah, J. Ribaut, O.A. Tanzarella, E. Porceddu, and M.E. Sorrells. 2001. Molecular linkage map for an intraspecific recombinant inbred population of durum wheat (Triticum turgidum L. var. durum). Theor. Appl. Genet. 102:177–186.

NCBI- National Center for Biotechnology Information Search database. 2014. http://www.ncbi.nlm.nih.gov. Accesses 02/2014

Nevo, E., and A. Beiles. 1989. Genetic diversity of wild emmer wheat in Israel and Turkey. Theor. Appl. Genet. 77: 421–455.

Nordborg, M., et al. 2005. The pattern of polymorphism in Arabidopsis thaliana. PLoS Biol. 3: e196.

Novaro, P., M.G. D'Egidio, B.M. Mariani, and P. Nardi. 1993. Combined effect of protein content and high-temperature drying systems on pasta cooking quality. Cereal Chem. 70:716-719.

Olmstead, A.L., and P.W. Rhode. 2011. Adapting North American wheat production to climatic challenges, 1839–2009. Proc. Natl. Acad. Sci. U. S. A. 108:480-485.

Oury, F.X., and C, Godin. 2007. Yield and grain protein concentration in bread wheat: how to use the negative relationship between the two characters to identify favourable genotypes? Euphytica. 157:45–57.

Ozkan, H., G. Willcox, A. Graner, F. Salamini, B. Kilian. 2011. Geographic distribution and domestication of wild emmer wheat (Triticum dicoccoides). Genet. Resour. Crop Evol. 58:11–53.

Parish, J.A., and N.J. Halse. 1968. Effects of light, temperature and the rate of desiccation on translucency in wheat grain. Aust. J. Agric. Res.19:365-372.

Parker, G.D., K.J. Chalmers, A.J. Rathjen, and P. Langridge. 1998. Mapping loci associated with flour color in wheat. Theor. Appl. Genetics. 97:238–245.

Parker, G.D., K.J. Chalmers, A.J. Rathjen, and P. Langridge. 1999. Mapping loci associated with milling yield in wheat (Triticum aestivum L.). Mol. Breeding. 5:561-568.

Patil, R.M., S.A. Tamhankar, M.D. Oak, A.L. Raut, B.K. Honrao, V.S. Rao, S.C. Misra. 2013. Mapping of QTL for agronomic traits and kernel characters in durum (Triticum durum Desf.). Euphytica. 190:117-129.

Paulsen, G.M., and J.P. Shroyer. 2008. The early history of wheat improvement in the Great Plains. Agron. J. 100:S70-S78.

Paux, E., S. Faure, F. Choulet, D. Roger, V. Gauthier, J.-P. Martinant, P. Sourdille, F. Balfourier, M.-C. Le Paslier, A. Chauveau, M. Cakir, B. Gandon, and C. Feuillet. 2010. Insertion site-based polymorphism markers open new perspectives for genome saturation and marker-assisted selection in wheat. Plant Biotechnol. J. 8:196-210.

Payne, P.I., E.A. Jackson, L.M. Holt. 1984. The association between g-gliadin 45 and gluten strength in durum wheat varieties: a direct causal effect or the result of genetic linkage? J. Cereal Sci. 2:73–81.

Payne, P.I., G.J. Lawrence. 1983. Catalogue of alleles for the complex loci, Glu-A1, Glu-B1 and Glu-D1 which code for high-molecular-weight subunits of glutenin in hexaploid wheat Cereal Res. Comm. 11:29–35.

Peleg, Z., T. Fahima, A.B. Korol, S. Abbo, and Y. Saranga. 2011. Genetic analysis of wheat domestication and evolution under domestication. J. Exp. Bot. 62:5051-5061.

Peña, R.J., J. Zarco-Hernandez, A. Amaya-Celis, A. Mujeeb-Kazi. 1994. Relationships between chromosome 1B-encoded glutenin subunit compositions and bread-making quality characteristics of some durum wheat (Triticum turgidum) cultivars. J. Cereal Sci. 19:243-249.

Peng, J.H., Y. Ronin, T. Fahima, M.S. Roder, Y.C. Li, E. Nevo, and A. Korol. 2003. Domestication quantitative trait loci in Triticum dicoccoides, the progenitor of wheat. Proc. Natl. Acad. Sci. 100:2489-2494.

Pinthus, M.J. 1973. Lodging in wheat, barley and oats: the phenomenon, its causes and preventative measures. Adv. Agron. 25:209–263.

Posner, E.S. 2009. Wheat flour milling. Pages 119-152 in: Wheat: chemistry and technology. K. Khan, P.R. Shewry, eds. St. Am. Assoc. Cereal Chem.: Paul, MN.

Pozniak, C.J., R.E. Knox, F.R. Clarke, and J.M. Clarke. 2007. Identification of QTL and association of a phytoene synthase gene with endosperm color in durum wheat. Theor. Appl. Genet. 114:525-537.

Prashant, R., N. Kadoo, C. Desale, P. Kore, H. Singh, P. Chhuneja, V. Gupta. 2012. Kernel morphometric traits in hexaploid wheat (Triticum aestivum L.) are modulated by intricate QTL x QTL and genotype x environment interactions. J. Cereal Sci. 56:432–439.

Prasad, M., N. Kumar, P.L. Kulwal, M.S. Roder, H.S. Balyan, H.S. Dhaliwal, J.K. Roy, P.K. Gupta. 2003. QTL analysis for grain protein content using SSR markers and validation studies using NILs in bread wheat. Theor. Appl. Genet. 106:659–667.

Prat, N., C. Guilbert, U. Prah, E. Wachter, B. Steiner, B. Langin, O. Robert, and H. Buerstmayr. 2017. QTL mapping of Fusarium head blight resistance in three related durum wheat populations. Theor. Appl. Genet. 130:13-27.

Raman, R., H. Allen, H. Raman, S. Diffey, P. Martin, K. McKelvie. 2009. Localisation of quantitative trait loci for quality attributes in a doubled haploid population of wheat (Triticum aestivum L.). Genome. 52:701–715.

Rasul, G., D.G. Humphreys, A. Brûlé-Babel, C.A. McCartney, R.E. Knox, R.M. DePauw, and D.J. Somers. 2009. Mapping QTLs for pre-harvest sprouting traits in the spring wheat cross “RL4452/AC Domain.”. Euphytica. 168:363–378.

Reif, J.C., H.P. Maurer, V. Korzun, E. Ebmeyer, T. Miedaner, T. Wurschum. 2011. Mapping QTL with main and epistatic effects underlying grain yield and heading time in soft winter wheat. Theor. Appl. Genet. 123:283–292.

Reimer, S.O., C.J. Pozniak, F.R. Clarke, J.M. Clarke, D.J. Somers, R.E. Knox, A.K. Singh. 2008. Association mapping of yellow pigment in an elite collection of durum wheat cultivars and breeding lines Genome. 51:1016–1025.

Rharrabti, Y., C. Royo, D. Villegas, N. Aparicio, L.F. Garci′a del Moral. 2003. Durum wheat quality in Mediterranean environments I. Quality expression under different zones, latitudes and water regimes across Spain. Field Crop Res. 80:123–131.

Rharrabti, Y., D. Villegas, C. Royo, V. Martos-Nunez, L.F. Garcıa del Moral. 2003. Durum wheat quality in Mediterranean environments II. Influence of climatic variables and relationships between quality parameters. Field Crops Res. 80:133-140.

Rharrabti, Y., L.F. Garcıa del Moral, D. Villegas, and C. Royo. 2003. Durum wheat quality in Mediterranean environments III. Stability and comparative methods in analyzing G x E interaction. Field Crops Res. 80:141-146.

Röder, M.S., V. Korzun, K. Wendehake, J. Plaschke, M.H. Tixier, P. Leroy, and M.W. Ganal. 1998. A microsatellite map of wheat. Genetics. 149:2007–2023.

Roncallo, P.F., G.L. Cervigni, C. Jensen, R. Miranda, A.D. Carrera, M. Helguera, V. Echenique. 2012. QTL analysis of main and epistatic effects for flour colour traits in durum wheat. Euphytica. 185:77–92.

Royo, C., E.M. Elias, F.A. Manthey. 2009. Durum wheat breeding. Pages 199-226 in: Cereals—Handbook of plant breeding. M.J. Carena ed. Springer: Berlin.

Ruan, Y., A. Comeau, F. Langevin, P. Hucl, J. Clarke, A. Brule-Babel, and C. Pozniak. 2012. Identification of novel QTL for resistance to Fusarium head blight in a tetraploid wheat population. Genome. 55:853-864.

Ruiz, M., and J.M. Carrillo. 1995. Separate effects on gluten strength of Gli-1 and Glu-3 prolamin genes on chromosomes 1A and 1B in durum wheat J. Cereal Sci. 21:137–144.

Russo, M. A., D.B.M. Ficco, G. Laidò, D. Marone, R. Papa, A. Blanco. 2014. A dense durum wheat × T. dicoccum linkage map based on SNP markers for the study of seed morphology. Mol. Breed. 34:1579–1597.

Sakamoto, S. 1973. Patterns of phylogenic differentiation in the tribe Triticeae. Seiken Ziho. 24:49-58.

Sarker, Z.I., A.K.M. Shamsuddin, L. Rahman, and R. Ara. 2007. Gene actions of traits contributing to lodging resistance in wheat (Triticum aestivum L.). Bangladesh J. Plant Breed. Genet. 20:23-30.

Schuler, S.F., R.F. Bacon, E.E. Gbur. 1994. Kernel and spike character influence on the test weight of soft red winter wheat. Crop Sci. 34:1309–1313.

Shewry, P.R., A.S. Tatham, F. Barro, P. Barcelo, and P. Lazzeri. 1995. Biotechnology of breadmaking – unraveling and manipulating the multi-protein gluten complex. Bio-Technol. 13:1185–1190.

Si, H.Q., Z.L. Zhou, X.B. Wang, C.X. Ma. 2012. A novel molecular marker for the polyphenol oxidase gene located on chromosome 2B in common wheat. Mol. Breeding. 30:1371–1378.

Singh, A., S. Reimer, C.J. Pozniak, F.R. Clarke, J.M. Clarke, R.E. Knox, and A. Singh. 2009. Allelic variation at Psy1-A1 and association with yellow pigment in durum wheat grain. Theor. Appl. Genet. 118:1539-1548.

Sissons, M. 2008. Role of durum wheat composition on the quality of pasta and bread. Food. 2:75-90.

Sissons, M.J., N.E. Egan, M.C. Gianibelli. 2005. New insights into the role of gluten on durum pasta quality using reconstitution method. Cereal Chem. 82:601–608.

Snape, J.W., K. Butterworth, E. Whitechurch, A.J. Worland. 2001. Waiting for fine times: genetics of flowering time in wheat. Euphytica. 119:185–90.

Soleimani V.D., B.R. Baum, D.A. Johnson. 2002. AFLP and pedigree-based genetic diversity estimates in modern cultivars of durum wheat [Triticum turgidum L. subsp. durum (Desf.) Husn.]. Theor. Appl. Genet. 104:350–357.

Sourdille, P., M.R. Perretant, G. Charmet, P. Leroy, M.F. Gautier, P. Joudrier, J.C. Nelson, M.E. Sorrells, M. Bernard. 1996. Linkage between RFLP markers and genes affecting kernel hardness in wheat. Theor. Appl. Genet. 93:580–586.

Stapper, M., Fischer, R.A. 1990. Genotype, sowing date and plant spacing influence on high-yielding irrigated wheat in Southern New South Wales. II. Growth, yield and nitrogen use. Aust. J. Agric. Res. 41:1021–1041.

Stenvert, N.L. 1972. The measurement of wheat hardness and its effect of milling characteristics. Aust. J. Exp. Agric. Anim. Husb. 12:159–164.

Sun, Y., J. Wang, J.H. Crouch, Y. Xu. 2010. Efficiency of selective genotyping for genetic analysis of complex traits and potential applications in crop improvement. Mol. Breed. 26:493–511.

Suprayogi, Y., C.J. Pozniak, F.R. Clarke, J.M. Clarke, R.E. Knox, A.K. Singh. 2009. Identification and validation of quantitative trait loci for grain protein concentration in adapted Canadian durum wheat populations. Theor. Appl. Genet. 119:437–448.

Taghouti, M., F. Gaboun, N. Nsarellah, R. Rhrib, M. Kamar, F. Abbad-Andaloussi, and S.M. Udupa. 2010. Genotype x Environment interaction for quality traits in durum wheat cultivars adapted to different environments. African J. Biotechnol. 9:3054-3062.

Terasawa ,Y., M. Ito, T. Tabiki, K. Nagasawa, K. Hatta, Z. Nishio. 2016. Mapping of a major QTL associated with protein content on chromosome 2B in hard red winter wheat (Triticum aestivum L.). Breed. Sci. 66:471-480.

Thuillet, A.C., T. Bataillon, S. Poirier, S. Santoni, J.L. David. 2005. Estimation of long-term effective population sizes through the history of durum wheat using microsatellite data. Genetics. 169:1589–1599.

Tommasini, L., T. Schnurbusch, D. Fossati, F. Mascher, B. Keller. 2007. Association mapping of stagonospora nodorum blotch resistance in modern European winter wheat varieties. Theor. Appl. Genet. 115:697–708.

Troccoli, A., Borrelli, G. M., De Vita, P., Fares, C., and Di Fonzo, N. 2000. Durum wheat quality: A multidisciplinary concept. J. Cereal Sci. 32:99-113.

Tsilo, T.J., G.A. Hareland, S. Simsek, S. Chao, J.A. Anderson. 2010. Genome mapping of kernel characteristics in hard red spring wheat breeding lines. Theor. Appl. Genet. 121:717-730.

Tuberosa, R., and S. Salvi. 2006. Genomics-based approaches to improve drought tolerance of crops. Trends in Plant Sci.11:405-412.

Uauy, C., A. Distelfield, T. Fahima, A. Blechl, and J. Dubcovsky. 2006. A NAC gene regulating senescence improves grain protein, zinc and iron content in wheat. Sci. 413:1298–1301.

Uddin, M.S., and A. Boerner. 2008. Genetic Diversity in Hexaploid and Tetraploid Wheat Genotypes Using Microsatellite Markers. Plant Tissue Cult. Biotech. 18:65–73.

van Poecke, R.M.P., M. Maccaferri, J. Tang, H.T. Truong, A. Janssen, N.J. van Orsouw, S. Salvi, M.C. Sanguineti, R. Tuberosa, E.A.G. van der Vossen. 2013. Sequence-based SNP genotyping in durum wheat. Plant Biotechnol. J. 11:809–817.

Waddington, S.R., J.K. Ransom, M. Osmanzai, and D.A. Saundres, 1986. Improvement in the yield potential of bread wheat adapted to northwest Mexico. Crop Sci. 26:698-703.

Wang, H., K.P. Smith, E. Combs, T. Blake, R.D. Horsley, G.J. Muelbauer. 2012. Effect of population size and unbalanced data sets on QTL detection using genome-wide association mapping in barley breeding germplasm. Theor. Appl. Genet.124:111-124.

Wang, S., D. Wong, K. Forrest, A. Allen, S. Chao, B.E. Huang, M. Maccaferri et al. 2014. Characterization of polyploid wheat genomic diversity using a high-density 90 00 single nucleotide polymorphism array. Plant Biotechnol. J. 12:787–796.

Wang, Y., C. Hao, J. Zheng, H. Ge, Y. Zhou, Z. Ma, X. Zhang. 2014. A haplotype block associated with thousand-kernel weight on chromosome 5DS in common wheat (Triticum aestivum L.). J. Integr. Plant Biol. 57:662–672.

Wardlaw, I. 1990. The control of carbon partitioning in plants. New Phytol. 116:341-381.

Wilhelm, E.P., M.I. Boulton, T.E.S. Barber, A.J. Greenland, W. Powell. 2013. Genotype analysis of the wheat semi-dwarf Rht-B1b and Rht-D1b ancestral lineage. Plant Breed. 132:539–545.

Worland, A.J., V. Korzun, M.S. Röder, M.W.Ganal, and C.N. Law. 1998. Genetic analysis of the dwarfing gene Rht8 in wheat. Part II. The distribution and adaptive significance of allelic variants at the Rht8 locus of wheat as revealed by microsatellite screening. Theor. Appl. Genet. 96:1110–1120.

Worland, T., and J.W. Snape. 2001. Genetic basis of worldwide wheat varietal improvement. Pages 59-100 in: The World Wheat Book, a History of Wheat Breeding. A.P. Bonjean and W.J. Angus, eds. Lavoisier Publishing, Paris.

Xiao, S.H., Z.H. He. 2003. Wheat yield and end use quality improvement in China. Pages 497-542 in: Chinese Wheat Improvement and Pedigree Analysis. Q.S. Zhuang, eds. China Agricultural Press, Beijing.

Xu, Y., and J.H. Crouch. 2008. Marker-assisted selection in plant breeding: From publications to practice. Crop Sci. 48:391–407.

Yamazaki, W.T., and L.W. Briggle. 1969. Components of test weight in wheat. Crop Sci. 9:331-342.

Yildirim, A., O. Ates Sonmezoglu, S. Gokmen. 2011. Determination of genetic diversity among Turkish durum wheat landraces by microsatellites. Afr. J. Biotechnol. 10:3915–3920.

Yoon, B.S., Brorsen, B.W., Lyford, C.P. 2002. Value of increasing kernel uniformity. J. Agric. Resour. Econ. 27:481–494.

Zeller, K.A., R.L. Bowden, and J.F. Leslie. 2004. Population differentiation and recombination in wheat scab populations of Gibberella zeae from the United States. Mol. Ecol. 13:563-557.

Zhai, S., Z. He, W. Wen et al. 2016. Genome-wide linkage mapping of flour color-related traits and polyphenol oxidase activity in common wheat. Theor. Appl. Genet.129:377-394.

Zhang, J., J. Chen, B.C. Bowman. 2014. Association mapping of Hagberg falling number in hard white spring wheat. Crop Sci. 54:1243–1252.

Zhang, L.P., X.X. Ge, Z.H. He, D.S. Wang, J. Yan, X.C. Xia, M.W. Sutherland. 2005. Mapping QTLs for polyphenol oxidase activity in a DH population from common wheat. Acta Agron. Sin. 31:7–10.

Zhang, W., S. Chao, F. Manthey, O. Chicaiza, J.C. Brevis, W. Echenique, J. Dubcovsky. 2008. QTL analysis of pasta quality using a composite microsatellite and SNP map of durum wheat. Theor. Applied Genetics. 117:1361–1377.

Zohary, D., and M. Hopf. 2000. Domestication of plants in the old world. Oxford University Press, Oxford.

Zondervan, K.T., and L.R. Cardon. 2004. The complex interplay among factors that influence allelic association. Nat. Rev. Genet. 5:89-100.

CHAPTER 2. GENETIC DIVERSITY OF DURUM

Abstract

Genetic diversity studies in breeding programs are important to identify parental lines for hybridization and introgression of desirable alleles into elite germplasm. Single nucleotide polymorphism (SNP) markers are considered effective in detecting genetic diversity in crop plants. In the present study, genetic diversity and genetic structure were investigated using 283 genotypes from the durum wheat breeding program at North Dakota State University (NDSU), using Illumina's Infinium 90K SNP assay. The genetic diversity analysis indicated that the population was structured according to its breeding history. The four major principal components (PCs) accounting for most of the genetic variation were identified with 3,110 SNP markers. Genetic analysis using 1,308 SNP markers with STRUCTURE also divided the genotypes into four major clusters. The maximum-likelihood tree analysis placed the genotypes into clusters and sub-clusters similar to those observed with the PC and STRUCTURE analyses. Genetic diversity analysis (H_T = 0.334) showed adequate levels of genetic variation. Genetic variation within clusters (H_S = 0.249) was shown to be greater than among clusters (D_{ST} = 0.085). The results from this study could be useful for durum wheat breeding programs to assess the genetic diversity of their breeding materials. This will help in breeding efforts to broaden the genetic base and select lines for crossing as well as for genetic and genomic studies to facilitate the combination of desirable alleles.

Introduction

Wheat breeding programs around the world are working to improve grain yield as well as quality, disease resistance, and agronomic performance. Knowledge of the genetic diversity within a germplasm collection can aid the selection of parents for crossing, for establishing

heterotic groups and can have a significant impact on the improvement of crops (Abdellatifa and AbouZeib, 2011). Assessing the extent and nature of genetic variation in durum wheat (*Triticum turgidum* L. var. *durum* Desf.) is also important to breeding and genetic resource conservation programs.

Molecular markers are widely used in wheat (*Triticum aestivum* L.) breeding programs to assess the genetic diversity available to the breeder (Barbosa-Neto et al., 1996; Das and Misra, 2010; Fu et al., 2006). Single nucleotide polymorphisms (SNPs) are the most abundant type of molecular markers, located every 100-300 bp in plants (Edwards et al., 2007). Recent developments in next-generation sequencing technology have significantly facilitated the identification of SNPs in wheat. Single nucleotide polymorphism markers are especially effective in detecting genetic diversity because of the large number of loci that are tagged (Ren et al., 2013). Molecular-marker profiles account for the effects from selection and genetic drift that occur over breeding cycles, making it possible to portray the relationship among genotypes at a very accurate level (Varshney et al., 2005).

North American hexaploid and tetraploid wheat breeding germplasm was developed from a relatively narrow genetic background (Royo et al., 2009; Charmet, 2011). Due to modern breeding effects based on the strict industrial quality standards for durum wheat, genetic diversity in wheat has been reduced (DeVita et al., 2007; Maccaferri et al., 2005). Haudry et al. (2007) reported that durum wheat elite germplasm has the lowest diversity among cultivated wheat ($\pi = 0.0004 \times 10^{-3}$). This narrow genetic diversity is a problem in breeding for adaptation to biotic stresses, such as diseases, and abiotic stresses, such as drought or salt tolerance (Autrique et al., 1996).

Over 67% of durum wheat hectares in the United States are located in North Dakota, and about 32% in Montana. Over 80% of hectares in North Dakota, Montana, and South Dakota are planted with durum wheat cultivars developed at NDSU. Therefore, assessing the genetic diversity and genetic structure in the current breeding program at NDSU will aid in better understanding the breeding history and dynamics within the breeding program. The information gained will also help to determine new selection strategies for germplasm/cultivar development.

Materials and Methods

Plant material

The present study used a collection of 283 NDSU durum wheat cultivars and inbred lines ($F_{5:9}$) tested in the Uniform Regional Durum Nursery (URDN) from 1997 to 2014. All advanced ($F_{4:6}$ and higher generations) genotypes in the breeding program were evaluated annually at single or multiple locations in various types of trials. Thirty-two of the most advanced elite experimental lines and cultivars were selected and extensively evaluated annually in URDN trials at multiple locations for agronomic, quality, and disease-resistance traits to select the most promising lines as future cultivars or parental lines for crossing.

Genomic DNA extraction and genotyping

Four seeds from each genotype were planted into potting mix in the greenhouse in the fall of 2014. Three young leaves from each genotype were harvested and sent to the USDA-ARS Cereal Crops Genotyping Laboratory in Fargo, North Dakota, for DNA isolation and genotyping. The extracted DNA samples were genotyped with the Illumina iSelect BeadChip platform containing 90,000 SNPs using the Illumina Infinium Assay protocol (Wang et al., 2014). The genotypes were called using the diploid version of Genome Studio software (Wang et al., 2014). The default settings of FastPHASE 1.3 software were used to impute missing loci for the dataset

using the "likelihood"-based imputation (Scheet and Stephens, 2006). The heterozygotes were considered missing data. A total of 8,082 SNP markers were polymorphic.

Data analysis

The population structure was determined using a models-based approach as implemented in the software STRUCTURE (Pritchard et al., 2000) with burnin/iteration combinations (50,000/20,000) and population (K) size 2 to 10 after MAF ≥ 0.5 and LD = 0.01 in an admixture model. Out of a total of 8,082 SNP markers, the total number of SNP markers used was 1,308. To ensure consistent results, 10 independent runs were conducted. The delta k approach of Evanno et al. (2005) was used to determine the uppermost hierarchical level of structure. F_{ST}, a measure of population differentiation, was calculated in STRUCTURE using the estimated sub-populations. The population structure was also determined using principal component analysis (PCA) with JMP Genomics 6.0 Software for all 8,082 SNP markers with MAF ≥ 5%, resulting in 3,110 SNP markers. After MAF ≥ 0.5 and LD = 0.01 a total of 1,308 markers were used for cluster analysis. Cluster analysis was conducted via the maximum-likelihood tree method using SNPhylo software followed by bootstrap analysis with 1,000 permutations. The tree was displayed with FigTree v1.4.3 software (Fig. 2-1).

Genetic diversity was estimated with the total diversity (H_T) using POGENE version 1.32 averaging two iteration of 500 random SNP markers (Nei, 1973; Yeh and Bolyle, 1997). The coefficient of genetic differentiation [the proportion of total variation distributed between clusters (G_{ST})] was calculated as D_{ST}/H_T, where D_{ST} is the genetic diversity between clusters calculated as H_T-H_S, where H_S is the mean genetic diversity within clusters. Genetic distances between clusters were calculated according to Nei's genetic distance (Nei, 1972).

Results and Discussion

During the domestication period, about 84% of the diversity in durum wheat was lost (Baloch et al., 2017). To meet the needs of farmers, millers, producers, consumers, and exporters, very strict requirements for durum wheat cultivars have been applied, leading to a narrow germplasm with these fixed desirable characteristics. Breeder-adapted germplasm collections generally are composed of high-yielding, disease-resistant, and high-quality cultivars and advanced lines that are intercrossed in the search for new gene combinations, leading to improved durum wheat cultivars. However, narrow genetic diversity in the elite germplasm may jeopardize the continued ability to improve crops.

Knowledge of genetic diversity and population structure is important for understanding the extent of genetic variability in plant breeding programs. Comprehensive knowledge about genetic diversity within breeding programs can remarkably influence the maintenance and usage of durum germplasm and facilitate breeding methods and diversification approaches. The genetic diversity can now easily be studied using molecular markers. Third-generation markers provide high resolution, high frequency, cost efficiency, and are a robust alternative to other types of markers.

In this study, a summary of genetic diversity within the NDSU durum wheat breeding program was obtained using a panel of 283 cultivars and advanced breeding lines. Cluster analysis using the maximum-likelihood method divided the 283 genotypes into two clusters. The small cluster (I – worldwide parent collection) consisted of the European and Tunisian lines, while the large cluster (II – core collection) consisted of the rest of the cultivars and genotypes (Fig. 2-1). Cluster II was further divided into six smaller sub-clusters based on the shared nodes supported by bootstrap values (Fig. 2-1).

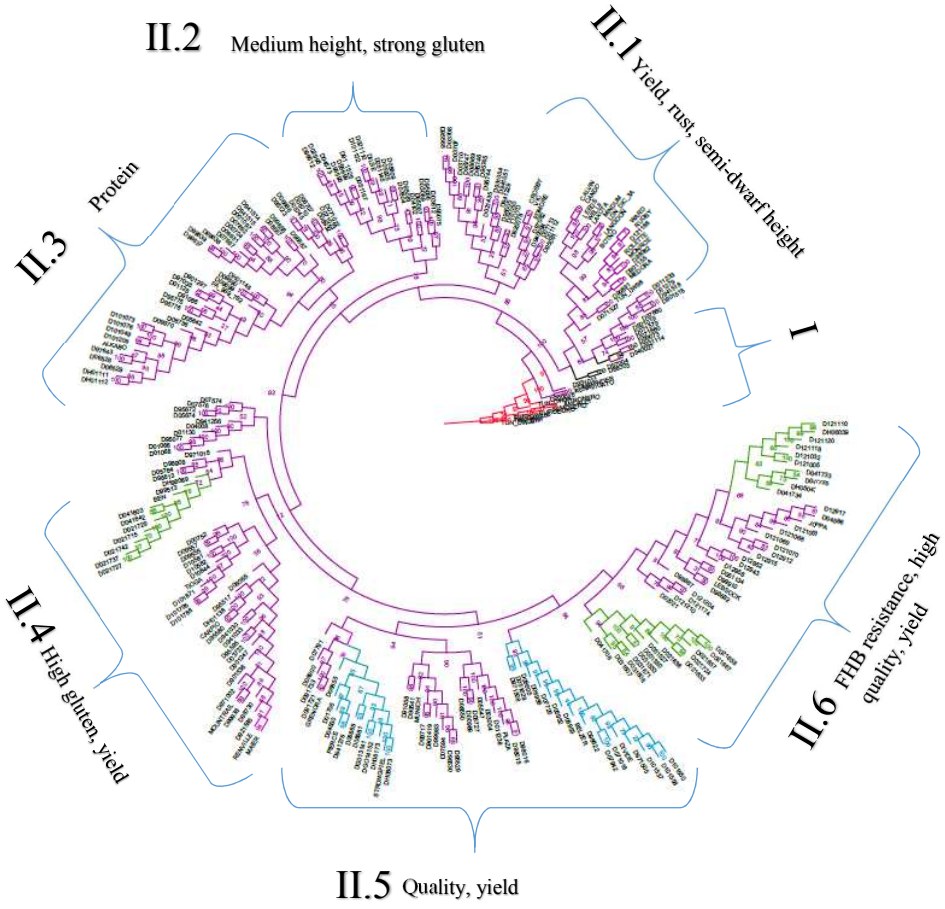


Figure 2-1. Rooted maximum-likelihood tree. Two clusters were identified as I and II. Six sub-clusters were identified in cluster II, designated as II.1, II.2, II.3, II.4, II.5, and II.6, according to the breeding purpose/phenotypic performance. Red, purple, green, and blue colors correspond to cluster colors identified with PCA and STRUCTURE analysis.

Cultivars developed and released from 1956 to 1980 tend to cluster together, composing sub cluster II.1, due to their narrow genetic background (Fig. 2-1). Sub-cluster II.1 includes 'Botno' (1973), 'Rugby' (1975), and 'Ward' (1974) developed from multiple crosses, all of which included 'Langdon' and 'Wells' as parents (Quick et al., 1974; Quick et al., 1974-2; Quick et al., 1975). These cultivars are in a cluster of genotypes in the likelihood tree that includes 'Mindum', which was one of the parents of 'Langdon' and a leading cultivar from 1920 to 1940 (Sibbit and Harris, 1948). Until about 1956, the main breeding goals were to develop durum

cultivars resistant to stem rust. Starting in 1956, the semi-dwarf genes were introduced into North America, resulting in 'Cando' (1975), 'Calvin' (1978), 'Lloyd' (1983), and 'Plaza' (1999) (Quick et al., 1976; Quick et al., 1979; Cantrell et al., 1984; Elias et al., 2001). Semi-dwarf genotypes were extensively intercrossed with existing germplasm, resulting in clusters that were not clearly defined. The yields of semi-dwarf cultivars were not as high as expected, so in the 1970s several high-yielding, medium-height cultivars were released, including 'Rollette' (1971), 'Ward' (1972), 'Rugby' (1973), 'Crosby' (1973), and 'Botno' (1973) (Quick et al., 1974a, 1974b, 1975). Concurrently, a shift to develop stronger gluten cultivars started in North America. One of the Italian strong-gluten cultivars, 'Cappelli', was used to introduce genes to the North American genepool (Quick et al., 1980). 'Edmore' (1976) was the first strong-gluten cultivar released by NDSU, followed by 'Vic' (1979), 'Monroe' (1985), and 'Renville' (1988) (Fig. 2-1). Sub-cluster II.2 represents the beginning of the development of medium-height lines with strong gluten. Bootstrapping provides assessments of 'confidence" for each cluster and sub-cluster observed in a likelihood tree. High values (expressed in %) of bootstraps among clusters in the maximum-likelihood tree determine clear differentiation between clusters and sub-clusters. For example, the branching between the large cluster (II-core collection) and small cluster (I – worldwide parents collection), which consists of the European and Tunisian lines, has 100% bootstrap support; in other words, the pattern was observed 100% of the time (Fig. 2-1). Bootstrap values among sub-clusters are high as well, and differentiation of the genotypes among them is determined (Fig. 2-1).

Although some germplasm was introduced from CYMMIT and Europe, most of the cultivars that have been developed by the NDSU breeding program have over 95% North American germplasm in their background (personal communication with Dr. Elias).

Four clusters were defined in PCA (Fig. 2-2).

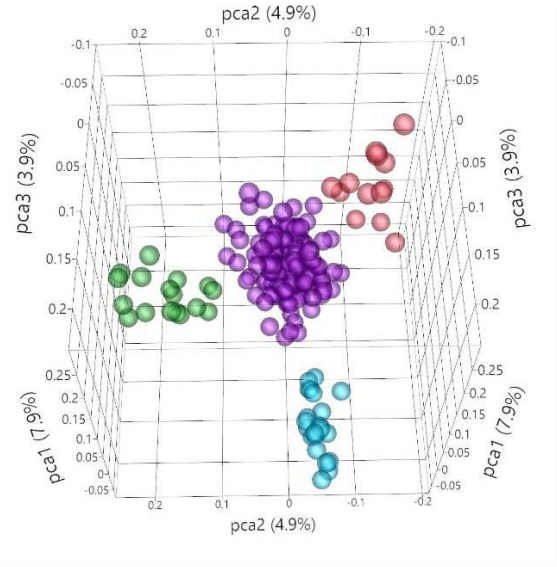


Figure 2-2. Estimate of the number of clusters according to PCA. The red cluster corresponds to cluster I in the maximum-likelihood tree and in STRUCTURE. The green cluster corresponds to the green sub-cluster in the maximum-likelihood tree and in STRUCTURE. The blue cluster corresponds to the blue sub-cluster in the maximum-likelihood tree and in STRUCTURE.

Four clusters were used from STRUCTURE at K=4 as due to the observed distribution of the clusters with a minimum admixture in each. In both the STRUCTURE analysis and PCA, one unique cluster included the lines that were developed and evaluated in 2002, 2003, and 2004, where Sumai-3, a spring wheat cultivar known for its resistance to Fusarium head blight (FHB) caused by *Fusarium graminearum* Schwabe (telomorph *Gibberella zea* (Schwein.) Petch., was used (the green color in the STRUCTURE, PCA, and likelihood-tree analyses) (Figs. 2-2 and 2-3). Another cluster evident in both the STRUCTURE analysis and PCA includes genotypes that have 'Belzer', 'Divide', and 'Strongfield' cultivars in their pedigrees (the blue color in the STRUCTURE, PCA, and likelihood-tree analyses). 'Belzer', 'Strongfield' and 'Divide' have very high-quality characteristics. Therefore, progeny from the multiple crosses with those cultivars form a unique cluster (Figs. 2-2 and 2-3). The small cluster (the red color in the

STRUCTURE, PCA, and likelihood-tree analyses) in the STRUCTURE analysis and PCA consisted of the European and Tunisian lines. The largest cluster was purple in the STRUCTURE analysis and PCA and included the core collection and a large admixture due to extensive interbreeding within the program (the purple color in the STRUCTURE, PCA, and likelihood-tree analyses).

The four durum wheat clusters showed a total genetic diversity (H_T) ranging from 0.156 (the blue cluster, where Sumai-3 was introduced for FHB resistance) to 0.348 (the red cluster that included the European and Tunisian lines) (Table 2-1). The relatively low value for genetic diversity among all genotypes ($H_T = 0.334$) and the low value for genetic diversity among the five clusters ($D_{ST} = 0.085$) resulted in a genetic differentiation value (G_{ST}) of 0.254, indicating that genetic variation was relatively low between clusters (25% of the variability), while most diversity lies within the clusters (75%). Nei's genetic distance indicates a larger distance between cluster II (red) and the other clusters, which corresponds to that found in the maximum-likelihood tree (Table 2-2).

Table 2-1. Genetic diversity of the four clusters defined with STRUCTURE analysis and PCA

	Genotypes	H_T	H_S	D_{ST}	G_{ST}
Total	283	0.334	0.249	0.085	0.254
Cluster I (red)	11	0.348			
Cluster II (green)	23	0.156			
Cluster III (blue)	25	0.247			
Cluster IV (purple)	224	0.246			

Total genetic diversity (H_T), genetic diversity within clusters (H_S), genetic diversity between clusters (D_{ST}), and coefficient of genetic differentiation (G_{ST}) calculated from SNP markers data according to the STRUCTURE analysis (A-1 Table).

Table 2-2. Genetic distances (Nei, 1972) between the clusters defined by STRUCTURE analysis

Clusters	I (red)	II (green)	III (blue)
II (green)	0.284	***	
III (blue)	0.228	0.119	***
IV (purple)	0.206	0.0.54	0.060

'Cluster II' (red color) represents the worldwide parents collection; 'cluster IV' (green color) represents crosses where Sumai-3 was used; and 'Cluster III' (blue color) represents the collection of crosses involving 'Divide', 'Belzer', and 'Strongfield' for quality purposes.

Therefore, due to the narrow genetic base, the structure defined using the maximum-likelihood tree is not clear-cut, except for a differentiation between the North American core collection germplasm and the Tunisian/European lines (the worldwide parents collection), which clustered together.

The two major clusters defined with the maximum-likelihood tree are in accordance with the delta K obtained using STRUCTURE analysis, which shows that the most divergent STUCTURE groups are defined at k=2 (Evanno et al., 2005). As more subpopulations (K) are applied to the STRUCUTRE analysis, the more clusters coincide with clusters observed in the maximum-likelihood tree, where clustering is based on a close relationship of the genotypes. However, each population defined with STRUCTURE includes a large percentage of admixture, suggesting the interrelatedness of the breeding material within the program.

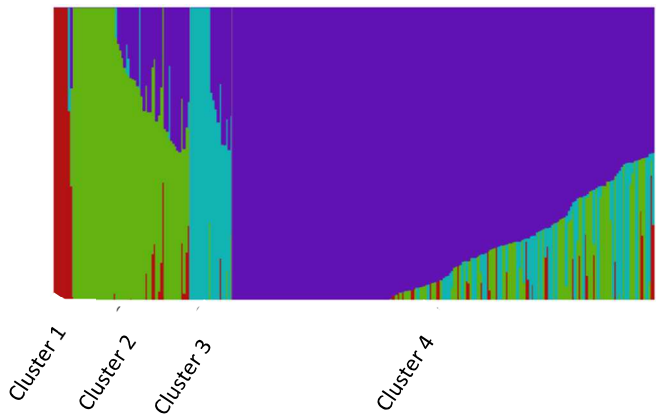


Figure 2-3. Analysis of the genetic structure of the germplasm collection. Each individual is represented by a color bar, with the length proportional to the estimated membership in each of the four clusters. 'Cluster 1' (red color) represents the worldwide parents collection; 'Cluster 2' (green color) represents crosses where Sumai-3 was used; 'Cluster 3' (blue color) represents the collection of crosses involving 'Divide', 'Belzer', and 'Strongfield' for quality purposes; and 'Cluster 4' (purple color) represents the core durum wheat germplasm collection.

The durum breeding program at NDSU is in a unique position in that it develops most of the cultivars grown in the Midwest, accounting for the majority of durum wheat production in the United States. To date, most of the studies on genetic diversity in durum wheat and other crops were conducted with landraces and small samples of modern durum wheat cultivars from different breeding programs (Ruiz et al., 2012; Laido et al., 2013; Maccaferri et al., 2003; Soriano et al., 2016). Most of these studies used AFLP, SSR, and DArT markers to conduct their research. The value estimated for the genetic diversity of the NDSU breeding collection (0.334) was lower than the values reported by previous studies involving durum wheat landrace and modern cultivar collections, which had mean values between 0.48 to 0.68 (Teklu et al., 2006; Altintas et al., 2007; Henkrar et al., 2016; Soriano et al 2016). This may be due to the size of and differences within the collections as well as the marker types used. For example, Soriano et al. (2016) reported genetic diversity among a set of 30 modern cultivars from Spain, Italy, France, and the United States to be 0.48 using SSR markers. Similar to this study's results, Ruis et al. (2012) reported a genetic diversity of 0.33 using SSR and DArT markers. These findings

suggest that within specific breeding programs, genetic diversity is narrower than within the larger landraces collection due to the use of adapted genetic material for the region of breeding.

Conclusion

The current study aimed to understand the genetic structure of a collection of durum wheat genotypes and the relationship between them. Methodologies used to determine the structure proved useful for the purposes of the study. Based on the data from 3,110 SNP markers, PCA assigned all genotypes into four clusters determined by genetic relatedness. Similar clustering was observed with STRUCTURE analysis at K=4 using 1,308 SNP markers. In both analysis, four clusters were identified and defined as: (1) the core collection, (2) the worldwide parents collection, (3) sumai-3 crosses, and (4) quality clusters. The rooted maximum-likelihood tree based on 1,308 SNP markers coincided in essence with the clusters obtained by PCA and STRUCTURE analysis. The maximum-likelihood tree and STRUCTURE software were helpful in differentiating genotypes based on the release period of varieties and breeding purposes. Genetic diversity analysis ($H_T = 0.334$) indicates the availability of adequate genetic diversity in the durum wheat breeding collection despite rigorous selection pressure to meet strict industry and market requirements. Twenty-five percent of differentiation among clusters exists (G_{ST}). These results can be helpful in accelerating wheat improvements by addressing the patterns of genetic variation within durum wheat so that breeders can maximize the level of variation present among clusters and sub-clusters by crossing cultivars with greater genetic distance. To increase genetic diversity within breeding programs, new germplasm from outside the region, such as from breeding programs in Italy, Spain, France, and North Africa, can be considered for crossing. It is always important, however, to maintain the beneficial quality alleles the breeding program built over the years.

References

Abdellatifa, K.F., and H.M. AbouZeib. 2011. Assessment of genetic diversity of Mediterranean bread wheat using randomly amplified polymorphic DNA (RAPD) markers. J. Gen. Eng. Biotechnol. 9:157-163.

Altıntaş, S., F. Toklu, S. Kafkas, B. Kilian, A. Brandolini, H. ÖZkan. 2007. Estimating Genetic Diversity in Durum and Bread Wheat Cultivars from Turkey using AFLP and SAMPL Markers. Plant Breeding. 127:9-14.

Autrique, E., M.M. Nachit, P. Monneveux, S.D. Tanksley, and M.E. Sorrells.1996. Genetic diversity in durum wheat based on RFLPs, morphophysiological traits, and coefficient of parentage. Crop Sci. 36:735-742.

Barbosa-Neto, J.F., M.E. Sorrells, and G. Cisar. 1996. Prediction of heterosis in wheat using coefficient of parentage and RFLP-based estimates of genetic relationship. Genome. 39:1142-1149.

Cantrell, R.G., J.S. Quick, J.W. Dick, and J.D. Miller. 1984. Registration of Lloyd durum wheat. Crop Sci. 24:386.

Cantrell, R.G., J.W. Dick, and J.D. Miller. 1989. Registration of 'Renville' durum wheat. Crop Sci. 29:1329-1330.

Cantrell, R.G., J.W. Dick, J.D. Miller, and J.S. Quick. 1986. Registration of Monroe durum wheat. Crop Sci. 26:200-201.

Charmet, G. 2011. Wheat domestication: lessons for the future. C. R. Biol. 334:212–220.

Das, S., and R.C. Misra. 2010. Assessment of genetic diversity among finger millet genotypes using Rapd markers. Indian J. Agric. Res. 44:112-118.

De Vita, P., O. Li Destri Nicosia, F. Nigro, C. Platani, C. Riefolo, N. Di Fonzo, L. Cattivelli. 2007. Breeding progress in morpho-physiological, agronomical and qualitative traits of durum wheat cultivars released in Italy during the 20th century. Eur. J. Agron. 26:39–53.

Edwards, D., J.W. Forster, D. Chagne, and J. Batley. 2007. What is SNPs? Pages 41-52 in: Association Mapping in Plants. N.C. Oraguzie, E.H.A. Rikkerink, S.E. Gardiner, and H.N. de Silva, eds. Springer: Berlin, Germany.

Elias, E.M., J.D. Miller, and F.A. Manthey. 2001. Registration of 'Plaza' Durum Wheat. Crop Sci. 41:2008.

Evanno, G., S. Regnaut, J. Goudet. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. Mol. Ecol. 14:2611–2620.

Fu, Y.B., G.W. Peterson, J.K. Yu, L. Gao, J. Jia, K.W. Richards. 2006. Impact of plant breeding on genetic diversity of the Canadian hard red spring wheat germplasm as revealed by EST-derived SSr markers. Theor. Appl. Genet. 112:1239-1247.

Haudry, A., A. Cenci, C. Ravel, T. Bataillon, D. Brunel, C. Poncet, I. Hochu, S. Poirier, S. Santoni, S. Glémin, J. David. 2007. Grinding up wheat: a massive loss of nucleotide diversity since domestication. Mol. Biol. Evol. 24:1506–1517.

Henkrar, F., E.H. Jamal, O. Hassan, N. Nsarellah, D. Iraqi, N. Bendaou, and S.N. Udupa. 2016. Genetic diversity reduction in improved durum wheat cultivars of Morocco as revealed by microsatellite markers. Genetics and Plant Breeding. 73:134-141.

Laidò, G., G. Mangini, F. Taranto, A. Galadeta, A. Blanco, L. Cattivelli et al. 2013. Genetic diversity and population structure of tetraploid wheats (Triticum turgidum L.) estimated by SSR, DarT and pedigree data. PloS ONE. 8:e67280.

Maccaferri, M., M.C. Sanguineti, E. Noli, R. Tuberosa. 2005. Population structure and long-range linkage disequilibrium in a durum wheat elite collection. Mol. Breeding. 15:271–290.

Maccaferri, M., M.C. Sanguineti, P. Donini, R. Tuberosa. 2003 Microsatellite analysis reveals a progressive widening of the genetic basis in the elite durum wheat germplasm. Theor. Appl. Genet. 107:783–797.

Nei, M. 1972. Genetic distance between populations. Am. Naturalist. 106:283–292.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. 70: 3321–3323.

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

Quick, J.S., B.J. Donnelly, J.D. Miller. 1980. Registration of Edmore durum wheat. Crop Sci. 20:827.

Quick, J.S., B.J. Doonelly, J.D. Miller. 1980. Registration of 'Vic' Durum Wheat. Crop Sci. 20:827.

Quick, J.S., D.E. Walsh, K.L. Lebsock, J.D. Miller. 1974a. Registration of Ward durum wheat. Crop Sci. 14:607-608.

Quick, J.S., D.E. Walsh, K.L. Lebsock, J.D. Miller. 1974b. Registration of Botno durum wheat. Crop Sci. 14:607-608.

Quick, J.S., D.E. Walsh, K.L. Lebsock, J.D. Miller. 1975. Registration of Rugby durum wheat. Crop Sci. 15:604.

Quick, J.S., J.D. Miller, B.J. Donnelly. 1976. Cando, North Dakota's first semidwarf durum. ND Farm Res. 33:15-18.

Ren, J., D. Sun, L. Chen, F.M. You, J. Wang, Y. Peng, E. Nevo, D. Sun, M.C. Luo, J. Peng. 2013. Genetic diversity revealed by single nucleotide polymorphism markers in a worldwide germplasm collection of durum wheat. Int. J. Mol. Sci. 14:7061-7088.

Royo, C., E.M. Elias, F.A. Manthey. 2009. Durum wheat breeding. Pages 199-226 in: Cereals—Handbook of plant breeding. M.J. Carena ed. Springer: Berlin, Germany.

Ruiz, M., P. Giraldo, C. Royo, D. Villegas, M.J. Aranzana et al. 2012. Diversity and genetic structure of a collection of Spanish durum wheat landraces. Crop Sci 52:2262–2275.

Sibbitt, L.D., and R.H. Harris. 1948. New durum wheats are high in macaroni quality. Bimonthly bulletin, XI. 2:45-48

Varshney, R.K., A. Graner, M.E. Sorrells. 2005. Genic microsatellite markers in plants: features and applications. TRENDS in Biotechnol. 23:48-55.

Yeh, F.C. and T.J.B. Boyle. 1997. Population genetic analysis of co-dominant and dominant markers and quantitative traits. Belgian J. Bot. 129:157.

CHAPTER 3. ASSOCIATION MAPPING FOR AGRONOMIC AND QUALITY TRAITS

Abstract

Developing new durum wheat [(*Triticum turgidum* L. var. *durum* (Desf.)] cultivars with superior agronomic and quality characteristics is a key objective of the North Dakota State University (NDSU) durum wheat breeding program. The quantitative nature of important target traits, combined with environmental effects, makes it difficult to bring the desirable improvement in durum wheat to meet the expectations of all the stakeholders involved in the durum wheat industry. Genetic studies of important durum traits can facilitate cultivar development using genomic tools. With an objective to identify molecular markers for marker-assisted breeding (MAB), the present study attempted to identify marker-trait associations for six agronomic and 29 quality traits using a genome-wide association study (GWAS) mapping approach. The association mapping panel consisted of 243 cultivars and advanced durum wheat breeding lines developed during the last 20 years. The lines were genotyped using Illumina's 90K Infinium SNP assay; a total of 4,196 polymorphic markers were used for GWAS. The study used two types of phenotypic datasets, a historic unbalanced dataset belonging to a total of 80 environments collected over a period of 16 years and a balanced dataset collected from two environments, to identify the applicability of historic unbalanced phenotypic data for GWAS analysis. A total of 292 QTL were identified for agronomic and quality traits, with 10 QTL showing major effects (R^2 >15%). Over 45% of QTL for agronomic traits were present in both the unbalanced and balanced datasets, with about 50% of those present in both environments in the balanced dataset. Similarly, for quality traits measured in both datasets, more than 50% of the QTL were identified in both datasets. QTL identified in only one of the datasets could be explained by low heritable traits, which are largely influenced by the environment. Genome-

wide association mapping studies identified several candidate markers for use in marker-assisted selection (MAS) for plant height, gluten strength, distribution of small kernels, polyphenol oxidase (PPO) activity, and grain yield. The molecular markers associated with the important traits could be extremely useful in expediting the process of developing improved durum wheat cultivars.

Introduction

Durum wheat is one of the most important crops in the world. Annual worldwide durum wheat production is estimated to be around 36 million tons, with approximately, 2.5 million tons produced in the United States (NASS, 2016). North Dakota is the largest durum-producing state in the country. It accounts for more than 50% of the total U.S. production, which is worth more than $300 million per year (NASS 2016). The market price of durum wheat is generally higher compared to other wheat classes, which makes durum wheat attractive for growers; however, a number of abiotic and biotic constraints present a challenge for durum production. Apart from grain yield and disease resistance, quality is a strong criterion of durum wheat variety selection. Grain quality parameters start with visual appearance, test weight (TWT), 1000-kernel weight (TKW), vitreousness (VIT), weather damage (falling number), and protein quantity and quality. Durum wheat is the hardest of all wheats (Miller et al., 1982). During milling, durum wheat endosperm is ground into coarse particles called semolina, which is ideal for making pasta and couscous. Some of the important semolina quality parameters include ash content, color, and particle-size distribution. One of the most important dough quality parameters is gluten strength. Gluten strength can affect such pasta characteristics as tolerance to overcooking, reduced stickiness, and minimal loss of solids during cooking.

Phenotypic evaluation for all important quality traits in not feasible in early generations due to the large number of lines and limited resources in most breeding programs. Moreover, the availability of grains per line is limited in early generations. Marker-assisted selection provides an opportunity for breeders to identify superior lines in early generations, thus saving significant resources and speeding up the process of cultivar development. Marker-assisted selection can benefit traditional breeding because: 1) it can be applied earlier in the life cycle of the plant, 2) it can be cost effective compared with conventional trait evaluation procedures, and 3) it can aid in selecting genotypes for traits that are difficult to select based on phenotype. However, the first requirement for MAS is to know the genetics of the target trait and identify associated markers. Moreover, the benefits of MAS can be best realized when the markers are tightly linked to major QTL that explain high phenotypic variation for the trait under consideration (Nakaya and Isobe, 2012).

Many important traits are influenced by multiple QTL, their interaction, the environment, and the interaction between QTL and environment. For dissecting these complex traits, two genetic tools are most commonly used: linkage analysis (QTL mapping) and association mapping (AM). Both approaches use phenotypic and genotypic data to identify molecular markers (QTL) linked to traits of interest for potential use in MAS. Compared to the traditional QTL mapping approach, AM uses a diverse set of germplasm without the need to develop a bi-parental mapping population and provides a higher genetic resolution than bi-parental populations. To date, AM has been carried out in many crops to identify QTL of interest (Tadesse et al., 2015; Muqaddasi et al., 2017).

Although a large number of QTL have been identified in many crops, the majority have not been used in breeding programs (Bernardo, 2008). Since the beneficial QTL allele can

already be fixed in the breeding programs, the marker does not have predictive value in the breeding germplasm and, consequently, cannot very useful. Therefore, detecting QTL using breeding lines would greatly improve the usefulness of QTL detection for MAS.

During the process of developing germplasm and cultivars in a breeding program, a large amount of phenotypic data is routinely collected for selected breeding lines in advanced generations. Typically, none of these data is used to gain a better understanding of the underlying genetics of traits. With relatively cheap genotyping techniques, the AM approach for identifying QTL could be even more cost-effective if a large amount of phenotypic data routinely collected by breeding programs could be used. However, each year, only a small number of advanced breeding lines are evaluated. Moreover, a few of those are replaced with new lines over the next few years, resulting in data from a large collection of advanced breeding lines comprised of genotypes that are evaluated in different years and locations, generating unbalanced data. This unbalanced data poses a serious challenge for its use in genomic studies. The comparison of unbalanced and balanced data results will aid in examining the utility of an unbalanced dataset for the genetic dissection of important traits. The objectives of this study were to identify: 1) genomic regions associated with agronomic and quality traits, 2) associated markers suitable for MAS, and 3) whether historic unbalanced data is suitable for AM analysis and QTL identification.

Materials and Methods

Plant material and field evaluation

The experimental population consisted of 243 durum wheat cultivars and inbred lines ($F_{5:9}$) entered into the Uniform Regional Durum Nursery (URDN) from 1997 to 2014 (except in 2010 and 2011 due to severe weather conditions). These cultivars and inbred lines were chosen

based on the phenotypic data available for the agronomic and quality traits routinely collected over the years. The historic unbalanced data was collected on 243 lines from 1997 until 2014. Each year, a subset of those lines was evaluated in five locations in North Dakota (Williston, Minot, Langdon, Carrington, and Prosper) in a randomized complete block design (RCBD) with four replications. Some of the lines were replaced each year with new inbred lines and evaluated in the same manner. In 2015, a panel of 256 lines was evaluated together in two locations (Prosper and Langdon) to collect balanced data for the traits under study. Genotypes were planted in a 16 × 16 lattice design with two replications. For both datasets, individual plots consisted of four, 2 m-long rows spaced 0.3 m apart. Plots were harvested with a plot combine (HEGE 140), and the grain was collected in individual sacks. The grain was dried to approximately 13% moisture content and stored at 16ºC until further processing. Agronomic and quality tests were performed in the Durum Wheat/Pasta Processing Laboratory at NDSU.

Data collection

Sample preparation

Post-harvest cleaning for yield evaluation was conducted using the Carter-Day Dockage Tester (Simon-Carter-Day Company, Minneapolis, Minnesota) configured with a number 25 riddle and a number 2 top sieve and number 2 bottom sieve.

Before all other tests were conducted and after yield was recorded, samples were cleaned again using the Carter-Day Dockage Tester configured with a number 25 riddle, a number 8 top sieve, and a number 2 bottom sieve.

Subsamples from the total amount of seed available were taken and ground into whole-wheat flour using a Udy Cyclone Mill (UDY Corporation., Boulder, Colorado) fitted with a 60-mesh sieve and stored in plastic bags at 4°C until tests were performed. Ash content, protein

content, sodium dodecyl sulphate (SDS)-microsedimentation value, falling number, and total yellow pigment content tests were performed using the whole-wheat samples.

Subsamples for milling were taken from the available grain. Prior to milling, all subsamples were tempered in two stages based on grain moisture. In the first stage, they were tempered to 12.5% moisture for at least 72 hours; and in the second stage, they were tempered to 15% 24 hours before milling. Durum wheat was milled into semolina using a Quadramat Jr. Mill (C.W. Brabender Instruments, Inc., South Hackensack, New Jersey) for the historic unbalanced data as well as for the balanced data according to AACC method 26-50.01 (AACC International 2010) and the NDSU Cereal Sciences Department Lab protocol. Semolina samples were kept at 4°C until further analysis.

Phenotypic data was recorded for the agronomic and quality traits discussed in the following sections.

Yield

Yield data was collected at grain maturity. Grain yield was determined by harvesting each plot with a plot combine (HEGE 140) and collecting the grain in individual sacks. Grain was dried in a forced-air dryer to approximately 13% moisture. The weight of the grain from each plot was recorded in grams and then converted to kg ha^{-1}.

Days to heading

The number of days to heading was recorded as the number of days from planting until the day when 50% of the spikes had completely emerged from 50% of the plants in an individual plot.

Plant height

Plant height was estimated as the distance in centimeters from the ground to the tip of the spikes, excluding the awns. Plant height was expressed in centimeters (cm).

Stem lodging

Stem lodging was evaluated at physiological maturity using a categorical scale from 0 (all plants in the plot completely upright) to 10 (all plants in the plot completely lodged), depending on the deviation of the plants from the vertical position and on the percentage of lodged plants per plot. For example, a score of 5 was recorded when more than 50% of the plants in a plot were lodged at an angle of at least 45^{o} from the vertical position.

Leaf disease

Foliar leaf disease was evaluated at heading time based on the severity of the lesion area from tan spot [caused by *Pyrenophora tritici-repentis* (Died.) Drechs], *Septoria tritici* blotch (STB) [caused by *Mycospharella graminicola* (Fückl) J. Schröt. in Cohn], and *Septoria nodorum* blotch (SNB) [caused by *Parastagnospora nodorum* (Berk.) Castellani & E.G. Germano]. A categorical scale from zero (no disease pressure) to 10 (high disease pressure) was used to assess the relative percentage of plants affected.

Test weight, 1000-kernel weight

Test weight (TWT) was determined according to AACC method 55-10.01 (2010). Thousand-kernel weight (TKW) was determined by counting the number of kernels in 10 g of clean grain using an electronic seed counter (Seedburo Equipment Company, Chicago, Illinois).

Kernel size distribution

Kernel size distribution was performed using a sieve shaker following the NDSU Cereal Sciences Department Lab protocol. Kernels remaining on the top sieve Tyler No.7 (2.92 mm

opening) were classified as ‘large’; those that remained on the middle sieve Tyler No. 9 (2.24 mm opening) were classified as ‘medium’. Kernels passing through the second sieve were classified as ‘small’.

Vitreousness

Grain vitreousness was determined by cutting 100 kernels taken at random transversally with a farinator (grain splitter) and identifying those not fully vitreous according to the appearance of the sectional areas of the endosperm.

Ash content, moisture content, falling number, and grain protein content

Approved methods (AACC International, 2010) were used to determine ash content (method 08-01.01), moisture content (method 44-15.02), and falling number (method 56-81.03). Protein content was determined using an Infratec 1226 Whole Grain Analyzer (FOSS Tecator, Höganas, Sweden).

Semolina extraction

Semolina extraction was expressed as a percentage weight per weight (w/w) of semolina from tempered durum wheat samples.

Semolina protein

Semolina protein content was determined using AACC method 39-25.01 adapted for the FOSS Infratec 1241 Grain Analyzer (Foss North America, Eden Prairie, Minnesota).

Semolina color

Semolina color was determined using the Minolta colorimeter CIEL CR410 (Hunter lab *L*, a, b). Value ‘*L*’ indicates lightness or brightness, value ‘b’ indicates yellowness, and value ‘a’ indicates ‘greenness’.

Semolina dough sheet color

A semolina dough sheet was made using a modified method described by Fu et al. (2011). A total of 30 grams of semolina was hydrated to 38% moisture at 45°C and mixed for one minute in a KitchenAid mixer (4.3 L KitchenAid CLASSIC Stand Mixer 5K45SS) at speed 4. After mixing, the dough was sheeted twice in a pair of sheeting rolls with a gap of 1 mm. The resulting dough sheet was folded twice and sheeted twice in a pair of sheeting rolls with a gap of 3 mm without folding. The smooth dough sheet was transferred to a plastic bag and stored in a closed drawer at room temperature. Color was measured on the dough sheet at intervals of 0.5 and 24 hours using a Minolta colorimeter CIEL CR410 (Hunter lab *L*, a, b).

Total yellow pigment

Total yellow pigment (TYP) content was determined using the water-saturated n-butanol AACC method 14-50.01 (2010) as modified by using 2 g of ground whole meal. Water-saturated n-butanol (10 mL) was added to 2 g of whole meal and shaken for two minutes. After resting 30 minutes, the extracts were centrifuged at 12,000 RPM for 10 minutes, and the supernatant was carefully transferred to cuvettes. Absorbance was measured using a spectrophotometer (Beckman Coulter DU 720 General Purpose UV/Vis Spectrophotometer) at a wavelength of 435.8 nm. Measurements per extracted sample were recorded and values averaged and converted to yellow pigment concentration (μg/g) using the extinction coefficient (1.6632) for β-carotene (Sims and Lepage, 1968).

Polyphenol oxidase

Polyphenol oxidase activity was determined using intact kernels as described by Anderson and Morris (2001) using AACC method 22-85.01 (2010). A 1.5-mL aliquot of 10 mM of L-DOPA (L-3,4-dihydroxyphenylalanine) containing 0.02%, v/v Tween-20 as a substrate in a

50 mM MOPS [3-(N-morpholino) propane sulfonic acid] buffer with a pH of 6.5 was added to five undamaged seeds in a 2-mL microcentrifuge tube. The tubes were placed on an orbital shaker (Glas-Col, Terre Haute, Indiana) and rotated for one hour at room temperature to allow the reaction to occur. Polyphenol oxidase activity was measured as the change in absorbance at 475 nm using a Beckman Coulter spectrophotometer (Beckman Coulter DU 720 General Purpose UV/Vis Spectrophotometer, Fullerton, California). Each sample was run in duplicate. The L-DOPA solution was made fresh daily.

Sodium dodecyl sulfate (SDS) – micro sedimentation

The SDS micro-sedimentation test was done using the method described by Dick and Quick (1983).

Mixograph

Semolina, 10 grams based on 14% moisture, were weighed and water was added based on the grain protein content using formula (Y=1.5X + 43.6) as described in AACC method 54-40A (2000), where Y = amount of water, mL) added to sample and X = protein content at 14% mb. The 10 g bowl mixograph (National Manufacturing, TMCO Division, Lincoln, Nebraska) was used to measure the dough mixing strength of semolina.Mixing tolerance was scored using a scale of one (weak) to eight (strong).

Wet gluten and gluten index

Wet gluten and gluten index were determined with the glutomatic instrument (Perten Instruments, Springfield, Illinois) using AACC method 38-12.02 (2010).

Semolina mixing and extrusion

Different approaches for semolina mixing and extrusion were used for the two datasets due to the amount of semolina available. For the historic unbalanced dataset, 1000 g of semolina

was hydrated and mixed in a Hobart C-100-T mixer equipped with a pastry knife agitator. The mixer was set on low speed for 10 seconds while distilled water was added, and then on high speed for 50 seconds. The semolina was mixed for an additional two minutes on high speed to ensure it reached a complete premix stage.

Processing was done in a semicommercial-scale pasta extruder (DeMaco, Melbourne, Florida) and extruded through an 84-strand 0.043" teflon spaghetti die. A jacketed extrusion tube (23 cm long x 4.4 cm inside diameter) was attached to the pasta extruder to allow a longer time for semolina hydration to minimize white specks in the spaghetti. Actual conditions for dough extrusion were a screw rotation speed at 28–29 rpm, a vacuum at 0.8-1.05 kg/cm^2, and a jacket temperature at 46–48°C. Room temperature and relative humidity were maintained at 25ºC and 40-45%, respectively.

For the 2015 balanced dataset, 300 grams of semolina was hydrated and mixed in a KitchenAid commercial mixer. The mixer was set on low speed for 10 seconds while distilled water was added and then on high for 50 seconds. The semolina was mixed for an additional two minutes on high to ensure it reached a complete premix stage. Processing was done using a commercial tabletop electric pasta machine (Arcobaleno, Lancaster, PA, model AEX18) and extruded through a 35-strand 1.09 mm teflon spaghetti die.

Spaghetti drying

Due to the large number of samples evaluated in 2015 and the amount of time needed for drying, the drying process was omitted for the 2015 balanced dataset. Thus, the following description of the drying process applies for the historic unbalanced dataset only. The extruded spaghetti was dried in a laboratory pilot-scale dryer (Standard Industries, Fargo, North Dakota)

on the low-temperature cycle with a total drying time of 18 hours. The low-temperature cycle typically has an 18-hour total drying time at 40°C (Yue et al., 1999).

Spaghetti cooking

Two spaghetti cooking times were used for the two datasets. In the historic unbalanced dataset, dry spaghetti (10 g) was broken into lengths of approximately 5 cm and placed in 300 ml of boiling water for 12 minutes.

However, preliminary research indicated that the optimum cooking time for the fresh spaghetti in the 2015 balanced dataset was four minutes. The optimum cooking time was determined using AACC method 66-50 (2000). Fresh spaghetti (10 g) was cut into lengths of approximately 5 cm and placed in 300 ml of boiling water for 4 minutes.

Cooked weight

After cooking, samples were rinsed off with distilled water in a Buchner funnel and drained. Spaghetti strips were weighed and reported in grams.

Cooking loss

Cooking loss (% weight of solids) was measured by evaporating the cooking water to dryness in a forced-air oven at 110°C overnight. The residue was weighed and reported as percentage of the dry spaghetti.

Cooked spaghetti firmness

Cooked firmness was measured using a plexiglass blade probe attached to a Texture Analyzer (Model TA-XT, Texture Technology Corporation, Scarsdale, New York) as described by Walsh and Gilles (1971). Five strands of cooked spaghetti were placed on a plexiglass plate and sheared at a 90-degree angle with a plexiglass tooth probe. A TA-XT2 texture analyzer was

used to calculate the area under the curve (g.cm), indicating the amount of work required to shear the cooked spaghetti (the CF score). The average of three CF scores was used to report CF.

Statistical analysis

The analysis of variance was conducted using the statistical analysis system (SAS) computer package version 9.3 (SAS Institute, Inc., 2014). The unbalanced historic dataset was analyzed using a mixed linear model (MLM) with Proc Mixed method III, where genotypes were the fixed effects and environments and replicates within environments were the random effects. The balanced dataset was analyzed using Proc GLM method III. Least square (LS) means were used for the analyses (Steele and Torrie, 1980).

The entry means plot-based heritability for all the traits were estimated using the method described by Holland et al. (2003). The variance and covariance parameters were calculated using the COVTEST and ASYCOV options of the MIXED procedure (SAS Institute, Inc., 2011), with environments and genotypes deemed random.

Trait correlations were calculated and plotted in R 3.0 (R Development Core Team, 2013) using cor.matrix and corrplot from the corrplot package (Wei and Wei, 2013). Correlation values were considered significantly different from zero at $P \leq 0.05$.

DNA isolation and SNP-marker genotyping and analysis

Four seeds from each genotype were planted into potting mix in the greenhouse in the fall of 2014. Three young leaf tissues from each genotype were harvested and sent to the USDA-ARS Cereal Crops Genotyping Laboratory in Fargo, North Dakota, for DNA isolation. The extracted DNA samples were genotyped using the Illumina 90K iSelect BeadChip platform, and the markers were called using the diploid version of Genome Studio software (Wang et al., 2014). FastPHASE 1.3 software with the default settings (Scheet and Stephens, 2006) was used

to impute missing loci data using a "likelihood"-based imputation. The heterozygotes were considered missing. Only markers having a minor allele frequency (MAF) > 0.01 were considered for further analysis.

Linkage disequilibrium for all pairwise comparisons between intrachromosomal SNP was computed and the genome-wide LD decay was estimated using JMP Genomics 8.1 software (SAS Institute, 2015). The LD was computed as the squared correlation coefficient (R^2) for each of the marker pairs. Genome-wide LD decay was estimated by plotting LD estimates (R^2) from all 14 chromosomes against the corresponding pairwise genetic distances (cM). Smoothing spline Fit (lambda = 338064.8) was applied to the estimate of LD decay.

Association mapping (AM) analysis

Association mapping analysis was done using JMP Genomics 8.1 software (SAS Institute, 2015; Zhao et al., 2007). Population structure (Q matrix), which can be defined as the differential relatedness among genotypes, was controlled with principal component analysis (PCA). The identity-by-state (IBS) matrix (K matrix) representing the proportion of shared alleles for all pairwise comparisons in each population was applied. Five regression models to analyze marker–trait association were generated: (1) naive, (2) kinship, (3) kinship plus population structure (the first two principle components (PCs) collectively explained 11.3% of genotype variation), (4) kinship plus population structure (the first three PCs collectively explained 15.46% of genotype variation), and (5) the kinship plus population structure (the first four PCs collectively explained 19.2% of genotype variation). The best model was determined according to the Bayesian Information Criterion (BIC), where the lowest BIC value is preferred (Ghosh et al., 2006; Zhang et al., 2010). The P-values of the selected models were later adjusted

by calculating the corresponding positive false discovery rate (pFDR) (Benjamini and Yekutieli, 2001). Marker–trait associations were considered significant at a pFDR ≤ 0.1.

Stepwise regression was performed on the significant SNPs associated with each trait to estimate the combined variation explained by the minimum number of markers that can be used for MAS (Mamidi et al., 2014; Gurung et al., 2014).

Results

Phenotypic data analyses of agronomic traits

Among the most important agronomic traits are yield (YLD), days to heading (DTH), height (HT), leaf disease (LD), and lodging (LDG). Some durum wheat quality traits can be considered both agronomic- and quality-related. Agronomic-related traits in this study's results include falling number (FN), test weight (TWT), 1000-kernel weight (TKW), and kernel size distribution, which includes the percentages of small kernels (SK), medium kernels (MK), and large kernels (LK). The statistical parameters for the agronomic-related traits are presented in Table 3-1. There were significant differences among genotypes for all the traits, as noted in the range.

Significant and consistent positive correlations for YLD were observed with TWT and TKW (Fig. 3-1). Similar correlations were reported by Bhatt (1973) and Dogan (2009). Test weight and TKW are indirect indicators of both YLD and milling quality. Yield, TWT, and TKW were also significantly correlated with kernel distribution. Yield, TWT, and TKW showed a significant positive correlation with LK, and a negative correlation with MK and SK (data not shown).

A consistent positive correlation was observed for plant HT with TWT and TKW, while the correlation of HT and YLD were not consistent, suggesting there are many factors that contribute to YLD (Richards, 1992). In the literature, the relationship between kernel size and

HT has been well observed and attributed to harvest index (Mohammadi et al., 2012). Plant HT showed a significant positive correlation with the distribution of LK, and a negative correlation with the distribution of MK and SK. Height was consistently negatively correlated with LD, and positively correlated with LDG. Taller plants often show lower levels of leaf infection compared to dwarf cultivars, which are nearer to inoculum sources (Eyal, 1971; Eyal et al., 1987). However, taller plants are more susceptible to lodging under unfavorable weather conditions, such as rains and winds. Reducing plant height and improving culm stiffness are indirect ways to improve plant lodging resistance in durum wheat (Keller et al., 1999).

Table 3-1. Statistical estimation of agronomic traits at each environment/trial.

Trait and Env.[a]	Mean	Std Dev[b]	Std Err Mean[c]	CV[d]	Min[e]	Max[k]
Yield, kg/ha						
Prosper	5337.2	757.5	33.5	14.1	2646.2	7084.9
Langdon	5534.5	629.1	27.8	11.3	4094.8	7335.0
Mean LP	5436.9	702.9	21.2	12.9	2646.2	7335.0
Unbalanced combined	3826.0	242.2	15.8	6.3	2873.3	4514.3
Lodging, 0-9						
Prosper	5.0	1.83	0.12	36.3	0.50	9.0
Langdon	1.3	1.51	0.10	116.3	0.00	8.0
Mean LP	3.1	1.36	0.09	42.9	0.25	8.5
Unbalanced combined	1.5	0.79	0.05	50.3	0.00	5.8
Leaf disease, 0-9						
Prosper	6.5	0.93	0.06	14.1	4.0	8.7
Langdon	4.1	0.80	0.05	19.4	2.0	6.0
Mean LP	5.3	0.71	0.05	13.2	3.5	7.1
Unbalanced combined	4.1	0.79	0.05	18.8	1.8	6.3
Height, cm						
Prosper	93.6	6.3	0.41	6.7	65.0	120.0
Langdon	93.6	5.6	0.36	5.9	71.5	120.0
Mean LP	93.6	5.6	0.36	6.0	68.2	120.0
Unbalanced combined	87.8	4.9	0.32	5.5	70.2	115.9
Days to heading						
Prosper	55.7	1.18	0.08	2.1	52.0	59.5
Langdon	64.2	1.40	0.09	2.1	61.0	68.0
Mean LP	60.0	1.15	0.07	1.9	56.5	63.5
Unbalanced combined	61.1	1.00	0.07	1.6	57.8	63.5

Table 3-1. Statistical estimation of agronomic traits at each environment/trial (continued)

Trait and Env.[a]	Mean	Std Dev[b]	Std Err Mean[c]	CV[d]	Min[e]	Max[k]
Falling number, sec						
Prosper	478.0	41.7	2.7	8.7	393.0	608.5
Langdon	461.8	57.0	3.6	12.5	322.0	606.0
Mean LP	469.9	42.4	2.7	9.0	370.7	586.2
Unbalanced combined	411.0	25.1	1.6	6.1	331.5	494.5
Test weight						
Prosper	56.8	1.68	0.11	2.9	50.02	60.6
Langdon	60.6	0.97	0.06	1.6	56.6	62.6
Mean LP	58.7	1.19	0.08	2.0	53.5	61.3
Unbalanced combined	60.2	0.76	0.05	1.2	57.4	62.1
Thousand-kernel weight						
Prosper	32.89	2.9	0.19	8.9	20.6	39.9
Langdon	39.7	2.5	0.16	6.3	30.4	45.9
Mean LP	36.3	2.4	0.16	6.7	26.4	41.6
Unbalanced combined	38.2	1.9	0.13	5.0	33.0	43.6
% Large kernels						
Prosper	39.9	12.1	0.79	30.5	3.3	70.4
Langdon	62.3	9.04	0.58	14.5	26.0	79.7
Mean LP	51.1	10.6	0.68	22.5	14.6	75.0
Unbalanced combined	50.3	7.6	0.50	15.2	23.6	69.4
% Medium kernels						
Prosper	55.1	9.8	0.64	17.8	28.5	77.3
Langdon	35.4	8.0	0.52	22.6	19.1	68.1
Mean LP	45.2	8.0	0.52	17.7	26.8	71.6
Combined	45.9	6.6	0.44	14.4	24.6	68.3
% Small kernels						
Prosper	4.9	3.1	0.20	62.5	0.98	28.6
Langdon	2.2	1.2	0.08	55.7	0.47	8.3
Mean LP	3.5	2.0	0.13	56.5	1.03	18.5
Unbalanced combined	4.6	1.4	0.09	30.6	0.56	10.3

[a]Identification of the agronomic traits and environment/experiment where it was evaluated: 'Prosper' stands for the balanced dataset grown at Prosper in 2015, 'Langdon' stands for the balanced dataset grown in Langdon in 2015, 'Mean LP' stands for the mean of the Langdon and Prosper experiments, and 'Unbalanced combined' stands for the unbalanced dataset that was combined over time and locations from 1998 until 2014. [b] Standard Deviation; [c] Standard Deviation Mean; [d] Coefficient of Variance (%); [e] Minimum observed phenotypic value; and [k] Maximum observed phenotypic value.

Correlations among agronomic-related traits showed that most agronomic traits are associated with each other based on phenotypic data analysis (Fig. 3-1). The ANOVA showed significant differences among genotypes for agronomic traits. Environment had a significant effect on most of the traits as indicated by a significant genotype by environment interaction (A1-A3 Tables).

Phenotypic data analysis of quality traits

Quality traits are subdivided into three categories: grain quality parameters, flour- and semolina-related characteristics, and pasta cooking-related traits. Grain quality parameters include such traits as grain protein concentration (WPROT) and kernel vitreousness (VIT). Flour- and semolina-related characteristics include milling extraction (total extraction, or TEXT) and semolina extraction (SEXT), semolina ash content (SASH), protein- and dough-related traits, total yellow pigment content, and the polyphenol oxidase activity. Pasta cooking-related traits are determined by cooked firmness, cooking loss, and cooked weight. Statistical parameters for quality traits are presented in Table 3-2. There were significant differences among genotypes for all the traits as noted in the range.

Grain quality parameters

The grain protein (WPROT) correlation with YLD was inconsistent over experiments/trials, suggesting dilution of the protein in starch during favorable growing conditions (Kibitie and Evans, 1984; Chee et al., 1998; Dogan, 2009; Kumar et al., 2017). Grain protein consistently positively correlated with kernel vitreousness (VIT) (Fig, 1).

Table 3-2. Statistical estimation of grain quality parameters at each environment/trial.

Trait and Env.[a]	Mean	Std Dev[b]	Std Err Mean[c]	CV[d]	Min[e]	Max[k]
Whole wheat protein						
Prosper	14.1	0.50	0.03	3.5	12.9	15.6
Langdon	15.3	0.46	0.03	3.0	13.8	16.4
Mean LP	14.7	0.41	0.03	2.7	13.3	15.8
Unbalanced combined	14.4	0.38	0.02	2.6	13.1	15.4
Vitreousness						
Prosper	70.5	10.2	0.66	14.4	43.0	90.5
Langdon	83.6	5.5	0.36	6.6	64.5	93.5
Mean LP	77.1	7.0	0.46	9.1	57.7	92.0
Unbalanced combined	90.2	4.0	0.26	4.4	62.6	100.0

[a]Identification of the agronomic traits and environment/ where it was evaluated: 'Prosper' stands for the balanced dataset grown at Prosper in 2015, 'Langdon' stands for the balanced dataset grown in Langdon in 2015, 'Mean LP' stands for the mean of the Langdon and Prosper experiments, and 'Unbalanced combined' stands for the unbalanced dataset that was combined over time and locations from 1997 until 2014. [b] Standard Deviation; [c] Standard Deviation Mean; [d] Coefficient of Variance (%); [e] Minimum observed phenotypic value; and [k] Maximum observed phenotypic value.

Flour- and semolina-related characteristics

Statistical parameters for flour- and semolina-related characteristics are presented in Table 3-3. There were significant differences among genotypes for these traits. Environment had a significant effect on most of the traits as indicated by the significant genotype by environment interaction (Tables A1-A3).

Total extraction and SEXT had no to low positive correlation with TWT, TKW, and kernel size distribution, which was unexpected (Fig. 3-1). However, SASH was negatively correlated with YLD, TWT, and TKW, while positively correlated with WPROT. This suggests a negative effect of SK on both YLD and milling. .

Total yellow pigment was highly positively correlated with color b† and highly negatively correlated with color L†, TWT, and TKW (Fig, 1). A negative correlation of TYP and color b† with TWT and TKW may be due to the effect of carotenoids being diluted in wheat endosperm.

Table 3-3. Statistical estimation of flour and semolina quality characteristics at each environment/trial.

Trait and Env.[a]	Mean	Std Dev[b]	Std Err Mean[c]	CV[d]	Min[e]	Max[k]
Total extraction						
Unbalanced combined	70.0	0.76	0.05	1.08	67.9	71.7
Semolina extraction						
Unbalanced combined	63.9	0.86	0.06	1.35	61.3	66.0
Sedimentation volume						
Prosper	57.3	7.2	0.47	12.6	24.5	77.0
Langdon	51.2	7.8	0.51	15.3	22.5	75.0
Mean LP	54.2	7.0	0.46	13.0	23.5	76.0
Unbalanced combined	50.9	5.8	0.38	11.4	25.1	67.4
Gluten index						
Prosper	65.2	17.4	1.13	26.8	0.99	96.6
Langdon	25.4	16.9	1.10	66.4	0.62	72.0
Mean LP	45.3	15.8	1.02	34.8	1.04	82.7
Unbalanced combined	57.0	13.8	0.93	24.2	6.88	92.5
Wet gluten						
Prosper	29.6	2.4	0.16	8.3	16.2	37.3
Langdon	37.0	2.3	0.15	6.2	29.5	42.4
Mean LP	33.3	2.0	0.13	6.2	23.5	39.5
Unbalanced combined	37.3	1.7	0.12	4.6	32.3	42.0
Glutork						
Prosper	1.80	0.15	0.01	8.5	1.02	2.1
Langdon	2.23	0.15	0.01	6.9	1.73	2.6
Mean LP	2.01	0.13	0.01	6.3	1.41	2.3
Mixograph						
Combined	6.2	0.76	0.05	12.1	2.4	8.2
Color L[†]						
Prosper	76.6	1.09	0.07	1.42	73.1	80.6
Langdon	74.6	0.77	0.05	1.03	72.5	78.0
Mean LP	75.6	0.76	0.05	1.01	73.3	79.3

Table 3-3. Statistical estimation of flour and semolina quality characteristics at each environment/trial (continued).

Trait and Env.[a]	Mean	Std Dev[b]	Std Err Mean[c]	CV[d]	Min[e]	Max[k]
Color a†						
Prosper	-2.05	0.53	0.03	0	-3.1	0.23
Langdon	-2.63	0.38	0.02	0	-3.5	-1.59
Mean LP	-2.34	0.40	0.03	0	-3.3	-0.78
Color b†						
Prosper	31.6	1.77	0.11	5.6	24.9	35.8
Langdon	30.0	1.83	0.12	6.0	21.7	34.4
Mean LP	30.8	1.72	0.11	5.5	23.3	34.9
Total yellow pigment						
Prosper	8.7	0.98	0.06	11.1	6.2	11.1
Langdon	8.0	0.86	0.06	10.6	5.7	10.1
Mean LP	8.4	0.89	0.06	10.5	6.2	10.5
Difference in color L† after 24 h						
Prosper	0.80	0.64	0.04	79.3	-0.42	2.87
Langdon	2.50	1.28	0.08	51.3	0.02	6.55
Mean LP	1.65	0.74	0.05	44.9	-0.02	4.16
Difference in color a† after 24 h						
Prosper	-0.14	0.21	0.01	0	-0.83	0.55
Langdon	-0.78	0.22	0.01	0	-1.43	-0.17
Mean LP	-0.46	0.18	0.01	0	-0.98	0.01
Difference in color b† after 24 h						
Prosper	1.38	0.74	0.05	53.3	-0.67	3.7
Langdon	2.34	0.95	0.06	40.7	-0.47	5.2
Mean LP	1.86	0.70	0.05	37.6	0.09	4.4
Pasta color						
Unbalanced combined	8.9	0.26	0.02	2.8	8.1	9.4
Polyphenol oxidase activity						
Prosper	0.10	0.127	0.008	124.5	0.03	0.60
Mean LP	0.11	0.111	0.007	95.2	0.03	0.53
Mean LP	0.109	0.118	0.008	107.5	0.03	0.53

[a]Identification of the agronomic traits and environment/experiment where it was evaluated: 'Prosper' stands for the balanced dataset grown at Prosper in 2015, 'Langdon' stands for the balanced dataset grown in Langdon in 2015, 'Mean LP' stands for the mean of the Langdon and Prosper experiments, and 'Unbalanced combined' stands for the unbalanced dataset that was combined over time and locations from 1998 until 2014. [b] Standard Deviation; [c] Standard Deviation Mean; [d] Coefficient of Variance (%); [e] Minimum observed phenotypic value; [k] Maximum observed phenotypic value. Columns with either balanced or unbalanced combined phenotypic data missing means the trait was evaluated in one of the experiments only.

Pasta cooking-related traits

Statistical parameters for pasta cooking-related traits are presented in Table 3-4. There were significant differences among genotypes for these traits. Environment had a significant effect on most of the traits as indicated by the significant genotype by environment interaction (Tables A-2 and A-3).

Grain protein (WPROT) was consistently positively correlated with VIT, WG, FIRM, and GLUT and negatively correlated with CWT and CLOSS (Fig. 3-1). Gluten strength, measured by GI, SDS, and MIXO, was significantly positively correlated with spaghetti firmness (FIRM), but neither SDS, MIXO, nor GI had a significant correlation with WPROT, CWT, or CLOSS (Fig. 3-1). These findings suggest that both protein quantity and quality/composition play an independent role in the end-use product (Ciaffi et al., 1991; Samaan et al., 2006).

The FN values of the durum wheat flour were negatively correlated to spaghetti cooking loss (CLOSS), indicating that low FN values (associated with high α-amylase activities) tended to increase the amount of residue in the spaghetti cooking water (Matsuo et al., 1982; Grant et al., 1993). Overall, the correlation analysis showed that the protein quantity and quality characteristics were associated with the cooking properties.

Table 3-4. Statistical estimation of cooking-related traits at each environment/trial.

Trait and Env.[a]	Mean	Std Dev[b]	Std Err Mean[c]	CV[d]	Min[e]	Max[k]
Fresh pasta firmness						
Prosper	227.6	19.9	1.29	8.76	167.3	279.4
Langdon	227.6	14.8	0.96	6.54	163.06	267.0
Mean LP	229.4	13.7	0.89	5.99	179.6	263.3
Dried pasta firmness						
Combined	5.9	0.33	0.02	5.5	4.5	6.8
Fresh pasta cooking loss						
Prosper	4.2	0.29	0.02	7.0	3.3	4.9
Langdon	3.7	0.29	0.02	7.6	2.9	4.5
Mean LP	3.9	0.20	0.01	4.9	3.4	4.5
Dried pasta cooking loss						
Combined	6.7	0.22	0.01	3.2	6.2	7.5
Fresh pasta cooked weight						
Prosper	19.5	0.89	0.06	4.5	9.7	20.9
Langdon	19.4	0.33	0.02	1.6	17.9	20.1
Mean LP	19.5	0.47	0.03	2.4	14.7	20.3
Dried pasta cooked weight						
Combined	30.3	0.33	0.02	1.08	28.5	31.5
Work to shear						
Prosper	17.6	1.7	0.12	10.1	13.1	22.9
Langdon	16.8	1.2	0.08	7.2	13.2	19.9
Mean LP	17.3	1.1	0.07	6.6	14.0	20.5

[a]Identification of the agronomic traits and environment/experiment where it was evaluated: 'Prosper' stands for the balanced dataset grown at Prosper in 2015, 'Langdon' stands for the balanced dataset grown in Langdon in 2015, 'Mean LP' stands for the mean of the Langdon and Prosper experiments, and 'Unbalanced combined' stands for the unbalanced dataset that was combined over time and locations from 1998 until 2014. [b] Standard Deviation; [c] Standard Deviation Mean; [d] Coefficient of Variance (%); [e] Minimum observed phenotypic value; and [k] Maximum observed phenotypic value. Columns with either balanced or unbalanced combined phenotypic data missing means the trait was evaluated in one of the experiments only.

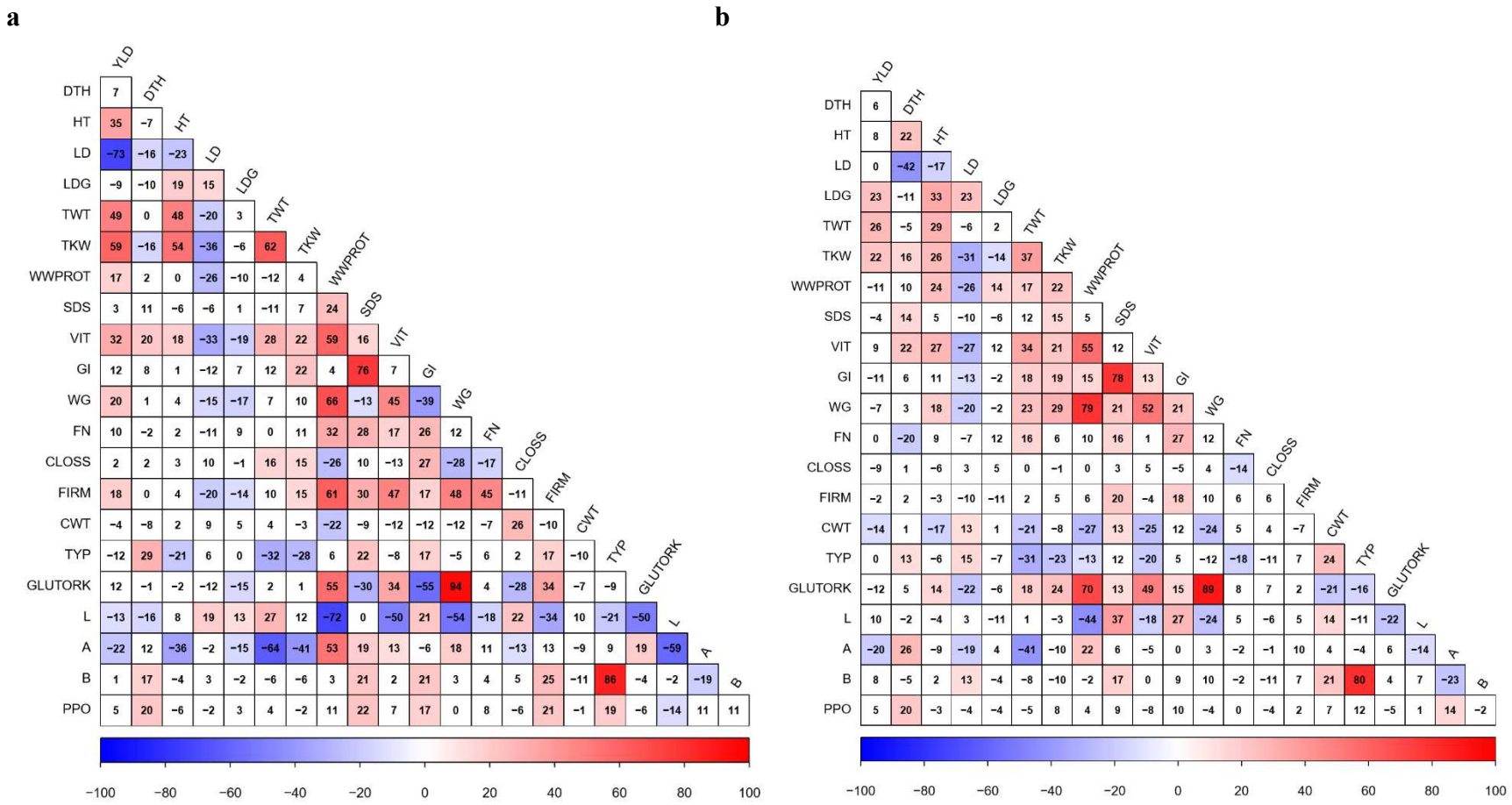


Figure 3-1. Correlation between traits and locations based on adjusted means. Data showing the relationship between agronomic and quality traits in (a) Prosper, (b) Langdon, (c) unbalanced combined, and (d) LP mean. YLD = yield, HT=height, LDG=lodging, LD=leaf disease, TWT=test weight, TKW=thousand kernel weight, WPROT=whole-wheat protein, SDS=sedimentation test, GI=gluten index, WG=wet gluten, FIRM=firmness, CLOSS=cooking loss, CWT=cooked weight, TYP=total yellow pigment, PPO=polyphenol oxidase, and FN=falling number. Cells with correlation values not significant at P-value < 0.01 are left blank.

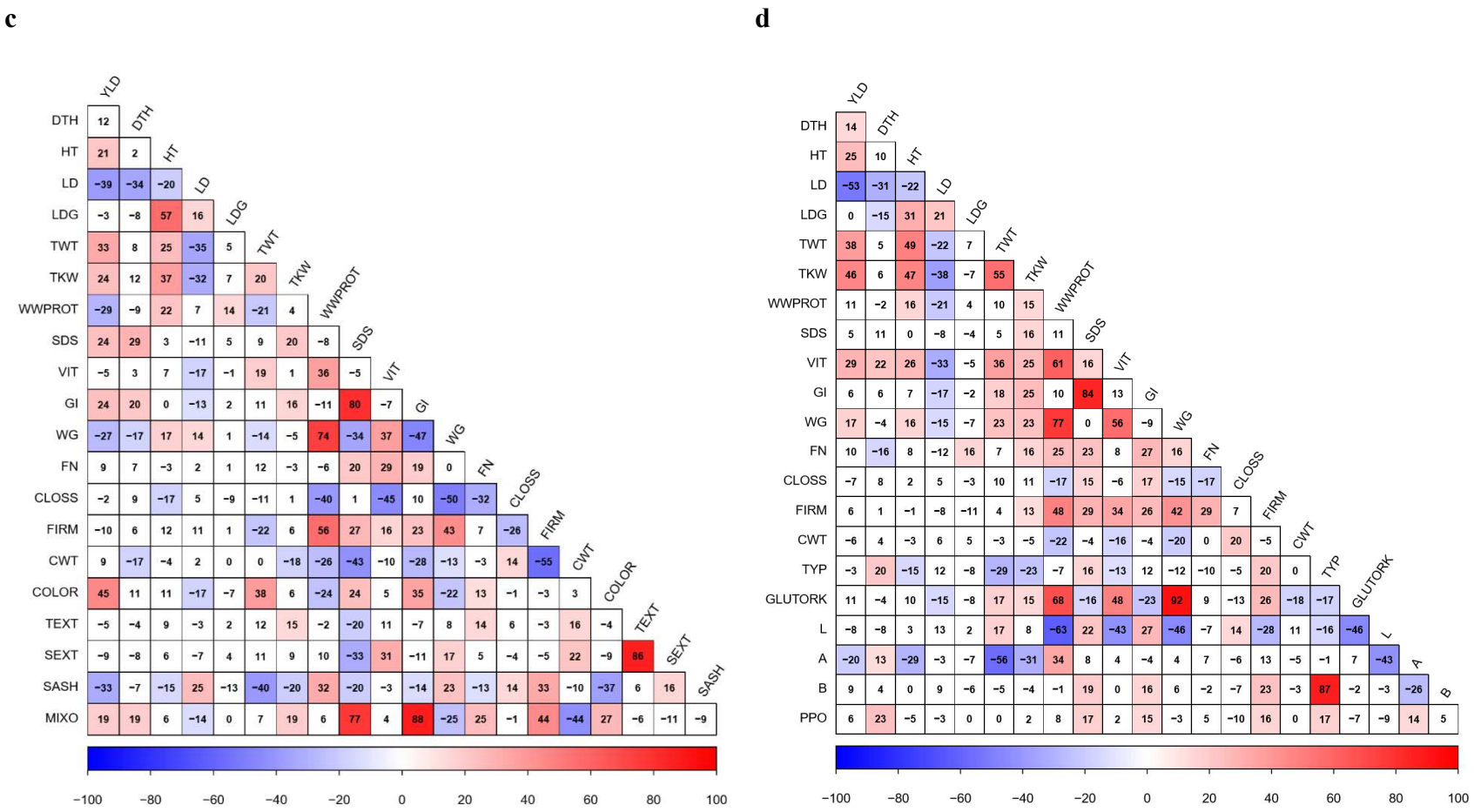


Figure 3-1. Correlation between traits and locations based on adjusted means (continued). Data showing the relationship between agronomic and quality traits in (a) Prosper, (b) Langdon, (c) unbalanced combined, and (d) LP mean. YLD = yield, HT=height, LDG=lodging, LD=leaf disease, TWT=test weight, TKW=thousand kernel weight, WPROT=whole-wheat protein, SDS=sedimentation test, GI=gluten index, WG=wet gluten, FIRM=firmness, CLOSS=cooking loss, CWT=cooked weight, TYP=total yellow pigment, PPO=polyphenol oxidase, and FN=falling number. Cells with correlation values not significant at P-value < 0.01 are left blank.

Marker properties and linkage disequilibrium analysis

A total of 4,196 SNP markers were selected for LD/AM after excluding the markers with MAF <1%, missing data points >10%, and markers with no genetic position on the consensus map. Markers were ordered according to the scaled map positions of the SNP marker based on the tetraploid wheat consensus map (Maccaferri et al., 2015). The LD decayed to 0.2 within 5 cM, on average (Fig. 3-2). Significantly-associated SNPs that were five or less cM apart and/or located between the pairwise LD (R^2) ≥ 0.7 were considered a single QTL.

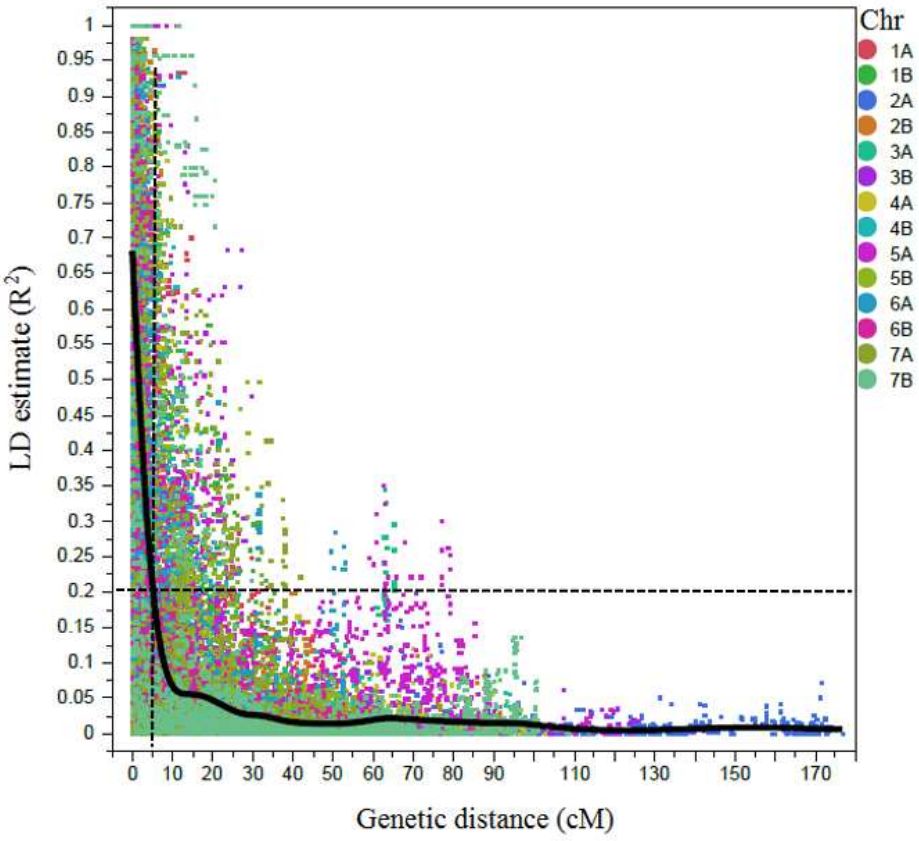


Figure 3-2. Scatter plot showing the linkage disequilibrium (LD) decay across the chromosomes (Chr) for 243 durum wheat genotypes. The genetic distance in centimorgan (cM) is plotted against the LD estimate (R^2) for pairs of SNPs. Smoothing Spline Fit, lambda=338064.8, R-Square 0.56902, and Sum of Squares Error 6804.7.

Population structure, kinship analysis, and regression model selection for AM

Population structure was inferred using PCA. Principal component analysis showed that two, three, and 10 PCs explain a cumulative 11.3, 15.4, and 26.8% of the genotype variation, respectively. The first three PCs clustered the collection into three subpopulations (Fig. 3-3).

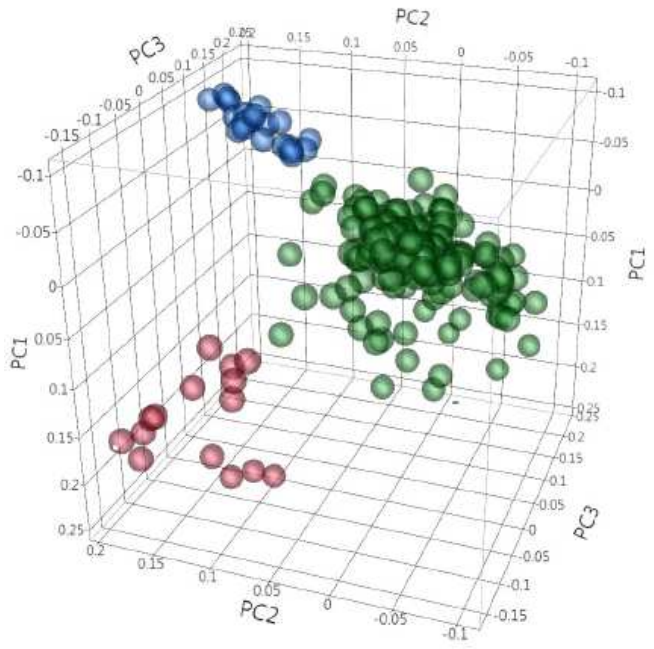


Figure 3-3. Principal component (PC) analysis obtained from 4,196 polymorphic SNPs, indicating the population structure in 243 durum wheat accessions. PC1, PC2, and PC3 explain 6.8, 4.5, and 4.1% of the variation, respectively. The colors represent three different clusters.

The familial relatedness was estimated using an identity-by-state matrix (**K** matrix), and kinship between accessions was calculated. A heat map of the marker-based **K** matrix is illustrated in Fig. 3-4. Some hotspots with related lines were observed on the heat map, suggesting intermediate familial relationships among genotypes. Accounting for the population structure and familial relationship between individuals in the AM analysis reduces the number of false-positive associations. Based on the BIC values of the five regression models tested, no single model fits best for all traits in different environments (Table 3-5). For instance, the results of the kinship model used for height were 4PC + kinship, 2PC + kinship, 3PC + kinship, and

2PC + kinship for Langdon, Prosper, mean Prosper, and Langdon data and unbalanced combined datasets, respectively (Table 3-5).

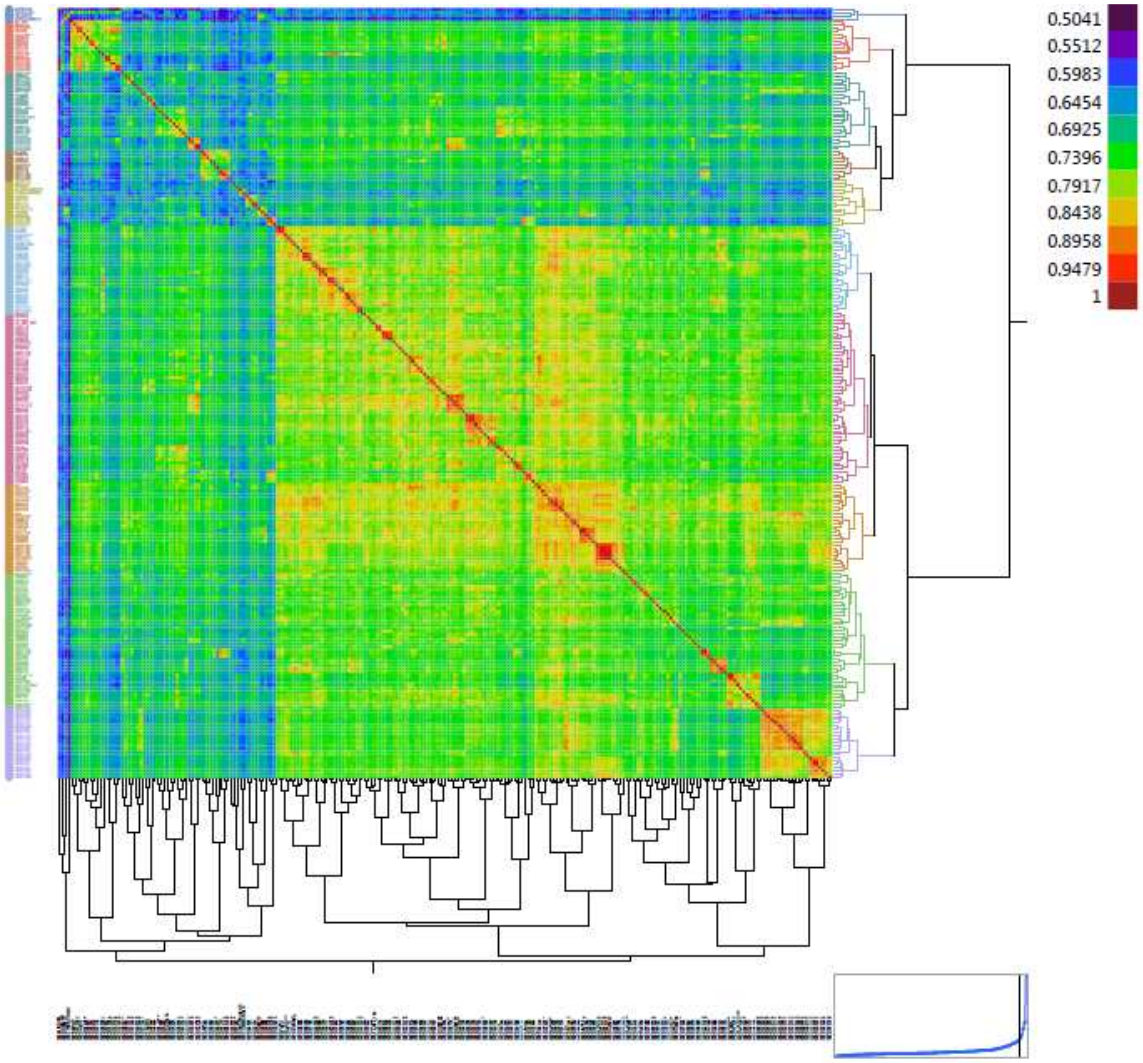


Figure 3-4. Heat map displaying the relationship matrix among durum wheat genotypes. The red diagonal represents a perfect relationship of each accession with itself. The symmetric off-diagonal elements represent the relationship measures (Identity-by-state) for pairs of accessions. The blocks of warmer colors on the diagonal show clusters of closely-related accessions.

Table 3-5. Bayesian information criterion (BIC) value for each model.

Trait	Naïve	Kinship	2PCs + Kinship	3PCs + Kinship	4PCs +Kinship
Langdon, 2015					
YLD	3076.8	3065.8	3032.7	3026.3	**3026.1**
HT	1512.2	1501.3	1453.7	1453.7	**1453.5**
DTH	850.3	839.3	831.9	830.4	**829.8**
LD	581.6	518.6	548.1	547.6	**513.5**
LDG	884.6	873.7	868	867.9	**865.9**
FN	2621.1	2610.1	2594.7	2594.7	**2593.9**
TWT	674.8	663.8	657.2	656.3	**633.1**
TKW	1131.1	1120.1	1110.7	**1073.5**	1073.5
LK	1740.5	1729.5	1721.9	1719.9	**1719.5**
MK	1684.6	1673.6	1666.8	1664.9	**1664.7**
SK	786.8	775.8	761.9	760	**757**
VIT	1507.5	1496.6	1481.6	**1435.2**	1471.3
WPROT	320.2	309.3	306.5	306.3	**297.3**
SPROT	382.6	371.6	364.9	364.8	**353.9**
SDS	1674.3	1663.4	1643.1	**1605.4**	1639.9
GI	2040.8	2029.8	2001.1	2000.4	**1997.9**
WG	1089.5	1078.6	1070	1066.8	**1053.8**
GLUT	-203.2	-214.2	-223.4	-224.1	**-228.7**
FIRM	1978.8	1967.9	1965.4	1964.8	**1958**
WTS	784.2	773.3	770.5	770.4	**767.5**
CLOSS	90.7	79.7	78.4	77.8	**70.2**
CWT	151.3	140.3	136.4	134.7	**126.6**
L	564.8	553.8	549.4	548.5	**548.3**
A	225.6	214.7	207.2	202.6	**191.6**
B	976.7	924.2	944.4	**920.7**	931.6
DIF_B	665.2	654.3	642.4	642.1	**640**
DIF_L	807.4	796.5	794.3	794.3	**793**
DIF_A	-25.2	-36.1	-39	-39	**-42.7**
PPO	-360.5	-371.4	-377.8	-379.2	**-379.4**
TYP	614.8	529.1	551.1	551.1	**519.7**
Prosper, 2015					
YLD	3246.3	3235.4	3229.5	**3222**	3222
HT	1569.1	1558.1	**1506.6**	1506.6	1506.6
DTH	766.6	755.6	**747.1**	747.1	747.1
LD	651.6	640.6	634.7	**631.1**	631.1
LDG	976.5	965.5	965.2	**959**	959
FN	2472.4	2461.4	**2452.4**	2452.4	2452.4
TWT	935.7	924.7	**914**	914	914
TKW	1202.5	1191.6	1177.4	**1176.2**	1176.2
LK	1883.3	1872.4	**1855.1**	1855	1855
MK	1781.4	1770.4	**1706.3**	1706.3	1706.3
SK	1229.2	1218.2	**1212.3**	1210.6	1210.6

Table 3-5. Bayesian information criterion (BIC) value for each model (continued).

Trait	Naïve	Kinship	2PCs + Kinship	3PCs + Kinship	4PCs +Kinship
VIT	1799.3	1788.3	1773.2	**1772.1**	1772.1
WPROT	354.9	**271.7**	321.6	319.8	319.8
SPROT	383.3	**282.3**	353.6	352.4	352.4
SDS	1634.2	1623.3	1613.2	**1611.8**	1611.8
GI	2055.6	2044.7	**1975.5**	1975.5	1975.5
WG	1121	1110	1109.5	**1107.1**	1107.1
GLUTORK	-207.2	**-256.2**	-219.6	-221.3	-221.3
FIRM	2119.1	2108.1	2100.4	**2100.3**	2100.3
WTSHEAR	967.4	956.5	**946.7**	946.7	946.7
CWT	630.2	619.2	619.2	**619**	619
CLOSS	104.3	93.4	91.6	**90.2**	90.2
L	730.1	719.1	716.7	**716.5**	716.5
A	381.6	370.6	369	**364.9**	364.9
B	961.9	884.3	899.6	**876.3**	876.3
DIF_B	542.9	531.9	518.4	**517.3**	517.3
DIF_L	472.3	461.3	456.4	**455**	455
DIF_A	-57.6	-68.5	-69.8	**-74.9**	-74.9
PPO	-299.6	**-337.5**	-321.2	-321.5	-321.5
TYP	679.5	**593**	600.3	600.2	600.2
Unbalanced combined data from 1997-2014					
YLD	1274.2	**1181.6**	1259.8	1249.3	1249.3
HT	1417.7	1406.8	**1404.1**	1404.1	1404.1
DTH	675.4	664.5	651.3	**650.9**	650.9
LD	504.3	493.7	491.8	**491.6**	491.6
LDG	515.7	505	503.1	**497.6**	497.6
FFN	2183.7	2172.7	2145.1	2144.7	**2143.3**
TWT	547.4	536.5	529.9	525.6	**517.5**
TKW	979.2	968.2	957.4	957.2	**955.3**
LK	1626.6	1615.7	1552	1552	**1551.8**
MK	1494	1483.2	1476	1476	**1475.8**
SK	836.3	825.4	800.5	**799.4**	799.4
VIT	1323.5	1312.6	1293.2	1292.6	**1292.1**
WPROT	217.5	206.5	205.5	205.1	**204.2**
TEXT	544.2	533.3	533.2	**531.6**	531.6
SEXT	603.9	593	590.3	590.3	**590.2**
SASH	-999.3	-1010.2	-1013.9	-1013.9	**-1014**
SPROT	156.5	145.6	143.4	143.4	**143.2**
SDS	1498.5	1487.5	1480.6	1480.3	**1477.5**
GI	1823.3	1812.5	1729	1787.2	**1786.7**
MIXO	545.6	449.1	508.7	507.4	**446.3**
FIRM	159.6	148.7	142.3	130.8	**90.4**
CWT	150.7	139.8	136.1	135.2	**123.2**
CLOSS	-29.5	-40.4	-69.4	-83.2	**-84.2**
COLOR	40.2	29.3	20.8	14.6	**-33.5**
Mean Langdon + Prosper					
YLD	3070.5	3059.5	3011.9	3005.9	**3005.8**
HT	1514.2	1503.2	1495.6	**1495.2**	1495.2
DTH	751.1	740.2	733.7	**731.7**	731.7
LD	522.8	511.8	495.2	492.2	**491.4**
LDG	833.9	823	810.4	810.4	**807.7**

Table 3-5. Bayesian information criterion (BIC) value for each model (continued).

Trait	Naïve	Kinship	2PCs + Kinship	3PCs + Kinship	**4PCs +Kinship**
FN	2480.3	2469.3	2466.8	2456.5	**2453.5**
TWT	773.3	762.3	742.9	734.6	**723.3**
TKW	1115.5	1104.5	1047.4	1044.2	**1042.7**
LK	1774.2	1763.2	1727.4	1722.7	**1721.4**
MK	1685.7	1674.7	1644.2	1638.8	**1638.1**
SK	1025.8	1014.8	971.1	969.8	**965.9**
VIT	1622.9	1611.9	1602.7	**1591.5**	1591.5
WPROT	263.2	252.3	247.6	232.4	**231.3**
SPROT	325.4	314.4	311.1	295.6	**293.8**
SDS	1623.2	1612.3	1606.6	1593.2	**1592.6**
GI	2007.7	1996.7	1981.9	**1967.7**	1967.7
WG	1038.4	1027.4	1017.7	1016.5	**1012.6**
GLUT	-292.7	-303.7	-311.6	-313.5	**-315.8**
FIRM	1940.5	1929.5	1919.5	1917.8	**1910.6**
WTSHEAR	754.2	743.2	733.7	727.6	**723.8**
LOSS	-85.8	-96.8	-98.1	-98.1	**-101.8**
CWT	328.7	317.7	317	316.8	**315.9**
L	559.8	548.9	544.7	**544**	544
A	246.2	235.3	217.6	209.6	**198.1**
B	947	936	905	897.9	**876.7**
DIF_B	517.8	506.9	486.2	485.3	**485.1**
DIF_L	545.7	534.8	527.2	525.8	**525.4**
DIF_A	-141.8	-152.8	-157.1	-157.1	**-157.4**
PPO	-334.9	-345.9	-357.6	-358.5	**-358.7**
TYP	633.2	**542.7**	577.8	569.4	560.3

†2PC, population structure matrix (Q matrix) based on the first two principal components; 3PC, population structure matrix (Q matrix) based on the first three principal components.
†Numbers in **bold** indicate the lowest BIC and best model for each trait. The best model was used to investigate single-nucleotide polymorphism–trait associations.

Association mapping

Agronomic-related traits

The majority of the agronomic traits of YLD, DTH, HT, LDG, LD, and FN have low heritability and are highly influenced by environmental factors (Tables A-1, A-2, and A-3). Grain-related characteristics, such as TWT, TKW, and kernel size distribution (LK, MK, and SK), also show low heritability and are highly influenced by environmental factors. In the present study, a total of 106 QTL for agronomic- and grain physical characteristics-related traits were detected across all 14 chromosomes. Most of the QTL reside on chromosomes 1B (13 QTL) and 4B (14 QTL).

For each of these traits, many QTL were identified, suggesting polygenic or quantitative genetic control. The R^2 ranged from 1 to 26%. Out of 106 QTL, 10 QTL for HT, SK, and TKW had a major effect ($R^2 > 10\%$). All major QTL except for two (one each for SK and TKW) were identified in both locations in the balanced datasets. Two major QTL, one associated with TKW on 4B 22-28.8 cM and another associated with SK on 3B 86.6-88.4 cM, were not detected in the unbalanced datasets. Out of 106 QTL, only 34 were detected in one of the datasets, while the majority of the QTL were detected in at least one of the locations in the balanced dataset and unbalanced dataset. The highest number of QTL was detected for HT (21 QTL), and the lowest for LD (5 QTL). All of the QTL for HT were consistent across the unbalanced and means of the balanced datasets. A major QTL for HT, explaining over 20% of the phenotypic variation, was detected on chromosome 4B (Table 3-6).

A total of nine QTL located on seven different chromosomes were detected for YLD. Six of them were detected in at least one of the locations (and the mean) in the balanced dataset and unbalanced dataset. The R^2 ranged from 0.6 to 8.6%. Out of 19 total QTL for TKW and TWT traits, 10 QTL were detected in both the mean balanced and unbalanced datasets, including one QTL with $R^2 > 8\%$ on chromosome 4B (26.4-28.8 cM). Twenty-nine QTL for kernel distribution were detected, with 19 QTL detected in the unbalanced dataset and at least one of the locations in the balanced dataset, while 8 QTL were present in both locations and the unbalanced dataset (Table 3-6). Out of the total number of QTL, 22 for YLD, HT, DTH, LDG, TWT, TKW, LK, SK, and FN with a cutoff level less than pFDR=0.1 were detected, with 12 QTL identified in both the unbalanced and balanced datasets (Supplementary table Excel).

Table 3-6. Significant markers associated with agronomic- and grain physical characteristics-related traits in the Prosper trial, the Langdon trial, and the unbalanced combined data set.

Traits and QTL	Chr	Other associated traits	Trials[a]	Position	-log10 (P-value)	R²¶
Days to heading						
	2B	-	I††, II†, III	44.5-44.8	5.45	9.0
	2B	PPO	I††, II†, III††	114.6-121.2	3.57	5.6
	5A	-	I††, II††, III††	55.2-59.2	2.95	3.8
	6A	FN, CLOSS, YLD, HT, DIF_B, B, L	I†, II††, III†	124.8-130.0	5.81	9.4
	6B	-	II††, III††	151.0-155.9	2.71	4.0
	7B	-	I††, II††	92.9-97.4	2.72	4.0
Height						
	1B	GI, SDS, MIXO, FIRM, CWT	I††, II†, III†	3	3.73	5.8
	1B	YLD, FN, LK, MK, LD, TKW	I††, III††	35.7	2.07	3.0
	1B	Dif_a, VIT	I††, III††	82.7-85.5	2.44	3.6
	2A	-	I††, II††, III†	1.6-1.9	3.45	5.4
	2A	-	I†, III†	207.1	4.12	6.5
	2B	LDG	II††, III†	127.6-132.8	3.46	5.5
	2B	TWT, Dif _a, LDG	II††, III††	172.0	2.43	3.5
	2B	-	I††, III††	189.8	2.70	4.1
	3B	SDS	II††, III†	51.9-57.2	3.39	5.4
	3B	YLD, TKW, TWT, LK, SK	I†, II†	88.2-88.4	3.91	6.2
	4A	DIF_B	I†, II†, III†	156.9-162.8	3.90	6.1
	4B	-	I††, II††, III†	3-4.8	3.71	14.6
	4B	TWT, SK, VIT	I, II, III	17.7-22.5	9.36	15.7
	4B	YLD, HT, CLOSS, A, DIF_BYLD, LK, MK, WG, A, CLOSS, TKW, LDG, TWT, SK, VIT, color	I, II, III	26.4-28.8	11.92	19.9
	5A	-	I††, II†, III††	80.9	4.73	7.6
	5B	YLD	I†, II††	189.2-193.4	4.11	6.5
	6A	-	II††, III††	62.0-62.1	2.42	3.5
	6A	YLD, DTH, DIF_B, L, A, B	I, II†, III†	124.0-126.5	8.0	13.2
	7A	-	II††, III††	133.8	2.88	4.3
	7B	FIRM	I††, II††	66.3-67.5	2.27	3.3
	7B	-	I††, II††, III††	102.6-104.4	2.88	4.3
Lodging						
	1A	SASH, FN	II††, III††	104.3	2.08	3.0
	1B	-	II††, III††	99.7	2.08	3.0
	2B	HT	II†, III†	126.6-131.6	3.51	8.6
	2B	-	III	137.9	5.28	9.5

Table 3-6. Significant markers associated with agronomic- and grain physical characteristics-related traits in the Prosper trial, the Langdon trial, and the unbalanced combined data set (continued).

Traits and QTL	Chr	Other associated traits	Trials[a]	Position	-log$_{10}$ (P-value)	R^2¶
	2B	HT, Dif_a, TWT	I††, II††, III†	175.7-181.6	3.24	5.6
	4A	A, B, Color	I††, II††	139.6-139.7	2.46	3.6
	4B	YLD, HT, CLOSS, A, DIF_BYLD, LK, MK, WG, A, CLOSS, TKW, HT, TWT, SK, VIT, color	II††, III	22.5-28.8	5.07	9.1
	5B	-	I††, III††	103.9-108.7	2.13	3.4
Leaf disease						
	1B	GI, SDS, CWT, MIXO	I††, III††	12.8-14.3	2.74	5
	1B	YLD, TKW, HT, FN, LK, MK	I†	38-4-38.8	4.18	6.6
	4B	-	II†	77.8-83.1	3.05	4.6
	5B	YLD, HT, SK	I†, II†	187.1-193.4	3.94	6.2
	6A	-	I††, II††	8.8-10.7	2.55	3.8
Yield						
	1B		I†	35.7-38.8	4.27	6.8
	2A	GLUT	II†	169.3	4.77	7.7
	2A	GI, WG	II††, III†	188.5-189.8	4.22	6.8
	3A	LK, MK, SK	I†, III†	79.3-80.9	3.69	5.8
	3B	HT, LK, SK, TWT, TKW,	III†	88.2-88.4	4.10	6.6
	4B	HT, LDG, LK, MK, SK, TWT, TKW, VIT, A, LDG, HT, CLOSS, A, DIF_BYLD, LK, MK, WG, A, CLOSS, TKW, HT, TWT, SK, VIT, color	I††, III†	22.5-28.8	3.8	6.1
	5B	LD, HT, SK	I, III††	187.1-190.9	5.32	8.6
	5B	HT, SK	I†, III††	192.5-194.7	4.21	6.7
	6A	HT, DTH, L, A, B, TWT, TKW, SK, DIF_B	I, III†	124.8-126.5	5.07	8.2
Falling number						
	1A	LDG, SASH	II††, III††	104.3-110.9	2.27	3.4
	1B	YLD, HT, MK, LK, TKW	II††, III††	35.7-37.2	2.89	4.6
	4B	TKW	II††, III††	97.5-98.9	2.59	3.9
	5B	VIT	III	145.2-149.0	5.34	1
	6A	YLD, HT, DTH, L, A, B, TWT, TKW, SK, DIF_B	I††, II†, III††	124.8-130	3.08	4.7
	6B	-	III	45.2	5.69	9.4
	6B	-	I†	83.4	3.75	5.9
	7A	GLUT	I††, II††, III††	75.2-79.9	2.54	3.8

Table 3-6. Significant markers associated with agronomic- and grain physical characteristics-related traits in the Prosper trial, the Langdon trial, and the unbalanced combined data set (continued).

Traits and QTL	Chr	Other associated traits	Trials[a]	Position	-log10 (P-value)	R^2¶
Test weight						
	2B		II†	53-60.1	3.42	5.3
	2B	LD, SASH	III†	174.8-175.6	3.17	5
	3B		I††, III†	86.6-88.4	3.25	5
	4A	TKW	I†, II†, III†	139.6-143.7	4.33	7
	4B	HT, SM, TKW, VIT, YLD	I, II†	17.7-22.5	4.06	6.4
	4B	HT, L, SM, TKW, VIT, YLD	I, II††, III††	26.4-28.8	5.24	8.5
	6A		I†, II††, III	124.8-126.5	5.00	8.2
Thousand-kernel weight						
	1B		II†, III††	28.8-33.7	3.58	5.6
	1B		I††, III††	50.2-50.6	2.94	4.5
	3B		I	88.2-88.4	5.62	9.1
	4A	TWT	I††, II††	134.4-139.7	2.51	4
	4B	SM, TWT, VIT, YLD	I, II††	22-28.8	6.52	11
	4B		III†	98.9	3.76	6
	5B		I††, III††	45-48.9	2.86	4.3
	6A		I††, II††, III††	64.1-67.9	2.85	4
	6A	SM, TKW, HT	I	124.8-127.1	5.81	9.5
	6B		II††, III††	65.5-67.1	2.29	3.3
	6B		I††, II††, III††	91.5-92.6	2.37	3.4
	7B		I††, II††	62.7-66.3	2.78	4.2
% Large kernel distribution						
	1B		I††, II††, III††	33.7-35.7	2.83	4.3
	1B	MED	I††, II††,III†	50.2-54.8	3.06	5
	3A		I†, II††	21.3	3.20	4.9
	3A	TKW, YLD, MED, SM	I††, II†	79.3-80.9	3.00	4.6
	3B	MED	II†, III†	48.9-51.9	3.53	5.5
	3B		I†	88.2-88.4	3.18	5
	4A	TKW, TWT	I††, II††	134.6-139.7	2.27	3.3
	4B	HT, YLD, VIT, TWT TKW, Color	I, II†	22.5-28.8	5.55	9
	6A		I†	125.6	3.39	5.3
	6B		I††, II†	155.9	4.71	8
	7A		I††, III††	172.9-176.2	2.64	4

Table 3-6. Significant markers associated with agronomic- and grain physical characteristics-related traits in the Prosper trial, the Langdon trial, and the unbalanced combined data set (continued).

Traits and QTL	Chr	Other associated traits	Trials[a]	Position	-log10 (P-value)	R^2¶
% Medium kernel distribution						
	1B		I††, II††, III††	33.7-35.7	2.64	4
	1B	L	I††, II††, III†	48.7-50.3	3.00	4.9
	3A	YLD, TKW, SM, L	I††, II††	79.3-80.9	2.56	3.8
	3B	L	II†, III††	48.9-51.9	3.72	5.6
	4B	HT, L, SM, TKW, YLD	I†, III††	22.5-28.8	4.36	6.9
	6B		II†	155.9	4.63	7.4
	7A		III†	90.9	3.37	5.6
% Small kernel distribution						
	3A	YLD, TKW, L, MED	I, II††, III†	79.3-79.5	6.21	10.1
	3B	-	I, II††	86.6-88.4	12.31	20.1
	4A	TWT, TKW, L, MED	I†, II†	41.3-41.6	3.15	4.8
	4A	TKW, TWT, LK	I, II†	139.6-143.7	5.37	8.7
	4B	HT, TWT, TKW	I, II†, III††	17.7-22.5	7.23	10
	4B	L, TWT, TKW, YLD, HT	I, II†, III†	26.4-28.8	6.87	11.2
	5B	HT, YLD	I†, III††	189.2-193.4	4.25	6.7
	6A	HT, TKW	I, II††, III††	124.0-139.9	14.59	23.6
	6B		I†††, II†	155.9	3.88	6.1
	7B		I†, III††	86.4-92.9	3.71	5.8
	7B		I†, III††	112.5	3.75	5.9

[a] I, Prosper trial; II, Langdon trial; III, unbalanced combined data set where an SNP marker was detected above the pFDR value

[c] cM, marker position on the consensus map

† SNP marker that in that trial (environment) was detected above $-\log_{10}$ (*P* value) of 3, but below the pFDR value

†† SNP marker that in that trial (environment) was detected above $-\log_{10}$ (*P* value) of 2, but below the pFDR value

¶R^2, proportion of phenotypic variation explained by the individual marker

Quality-related traits

Quality-related traits include 25 traits for both the unbalanced and balanced datasets. Most quality traits were interrelated, which was reflected in their significant correlations and that they shared a common locus (Fig. 3-1, Tables 3-7 to 3-11). Quality traits were divided into three groups, including grain quality, flour and semolina quality, and pasta cooking quality parameters.

Grain quality parameters include grain protein (WPROT) and VIT. Flour and semolina quality characteristics were further subdivided into milling quality (TEXT, SEXT, and SASH), protein and gluten strength-related parameters (SPROT, SDS, GI, WG, and GLUT), and semolina- and pasta color-related parameters. Cooking quality parameters were measured with FIRM, CLOSS, and CWT.

Table 3-7. Genetic location and single nucleotide polymorphism (SNP) loci significantly associated with durum kernel characteristics across three datasets.

Traits and QTL	Chr	Other associated traits	Trials[a]	Position	-log$_{10}$ (P-value)	R^2¶
Grain protein						
	5B	SPROT, Color	II††, III††	204.7-206.1	2.86	4.4
	7A	SPROT, WG	II†, III†	59.5-62.5	3.02	4.7
	7B	FIRM	I††, III††	62.2-62.7	2.31	3.4
Kernel vitreousness						
	1B		II††, III††	88.2-93.5	2.51	3.8
	1B		I††. III††	150.9-152.5	2.30	3.4
	2B		I††, III††	183.1-188.6	2.45	3.6
	3A		III†	41.6	3.60	5.7
	3B		I††, II††	100.15-100.9	2.54	3.8
	4B	HT, L, MED, SM, TKW, TWT, YLD, Color	II	17.7-22.5	5.63	9.1
	4B		II†	26.4-28.8	3.34	8.6
	5A		III†	14.3	4.09	6.6
	5B		III	2.8	5.87	9.8
	5B	SPROT, WPROT	III	135-1-135.6	3.35	5.3
	5B	-	III	140.5-145.2	9.74	16.3
	5B	FN	III	146.14-149	9.74	16.3
	6B	-	III	54.6-58	5.62	9.3
	7A		I††, II††	59.5-62.5	2.84	4.2

[a] I, Prosper trial; II, Langdon trial; III, unbalanced combined data set where an SNP marker was detected above the pFDR value

[c] cM, marker position on the consensus map

† SNP marker that in that trial (environment) was detected above –log$_{10}$ (*P* value) of 3, but below the pFDR value

†† SNP marker that in that trial (environment) was detected above –log$_{10}$ (*P* value) of 2, but below the pFDR value

¶R^2, proportion of phenotypic variation explained by the individual marker

Three QTL for WPROT and 14 QTL for VIT reside on all chromosomes except 1A, 2A, 4A, and 6A (Table 3-7). All QTL for WPROT were present in both means across environments in both the balanced and unbalanced dataset. Major QTL for VIT were found in the unbalanced dataset only on chromosome 5B (Table 3-7).

Milling-related traits are represented by TEXT, SEXT, and SASH and were evaluated in the unbalanced dataset only. Four QTL on chromosomes 2A, 2B, 4A, and 5A were associated with both TEXT and SEXT, while two QTL on chromosomes 6B and 7A were only detected for SEXT (Table 3-8). No common QTL were detected for TEXT or SEXT with TWT, TKW, and YLD.

Semolina protein- (SPROT) and five gluten strength-related traits were presented by 27 QTL across eight chromosomes (Table 3-9). Two major QTL on chromosome 1B were common for SDS, GI, and MIXO and consistent in two locations of both the balanced and unbalanced datasets.

Table 3-8. Genetic location and single nucleotide polymorphism (SNP) loci significantly associated with durum milling-related characteristics across three datasets.

Traits and QTL	Chr	Other associated traits	Trials[a]	Position	-log$_{10}$ (P-value)	R^2¶
Total extraction						
	2A		III††	145.9	2.23	3.3
	2B	SEXT	III	158.3-161.5	5.14	8.5
	4A		III†	25.8	3.18	5.0
	5A		III††	127.5-128.3	2.07	3.0
Semolina extraction						
	2A		III††	145.8-145.9	2.99	3.6
	2B	TEXT	III††	161.5	2.66	4.1
	4A		III†	25.2-25.8	3.10	4.8
	5A		III††	127.5-128.3	2.25	3.3
	6B	-	III††	114.5	2.16	3.2
	7A	TYP, B	III†	102.3-102.4	3.13	4.9
Semolina ash						
	1A	LDG, FN	III††	102.8-105.5	2.17	3.2
	4A		III†	39	3.13	4.9
	5B	Dif_a	III††	0.9-6.5	2.45	3.7
	6A		III†	45.2	3.12	4.9
	6A		III†	117.7-118.2	3.10	4.8
	6B		III††	131.8-135.2	2.11	3.1
	7A		III††	70.7	2.23	3.3
	7A		III††	136.4	2.06	3.3

[a] I, Prosper trial; II, Langdon trial; III, unbalanced combined data set where an SNP marker was detected above the pFDR value

[c] cM, marker position on the consensus map

† SNP marker that in that trial (environment) was detected above –log$_{10}$ (*P* value) of 3, but below the pFDR value

†† SNP marker that in that trial (environment) was detected above –log$_{10}$ (*P* value) of 2, but below the pFDR value

¶R^2, proportion of phenotypic variation explained by the individual marker

Table 3-9. Genetic location and single nucleotide polymorphism (SNP) loci significantly associated with durum protein- and gluten strength-related characteristics across three datasets.

Traits and QTL	Chr	Other associated traits	Trials[a]	Position	-log$_{10}$ (P-value)	R^2¶
Semolina protein						
	1B	WPROT	I††, III††	111.7	2.11	3.2
	5B	WPROT, Color	II††, III†	204.7-206.1	3.14	5.0
	6A	FIRM	I††, II†	111.9-113.5	3.33	5.1
	7A	WPROT, WG	I††, III	59.5-59.8	3.31	5.2
	7A	FIRM	II††, III††	114	2.64	3.3
Sedimentation volume						
	1A	GI, FIRM, CWT	I†, II†, III†	1.3-4.6	4.73	7.7
	1A		I††, II††	48.1-49.7	2.51	3.7
	1B	CWT, FIRM, GI, MIXO	I†, II†, III	0.3-6.1	5.2	8.7
	1B	CWT, FIRM, GI, MIXO	I†, II, III	15.2-15.7	6.11	1.0
	2B		I††, III††	169.3-170.9	2.49	3.7
	3A		I††, II††, III††	79.5	2.86	4.3
	3B		I††, II††	51.9-56.9	2.77	4.2
	3B		II††, III††	75.5-79.1	2.91	4.4
	4A		I††, II†	0	3.31	5.1
	6A	CLOSS	I††, II††	67.9-69.1	2.62	3.9
	7B		I†, II††	65.5	3.18	4.9
Gluten index						
	1A	SDS, FIRM, CWT	II†	1.3	3.67	5.7
	1B	CWT, SDS, FIRM, MIXO	I†, II††, III†	1.3-3	4.30	7.3
	1B	CWT, SDS, FIRM, MIXO, GLUT	I†, II†, III	12.8-15.7	6.59	11.0
	2A		I††, III††	189.8	2.47	3.9
	3A		I†, II††	170.1-176.9	3.49	5.1
	3B		I††,II†, III††	75.5-79.1	3.15	4.8
	6A		I†, II††, III††	67.9-72.4	3.47	5.4
	7B		I††, II††, III††	169.8-175.9	2.67	3.3

Table 3-9. Genetic location and single nucleotide polymorphism (SNP) loci significantly associated with durum protein- and gluten strength-related characteristics across three datasets (continued).

Traits and QTL	Chr	Other associated traits	Trials[a]	Position	-log10 (P-value)	R^2¶
Wet gluten						
	2A		I††, III††	186.2-189.8	2.86	4.3
	2B	GLUT	II†, III†	146.8	3.24	5.3
	2B		I	181.6	5.17	8.3
	4B		II†, III††	22.5-28.8	3.48	5.4
	4B		I††, III††	60	2.48	3.7
	6A		II†	112.9	3.80	6.0
	6B		II†	155.9	4.28	6.8
	7A	GLUT, SPROT, WPROT	I††, III††	59.5	2.69	4.3
Mixogram score						
	1B	CWT, FIRM, GI, SDS	III	3-6.1	5.72	9.5
	1B	CWT, FIRM, GI, SDS, GLUT	III	15.2-15.7	5.64	9.3
	2A		III††	197.6	2.49	3.8
	3B		III††	75.5-75.6	2.49	3.7
	7B		II†	169.8-173.1	3.51	5.6
Glutork						
	2A		I††, II††	169.3-171	2.67	4.0
	2B	WG	I†	181.6	4.56	7.3
	3A		II†	71.6-74.4	3.08	4.7
	6A		I††, II††	3-6.6	2.51	3.7
	6A		I††, II††	71.8-42.4	2.65	3.9
	6B		II†	155.9	3.04	4.6
	7A		II†	75.2	3.38	5.2

[a] I, Prosper trial; II, Langdon trial; III, unbalanced combined data set where an SNP marker was detected above the pFDR value

[c] cM, marker position on the consensus map

† SNP marker that in that trial (environment) was detected above $-\log_{10}$ (*P* value) of 3, but below the pFDR value

†† SNP marker that in that trial (environment) was detected above $-\log_{10}$ (*P* value) of 2, but below the pFDR value

¶R^2, proportion of phenotypic variation explained by the individual marker

A total of 48 QTL were identified for nine color traits measured in both datasets. 'Color' represents spaghetti color and was only measured in the unbalanced dataset. Two QTL for 'color' on chromosomes 4A and 4B were also detected for col_a, col_b, dif_b, and TYP measured in the balanced dataset (Table 3-10). A consistent common QTL for TYP and col_b was detected on chromosome 7A at 180-181.8 cM. One QTL on chromosome 7B was not only detected for TYP and col_b, but also for dif_b and the 'color', suggesting its involvement in color degradation during processing.

Table 3-10. Genetic location and single nucleotide polymorphism (SNP) loci significantly associated with color-related traits in durum wheat across three datasets.

Traits and QTL	Chr	Other associated traits	Trials[a]	Position	-log10 (P-value)	R^2¶
Color_a						
	1A	PPO	II††	6.6	2.69	4.0
	3A		II†	20.9	3.31	5.1
	3B		I	88.2-88.4	5.65	9.2
	4A	Color_b, color, Dif_b	I†, II†	139.2-143.7	4.73	7.6
	4B	Col_b24, color, dif_b, HT, SK	I†	22.5-28.8	3.64	5.7
	6A	Color_L, Dif_b	I†	124-125.6	4.51	7.2
Color_b						
	2B	TYP	II††	6.6-8.3	2.14	3.1
	4A	Color_a, color, dif_b	II††	139.7	2.50	3.7
	5A		I††, II††	52.9	2.46	3.6
	6A		I††, II††	124.8-129.4	2.50	3.7
	7A	TYP, Dif_b	I†, II††	180.3-181.8	3.31	5.1
	7B		I†, II††	195.9-196.5	4.37	6.9
Color_L						
	2A		I††, II††	189.8	2.91	4.4
	5A		I†	148.8	3.10	4.7
	6A		I††, II††	0.1-3.1	2.28	3.3
	6A	Color_a, Dif_b	I†	124.8	3.25	5.0
Color						
	4A		III†	23.7-25.6	3.04	4.7
	4A	Color_a, color_b, Dif_b	III†	139.7	3.56	5.6
	4B	HT, SK, color_a, color_b, Dif_b, TKW, VIT, YLD, LK, MK	III†	26.4-28.8	3.25	5.1
	6A		III†	90.6-95.9	3.13	4.9

Table 3-10. Genetic location and single nucleotide polymorphism (SNP) loci significantly associated with color-related traits in durum wheat across three datasets (continued).

Traits and QTL	Chr	Other associated traits	Trials[a]	Position	-log10 (P-value)	R²¶
	7B		III†	124.1	3.05	4.7
Difference in color a						
	1B		II†	87.8-89.1	3.70	5.8
	1B		II†	109-109.8	3.02	4.6
	2B		I†	171.1	3.33	5.1
	5A		II†	26.2-27.9	4.1	6.6
	6A	-	I	71.8-72.4	3.81	6.0
Difference in color b						
	3B		I††, II††	86.4-89.4	2.29	3.3
	4A		I†, II††	159.5	3.53	5.5
	4B	HT, SK, TKW, LK, MK, TWT, VIT, YLD	I†, II††	18.4-28.8	3.37	5.2
	4B		I††, II††	115.5	2.87	4.3
	5A		I††, II††	26.2-26.5	2.42	3.6
	6A	Color_a, Color_L	I†	124.8-126.5	3.23	5.0
	7B		I††, II††	112.5	2.77	4.2
	7B		I†, II†	120.4-123.2	3.16	4.8
	7B		I†, II††	138.3-140.4	3.54	5.5
Difference in color L						
	2B		I†, II†	17.7-19	4.22	6.7
	2B		I††, II††	183.1-189	2.48	3.7
	7A		II†	90.9-91.8	3.06	4.7
	7A		II†	184.1	3.16	4.8
	7B	PPO	II†	13.8	3.69	5.8
Total yellow pigment						
	2B	Color_b	I††, II††	8.3	2.27	3.4
	3B		I†	88.2-88.4	3.28	5.0
	4A		I††, II††	139.2	2.72	4.1
	7A	Color_b, Dif_b	I†, II††	180.3-181.8	3.22	5.0
	7B		I††, II††	132.9	2.12	3.0
	7B	-	I††, II††	187.5	2.54	3.8
	7B		I††, II††	196.5	2.39	3.5
Polyphenol oxidase activity						
	1A	A	I, II†	6.6	5.14	8.3
	1B		I††, II††	150.9-152	2.61	3.9
	2B		I, II†	113.2-118.7	5.14	8.3
	2B		I†, II	120.2-124.9	6.67	11.0

Table 3-10. Genetic location and single nucleotide polymorphism (SNP) loci significantly associated with color-related traits in durum wheat across three datasets (continued).

Traits and QTL	Chr	Other associated traits	Trials[a]	Position	-log$_{10}$ (P-value)	R^2¶
	2B		I††, II††	131.6	2.18	3.1
	3A		I†, II†	149.5-154.3	4.02	6.3
	3A		I, II	163.8-167.4	7.49	12.3
	3A		I, II	169.5-174.4	9.03	15.0
	3A		I, II	176.6-179.7	7.49	8.4
	3A		I, II	183.8-184	7.49	12.3
	3B		I, II	190.4	9.03	15.0
	3B		I, II	198.5-205.1	9.03	15.0
	5A	-	I††, II††	136.3-141.4	2.48	3.6
	5A		I††, II†	167.1-167.4	3.39	5.2
	5B		I†, II	63.4	5.95	9.7
	6A		I†, II†	105.7	3.84	6.0
	6B		I†, II†	27.1	5.15	8.3
	7B	Dif_L	I††, II††	13.8-15	2.77	4.2

[a] I, Prosper trial; II, Langdon trial; III, unbalanced combined data set where an SNP marker was detected above the pFDR value

[c] cM, marker position on the consensus map

† SNP marker that in that trial (environment) was detected above –log$_{10}$ (*P* value) of 3, but below the pFDR value

†† SNP marker that in that trial (environment) was detected above –log$_{10}$ (*P* value) of 2, but below the pFDR value

¶R^2, proportion of phenotypic variation explained by the individual marker

Pasta cooking parameters, such as cooked firmness (FIRM), cooking loss (CLOSS), work to shear (WTS), and cooked weight (CWT) were identified on fresh pasta in the balanced dataset from both the Langdon and Prosper locations grown in 2015. Pasta cooking parameters were taken on dry spaghetti in the unbalanced dataset. A total of 31 QTL were identified for cooking-related traits (Table 3-11). Nine QTL were identified in the unbalanced dataset only, 10 QTL at the Prosper and two QTL at the Langdon locations only, while six QTL were identified in both the mean balanced and unbalanced datasets. Five major QTL (R^2>10) were identified for CWT at the Prosper location only.

Table 3-11. Genetic location and single nucleotide polymorphism (SNP) loci significantly associated with pasta cooking-related properties across three datasets.

Traits and QTL	Chr	Other associated traits	Trials[a]	Position	-log$_{10}$ (P-value)	R^2¶
Firmness						
	1A	SDS, GI, CWT	III†	1.3-5.2	3.27	5.1
	1B	CWT, GI, MIXO, SDS	II††, III†	3-8.5	4.93	8.1
	3A	CWT	I††, II††	7.3-9	2.71	4.1
	3B	CWT, CLOSS	I†, II†	4.2-6	3.75	5.9
	3B	CWT, TWT, SK	I††, II††	79.1-86.9	2.51	3.8
	5A	CWT	III†	113.7	4.10	6.6
	6A	SPROT, WG	III†	113.5	4.24	6.9
	6B	TKW	III†	92.6	3.63	5.8
	7A	CWT, TYP, B, DIF_B, DIF_L	I††, II††	180.3-184.1	2.82	4.2
	7B	HT, TKW, SDS, WPROT	I††, III††	62.2-67.3	2.90	4.5
Cooking loss						
	2A		II†	151.2-154.6	3.99	6.3
	3A		I†	21.3	4.21	6.7
	3B	WTS	I††, II†	7.1-9.6	3.75	5.9
	4A		I††, III††	129.3, 134.6	2.45	3.6
	4B		III†	22.5-28.8	3.27	5.1
	4B	-	I	105.5-106	3.85	6
	6A	SDS	III†	68.3	3.73	6
	6A		II††, III††	127.1-130.0	2.56	3.8
	6B		II†	65.1-65.5	3.00	4.6
Work to shear						
	2B	-	I††, II††	153.4	2.32	3.4
	3B	CWT, FIRM	I†	6	4.43	7
	5A	FIRM	I†	114.9-115.8	3.20	4.9
	7A	FIRM, CWT	II†	184.1	3.02	4.6
Cooked weight						
	1B	FIRM, GI, MIXO, SDS	I†, III†	1.3-8.5	4.33	6.9
	1B	FIRM, GLUT, GI, SDS	III†	15.2	4.7	7.7
	1B		I	27.6	6.11	10
	1B		III†	50.3-54.8	3.47	5.5
	2A		III	143.2	4.90	8
	2B		I†	44.5-44.7	3.88	9.1
	2B	-	I†	53.4-56.4	3.40	5.3
	2B		I†, III†	80.6-84	4.77	7.8
	2B		I	124.5-129.8	8.17	14
	3A	CLOSS	I†, III††	9-12.5	3.85	6

Table 3-11. Genetic location and single nucleotide polymorphism (SNP) loci significantly associated with pasta cooking-related properties across three datasets (continued).

Traits and QTL	Chr	Other associated traits	Trials[a]	Position	-log10 (P-value)	R²¶
	3A		I	165.9-172.5	7.32	12
	3B		I	79.1	8.14	13
	3B	SMALL	I	86.6-89.4	6.64	11
	5A	-	I	134.5-140.6	3.33	5.1
	5A		I††, III††	153.6-156.5	2.61	3.9
	6B	FIRM, TKW	I†	92.6	3.37	5.2
	7A		I†	148.1	3.41	5.3
	7A		III†	180.3	3.26	5.1
	7B		I††, III††	62.7-65.5	2.47	3.6

[a] I, Prosper trial; II, Langdon trial; III, unbalanced combined data set where an SNP marker was detected above the pFDR value
[c] cM, marker position on the consensus map
† SNP marker that in that trial (environment) was detected above $-\log_{10}$ (*P* value) of 3, but below the pFDR value
†† SNP marker that in that trial (environment) was detected above $-\log_{10}$ (*P* value) of 2, but below the pFDR value
¶R^2, proportion of phenotypic variation explained by the individual marker

A high genetic variation (R^2 ranged between 7% and 10%) was identified on chromosomes 1B for FIRM and CWT and 2B for CWT in both the mean balanced and unbalanced datasets (Table 3-11). Differences in spaghetti type (dry vs. fresh) during cooking might account for the low number of shared QTL between the two datasets.

Pleiotropic QTL

Co-localized or closely linked QTL may help in improving several traits simultaneously when desirable alleles for each trait are contributed by the same parent. Many QTL for agronomic traits showed a pleiotropic effect and were associated with quality traits, including kernel size and weight, semolina ash content, PPO, semolina/spaghetti color, and gluten strength (Tables 6-11). Eighty out of 106 QTL for agronomic-related traits co-localized with other agronomic- and quality-related traits.

The genomic region on chromosome 4B harbored the largest number of co-localized QTL associated with multiple traits, including YLD, HT, LK, MK, SK, TWT, TKW, VIT, LDG, WG, and 'color' (Tables 6-11). Six out of eight genomic regions associated with YLD were also associated with kernel characteristics, especially the SK distribution trait. Only one minor QTL was detected for YLD associated with GI and WG in both the mean balanced and unbalanced datasets. Genomic regions on chromosomes 4A and 4B for TWT and TKW were associated with each other, as well as with HT, VIT, and YLD (Tables 6 and 7).

Three genomic regions on chromosomes 5B, 6A, and 7A were associated with WPROT, SPROT, WG, and FIRM. Three genomic regions on chromosomes 1A and 1B were associated with gluten strength, CWT, and FIRM. Considering that two independent QTL for GI and WPROT were both associated with FIRM may suggest that both gluten quantity and quality affect pasta FIRM. Out of 24 QTL for five cooking-related traits, six genomic regions harbored QTL for GI, SDS, and MIXO, while two harbored QTL for WPROT and WG.

Color-related traits were mainly measured in the balanced dataset, with only 'color' being measure in the unbalanced dataset. Two genomic regions on chromosomes 4A and 4B were pleiotropic and associated with color-related, agronomic, and kernel characteristic traits. Two genomic regions harbored QTL for TYP and color_b.

Discussion

Ongoing improvement of the agronomic and quality performance of genotypes is always a breeding objective in wheat breeding programs worldwide. Identifying genetic regions and dissecting the genetic makeup of key traits can enable wheat genotypes to be improved using MAS, which in turn can allow for a more efficient utilization of resources in breeding operations.

Agronomic-related traits

Six parameters related to the agronomic performance of durum wheat were studied, including YLD, DTH, HT, LDG, LD, and FN. Many studies have investigated the variation in DTH and flowering time due to their importance to plant adaptability in different growing environments (Zhang et al., 2009; Xu et al., 2005). In the present study, the small number of QTL that explained the correspondingly small amount of phenotypic variation suggests a low genetic variation for DTH. It also suggests that favorable genotypic variation is already fixed in breeding populations over time. The QTL for DTH identified on chromosome 5A were near the QTL for earliness p*er se* reported by Kato et al. (2002) and Chu et al. (2008) and associated with earliness *per se* genes that modify heading date independent of vernalization and photoperiod (Hoogendoorn, 1985). A major QTL for DTH was earlier reported on homologous group 2 (Heidari et al., 2012) near the QTL identified in the present study on chromosome 2B at 114.6-121.2 cM (Table 3-6). The significant negative correlation between DTH and LD may suggest that later heading genotypes avoid foliar diseases better. The lack of correlation between YLD and DTH and a single common genomic region harboring QTL for DTH and YLD may suggest that a nearly optimum balance for earliness and YLD potential has been attained in the NDSU durum wheat breeding program.

Several plant and grain characteristics controlled by several genes and highly influenced by environmental conditions determine grain YLD in wheat. Pfeiffer et al. (2000) suggested that a nearly optimal balance in YLD components had been reached in modern elite durum wheats under optimal growing conditions. Pleiotropic QTL on chromosomes 3A and 4B were identified for grain YLD and kernel characteristics, which were expected due to the positive correlation of YLD with TWT, TKW, and kernel size distribution. Semi-dwarf wheat varieties, with their low

susceptibility to LDG and high tillering ability and harvest index, were a major factor in the success of wheat production using a high fertilizer input (Brancourt_Hulmel et al., 2003). A major genomic region on chromosome 4B (17.7-28.8 cM) was identified near the known major dwarfing gene *Rht1* (Gale et al., 1995; Gale and Youssefian, 1985; Pearce et al., 2011). The *Rht1* gene is gibberellic acid-insensitive (GA), meaning that dwarf mutants of this type show a reduced response or complete insensitivity to GA application (Verma et al., 2005). The observed association between lower yield and reduced height was in line with other studies (Brandle and Knott, 1986; Singh et al., 2001).

In previous studies, grain yield QTL were reported on all chromosomes with the exception of chromosomes 3D and 5D (Huang et al., 2006; Kuchel et al., 2007; Kumar et al., 2007). The stable QTL on 3A in the present study was near the QTL for TKW, HT, kernel number per spikelet, and spike number per square meter identified by Shah et al. (1999) and Borner et al. (2002); this makes it a suitable marker for MAS for YLD.

Independent QTL for LDG on chromosomes 1B, 2B, and 5B suggest that development of LDG-tolerant genotypes could be possible without any negative effect on other traits. Keller et al. (1999) identified a QTL associated with LDG and culm stiffness on chromosome 5B at the 124-136 cM interval, which is near QTL for LDG identified in the present study.

The plant materials showed significant variation for FN. The QTL on chromosome 6B for FN was near an earlier-reported QTL for seed dormancy (Singh et al., 2010) and late maturity, α-amylase (LMA) (Kunert et al., 2007), which corresponds to the *Amy-1B* gene (Emebiri et al., 2010). QTL on chromosomes 1A and 5B for FN were also near QTL for seed dormancy reported in previous studies (Singh et al., 2010; Zhang et al., 2014).

Identification of major and minor QTL for agronomic-related traits and grain characteristics confirm their polygenic nature (Sun et al., 2009; Wu et al., 2015; Kumar et al., 2016). The identification of QTL for grain weight and size on all chromosomes except 5A also shows their wide distribution as reported in previous studies (Peng et al., 2003; Elouafi and Nachit, 2004; Blanco et al., 2006; Kumar et al., 2016). Major QTL associated with TWT, KWT, LK, MK, and SK were identified on chromosome 4B (17.7-28.8 cM). The same region was associated with HT and YLD, suggesting the pleiotropic effect of the dwarfing gene *Rht-B1*, which is located near the detected QTL (Cutbert et al., 2008; Singh et al., 2001). This genomic region on chromosome 4B was found highly syntenous to rice chromosome 3 (Kumar et al., 2016), which harbors two cloned genes for grain length, width, and YLD (Fan et al., 2006; Mao et al., 2013; Zhang et al., 2012).

Quality-related traits

Grain quality (WPROT and VIT), flour and semolina quality, and pasta cooking quality (FIRM, CWT, CLOSS, and WTS) parameters are discussed in the following sections.

Grain quality

Grain protein and kernel vitreousness (VIT) are important grain characteristics associated with many grain, flour, semolina, and pasta quality traits, including milling and pasta firmness. A high percentage of vitreous kernels maximizes semolina yield (Dexter et al., 1989). Protein content, which is one the most important quality traits in durum wheat, showed a positive correlation with kernel vitreousness in this study, which corresponds with the results of other studies (Sissons, 2004; Bilgin et al., 2010; Sieber et al., 2015). Vitreous areas of the endosperm are known to be higher in protein than mealy ones (Matsuo and Dexter, 1980; Dexter et al., 1994).

QTL for grain protein were previously identified on all chromosomes (Lie et al., 2012; Bogard et al., 2013; Echeverry-Solarte et al., 2016; Kumar et al., 2017). A limited number of QTL for grain protein in this study may be due to the low genetic diversity within the genotypes. The QTL on chromosome 7A was located near a QTL identified in earlier studies (Groos et al., 2003; Prasad et al., 2003, Sun et al., 2010). QTL on chromosomes 5B and 7B were not previously reported in durum wheat and may be novel for grain protein. No co-locations between QTL for grain yield and grain protein were observed in this study, suggesting the possibility of independent improvement of these traits (Blanco et al., 2012). Such QTL could be important for improving grain protein through MAS without an impact on yield.

One common QTL for VIT and WPROT was identified on chromosome 5B at 140.5-143.7 cM. There was a positive phenotypic correlation between the two traits in all environments, corresponding with the results of previous studies (Sissons, 2004; Pagnotta et al., 2005; Bilgin et al., 2010; Sieber et al., 2016). The relationship, however, remains controversial in the literature (Autran et al., 1986; Longin et al., 2013; Pinheiro et al., 2013).

Identification of major and minor QTL for traits affecting milling quality parameters confirm their polygenic nature (Kumar et al., 2016; Sun et al., 2009; Wu et al., 2015). The majority of the QTL were associated with two or more milling-related characteristics, suggesting that these loci either have a pleiotropic effect or are tightly linked (Kumar et al., 2016; Russo et al., 2014; Hessler et al., 2002). Similar results were reported in other studies and were also expected based on phenotypic analysis (Fig. 3-1).

Durum wheat flour- and semolina-related characteristics

Milling quality. The aim of the durum wheat grain milling process is to maximize semolina and minimize flour production through successive steps of grinding and sieving

(Posner, 2009). The process is complex as it depends on different factors, such as the moisture content of the grain, impurities and broken durum wheat kernels, the size and texture of the grain, and grain protein content (Gonzales, 1995).

Durum wheat flour- and semolina-related characteristics in this study included three traits: TEXT, SEXT, and SASH. Test weight and kernel weight are known to be positively correlated with SEXT; however, in this study, no significant correlation between these traits was observed (Marshall et al., 1986; Matsuo and Dexter, 1980; Bilgin et al., 2010). This could be due to the small genetic variation for SEXT in this study (Table 3-3). All of the QTL for TEXT were identified for SEXT as well. A QTL for semolina yield was identified earlier in the same region on chromosome 7A using diverse germplasm from around the world (Clarke et al., 2008). The QTL on chromosome 1B could be the same as the QTL reported by Zhang et al. (2008). The QTL on chromosomes 2B and 6B were not previously reported and could be novel.

Semolina ash is a complex trait; however, a decrease in grain weight always results in lower extraction rates or increased ash content (Breseghello and Sorrells, 2007; Brevis et al., 2010). Ash content affects pasta color and has been associated with LDG and FN on chromosome 1A and change in color a (Dif_a) on chromosome 5B. QTL for SASH on chromosomes 1B and 6A were also reported in earlier studies (Zhang et al., 2008).

Protein- and dough-related traits. The quantity and quality of gluten are considered the most important parameters for pasta production. Therefore, to maintain and enhance market share, selecting for gluten strength is one of the primary criterion in durum wheat breeding programs. This requires an understanding of the genetic control of this trait and an ability to exploit the possibilities for using MAS.

The protein- and dough-related characteristics measured in this study included semolina protein (SPROT), gluten index (GI), wet gluten (WG), sedimentation volume (SDS), mixogram score (MIXO), and glutograph (GLUT). Gluten index (GI), SDS, and MIXO were highly correlated in the present study (Fig. 3-1). While QTL for gluten strength in both tetraploid and hexaploid wheat have been identified on most wheat chromosomes, the major and most consistent QTL across environments are located on chromosome 1B (Conti et al., 2011; Patil et al., 2009; Elouafi et al., 2000; Kumar et al., 2013). This corresponds with the present study, where two consistent QTL measured with GI, SDS, and MIXO (explaining 7 to 11% of the phenotypic variation) and located 15 cM apart were detected on the distal part of the long arm of chromosome 1B. In addition, the other QTL for gluten strength on chromosome 1A explains about 6% of the phenotypic variation. Previous studies show that the high molecular weight glutenin subunits (HMW-GS) are particularly important for determining dough elasticity and correlated positively with dough baking quality (Anjum et al., 2007). Group 1 chromosomes harbor genes for glutenins subunits: HMW-GS loci (Glu-A1, Glu-B1, and Glu-D1) on their long arms (Payne and Lawrence, 1983) and LMW-GS loci (Glu-A3, Glu-B3, and Glu-D3) on their short arms (D'Ovidio and Masci, 2004). In the present study, QTL for GI, SDS, and MIXO detected on homologous chromosomes 1 were also co-located with QTL for FIRM. Although QTL for gluten strength parameters on chromosomes 3B and 7B showed a minor contribution toward phenotypic variation (3-6%), they were consistent across environments and datasets. Conti et al. (2011) also reported QTL for gluten strength on chromosome 3B. The QTL on chromosomes 6A and 7B were earlier reported by Patil et al. (2009) and Kumar et al. (2013).

The majority of the earlier studies identified QTL for protein quality and gluten strength using mainly the SDS method and a bi-parental mapping population. Association mapping

population and use of SNP markers can identify the location of the QTL for gluten strength more precisely.

In the present study, QTL for WG and GLUT were identified on chromosomes 1B, 2A, 2B, 6A, and 7A. They shared more common QTL for semolina protein content (SPROT) than gluten strength (GI, SDS, MIXO), suggesting the importance of protein quantity and quality for dough strength and pasta production (D'Egidio et al,, 1990; Sissons, 2008). Lower heritability of these traits suggests influence by the environment, which was also demonstrated by the inconsistency in QTL identification across datasets.

Color-related traits. Eight color-related traits were measured in the present study, including dough color $a^{\dagger}$, color $b^{\dagger}$, color $L^{\dagger}$ before and after 24 hours, indicated pigment loss (dif_a, dif_b, dif_L), dried pasta color ('color'), total yellow pigment (TYP) in whole-wheat flour, and polyphenol oxidase (PPO) in whole wheat flour. As expected, color-related traits were inherently correlated. Four overlapping QTL were identified for TYP and dough color $b^{\dagger}$ on chromosomes 2B and 7A. Dough color $a^{\dagger}$, dough color $b^{\dagger}$, pigment loss measured as difference in dough color $b^{\dagger}$ after 24 hours, and overall spaghetti color had overlapping QTL on chromosomes 4A, 4B, and 7A. Essentially all the QTL for dough color $b^{\dagger}$, $a^{\dagger}$, ,and TYP had been identified earlier (Hessler et al., 2002; Carrera et al., 2007; Garbus et al., 2009; Verlotta et al., 2010; Parker et al., 1998; Mares and Cambell, 2001; Diaye et al., 2017; Poznial et al., 2012 Pozniak et al., 2007; Zhang and Dubkovsky, 2008). Major QTL for yellowness were detected on chromosomes 7A and 7B for both common (Parker et al., 1998; Mares and Cambell, 2001) and durum wheat (Zhang et al., 2008).

In durum wheat, chromosome 4B has two Lpx-1 genes, *Lpx-B1.1* and *Lpx-B1.2*. Previous studies show that deletion of *Lpx-B1.1* is associated with a carotenoid pigment

degradation during pasta processing (Hessler et al., 2002; Carrera et al., 2007; Garbus et al., 2009; Verlotta et al., 2010). The semi-dwarfing gene *Rht-B1b* was found linked to *Lpx-B1.1* in durum wheat (Peng et al., 1999) (Diaye et al., 2017; Poznial et al., 2012). This study found the same QTL associated with pigment loss and HT on chromosome 4B in the region between 18.4 to 28.8 cM. Markers for pigment loss on chromosome 4B did not show an association with dough color b† and TYP, confirming that *Lpx-B1.1* deletion has an effect on LOX activity during processing, but not on initial semolina or pasta color (Borelli et al., 1999; Carrera et al., 2007).

The QTL on chromosomes 6A, 7A, and 7B showed no association with differences in grain size or shape, suggesting a more direct effect on the accumulation of carotenoid pigments. In tetraploid wheat, the distal region on chromosomes 7A and 7B has been associated with TYP (Pozniak et al., 2007; Zhang and Dubkovsky, 2008). QTL for color b† and TYP on chromosome 7B were in the same vicinity as the previously-reported *Phytoene synthase 1* locus (PSY-B1) (Pozniak et al., 2007).

For pasta color a† (green-red chromaticity), the four loci detected on chromosomes 1B, 2B, 4A, and 4B suggest complex genetic control of this trait. Considering the overlapping QTL on chromosome 3B, 4A, and 4B with QTL for color, col_b, and dif_b, as well as the negative correlation between dough color a† and b† (R=-0.26) and pigment loss as measured by color a† and b† (R=-0.63), a genetic linkage between these two traits could be suggested. Therefore, much effort should focus on breaking the LD to facilitate selecting against redness in dough color. This study's findings and those by Diaye et al. (2017) support the undesirable association between pasta redness and pasta yellowness. The positive correlation between dif_b and dif_L (r=0.31), as well as the negative correlation between dif_L and dif_a (-0.41) and dif_b and Dif_a (r=-0.63) and a single QTL on chromosome 6A associated with col_b and dif_b, may suggest an

indirect masking effect of col_a on col_L by directly influencing col_b, especially in semolina dough over time.

Previously-reported QTL for polyphenol oxidase (PPO) activity were independent from other color-related traits, suggesting that its effect on pasta quality is due to a browning reaction rather than the influence of semolina color components (Zhai et al., 2016). Two QTL were co-localized for PPO and dif_L on chromosome 7B and color_a on chromosome 1A. Correlation of the PPO trait with other agronomic or quality traits was not significant; however, suggesting a minor effect of co-localized QTL on color-related traits. A major QTL for PPO on chromosome 2B was located in a proximate region compared to the earlier-reported QTL by Beecher et al. (2012) and Si et al. (2012). In hexaploid wheat, PPO activity is mainly controlled by the genes located on chromosomes 2A and 2D (Zhang et al., 2005; He et al., 2007; Wang et al., 2009). Previously-reported major QTL for PPO on chromosome 2A in tetraploid wheat could not be identified in the present study (Watanabe et al., 2006). The major QTL for PPO on 3A and 3B seem to be novel and could be attributed to different sources of germplasm used in this study.

Cooking-related traits

QTL for four cooking-related traits (FIRM, CLOSS, WTS, and CWT) were overlapping with QTL for gluten strength on chromosome 1B and WPROT on chromosome 7B. Zhang et al. (2008) reported QTL for mixograph peak height and width near the QTL for firmness and cooking loss on chromosome 1B. Whole-wheat protein, GI, and SDS are positively correlated with pasta firmness and inversely correlated with CLOSS. Independent QTL for firmness on chromosomes 6A and 7A suggest that other parameters than protein quantity and quality affect pasta firmness and cooking loss. For instance, genes responsible for amylose synthesis are reported on chromosome 7A (Miura et al., 1999).

Applicability of using unbalanced historic phenotypic data for identification of QTL/marker trait associations

The collection of phenotypic data requires extensive effort and resources. Breeding programs collect large amounts of phenotypic data from advanced breeding lines every year for selection purposes. However, the number of such advanced breeding lines tested each year is relatively small, and some of these lines are replaced by other breeding lines, resulting in an unbalanced dataset of advanced breeding lines developed over time. Combining existing historic unbalanced dataset with affordable marker genotyping could be a reliable and useful method for genome-wide association studies (GWAS). This could save significant resources and provide useful information about marker trait associations in breeding programs.

The present study used data from 243 lines grown at different times over 16 years at five locations and found that unbalanced data could be efficiently used for QTL detection. Most of the major and minor QTL for agronomic and quality traits were identified using both the balanced and unbalanced datasets. Over 45% of QTL for agronomic traits were present in both the unbalanced and balanced datasets, with about 50% of those present in both locations in the balanced dataset. Similarly, over 50% of the quality traits measured were identified in both datasets. In most instances, the significance of the QTL in the unbalanced dataset (p-value) were commensurate with those in the balanced data set. These observations also corresponded with the phenotypic data, where all the correlations for agronomic and quality traits between the balanced and unbalanced datasets were above r=0.55 and over half of the traits above r=0.75 (SDS, GI, DTH, HT, TKW, MK, and LK). The correlations of traits related to cooking between the two datasets were a little bit lower, but still significant (r=20 for CLOSS and r=41 for FIRM),

which might be explained by the difference in cooking methods (dry spaghetti vs. fresh spaghetti cooking).

Mixograph data that measured gluten and dough strength were only available from the unbalanced dataset, however, QTL for MIXO detected in that dataset were located in the same genomic regions as those for gluten strength (measured with different approaches), which was present in both datasets.

A limited number of studies have compared the efficacy/accuracy of using unbalanced datasets for QTL detection with GWAS. Wang et al. (2012) estimated the optimum number of lines to be 384 and above for accurately predicting major and minor QTL. Dawson et al. (2013) used yield data collected from 168 locations over 17 years to show the usefulness of an unbalanced dataset for genomic prediction. While they used a unique set of genotypes each year, the genotypes in the present study overlapped. The phenotypic correlation of the traits was also comparable between the balanced and unbalanced datasets. Yield shows a significant positive correlation with HT, TWT, and kernel weight and a negative correlation with LD in both datasets. Gluten index was highly positively correlated with SDS and mixograph score in both datasets. In both datasets, the GI was positively correlated with FIRM, however, WG had a higher correlation with FIRM and WPROT. Based on correlations among traits in both datasets, common trends can clearly be observed.

Candidate markers for MAS

Applying markers in plant breeding enables selection of superior genotypes for the traits that are difficult and expensive to phenotype. It allows breeders to make decisions in early generations and advance superior genotypes in a timely manner. Marker assisted breeding allows breeders to hasten the process of transferring desirables alleles. In the present study,

some markers identified for gluten strength, HT, YLD, PPO, and SK were considered promising for MAS in the durum wheat breeding program.

Gluten strength

Gluten strength was measured using three different methods: SDS, GI, and mixogram score. Marker IWB70674 on chromosome 1B at 15.2 cM was significant and constant in all locations and both datasets. The pFDR value was low in the unbalanced dataset and LP dataset and present in both the Langdon and Prosper locations for GI and SDS. It was also detected by the mixogram score, which was only measured in the unbalanced dataset with a pFDR value less than 0.1. A single marker explained about 9-11% of phenotypic variation in the tests that measured gluten strength. The marker IWB70674 on 1B was most significant for gluten strength as it had a stable and major effect. The SNP IWB6234 detected for SDS could also be useful for MAS as it was detected in all datasets. Based on the information for these two markers, two haplotypes are present for gluten strength measured by SDS, GI, and the mixograph (Table 3-12).

Plant height

One of the strongest marker-trait associations was identified on chromosome 4B in the region 17.7 to 28.8 cM with markers in that genomic region being in high linkage disequilibrium. This genomic region was highly positively associated with YLD, TWT, TKW, and SK. The SNP marker IWB72203 had a pFDR value less than 0.01 in both datasets, as well as both locations, with the highest $R^2 = 19\%$; thus having the strongest association for height. The SNP marker IWA4773 on chromosome 6A also showed a low pFDR value and was consistent across environments. SNP marker IWB7419 on chromosome 1B had a slightly higher p-value; however, the $R^2 = 5\text{-}6\%$ and was consistent over environments.

Table 3-12. Phenotypic means and t-test p-values for population with various combinations of markers for gluten strength.

Marker applicable for MAS	Phenotype	Number of genotypes	Phenotypic mean	Range of the phenotypic trait	Actual nucleotide
IWB70626					
Unbalanced	High GI haplotype	179	60.2	22.8 - 92.5	G
	Low GI haplotype	42	43.9	9.2-69.5	T
t-test			**6.66236E-06**		
Mean PL					
	High GI haplotype	191	47.6	12.9 - 82.8	G
	Low GI haplotype	49	36.7	1.04-64.7	T
t-test			**5.44463E-05**		
IWB6234 and IWB70626					
Unbalanced	High GI haplotype	14	69.0	50.9-87.6	GC
	Low GI haplotype	35	44.8	9.2-69.5	TT
			7.45444E-07		
IWB70626					
Unbalanced	High SDS haplotype	187	52.3	39.0-67.4	C
	Low SDS haplotype	47	45.6	25.2-54.6	T
t-test			**1.15538E-12**		
Mean PL	High SDS haplotype	190	55.3	38.8-74.8	C
	Low SDS haplotype	49	50.1	23.5-65.5	T
t-test			**2.50915E-05**		
IWB6234 and IWB70626					
Unbalanced	High SDS haplotype	15	58.4	49.4-65.3	GC
	Low SDS haplotype	39	45.0	25.2-50.9	TT
t-test			**7.08108E-11**		
Mean PL	High SDS haplotype	15	66.3	57.3-74.8	GC
	Low SDS haplotype	39	49.0	23.5-61.8	TT
			5.93077E-10		
IWB70626					C
Unbalanced	High Mixo score	187	6.5	4.9-8.3	T
	Low Mixo Score	47	5.5	2.5-6.6	
t-test			**3.29702E-12**		
IWB6234 and IWB70626					
	High Mixo score	15	6.7	5.8-7.5	GC
	Low Mixo Score	39	5.4	2.5-6.6	TT
t-test			**2.758E-07**		

Table 3-13. Phenotypic means and t-test p-values for population with various combinations of markers for plant height.

Marker applicable for MAS	Phenotype	Number of genotypes	Phenotypic mean	Range of the phenotypic trait	Actual nucleotide
IWB72203					
Unbalanced	Tall plants	215	88.7	76.8-115.9	A
	Short plants	19	78.8	70.2-87.2	C
t-test			**1.362E-07**		
Mean PL					
	Tall plants	221	94.6	76.8-120.0	A
	Short plants	18	82.0	68.3-95.0	C
t-test			**3.11888E-06**		
IWA4773					
Unbalanced	Tall plants	230	88.0	70.2-116.0	A
	Short plants	4	78.1	75.0-81.7	C
t-test			**0.006898452**		
Mean PL	Tall plants	235	93.9	68.3-120.0	A
	Short plants	4	76.6	69.8-87.2	C
t-test			**0.019759805**		
IWB7419					
Unbalanced	Tall plants	5	99.6	90.6-116.0	G
	Short plants	229	87.6	70.2-99.6	A
t-test			**0.050899634**		
Mean PL	Tall plants	5	105.7	98-120.0	G
	Short plants	235	93.4	68.3-103.5	A
t-test			**0.088551171**		
IWB72203 and IWA4773					
Unbalanced	Tall plants	214	88.7	76.3-116.0	AA
	Short plants	3	76.9	75.0-79.7	CC
t-test			**0.012078517**		
Mean PL	Tall plants	220	94.6	76.8-120.0	AA
	Short plants	3	73	69.8-76.8	CC
t-test			**0.007815881**		

Using markers and their combinations (haplotypes), it is possible to select for height using MAS in cases where a desired height is needed or for prediction purposes.

Yield

Yield is a complex trait influenced by both genotype and environment. Markers on chromosomes 2A, 3B, and 5B in the present study were shown to be good candidates for MAS for YLD together improving it by 23% and 36 % in the combined and PL datasets, respectively. The yield trait on chromosome 2A was also associated with gluten strength; on chromosome 3B with HT, kernel size distribution, TWT, and TKW; and on chromosome 5B with LD, HT, and kernel distribution (Table 3-14).

Small kernel distribution

Three markers were significantly associated with the percentage of small kernel distribution (Table 3-15). The results for the unbalanced dataset were likely more pronounced due to a smaller genotypic effect and GxE interaction. Marker combinations used to identify low and high percent SK showed a lower t-test value as compared to using markers individually, suggesting the usefulness of several markers for MAS.

Table 3-14. Phenotypic means and t-test p-values for population with various combinations of markers for yield.

Marker applicable for MAS	Phenotype	Number of genotypes	Phenotypic mean	Range of the phenotypic trait	Actual nucleotide
IWB25863					
Unbalanced	High yields	3	63.2	60.7-66.1	C
	Low yields	231	56.9	42.8-67.1	T
t-test			**0.05612303**		
Mean PL					
	High yields	3	1753.3	1647.0-1869.4	C
	Low yields	236	1618.1	1049.5-2063.5	T
t-test			**0.167899275**		
IWB73293					
Unbalanced	High yields	230	57.1	48.0-67.2	T
	Low yields	4	49.1	42.8-54.9	C
t-test			**0.068940789**		
PL	High yields	235	1624.5	1183.5-2063.5	T
	Low yields	4	1337.7	1049.5-1565.5	C
t-test			**0.110380431**		
IWB40750					
Unbalanced	High yields	224	57.2	45.9-67.2	T
	Low yields	10	51.7	42.8-59.7	C
t-test			**0.00592538**		
Mean PL	High yields	229	1632.8	1315.9-2063.5	T
	Low yields	10	1321.5	1049.5-1729.9	C
t-test			**0.000375448**		
IWB25863 and IWB73293 and IWB40250					
Unbalanced	High yields	3	**63.2**	60.7-66.1	CTT
	Low yields	2	48.8	42.8-55	TCC
t-test			**0.238382221**		
Mean PL	High yields	3	1753.3	1647.01-1869.4	CTT
	Low yields	2	1122.31	1049.5-1195.1	TCC
t-test			**0.013666514**		

Table 3-15. Phenotypic means and t-test p-values for population with various combinations of markers for small kernel distribution.

Marker applicable for MAS	Phenotype	Number of genotypes	Phenotypic mean	Range of the phenotypic trait	Actual nucleotide
IWB7614 and IWB7940 and IWB72203					
Unbalanced	Low % SK	207	4.4	0.6-8.8	TAA
	All other lines	24	6.2	3.5-10.4	CGC
t-test			**4.33044E-05**		
Mean PL					
	Low % SK	215	3.2	1.0-10.5	TAA
	All other lines	24	6.6	2.9-18.5	CGC
t-test			**0.000283574**		

PPO activity

QTL strongly associated with PPO were identified on chromosomes 1A (IWA5150), 2B (IWA1488), 3A (IWB69399), and 3B (IWB23604) (Table 3-16). All SNP markers had low pFDR values. Due to the low minor allele frequency (MAF), the ability to detect QTL was low, but because the pFDR was low, this association is probably not a false positive.

Table 3-16. Phenotypic means and t-test p-values for population with various combinations of markers for PPO activity.

Marker applicable for MAS	Phenotype	Number of genotypes	Phenotypic mean	Range of the phenotypic trait	Actual nucleotide
IWA5150					
Mean PL	Low PPO	237	0.11	0.04-0.54	T
	All other lines	2	0.41	0.34-0.48	G
t-test			**0.140956287**		
IWA1488					
PL					
	Low PPO	235	0.10	0.04-0.54	A
	All other lines	4	0.43	0.038-0.048	G
t-test			**0.000149301**		
IWB69399					
Mean PL	Low PPO	229	0.10	0.038-0.048	A
	All other lines	10	0.38	0.06-0.54	G
			2.20191E-05		
IWB23604	Low PPO	231	0.10	0.038-0.490	T
Mean PL	All other lines	8	0.422	0.10-0.54	G
			0.000360083		
IWA1488 + IWB69399 + IWB23604					
Mean PL	Low PPO	228	0.95	0.038-0.480	AAT
	All other lines	11	0.43	0.10-0.54	GGG
			7.32631E-07		
IWA5150 + IWA1488 + IWB69399 + IWB23604					
Mean PL	Low PPO	225	0.09	0.04-0.50	TAAT
	All other lines	14	0.40	0.06-0.54	GGGG
t-test			**2.56311E-06**		

Conclusion

The present study combined a large amount of multi-environment unbalanced historic and balanced phenotypic data from durum wheat with high-density Infinium SNP marker data to identify marker-trait associations (MTAs) for agronomic and quality traits. The study also attempted to find out if the unbalanced data collected over the years could be useful for GWAS. Genome-wide association studies revealed that QTL/MTAs for agronomic and quality traits are distributed on all durum wheat chromosomes. A genomic region near Rht-B1 on 4BS showed a stable and pleiotropic effect on HT, TWT, TKW, kernel size distribution, and LDG and could be an important region for MAS to improve these traits. Haplotypes were proposed for MAS for HT and SK distribution; however, the TKW and TWT phenotypic distribution was not vast/diverse enough to identify candidate QTL for MAS. Markers associated with major QTL for gluten strength on chromosomes 1A and 1B; PPO activity on chromosomes 1A, 2B, 3A, and 3B; and yield on chromosomes 3A, 2B, and 5B could also be excellent candidates for MAS in durum wheat breeding programs. Yield is a highly complex trait, influenced by many factors. Genome-wide selection for improving yield is one of the promising approaches. Therefore, the haplotypes suggested for YLD improvement in this study can serve as a preliminary step for identifying promising genotypes in the early breeding stages through MAS. Common QTL detected in both the unbalanced historic and balanced datasets suggest the practicality of using unbalanced data from breeding trials for identifying MTAs. Breeders annually obtain phenotypic information on a large number of genotypes, which could be beneficial if combined to produce a large population size, eliminating the need to develop special population for identifying major QTL.

References

AACC. 2000. Approved Methods of the American Association of Cereal Chemists, 10th Ed., Methods (08–12), (38–12A), (44–15A), (46–30), and (54–40A), American Association of Cereal Chemists, St. Paul, MN.

American Association of Cereal Chemists. 1995. Approved Methods of the AACC, 9th ed. Method 08-01, approved April 1961, revised October 1981 and October 1986; Method 26-10A, approved April 1961, final approval October 1988, revised October 1994; Method 26-41, approved April 1961, revised October 1994; Method 44-40, approved April 1961, revised October 1976, October 1982, revised October 1994; Method 46- 30, approved November 1995; Method 66-41, approved October 1984, reviewed October 1994; Method 76-31, approved January 1995. Am. Assoc. Cereal Chem.: St. Paul, MN.

Anjum, F.M., M.R. Khan, A. Din, M.Saeed, I. Pasha, M.U. Arshad. 2007. Wheat gluten: high molecular weight glutenin subunits-structure, genetics, and relation to dough elasticity. J. Food Sci. 72:56-63.

Autran, J.C., J. Abecassis, P. Feillet. 1986. Statistical evaluation of different technological and biochemical tests for quality assessment in durum wheats. Cereal Chem. 63:390–394.

Beecher, B., D. Skinner. 2011. Molecular cloning and expression analysis of multiple polyphenol oxidase genes in developing wheat (Triticum aestivum) kernels. J. Cereal Sci. 53:371–378.

Beecher, B.S., A.H. Carter, D.R. See. 2012. Genetic mapping of new seed-expressed polyphenol oxidase genes in wheat (Triticum aestivum L.). Theor. Appl. Genet. 124:1463–1473.

Benjamini, Y., and D. Yekutieli, 2001. The control of the false discovery rate in multiple testing under dependency. Ann. Stat. 29:1165-1188.

Bhatt, G.M. 1973. Significance of path coefficient analysis in determining the nature of character association. Euphytica. 22: 338–343.

Bilgin, O., K.Z. Korkut, I. Baser, O. Daglioglu, I. Öztürk, A. Balkan, T. Kahraman. 2010. Variation and heritability for some semolina characteristics and grain yield and their relations in durum wheat (Triticum durum Desf.). World J. Agric. Sci. 6:301–308.

Blanco, A., G. Mangini, A. Giancaspro, S. Giove, P. Colasuonno, R. Simeone, A. Signorile, P. De Vita, A.M. Mastrangelo, L. Cattivelli, A. Gadaleta. 2012. Relationships between grain protein content and grain yield components through quantitative trait locus analyses in a recombinant inbred line population derived from two elite durum wheat cultivars. Mol. Breed. 30:79–92.

Blanco, A., R. Simeone, A. Cenci, A. Gadaleta, O.A. Tanzarella, E. Porceddu, S. Salvi, R. Tuberosa, G. Figliuolo, P. Spagnoletti, M.S. Röder, V. Korzun. 2004. Extension of the "Messapia x dicoccoides" linkage map of Triticum turgidum (L.). Thell. Cell Mol. Biol. Lett. 9:529–541.

Blanco, A., R. Simeone, A. Gadaleta. 2006. Detection of QTLs for grain protein content in durum wheat. Theor. Appl. Genet. 112:1195–1204.

Bogard, M., V. Allard, P. Martre, E. Heumez, J.W. Snape, S. Orford, S. Griffiths, O. Gaju, J. Foulkes, J.L. Gouis. 2013. Identifying wheat genomic regions for improving grain protein concentration independently of grain yield using multiple inter-related populations. Mol. Breed. 31:587–599.

Börner, A., E. Schumann, A. Furste, H. Coster, B. Leithold, M.S. Roder, W.E. Weber. 2002. Mapping of quantitative trait loci determining agronomic important characters in wheat (Triticum aestivum L.). Theor. Appl. Genet. 105:921–936.

Borrelli, G.M., A. Troccoli, N. Di Fonzo, C. Fares. 1999. Durum wheat lipoxygenase activity and other quality parameters that affect pasta color. Cereal Chem. 76:335–340.

Brancourt-Hulmel, M., G. Doussinault, C. Lecomte, P. Bérard, B.Le Buanec, and M. Trottet. 2003. Genetic improvement of agronomic traits of winter wheat cultivars released in France from 1946 to 1992. Crop Sci. 43: 37–45.

Brandle, J.E., and D.R. Knott. 1986. The effect of a gene for semidwarfism (Rht1) on various characters in a spring wheat cross. Can. J. Plant Sci. 66: 529–533.

Breseghello, F., and M.E. Sorrells. 2007. QTL analysis of kernel size and shape in two hexaploid wheat mapping populations. Field Crops Res. 101:172-179.

Brevis, H.C., C. Morris, F. Manthey, and J. Dubkovsky. 2010. Effect of the grain protein content locus Gpc-B1 on bread and pasta quality. J. Cereal Sci. 51:357-365.

Carrera, A., V. Echenique, W. Zhang, M. Helguera, F. Manthey, A. Schrager, A. Picca, G. Cervigni, J. Dubcovsky. 2007. A deletion at the Lpx-B1 locus is associated with low lipoxygenase activity and improved pasta color in durum wheat, Triticum turgidum ssp. durum. J. Cereal Sci. 45:67-77.

Chu, C.G., S.S. Xu, T.L. Friesen, and J.D. Farris. 2008. Whole genome mapping in a wheat doubled haploid population using SSRs and TRAPs and the identification of QTL for agronomic traits. Mol. Breeding. 22:251-266.

Ciaffi, M., S. Benedettelli, B. Giorgi, E. Porceddu, D. Lafiandra. 1991. Seed Storage Proteins of Triticum turgidum ssp. dicoccoides and Their Effect on the Technological Quality in Durum Wheat. Plant Breeding. 107:309–319.

Clarke, F.R., J.M. Clarke, C.J. Pozniak, D.J. Somers, J.E. Dexter, R.E. Knox, S. Reimer, T.N. McCaig, and A.K. Singh. 2008. Association mapping of semolina yield in diverse durum wheat germplasm.

Conti, V., P.F. Roncallo, V. Beaufort, G.L. Cervigni, R. Miranda, C.A. Jensen, V.C. Echenique. 2011. Mapping of main and epistatic effect QTLs associated to grain protein and gluten strength using a RIL population of durum wheat. J. Appl. Genet. 52:287-298.

Cuthbert, J.L., D.J. Somers, A.L. Brule-Babel, P.D. Brown, G.H. Crow. 2008. Molecular mapping of quantitative trait loci for yield and yield components in spring wheat (Triticum aestivum L.). Theor. Appl. Genet. 17:595–608.

D'Egidio, B.M. Mariani, S. Nardi, P. Novaro, R. Cubadda. 1990. Chemical and technological variables and their relationships. A predictive equation for pasta cooking quality. Cereal Chem. 67:275-281.

D'Ovidio, R., and S. Masci. 2004. The low-molecular-weight glutenin subunits of wheat gluten. Cereal Sci. 39:321-339.

Dawson, J.C., J.B. Endelman, N. Heslot, J. Crossa, J. Poland, S. Dreisigacker. 2013. The use of unbalanced historical data for genomic selection in an international wheat breeding program. Field Crops Res. 154:12–22.

Dexter, J.E., B.A. Marchylo, A.W. MacGregor, R. Tkachuk. 1989. The structure and protein composition of vitreous, piebald and starchy durum wheat kernels. Cereal Sci. 10:19–32.

Dexter, J.E., D.G. Martin, G.T. Sadaranganey, J. Michaelides, N. Mathieson, J.J. Tkac. 1994. Pre-processing: Effects on durum wheat milling and spaghetti-making quality. Cereal Chem. 71:10-16.

Dogan, R. 2009. The Correlation and path analysis for yield and some of yield components of durum wheat (Triticum turgidum var. durum) in west Anatolia conditions. Pak. J. Bot. 41:1081-1089.

Echeverry-Solarte, M., A. Kumar, S. Kianian, S. Simsek, M.S. Alamri, E.E. Mantovani, P.E. McClean, E.L. Deckard, E. Elias, B. Schatz, S.S. Xu, M. Mergoum. 2015. New QTL alleles for quality-related traits in spring wheat revealed by RIL population derived from supernumerary × non-supernumerary spikelet genotypes. Theor. Appl. Genet. 128:893–912.

Elouafi, I., and M.M. Nachit. 2004. A genetic linkage map of the Durum x Triticum dicoccoides backcross population based on SSRs and AFLP markers, and QTL analysis for milling traits. Theor. Appl. Genet. 108: 401–413.

Elouafi, I., M.M. Nachit, A. Elsaleh, A. Asbati, D.E. Mather. 2000. QTL-mapping of genomics regions controlling gluten strength in durum (Triticum turgidum L. var. durum). Options Mediterraneennes. 40:505-509.

Emebiri L.C., J.R. Oliver, K. Mrva, D. Mares. 2010. Association mapping of late maturity α-amylase (LMA) activity in a collection of synthetic hexaploid wheat. Mol. Breeding. 26:39-49.

Eyal, Z. 1971. The kinetics of pycnidiospore liberation in Septoria tritici. Can. J. Bot., 49:1095-1099.

Eyal, Z.E. 1981. Integrated control of Septoria diseases of wheat. Plant Dis. 65:763-768.

Gale, M.D., and S. Youssefian. 1985. Dwarfing genes in wheat. Pages 1-35 in: Progress in plant breeding. G.E. Russel ed. Butterworth and Co.: London, U.K.

Gale, M.D., Atkinson, M.D., Chinoy, C.N., Harcourt, R.L., Jia, J., Li, Q.Y., et al. 1995. Genetic maps of hexaploid wheat. In Proceedings of the 8th International Wheat Genetics Symposium, Beijing, China, 20–25 July, 1993. Edited by Z.S. Li and Z.Y. Xin. China Agricultural Scientech Press, Beijing, China. pp. 29–40.

Garbus, I., A.D. Carrera, J. Dubcovsky, V. Echenique. 2009. Physical mapping of durum wheat lipoxygenase genes. Cereal Sci. 50:67–73.

Ghosh, J.K., M. Delampady, and T. Samanta. 2006. An Introduction to Bayesian Analysis: Theory and Methods. Springer-Verlag, New York.

González, A.T.J. 1995. Milling process of durum wheat. Pages 43-51 in: Durum wheat quality in the Mediterranean region. N. Di Fonzo, F. Kaan, M. Nachit eds. CIHEAM: Zaragoza, Spain.

Grant, L.A., J.W. Dick, D.R. Shelton. 1993. Effects of drying temperature, starch damage, sprouting, and additives on spaghetti quality characteristics. Cereal Chem. 70:676–684.

Groos, C., N. Robert, E. Bervas, G. Charmet. 2003. Genetic analysis of grain protein content, grain yield and thousand-kernel weight in bread wheat. Theor. Appl. Genet. 106:1032–1040.

Gurung, S., S. Mamidi, J.M. Bonman, M. Xiong, G. Brown-Guedira, and T.B. Adhikari. 2014. Genome-wide association study reveals novel quantitative trait loci associated with plant immunity against major leaf spot diseases of spring wheat. PLoS ONE 9:e108179.

He, X.Y., Z.H. He, L.P. Zhang et al. 2007. Allelic variation of polyphenol oxidase (PPO) genes located on chromosomes 2A and 2D and development of functional markers for the PPO genes in common wheat. Theor. Appl. Genet. 115:47–58.

Hessler, T.G., M.J. Thomson, D. Benscher, M.M. Nachit, M.E. Sorrells. 2002. Association of a lipoxygenase locus, Lpx-B1, with variation in lipoxygenase activity in durum wheat seeds. Crop Sci. 42:1695–1700.

Huang, X.Q., S. Cloutier, L. Lycar, N. Radovanovic, G. Humphreys, J.S. Noll, D.J. Somers, P.D. Brown. 2006. Molecular detection of QTLs for agronomic and quality traits in a doubled haploid population derived from two Canadian wheats (Triticum aestivum L.). Theor. Appl. Genet. 113:753–766.

Kato, K., H. Miura, S. Sawada. 2002. Characterization of QEet.ocs-5A.1, a quantitative trait locus for ear emergence time on the chtomosome 5AL. Plant Breeding. 121:389-393.

Keller, M., C. Karutz, J.E. Schmid, P. Stamp, M. Winzeler, B. Keller, M.M. Messmer.1999. Quantitative trait loci for lodging resistance in a segregating wheat x spelt. Population. Theor. Appl. Genet. 98:1171-1182.

Knott, D.R., and G. Gebeyehou. 1987. Relationships between the lengths of the vegetative and grain filling periods and agronomic characters in three durum wheat crosses. Crop Sci. 27:857-860.

Kuchel, H., K.J. Williams, P. Langridge, H.A. Eagles, S.P. Jefferies. 2007. Genetic dissection of grain yield in bread wheat I QTL analysis. Theor. Appl. Genet. 115:1029-1041.

Kumar, A., S. Jain, E.M. Elias, M. Ibrahim. 2017. An overview of QTL identification and marker assisted selection for grain protein content in wheat. In: R.S. Sengar, A. Singh, eds. Eco-friendly Agro-biological Techniques for Enhancing Crop Productivity. Springer. (in press).

Kumar, N., P.L. Kulwal, H.S. Balyan, P.K. Gupta. 2007. QTL mapping for yield and yield contributing traits in two mapping populations of bread wheat. Mol. Breed. 19:163–177.

Kunert, A., A.A. Naz, O. Dedeck, K. Pillen, J. Leon. 2007. AB-QTL analysis in winter wheat: I. Synthetic hexaploid wheat (T. turgidum ssp. dicoccoides · T. tauschii) as a source of favourable alleles for milling and baking quality traits. Theor. Appl. Genet. 115:683–695.

Li, F., J. Cui, A.M. Ding, C.H. Zhao, X.Q. Wang, L. Wang, Y.G. Bao, X.L. Qi, X.F. Li, J.R. Gao, D.S. Feng, H.G. 2012. Wang, QTL detection of seven quality traits in wheat using two related recombinant inbred line population. Euphytica. 183:207–226.

Longin, C.F.H., A.-N. Sieber, J.C. Reif. 2013. Combining frost tolerance, high grain yield and good pasta quality in durum wheat. Plant Breed. 132:353–358.

Mamidi, S., S. Chikara, R.J. Goos, D.L. Hyten, D. Annam et al. 2011 Genome-wide association analysis identifies candidate genes associated with iron deficiency chlorosis in soybean. The Plant Genome 4:154–164.

Mares, D.J., and A.W. Campbell. 2001. Mapping components of flour and noodle colour in Australian wheat. Aust. J. Agric. Res. 52:1297–309.

Marshall, D.R., D.J. Mares, H.J. Moss, and F.W. Ellison. 1986. Effects of grain shape and size on milling yields in wheat. II. Experimental studies. Aust. J. Agric. Res. 37:331-342.

Matsuo, R.R., J.E. Dexter. 1980. Relationship between some durum wheat characteristics and semolina milling properties. Can. J. Plant Sci. 60:49-53.

Matsuo, R.R., J.E. Dexter, F.G. Kosmolak, and D. Leisle. 1982. Statistical evaluation of tests for assessing spaghetti-making quality of durum wheat. Cereal Chem. 59:222–228.

Miller, B.S., S. Afeork, Y. Pomeranz, B.L. Bruinsma, and G.D. Booth. 1982. Measuring hardness of wheat. Cereal Foods World. 27:61-64.

Miura, H., Araki, E., and Tarui, S. 1999. Amylose synthesis capacity of the three Wx genes of wheat cv. Chinese Spring. Euphytica. 108:91-95.

Mohammadi, M., P. Sharifi, R. Karimizadeh, M. Kazem, M.K. Shefazadeh. 2012. Relationships between grain yield and yield components in bread wheat under different water availability (dryland and supplemental irrigation conditions). Not. Bot. Hortic. Agrobo. 40:195-200.

Muqaddasi, Q.H., J.C. Reif, Z. Li, B.R. Basnet, S. Dreisigacker, M.S. Röder. 2017. Genome-wide association mapping and genome-wide prediction of anther extrusion in CIMMYT spring wheat. Euphytica. 213:73.

Nakaya, A., and S.N. Isobe. 2012. Will genomic selection be a practical method for plant breeding? Ann. Bot. 110:1303–1316.

NASS-National Agricultural Statistics Service. 2016. National statistics: wheat, durum, North Dakota. United States Department of Agriculture. http://usda.mannlib.cornell.edu/

Pagnotta, M.A., A. Blanco, A. Gadaleta, C. Fares. 2005. Functional determinants of grain quality. Pages 483-527 in: Durum Wheat Breeding – Current Approaches and Future Strategies. C. Royo, M.M. Nachit, N. Di Fonzo, J.L. Araus, W.H. Pfeiffer, G.A. Slafer, eds. Food Products Press: Binghampton, NY.

Parker, G.D., K.J. Chalmers, A.J. Rathjen, P. Langridge. 1998. Mapping loci associated with flour colour in wheat (Triticum aestivum L.). Theor. Appl. Genetics. 97:238–245.

Patil, R.M., M.D. Oak, S.A. Tamhankar, V.S. Rao. 2009. Molecular mapping of QTLs for gluten strength as measured by sedimentation volume and mixograph in durum wheat (Triticum turgidum L. ssp durum). Cereal Sci. 49:378-386.

Payne, P.I., and G.J. Lawrence. 1983. Catalogue of alleles for the complex loci, Glu-A1, Glu-B1 and Glu-D1 which code for high-molecular-weight subunits of glutenin in hexaploid wheat. Cereal Res. Comm. 11:29-35.

Peng, J., Y. Ronin, T. Fahima, M.S. Roder, Y. Li et al. 2003 Domestication quantitative trait loci in Triticum dicoccoides, the progenitor of wheat. Proc. Natl. Acad. Sci. 100: 2489–2494.

Peng, Z.S., Z.X. Su, K.C. Cheng. 1999. Characterization of dwarf trait in the tetraptoid wheat landrace, Aiganfanma. Wheat Inf. Serv. 89:7–12.

Pfeiffer, W.H., K.D. Sayre, and M.P. Reynolds. 2000. Enhancing genetic grain yield potential and yield stability in durum wheat. Pages 83– 93 in: Options Mediterraneennes, Serie A, Seminaires Mediterraneens. Vol. 40. C. Royo, M.M. Nachit, N. di Fonzo and J.L., eds. Araus. IAMZ: Zaragoza, Spain.

Pinheiro, N., R. Costa, A.S. Almeida, J. Coutinho, C. Gomes, B. Maçãs. 2013. Durum wheat breeding in Mediterranean environments - influence of climatic variables on quality traits. Emir. J. Food Agric. 25:962–973.

Posner, E.S. 2009. Wheat flour milling. Pages 119-152 in: Wheat chemistry and technology. K. Khan, P.R. Shewry, eds. American Association of Cereal Chemistry: St. Paul, MN.

Pozniak, C.J. Clarke, F. Clarke. 2012. Potential for detection of marker—trait associations in durum wheat using unbalanced, historical phenotypic datasets. Mol. Breeding. 2012:1-14.

Pozniak, C.J., R.E. Knox, F.R. Clarke, J.M. Clarke. 2007. Identifcation of QTL and association of a phytoene synthase gene with endosperm colour in durum wheat. Theor. Appl. Genet. 114:525–537.

Prasad, M., N. Kumar, P.L. Kulwal, M.S. Roder, H.S. Balyan, H.S. Dhaliwal, J.K. Roy, P.K. Gupta. 2003. QTL analysis for grain protein content using SSR markers and validation studies using NILs in bread wheat. Theor. Appl. Genet. 106:659–667.

Richards R.A.1992. The effect of dwarfing genes in spring wheat in dry environments. I. Agronomic characteristics. Austr. J. Agric. Res. 43:517–527.

Samaan, J., G.H. El-Khayat, F.A. Manthey, M.P. Fuller, and C.S. Brennan. 2006. Durum wheat quality: II. The relationship of kernel physicochemical composition to semolina quality and end product utilisation. Int. J. Food Sci. Technol. 41:47–55.

Scheet, P., and M. Stephens. 2006. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. Am. J. Hum. Genet. 78: 629–644.

Shah, M.M., K.S. Gill, P.S. Baenziger, Y. Yen, S.M. Kaeppler, H.M. Ariyarathne. 1999. Molecular mapping of loci for agronomic traits on chromosome 3A of bread wheat. Crop Sci. 39:1728–1732.

Sieber, A.-N., T. Wurschum, C.F.H. Longin. 2015. Vitreosity, its stability and relationship to protein content in durum wheat. Cereal Sci. 61:71-77.

Singh, R., M. Båga, M.P. Hucl, R. Chibbar. 2010. Identification of genomic regions associated with seed dormancy in white-grained wheat. Euphytica.174:391-406.

Singh, R.P., J. Huerta–Espino, S. Rajaram, J. Crossa. 2001. Grain yield and other traits of tall and dwarf isolines of modern bread and durum wheats. Euphytica. 119:241–244.

Sissons, M. 2008. Role of durum wheat composition on the quality of pasta and bread. Food. 2:75-90.

Sun, X.-Y., K. Wu, Y. Zhao, F.-M. Kong, G.-Z. Han, H.-M. Jiang, X.-J. Huang, R.-J. Li, H.-G. Wang, S.-S. Li. 2009. QTL analysis of kernel shape and weight using recombinant inbred lines in wheat. Euphytica. 165:615-624.

Suprayogi, Y, C.J. Pozniak, F.R. Clarke, J.M. Clarke et al. 2009. Identification and validation of quantitative trait loci for grain protein concentration in adapted Canadian durum wheat populations. Theor. Appl. Genet. 119:437–448.

Tadesse, W., F.C. Ogbonnaya, A.Jighly, M. Sanchez-Garcia, Q. Sohail, S. Rajaram, M. Baum. 2015. Genome-wide association mapping of yield and grain quality traits in winter wheat genotypes. PLoS One. 10:e0141339.

Verlotta, A., V.D. Simone, A.M. Mastrangelo, L. Cattivelli, R. Papa, D. Trono. 2010. Insight into durum wheat Lpx-B1: a small gene family coding for the lipoxygenase responsible for carotenoid bleaching in mature grains. BMC Plant Biol. 10:263-269.

Verma, V., A.J. Worland, E.J. Sayers, L. Fish, P.D.S. Caligari, J.W. Snape. 2005. Identification and characterization of quantitative trait loci related to lodging resistance and associated traits in bread wheat. Plant Breed. 124:234–241.

Wang, H., K.P. Smith, E. Combs, T. Blake, R.D. Horsley, and G.J. Muehlbauer. 2012. Effect of population size and unbalanced data sets on QTL detection using genome-wide association mapping barley breeding germplasm. Theor. Appl. Genet. 124:111–124.

Wang, X.B., C.X. Ma, H.Q. Si et al. 2009. Gene markers for grain polyphenol oxidase activity in common wheat. Mol. Breed. 23:163–170.

Watanabe, N., A.S. Akond, M.M. Nachit. 2006. Genetic mapping of the gene affecting polyphenol oxidase activity in tetraploid durum wheat. J. Appl. Genet. 47:201–205.

Wu, Q.H., Y.X. Chen, S.H. Zhou, L. Fu, and J.J. Chen. 2015. High-density genetic linkage map construction and QTL mapping of grain shape and size in the wheat population Yanda1817 × Beinong6. PLoS One. 12:e0118144.

Xu, X., G. Bai, B. Carver, and E. Shaner. 2005. A QTL for early heading in wheat cultivar Suwon 92. Euphyt. 146:233-237.

Yue, P., P. Rayas-Duarte, E. Elias. 1999. Effect of drying temperature on physicochemical properties of starch isolated from pasta. Cereal Chem. 76:541–547.

Zhang J., J. Chen, B.C. Bowman, K. O'Brien, J.M. Marshall, J.M. Bonman. 2014. Association mapping of Hagberg falling number in hard white spring wheat. Crop Sci. 54:1243–1252.

Zhang, K., J. Tian, L. Zhao, B. Liu, G. Chen. 2009. Detection of quantitative trait loci for heading date based on the doubled haploid progeny of two elite Chinese wheat cultivars. Genetica. 135:257-265.

Zhang, L.P., X.X. Ge, Z.H. He, D.S. Wang, J. Yan, X.C. Xia, M.W. Sutherland. 2005. Mapping QTLs for polyphenol oxidase activity in a DH population from common wheat. Acta Agron. Sin. 31:7–10.

Zhang, W., J. Dubcovsky. 2008. Association between allelic variation at the Phytoene synthase 1 gene and yellow pigment content in the wheat grain. Theor. Appl. Genet. 116:635–45.

Zhang, Z., E. Ersoz, C. Lai, R. Todhunter, H. Tiwari, M. Gore, P. Bradbury, et al. 2010. Mixed linear model approach adapted for genome-wide association studies. Nature genetics. 42: 355-360.

APPENDIX

Table A-1. Unbalanced combined data for agronomic traits measured in RCBD over 5 locations and 19 years.

Trait	MS									H¶	SE#
	G	Y	L	Y*L	Rep (Y*L)	Y*G	L*G	G*Y*L	Error		
YLD	288.65**	33021*	405434**	16496**	223.44**	77.41*	106.55**	65.87**	31.92	0.096	0.018
HT	4015.26	9910.28*	210039	4860.97**	143.752**	30.27	104585**	30.21**	18.10	0.539	0.025
DTH	28.12	5698.48**	27814**	1223.38**	5.69**	23.48**	2.17**	1.46**	0.76	0.476	0.028
LD	5.57**	80.52	374.79*	44.84**	4.55**	1.95	3.40**	2.28**	0.83	0.168	0.029
LDG	19.68**	124.34	3.55	54.46**	6.04**	2.43	64.64**	2.09**	0.90	0.202	0.029

*, ** Significance at $P < 0.05$, 0.01, respectively

¶ Heritability on plot basis calculated for the genotypes. # Standard error for heritability

Table A-2. Unbalanced combined data for quality traits.

Trait	MS					H¶	SE#
	G	L	Rep(L)	G*L	Error		
GI	3531.53**	6918.65**	3226.44**	116.41**	91.77	0.655	0.021
WG	44.04**	766.45*	773.30**	7.69**	5.43	0.291	0.025
TKW	39.93**	1430.98**	845.01**	5.05**	4.35	0.389	0.025
TWT	6.73**	191.06**	136.88**	1.31**	0.90	0.309	0.025
VIT	47310**	32988**	300035**	16500	152094	0.223	0.021
WPROT	1.92**	87.19*	88.74**	0.41**	0.35	0.209	0.020
SPROT	1.45**	82.62*	79.82**	0.36**	0.27	0.194	0.020
TEXT	6.96**	546.61**	89.63**	1.24*	1.07	0.288	0.024
SEXT	8.84**	405.46**	79.43**	1.28	1.20	0.338	0.024
SASH	0.008**	0.313**	0.233**	0.002**	0.002	0.224	0.022
Color	0.62**	14.42**	9.59**	0.12**	0.10	0.272	0.025
MIXO	12.09**	8.95	11.54**	0.54**	0.42**	0.551	0.023
CWT	1.87**	37.84**	24.67**	0.72**	0.60	0.088	0.011
CLOSS	0.54**	13.48	18.44**	0.21**	0.16	0.116	0.017
FIRM	1.85**	136.52*	135.37**	0.28**	0.22	0.262	0.022
LK	612.69**	19047.00**	14882.00**	63.17*	55.36	0.446	0.024
MK	480.25**	14540.00**	11904.00**	57.17*	50.91	0.103	0.025
SK	19.73**	548.75**	461.29**	5.74**	4.38	0.178	0.020
SDS	643.98**	2636.55**	1953.89**	28.72**	21.77	0.574	0.023
FN	9310.93**	405182.00**	204655.00**	1933.29**	1386.54	0.237	0.021

*, ** Significance at $P < 0.05$, 0.01, respectively

¶ Heritability on plot basis calculated for the genotypes. # Standard error for heritability

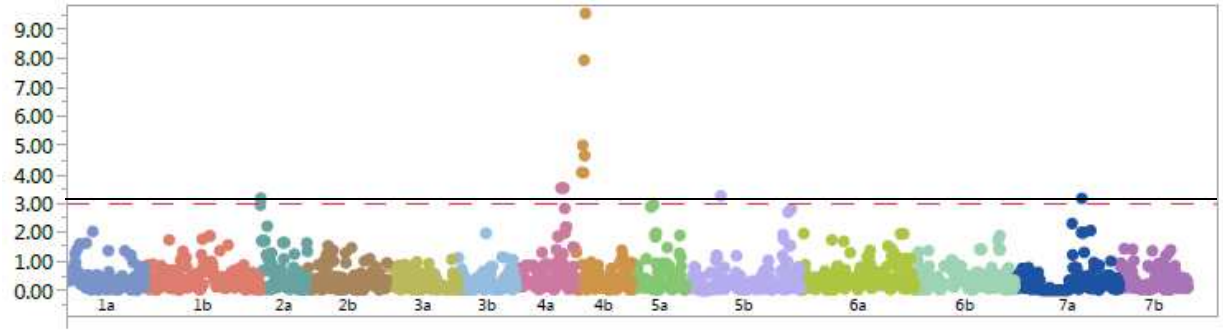
Table A-3. Balanced data for agronomic and quality traits.

Trait	MS					H¶	SE#
	L	Rep(L)	G	G*L	Error		
YLD	834397.83**	389897.53**	87005.17**	51552.76**	15951.47	0.207	0.048
HT	0.64	108.18**	127.59**	16.65	14.05	0.642	0.028
DTH	18900.69**	7.17**	6.37**	1.52**	0.99	0.487	0.036
LD	1507.75**	0.35	2.21**	0.99**	0.48	0.295	0.043
LDG	3518.55**	64.73**	7.61**	3.82**	2.26	0.236	0.042
FN	75455.91**	311858.10**	7185.25**	2882.20**	1610.02	0.322	0.042
VIT	43076.09**	14301.21**	205.03**	72.65**	45.93	0.358	0.040
TWT	3471.20**	137.55**	5.76**	2.18**	0.50	0.400	0.046
TKW	11597.96**	227.56**	24.11**	6.49**	3.19	0.476	0.038
WPROT	382.30**	7.05**	0.69**	0.26**	0.10	0.385	0.043
SPROT	570.61**	5.66**	0.86**	0.27**	0.10	0.447	0.042
LK	124362.67**	85.63	370.71**	86.91**	32.01	0.545	0.037
MK	96229.79**	77.85*	257.10**	64.67**	25.09	0.518	0.038
SK	1801.29**	11.24**	15.84**	5.84**	0.95	0.426	0.047
GI	405220.08**	6920.06**	1000.44**	176.54**	78.81	0.619	0.032
SDS	9123.01**	5242.92**	210.49**	28.01**	16.62	0.670	0.027
GLUT	47.78**	0.61**	0.07**	0.03**	0.01	0.297	0.056
WG	14402.79**	235.27**	18.30**	6.49**	2.16	0.411	0.043
L*	1030.62**	13.72**	2.29**	1.30**	0.75	0.195	0.042
a*	77.01**	6.09**	0.66**	0.22**	0.11	0.403	0.041
b*	645.64**	18.65**	11.63**	1.19**	0.75	0.730	0.023
Dif_L	585.54**	47.02**	7.45	7.29	6.60	0.005	0.033
Dif_a	97.65**	0.38**	0.13**	0.08*	0.06	0.169	0.038
Dif_b	219.48**	0.38	1.90**	0.95**	0.64	0.232	0.041
TYP	128.28**	7.16**	3.33**	0.24**	0.14	0.803	0.018
PPO	0.0486**	0.0026	0.0568**	0.0018**	0.0011	0.905	0.009
CLOSS	52.68**	38.04**	0.16	0.18**	0.14	0.103	0.045
CWT	2.79	8.81**	0.85	0.85	0.80	0.156	0.044

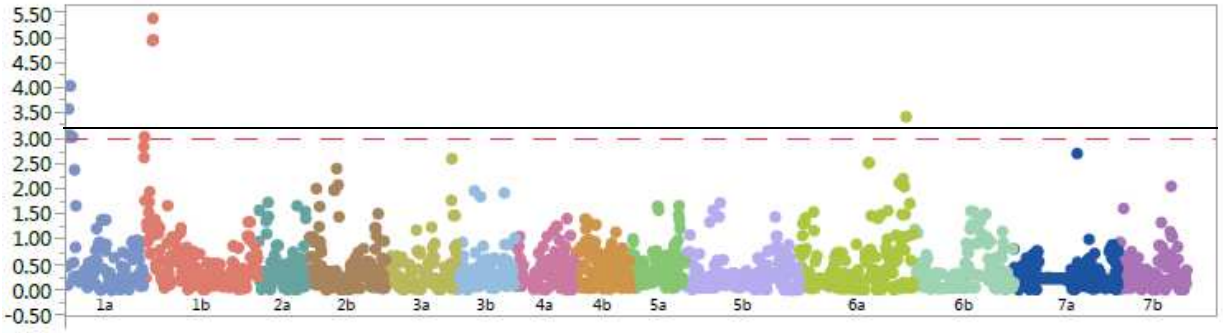
*, ** Significance at $P < 0.05$, 0.01, respectively

¶ Heritability on plot basis calculated for the genotypes. # Standard error for heritability

A



B



C

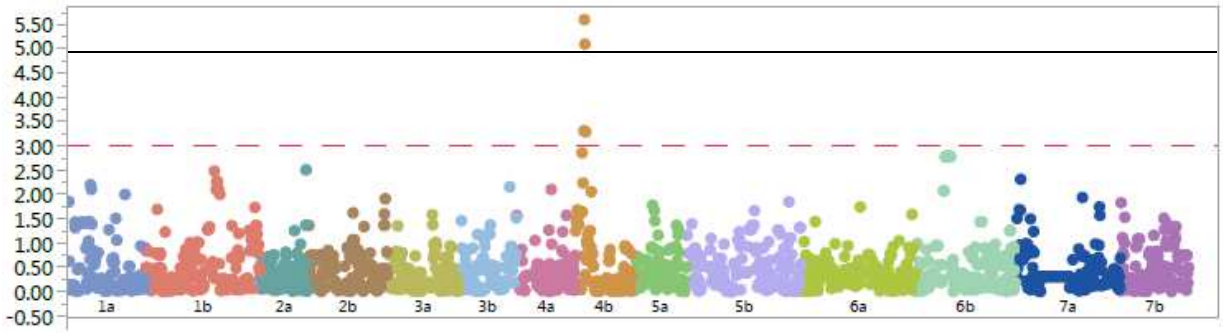
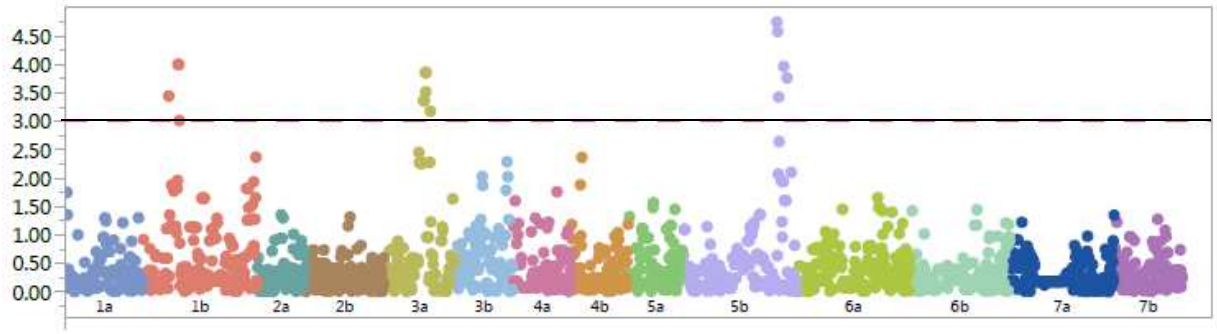
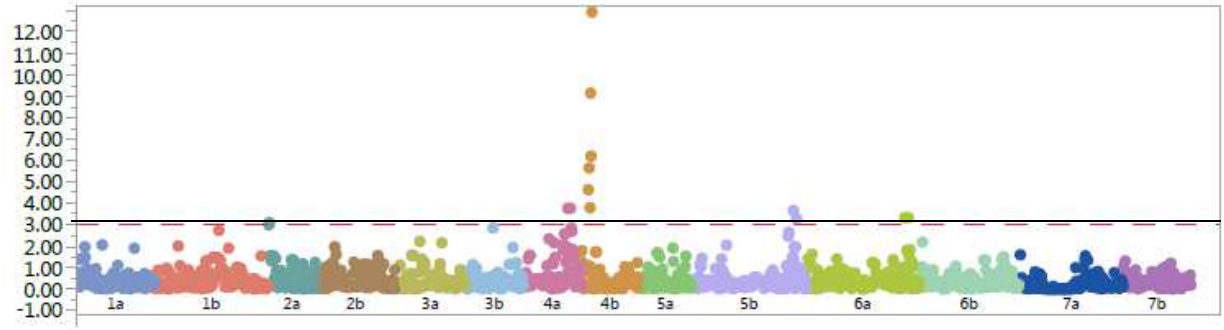


Figure A-1. Manhattan plots showing P-values for single-nucleotide polymorphism (SNP) markers associated with (A) plant height; (B) sedimentation value (SDS); (C) kernel vitreousness in Langdon, North Dakota in 2015. The horizontal dotted red line indicates significant threshold at P-value = 0.001. The black horizontal line indicates significant threshold at positive false discovery rate (pFDR) = 0.1.

A



B



C

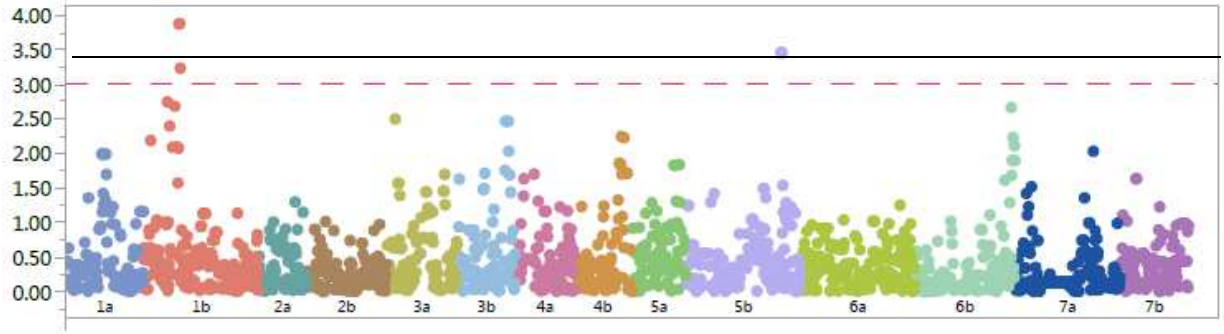
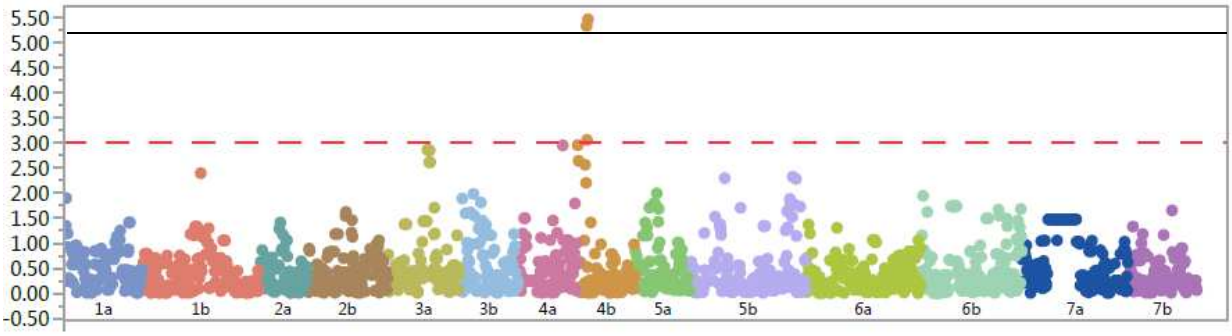
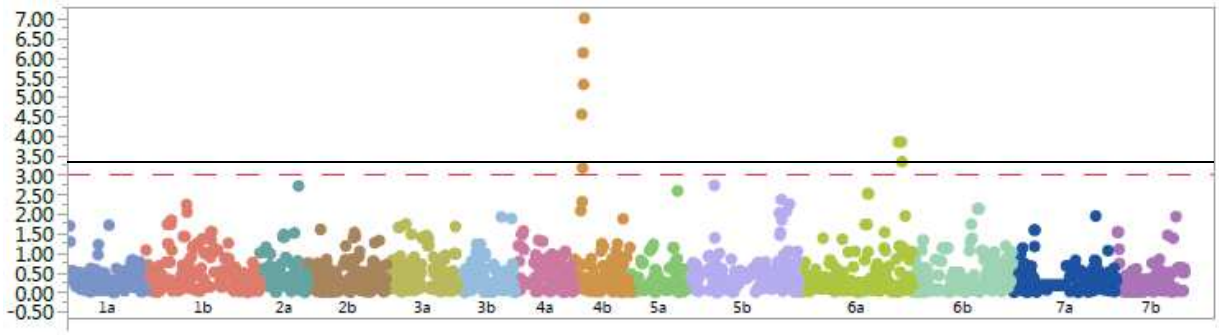


Figure A-2. Manhattan plots showing P-values for single-nucleotide polymorphism (SNP) markers associated with (A) yield; (B) plant height; (C) leaf disease; (D) test weight; (E) 1000-kernel weight; (F) percent of large kernels; (G) percent of medium kernels; (H) percent of small kernels; (I) polyphenol oxidase activity (PPO); (J) sedimentation value (SDS); (K) gluten index; (L) Difference in A color; (M) spaghetti cooked weight; and (N) cooking loss in Prosper, North Dakota in 2015. The horizontal dotted red line indicates significant threshold at P-value = 0.001. The black horizontal line indicates significant threshold at positive false discovery rate (pFDR) = 0.1.

D



E



H

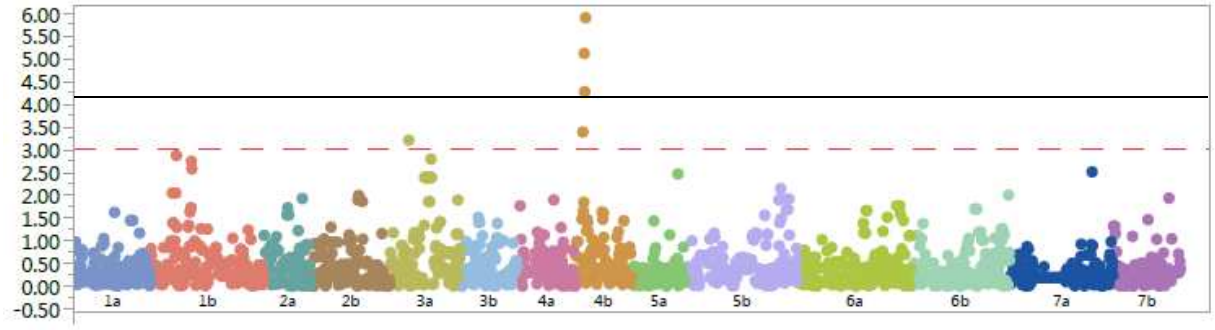
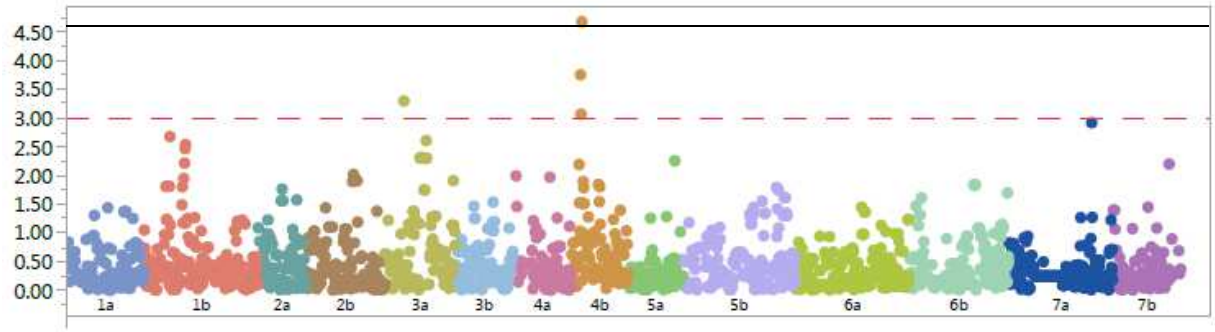
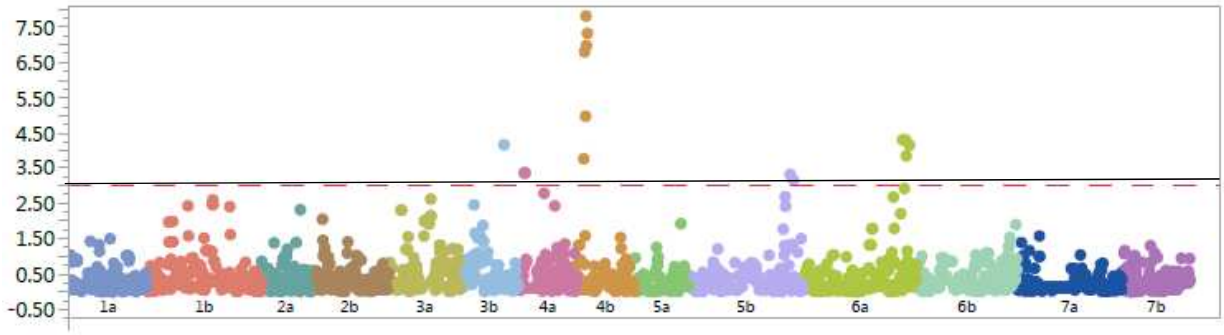


Figure A-2. Manhattan plots showing P-values for single-nucleotide polymorphism (SNP) markers associated with (A) yield; (B) plant height; (C) leaf disease; (D) test weight; (E) 1000-kernel weight; (F) percent of large kernels; (G) percent of medium kernels; (H) percent of small kernels; (I) polyphenol oxidase activity (PPO); (J) sedimentation value (SDS); (K) gluten index; (L) Difference in A color; (M) spaghetti cooked weight; and (N) cooking loss in Prosper, North Dakota in 2015 (continued). The horizontal dotted red line indicates significant threshold at P-value = 0.001. The black horizontal line indicates significant threshold at positive false discovery rate (pFDR) = 0.1.

G



H



I

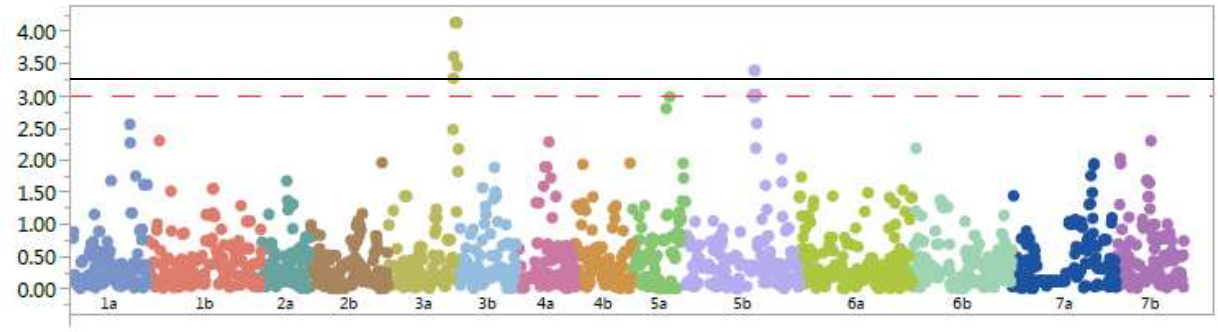
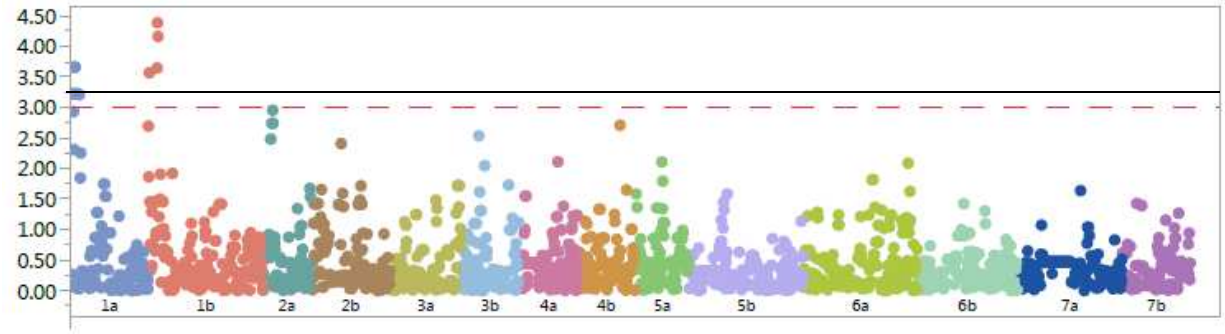
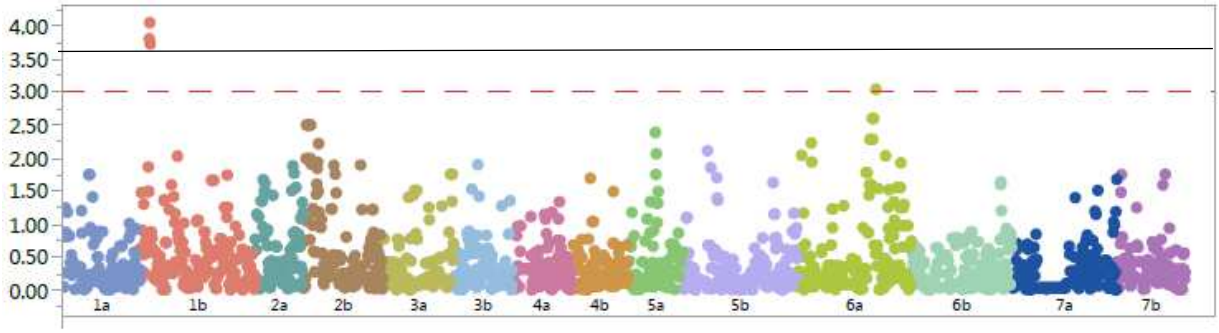


Figure A-2. Manhattan plots showing P-values for single-nucleotide polymorphism (SNP) markers associated with (A) yield; (B) plant height; (C) leaf disease; (D) test weight; (E) 1000-kernel weight; (F) percent of large kernels; (G) percent of medium kernels; (H) percent of small kernels; (I) polyphenol oxidase activity (PPO); (J) sedimentation value (SDS); (K) gluten index; (L) Difference in A color; (M) spaghetti cooked weight; and (N) cooking loss in Prosper, North Dakota in 2015 (continued). The horizontal dotted red line indicates significant threshold at P-value = 0.001. The black horizontal line indicates significant threshold at positive false discovery rate (pFDR) = 0.1.

J



K



L

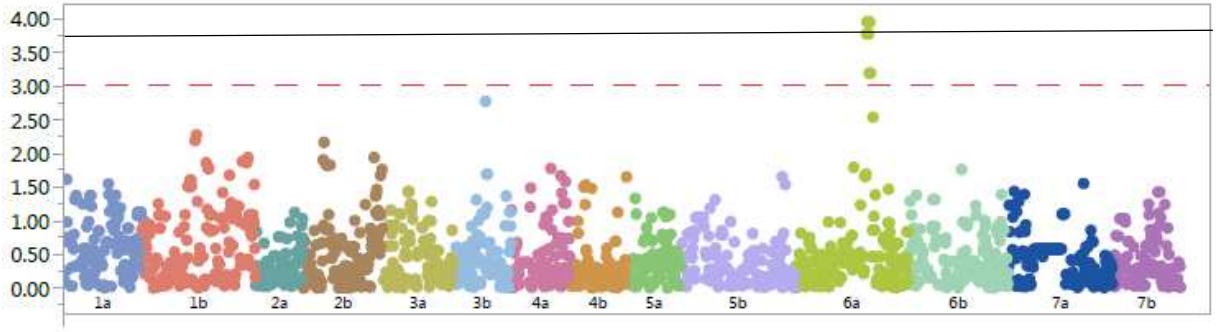
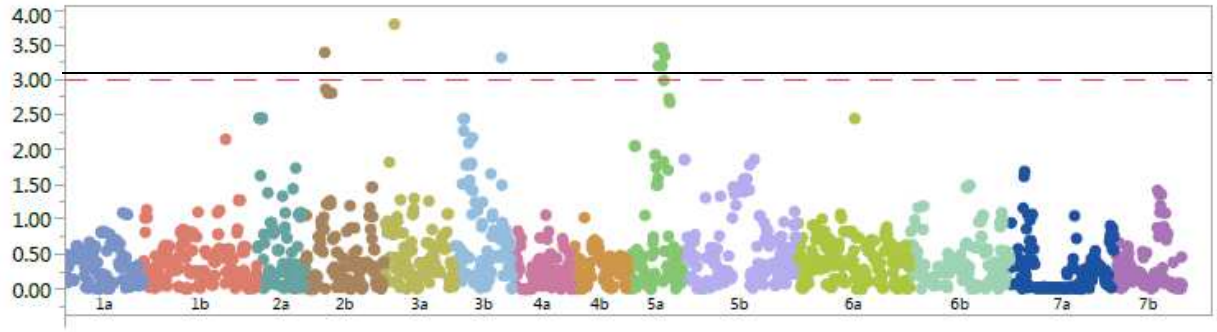


Figure A-2. Manhattan plots showing P-values for single-nucleotide polymorphism (SNP) markers associated with (A) yield; (B) plant height; (C) leaf disease; (D) test weight; (E) 1000-kernel weight; (F) percent of large kernels; (G) percent of medium kernels; (H) percent of small kernels; (I) polyphenol oxidase activity (PPO); (J) sedimentation value (SDS); (K) gluten index; (L) Difference in A color; (M) spaghetti cooked weight; and (N) cooking loss in Prosper, North Dakota in 2015 (continued). The horizontal dotted red line indicates significant threshold at P-value = 0.001. The black horizontal line indicates significant threshold at positive false discovery rate (pFDR) = 0.1.

M



N

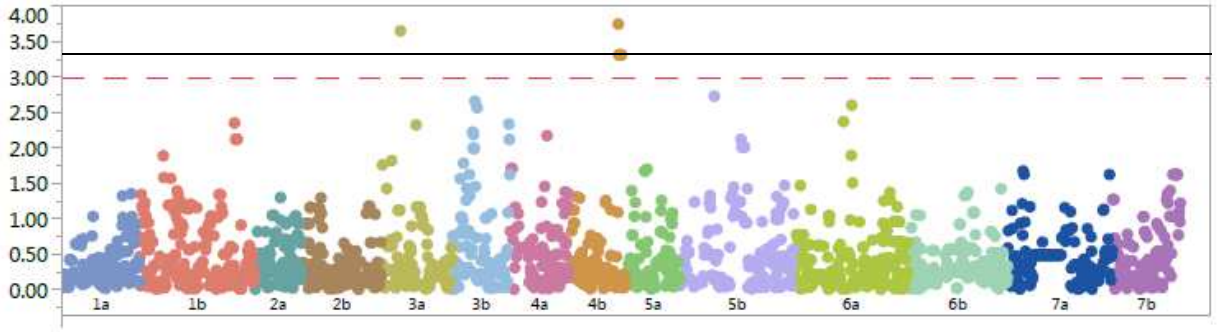
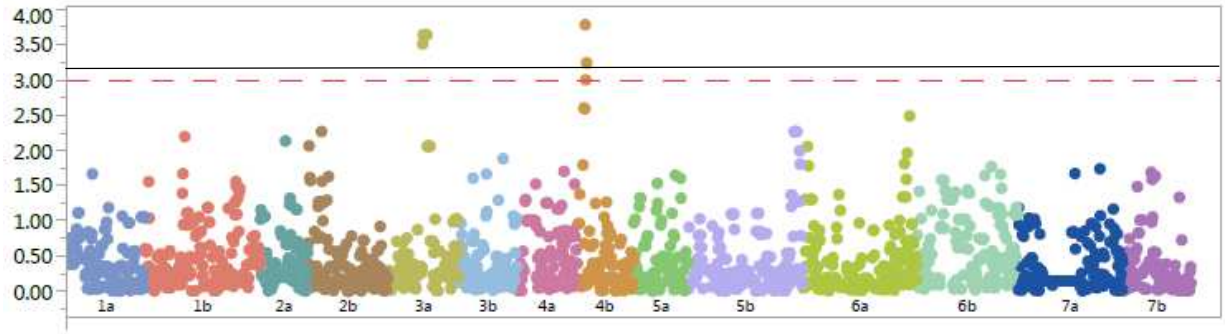
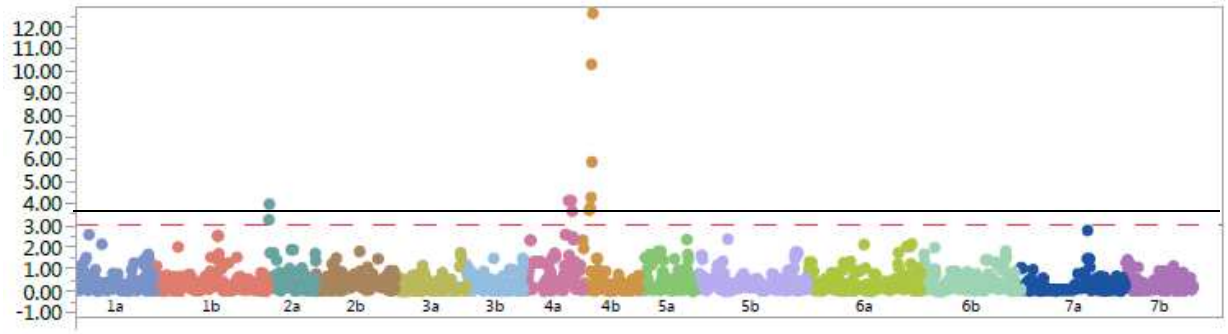


Figure A-2. Manhattan plots showing P-values for single-nucleotide polymorphism (SNP) markers associated with (A) yield; (B) plant height; (C) leaf disease; (D) test weight; (E) 1000-kernel weight; (F) percent of large kernels; (G) percent of medium kernels; (H) percent of small kernels; (I) polyphenol oxidase activity (PPO); (J) sedimentation value (SDS); (K) gluten index; (L) Difference in A color; (M) spaghetti cooked weight; and (N) cooking loss in Prosper, North Dakota in 2015 (continued). The horizontal dotted red line indicates significant threshold at P-value = 0.001. The black horizontal line indicates significant threshold at positive false discovery rate (pFDR) = 0.1.

A



B



C

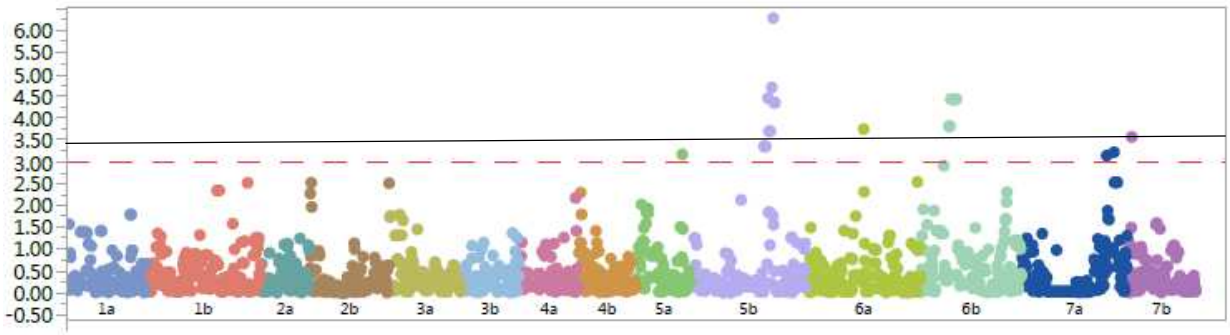
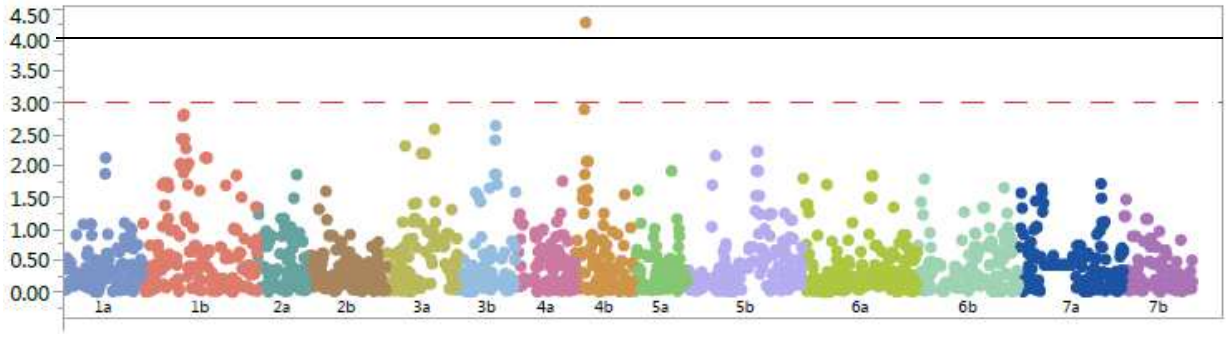
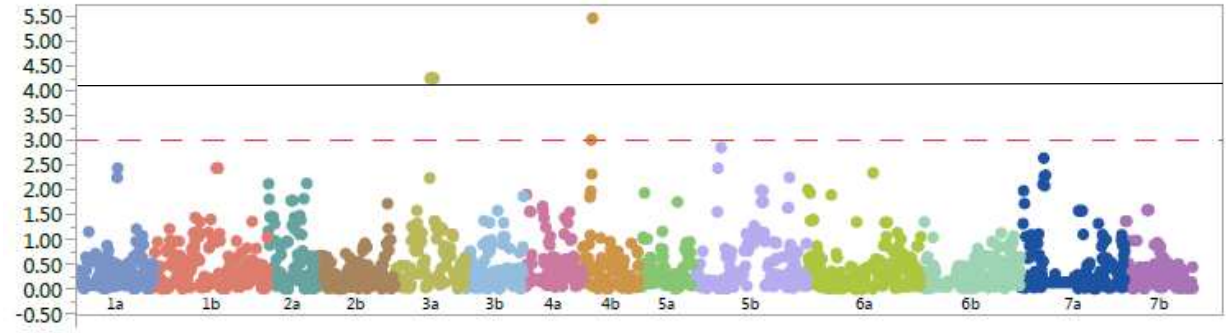


Figure A-3. Manhattan plots showing P-values for single-nucleotide polymorphism (SNP) markers associated with (A) yield; (B) height; (C) kernel vitreousness; (D) percent of large kernels; (E) percent of small kernels; (F) total milling extraction; (G) sedimentation value (SDS); (H) gluten index; (I) mixograph score; (J) falling number weight; (L) cooked spaghetti firmness; and (M) cooking loss in combined data from 1997-2014. The horizontal dotted red line indicates significant threshold at P-value = 0.001. The black horizontal line indicates significant threshold at positive false discovery rate (pFDR) = 0.1.

D



E



F

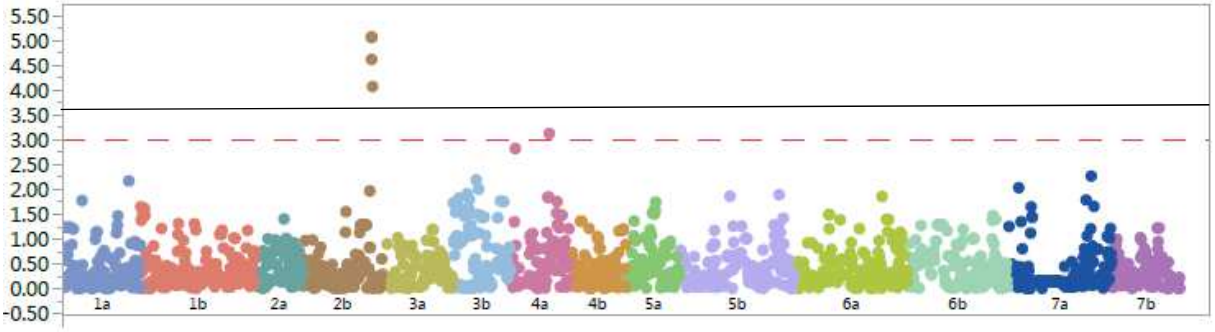
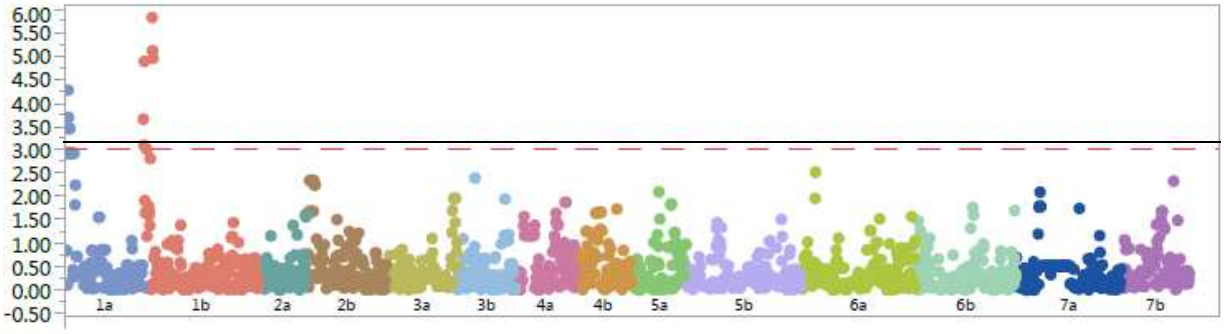
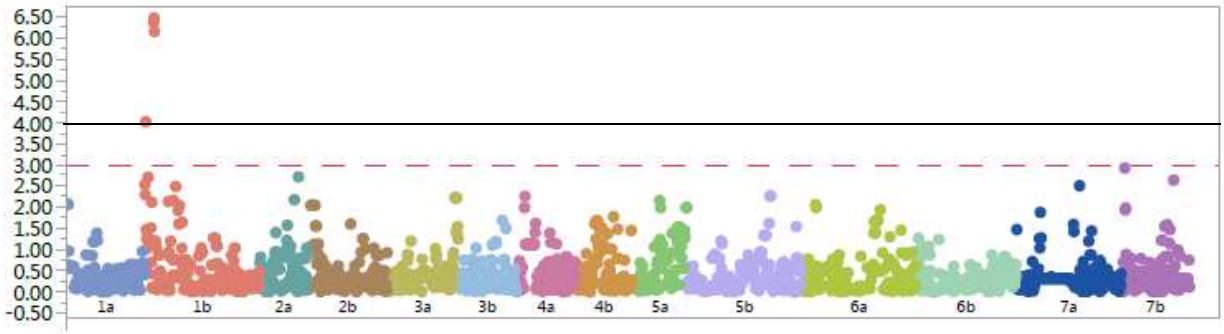


Figure A-3. Manhattan plots showing P-values for single-nucleotide polymorphism (SNP) markers associated with (A) yield; (B) height; (C) kernel vitreousness; (D) percent of large kernels; (E) percent of small kernels; (F) total milling extraction; (G) sedimentation value (SDS); (H) gluten index; (I) mixograph score; (J) falling number weight; (L) cooked spaghetti firmness; and (M) cooking loss in combined data from 1997-2014 (continued). The horizontal dotted red line indicates significant threshold at P-value = 0.001. The black horizontal line indicates significant threshold at positive false discovery rate (pFDR) = 0.1.

G



H



I

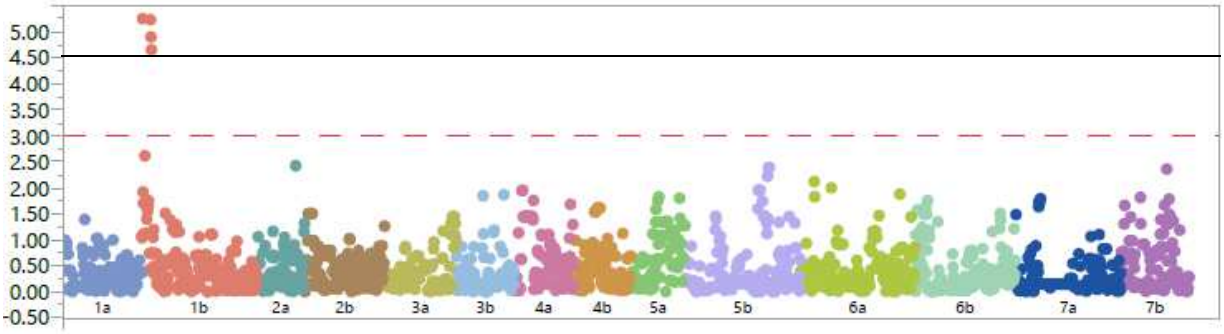
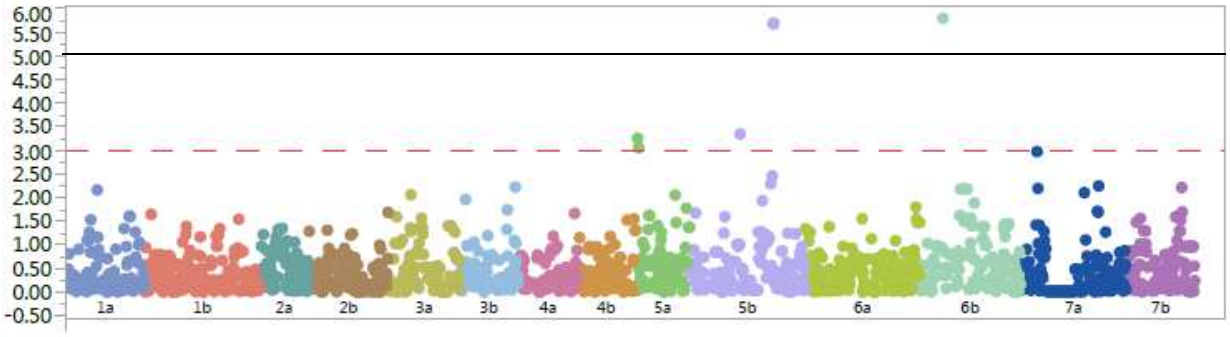
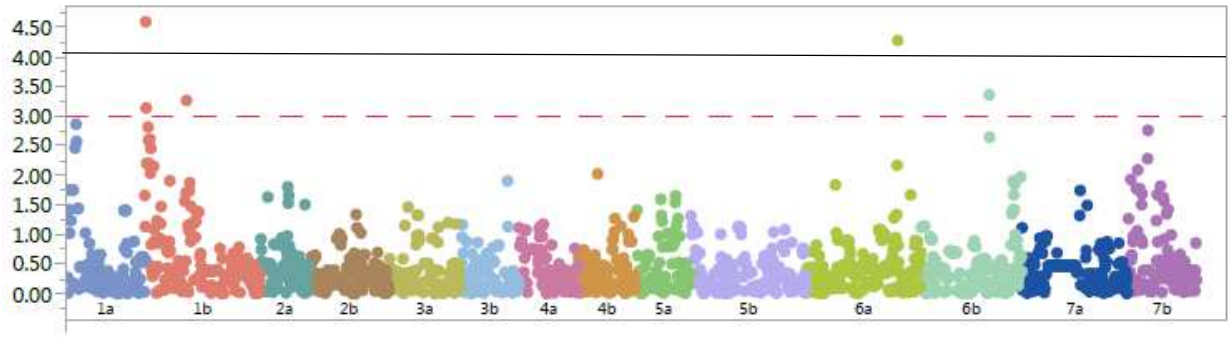


Figure A-3. Manhattan plots showing P-values for single-nucleotide polymorphism (SNP) markers associated with (A) yield; (B) height; (C) kernel vitreousness; (D) percent of large kernels; (E) percent of small kernels; (F) total milling extraction; (G) sedimentation value (SDS); (H) gluten index; (I) mixograph score; (J) falling number weight; (L) cooked spaghetti firmness; and (M) cooking loss in combined data from 1997-2014 (continued). The horizontal dotted red line indicates significant threshold at P-value = 0.001. The black horizontal line indicates significant threshold at positive false discovery rate (pFDR) = 0.1.

J



L



M

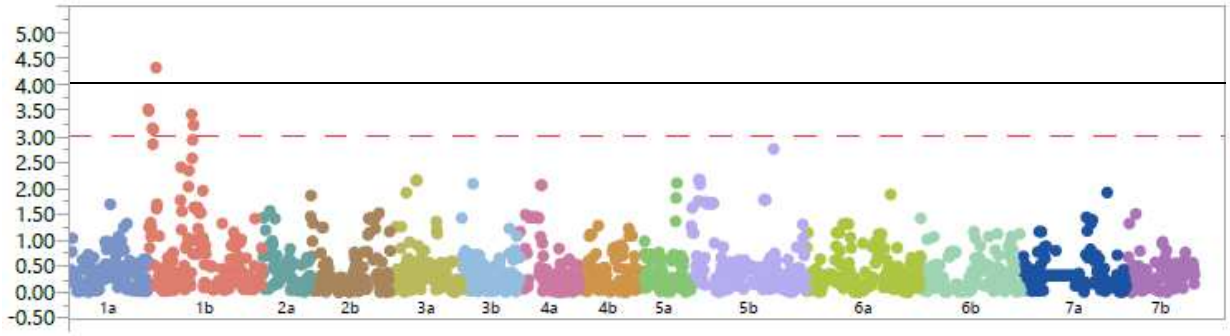


Figure A-3. Manhattan plots showing P-values for single-nucleotide polymorphism (SNP) markers associated with (A) yield; (B) height; (C) kernel vitreousness; (D) percent of large kernels; (E) percent of small kernels; (F) total milling extraction; (G) sedimentation value (SDS); (H) gluten index; (I) mixograph score; (J) falling number weight; (L) cooked spaghetti firmness; and (M) cooking loss in combined data from 1997-2014 (continued). The horizontal dotted red line indicates significant threshold at P-value = 0.001. The black horizontal line indicates significant threshold at positive false discovery rate (pFDR) = 0.1.