IDENTIFICATION OF QUANTITATIVE TRAIT LOCI ASSOCIATED WITH A LOW CADMIUM UPTAKE GENE IN DURUM WHEAT (*TRITICUM TURGIDUM* L. VAR. *DURUM*)

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

By

Wesam Ali AbuHammad

In Partial Fulfillment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

Major Department:
Plant Sciences

October 2013

Fargo, North Dakota

North Dakota State University
Graduate School

Title

Identification of Quantitative Trait Loci Associated with a Low Cadmium Uptake Gene In Durum Wheat (*Triticum turgidum* L. var. *Durum*)

By
Wesam Ali AbuHammad

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 Elias
Chair

Dr. Frank Manthey

Dr. Mohamed Mergoum

Dr. Seung Won Hyun

Approved:

11/04/2013
Date

Dr. Richard Horsley
Department Chair

ABSTRACT

The main objective of durum wheat breeding is to identify lines that are low in Cadmium accumulation by using linkage mapping. 178 recombinant inbred lines (RILs) derived from across between Grenora × Haurani were used to identify QTL for Cadmium (Cd) tolerance. A total of 1,132 polymorphic loci (1,111 SNP and 21SSR loci) were used for linkage map analysis. Among these, 330 (29 percent) markers were successfully integrated into the linkage maps at a LOD score of 3.0. The linkage map had a total genetic distance of 720.2 cM with an average distance of 2.2 cM between adjacent markers distributed on 14 chromosomes. Quantitative trait loci analysis was conducted using composite interval mapping (CIM). A Single putative QTL associated with Cd uptake was detected on Chromosome 5B. This QTL increased Cd uptake by the presence of alleles from Grenora. The additive effect of the Grenora allele was 0.02 mg/kg. The variation in Cd accumulation explained by this QTL (r^2) was 54.3 percent.

Association mapping was also used to analyze two durum wheat collections consisting of advanced breeding lines from the North Dakota wheat breeding program to discover markers associated with the amount of Cd accumulated in the grains. For the 2009 collection, one major QTL that explains 3 percent of phenotypic variation was identified on Chromosome 2B at 7.25 cM which could cosegregate with a height locus identified earlier. For the 2010 collection, one QTL that explains 34 percent of phenotypic variation was identified on Chromosome 5B at 165.7cM and one other at 178.3cM that explains 27 percent of phenotypic variation. Because of the complexity of Cd accumulation in wheat grains, the identification of additional QTL will require a better coverage of markers and a larger collection of genotypes. This will help breeding for low Cd using MAS.

DEDICATION

This dissertation is dedicated to my wife Marina, my parents Ali and Amal and my brothers and sisters for their endless love, support and encouragement

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AM.....Association mapping

Cd.....Cadmium

CIM.....Composite interval mapping

CV.....Coefficient of variation

DNA.....Deoxyribonucleic acid

DTH.....Days to heading

GLM.....General linear model

IBS.....Identical by state

K.....Kinship

LD.....Linkage disequilibrium

LOD.....Logarithm of the odds

LSD.....Least significant difference

MAF.....Minor allele frequency

MAS.....Marker assisted selection

MLM.....Mixed leaner model

MSD.....Mean square difference

NIL.....Near inbred lines

PCA.....Principle component analysis

PFDR.....Positive false discovery rate

QTL.....Quantitative trait loci

RIL.....Recombinant inbred lines

SAS.....Statistical analysis system

SNP.....Single nucleotide polymorphism

SSR.....Single sequence repeats

TWT.....Test weight

YLD.....Yield

GENERAL INTRODUCTION

This dissertation contains four chapters. Chapter I includes a literature review. Chapters II and III are written to be submitted for publication. Both chapters include an abstract, introduction, materials and methods, and results and discussion. The references at the end of each chapter are specific for that chapter. Chapter IV includes a general research conclusion. Repetitions and/or similarities among Chapters occur because the same genetic and statistical tools were used.

CHAPTER I. LITERATURE REVIEW

Durum wheat

Wheat grown in the United States has had a major effect on the U.S. economy. The U.S. total production of durum wheat in 2007 was roughly 2.0 million metric tons. North Dakota had the largest amount of durum wheat production with approximately 1.2 million metric tons from the total U.S. production for the 2012 growing season (NASS, 2012).

Durum wheat (*Triticum turgidum* L. var. *durum*) is unsuitable for manufacturing yeast-leavened breads because of the poor gluten complex that is formed when water is added and mixed with durum wheat flour. However, it is the ideal wheat for pasta production due to its unique color, flavor, and cooking quality (Feillet and Dexter, 1996). Unlike other wheats, durum wheat is milled into a coarse granular product called semolina. The semolina should contain less than 3 percent flour and be free of germ and bran coats. Semolina is used to produce pasta, the general term for food such as macaroni, noodles, and spaghetti.

History of durum wheat in the United States

In the early 1850s, the U.S. Department of Agriculture (USDA) started dispensing durum wheat to farmers. The early durum wheat cultivars introduced in the United States were grown in limited quantities due to the lack of adaptation in most areas (Joppa *et al.,* 1988). In late 1894, the USDA established a strong base of studying wheat cultivars and production in the United States. The outcome of this work suggested that high yield, drought tolerance, and early maturity cultivars were needed (Joppa and Williams, 1988). In 1900, Mark Carleton traveled to Russia and collected many types of wheat. These wheats were increased at the Kansas Agricultural Experiment Station and were planted across the southern plains. Carleton also had this ‘other’ wheat from Russia that nobody wanted. This ‘other’ wheat, “goose wheat” as called by farmers, was very hard to mill, and its flour did not fit American tastes. The next decades Carlton spent

all his efforts trying make this new wheat acceptable to farmers (Joppa and Williams, 1988). Later, when devastating diseases of rust returned to the northern plains, the goose wheat proved more leaf rust resistant than any other cultivars, which made the farmers accept the new wheat. Today Carleton's goose wheat has become the foundation of a major commodity industry on the northern plains. The goose wheat was Kubanka durum, and Mark Carleton became the father of the durum industry on the northern plains (Joppa and Williams, 1988).

The classification of durum wheats

Wheat class and physical condition are important factors for determining wheat milling potential and end-product quality (Dexter and Edwards, 1998). Durum wheat is classified as a hard wheat, having high protein or low protein; and after adding water to the flour, the dough can be classified as having strong gluten or weak gluten (Dexter, 1993). The class of U.S. durum wheat includes all cultivars of amber durum wheat and is divided into three subclasses: hard amber durum wheat (with 75 percent or more vitreous kernels of amber color), amber durum wheat (with 60 percent or more, but less than 75 percent, hard and vitreous kernels of amber color), and durum wheat (with less than 60 percent hard and vitreous kernels of amber color).

Cadmium in soil

Cadmium (Cd) is a biotoxic heavy metal present in low concentrations in the soil worldwide. This toxic heavy metal is not known to be necessary to the life processes of any living organism. Controlling heavy metals availability in plants is altered by its concentration in soils (Bruemmer *et al.,* 1986; Gerritse *et al.,* 1983). The solubility and bioavailability of heavy metal ions is affected by soil pH (Street *et al.,* 1978), clay mineral content (Herms and Brummer, 1984), monoxide concentration (King, 1988), and organic matter content in soil (McClean, 1976). Uncontaminated soils usually contain between 0.01 to 30 μg Cd g^{-1} with a typical range

of 0.06 to 0.5 µg Cd g^{-1} (Haghiri, 1974). Alloway (1995) reported that Cd is found in soil solutions as Cd^{+2}, which may bond with other elements to form new complexes in the soil solution. Cadmium migration to deeper soil layers and to underground water by leaching is usually more than other heavy metals (Breslin, 1999). Soil properties and adsorption isotherms are related to Cd migration modeling (Filius *et al.,* 1998).

Cadmium in plants

Cadmium distribution and accumulation differ among plant species and cultivars (John, 1973). McLaughlin *et al.* (1994) reported that Cd concentration among potato differed up to 50 percent in tuber. Cadmium may be accumulated either in the roots or in the leaves. Cadmium concentration in the leaves was found to be higher than Cd accumulation in the roots (Bingham, 1979; Jalil *et al.,* 1994b). Conversely, Guo and Marschner (1995) studied the level of Cd concentration in beans and rice. They found that the concentration of Cd was very high in the roots compared to other parts of the plant.

Other than the differences in Cd concentration among species, varietal differences have been investigated within a species. Eriksson (1990b) found that certain wheat cultivars grown within the same region had higher Cd concentrations than others grown within the same region. In a different experiment, 20 soybean cultivars grown in the greenhouse on three types of soil indicated Cd concentration differences among cultivars (Boggess *et al.,* 1978).

Cadmium toxicity in plants can be observed as interveinal chlorosis. However, this cannot be generalized since the deficiency of several nutrients and nonessential elements could generate analogous symptoms (Bingham, 1979).

Cadmium and zinc interaction

In soils, Cd is found associated with other metals such as zinc (Zn), lead (Pb), and copper (Cu) (Elinder, 1992). Zinc and Cd are naturally occurring trace metals that are ubiquitous in soils. Both elements undergo similar chemical and physiological processes in plants and soils (Jalil *et al.*, 1994a). The interaction between Cd and Zn was investigated on soils contaminated by Cd (soil Cd content < 1 mg kg^{-1} DW) (Oliver *et al.*, 1994). They found that Zn fertilizer application up to 5 kg Zn ha^{-1} in sulfate form decreased wheat grain Cd concentration up to 50 percent. However, the effect of Zn fertilizer decreased with time. The results suggested that the availability of applied Zn depended on soil reactions. The loss of root membrane integrity or the release of phytosiderophores (chelate compounds that sequester iron) by wheat roots under nutrient deficiency conditions was believed to be involved in the process (Tiller *et al.*, 1997).

Zinc and cadmium are similar physicochemically in animals. The interaction of both elements was studied by Parizek (1957) and Gunn *et al.* (1961). They observed less testicular injury of rats compared with the subcutaneous administration of Cd alone. Gunn *et al.* (1963) reported that the administration of Zn alone can prevent Cd-induced testis tumors. Nordberg (1972) suggested that both Zn and Cd are bonded together of a molar ratio of 1:1 to form a metallothionein. The latter was believed to cause the direct increase of both elements.

Plant breeding efforts to reduce Cd accumulation

Genetic selection

Selection of low Cd genotypes might be a good solution for reducing the level of Cd in the diet. Li *et al.* (1993) produced new lines of sunflower that show low Cd uptake. Likewise, the same action can be taken when working with other crops such as durum wheat, which can be used in different types of end products with economical importance worldwide. Penner *et al.*

(1995) identified markers for a single gene governing low Cd uptake in durum wheat cultivars grown in Canada. Buckley *et al.* (1997) suggested that the differences in Cd concentration in durum isoline seeds may be related to the accumulation of Cd in the root system.

Site selection

The level of Cd in soil differs by location (Andersson and Pettersson, 1981). Soil characteristics such as Cd concentration, hydrogen concentration (pH), salinity, organic matter content, and soil texture maybe the reason for the geographical differences (Grant *et al.*, 1998). Moreover, atomospheric deposition and fertilizer use may also play a role in these differences (Grant *et al.*, 1998). McLaughlin *et al.* (1997) determined that the conductivity of applied irrigation water was correlated with Cd content of potatoes. Li *et al.* (1995) studied the effect of soil chloride level on Cd concentration in sunflower kernels. They reported that when sunflower are grown on soil with low pH and high Cl concentrations the Cd concentration was high. Grant *et al.* (1998) studied Cd accumulation in crops. They concluded that the availability of reliable methods for predicting Cd phytoavailability is necessary for accurate prediction of Cd concentration across regions.

Fertilizer management

Cadmium uptake by plants may be increased by using fertilizers. Soil type affects the level of accumulation of Cd on soils. Eriksson (1990) studied the factors influencing absorption and plant uptake of Cd from agricultural soils. He found that due to the lower Cd absorption ability, the process in which one substance takes up or holds another by either absorption or adsorption, and lower average soil water content in sandy soils, fertilizers affect the accumulation of Cd more than in clay soils. Reducing the long-term accumulation of Cd in the soil can be achieved by reducing the amount of Cd in phosphorus fertilizer (Grant *et al.*, 1998).

Cadmium concentration in the plant can increase also by using the nitrogen fertilizers which do not naturally contain any significant levels of Cd (Andersson, 1976). In similar study, Lorenz *et al.* (1994) reported that the application of nitrogen fertilizer led to a significant increase of Cd concentration in soil solutions and uptake by plants. Moreover, potassium fertilizers also may influence Cd concentration of plants. In contrast to the findings of McLaughlin *et al.* (1995), Sparrow *et al.* (1994) reported that the application of potassium in the form of KCl elevated Cd content of potato tubers more than the application in the form of K_2SO_4.

Marker-assisted selection

Segregating plants that contain appropriate combinations of genes play a critical role in plant breeding programs (Weeden *et al.,* 1994). Indirect selection methods such as MAS rely on markers that are tightly linked to the gene of interest (Bernardo, 1998). Plant breeders may use specific DNA marker alleles as a tool to identify the gene of interest when markers that are tightly linked to the quantitative trait loci (QTL) have been identified. This can be done before the field evaluation of a large number on individuals (Michelmore, 1995). The value of markers in the selection process can be evaluated by the degree of linkage between the marker and the gene on interest, segregation of both the target gene and the marker, the degree of linkage disequilibrium in the population to be selected, and the linkage phase between the gene of interest and the marker (Weber and Wricke, 1994). Marker-assisted selection has been a useful tool in evaluating traits that are very difficult or expensive to phenotype (Yousef and Juvik, 2001). It also speeds up the selection process in plant breeding programs by identifying different genotypes that carry the gene of interest (Ahmad, 2000). Genomic tools have been used in plant breeding to identify and characterize all possible genetic variations of germplasm resources

(Tanksley *et al.*, 1989) and to enhance the target trait by using the genetic transformation and molecular markers techniques (Dudley, 1993).

Advantages of MAS over phenotypic selection

Using MAS may increase the efficiency and effectiveness of selection in plant breeding programs. By using a flanking marker, the efficiency of marker-assisted selection can be increased, and the target gene would be easily identified compared to the use of a single marker (Tanksley, 1983). Only one crossover event is required to separate the QTL from a single marker. Two crossover events, one on each side of the QTL, would be required to separate a QTL from both flanking markers. Double crossovers are less likely than single crossovers, so the probability that the marker-QTL association will be maintained is higher with flanking markers. Van Berloo and Stam (1998) reported that MAS can be more efficient when markers are linked to the QTL in the coupling phase. Coupling phase provides greater proportion of homozygous resistant selections. Several factors may affect the efficiency of MAS. These factors include, but are not limited to: 1) the genetic distance between the marker and the QTL, 2) the population size, 3) the heritability of the trait (Lande and Thompson, 1990), 4) the mode of inheritance of the trait (Stuber *et al.*, 1999), and 5) when a few QTL explain a large portion of the genotypic or phenotypic variance (Moreau *et al.*, 1998).

Molecular markers have been a useful tool in plant breeding programs to make selection in early generations before maturity and to decrease the size of a segregating population effectively. Marker-assisted selection provides alternative solutions to prevent any crop from contamination by heavy metals. Molecular marker techniques may play a role in developing of low Cd food crops by gene transformation across species (Wagner, 1993; Penner *et al.*, 1995).

Marker systems have been applied to genomic mapping studies such as restriction fragment length polymorphisms (RFLP), randomly amplified polymorphic DNA (RAPD), simple sequence repeats (SSR) or microsatellites, and amplified fragment length polymorphisms (AFLP) (Yang *et al.*, 1994). Most of the previous molecular marker analysis methods are low throughput due to their dependence on gel electrophoresis (Jaccoud *et al.*, 2001). With the later-discovery of single nucleotide polymorphisms (SNPs), markers became abundant and had uniform distribution throughout a genome, thus becoming the markers of choice (Gupta *et al.*, 2008). Single nucleotide polymorphisms (SNPs) are biallelic genetic markers and are considered as high throughput markers which offer a low-cost and robust system.

Linkage disequilibrium

Linkage disequilibrium (LD) can be described as a nonrandom association of alleles at different loci. Jannink and Walsh (2002) referred to LD as "gametic phase disequilibrium" (GPD) or "gametic disequilibrium" (GLD). This can be used in the literature that describes the same nonrandom association of haplotypes within distantly related populations assuming Hardy-Weinberg equilibrium (HWE).

The difference between expected haplotype frequency and observed frequency based on allele frequencies is defined as D.

$$D = P_{AB} - P_A P_B$$

where P_{AB} is the frequency of gamete AB and P_A and P_B are the frequency of the allele A and B, respectively.

In the absence of forces, recombination breaks down the LD through random mating. This can be defined as D_t.

$$D_t = D_o(1 - r)^t$$

In this formula the D_t is the remaining LD between two loci after t generations of random mating from the original D_o.

LD measurements were proposed by several statistics. Such measurements differ due to small sample sizes and allele frequencies (Hedrick, 1987). To quantify LD, both the standard measure of LD (D`) (Lewontin, 1964) and the correlation between a pair of loci ($r^2$) (Hill and Robertson, 1968) have been widely used. The D` and r^2 have the following formula for two bi-allelic loci:

$$D' = \frac{|D|}{D_{\max}}$$

Where

D_{max} = min ($P_A P_b$,$P_a P_B$) if D > 0

D_{max} = min ($P_A P_B$,$P_a P_b$) if D < 0

$$r^2 = \frac{D^2}{P_A P_a P_B P_b}$$

This formula is useful because the value of r^2 ranges from 0 when the loci are in complete linkage equilibrium to 1 when the loci are in complete linkage disequilibrium. The range of D is determined by the allele frequency. This feature is considered undesirable; and, for this reason, the D' statistic was developed to partially normalize the D value with respect to the maximum value possible for the allele frequencies (Hill and Robertson, 1968). Researchers often provide a graph showing the relationship of r^2 to either genetic or physical distance. Typically the r^2 vs distance is calculated using non-linear regression. Several factors can affect LD; these include, but are not limited to: 1) mating system, 2) recombination rate, 3) population structure, 4) population history, 5) genetic drift, 6) directional selection, and 7) gene fixation (Gaut and Long 2003). On the other hand, the power of association analysis is affected by 1) the

patterns of LD, 2) the extent of LD in the genome, and 3) the variation in LD from one population to another (Gaut and Long 2003).

Approaches of association mapping

Candidate gene approach

Candidate genes are selected based on the information provided by the mutational and linkage analysis of the trait of interest. This approach is low in cost, based on hypothesis, and is trait-specific (Zhu *et al.,* 2008). It consists of two subsets of molecular markers. The first subset of genetic markers is selected to estimate the population structure accurately. The second subset of markers targets regions in the traits of interest of the population's genome (Jay, 2011).

Genome-wide scan approach

This approach involves screening a population of selected germplasms with large number of molecular markers spread throughout the genome to estimate the population structure and to search for casual genetic variation (Jay, 2011). It does not require prior information regarding candidate genes (Zhu *et al.,* 2008). The coverage of markers is based on resources available and the extent of LD (Jay, 2011). Genome-wide association mapping reveals informative allelic variation and validates QTL previously discovered by traditional mapping (Rafalski, 2002). Genome-wide scan is a better approach in self-pollinated species than in cross-pollinated species because in cross-pollinated species LD decays faster and a large number of markers is required to reveal useful polymorphisms (Flint-Garcia *et al.,* 2003). However, in self-pollinated species such as durum wheat, LD decay is slower and fewer markers are required to reveal a significant association with a QTL (Wang, 2005).

Association mapping models

A linear-mixed model approach has been the model of choice in association mapping studies. To account for population stratification, the linear-mixed model combines fixed and random effects (Yu *et al.,* 2006). Henderson (1975) proposed a general model of associating a marker and trait:

$$y = X\beta + Zu + e$$

where

y is a vector of observations from the population

β is an unknown vector of fixed effects including the molecular marker and population structure information

u is the random, unknown, additive background genetic effects

e is the residual error in the model

X and *Z* are known incidence matrices

X also called the Q matrix and consists of the actual genotypic data and a matrix containing values and is used to estimate the population structure. *Z* also called the K matrix and consists of a pair-wise comparison matrix or kinship matrix and is used to estimate the covariance of each individual's genetic background effect. The appropriate choice of the right model (complete or reduced) for the analysis is based on the genetic diversity of a population.

Population structure can be estimated by using different methods such as Bayesian method implemented in the software STRUCTURE (Pritchard *et al.,* 2000), the principal components analysis (PCA), and the two multivariate analyses (Price *et al.,* 2006), and multi-dimensional scaling (Li *et al.,* 2010). The kinship matrix can be used to estimate the variance/covariance matrix associated with the random effects of the background markers,

excluding the marker being tested (Jay, 2011). The kinship (K) matrix can be developed by several methods such as the SPAGeDi software that can calculate up to nine different kinship coefficients (Loiselle *et al.,* 1995) and pedigree information that estimates the genetic correlation among individuals in the population (Colonna *et al.,* 2009).

References

Ahmad, M. (2000). Molecular marker-assisted selection of HMW glutenin alleles related to wheat bread quality by PCR-generated DNA markers. Theoretical and Applied Genetics. 101: 892-896.

Alloway, B. J. 1995. Cadmium. P. 122-151. In: B. J. Alloway (ed). Heavy metals in soils. Blackie Academic & Professional. London, United Kingdom.

Andersson, A. 1976. On the influence of manure and fertilizers on the distribution and amounts of plant available Cd in soils. Swed. J. Agric. Res. 6: 27-36.

Andersson, A., and O. Pettersson. 1981. Cadmium in Swedish winter wheat: regional differences and their origin. Swed. J. Agric. Res. 11: 49-55.

Bernardo, R. 1998. A model for marker-assisted selection among single crosses with multiple genetic markers. Theor. Appl. Genet. 97: 473-478.

Bingham, F. 1979. Bioavailability of Cd to food crops in relation to heavy metal content of sludge-amended soil. Environ. Health Perspective 28: 39-43.

Boggess. S., S. Willavize, and D. Koeppe. 1978. Differential responses of soybean varieties to soil cadmium. Agron. 70: 756-760.

Breslin, V. T. 1999. Retention of metals in agricultural soils after amending with MSW and MSW biosolids compost. Water Air Soil Pollut. 109: 163-178.

Bruemmer, G., J. Gerth, and U. Herms. 1986. Heavy metal species, mobility and availability in soils. Z. Pflanzenemaehr. Bodenkd. 149: 382-398.

Buckley, W., K. Buckley, and C. Grant. 1997. Adsorption, absorption and translocation of cadmium in high-cadmium and low-cadmium accumulating lines of durum wheat. Pages 129-130 in I. K. Iskandar, S. E. Hardy, A. C. Chang, and G. M. Pierzynski, eds. Proc. Fourth International Conference on the Biogeochemistry of Trace Elements, Berkeley, CA.

Colonna V., T. Nutile, R.R. Ferrucci, G. Fardella, M. Aversano, G. Barbujani, and M. Ciullo.

2009. Comparing population structure as inferred from genealogical versus genetic information. European Journal of Human Genetics 17: 1635-1641.

Dexter, J. E., and N. M. Edwards. 1998. The implications of frequently encountered grading factors on the processing quality of durum wheat. AOM Bull. October: 7165-7171.

Dexter, J. E. 1993. Quality requirements of western Canadian bread wheats. Pages 697-722 in: Grain and Oilseeds: Handling, Market Processing. 4th Edition. E.J. Bass, editor. Canadian International Grains Institute, Winnipeg, Canada.

Dudley, J. 1993. Molecular markers in plant improvement: Manipulation of genes affecting quantitative traits. Crop Sci. 33: 660-668.

Elinder, C. 1992. Cadmium as an environmental hazard. In: 'Cadmium in the Human Environment: Toxicity and Carcinogenicity' (Eds. G.F. Nordberg, R.F.M. Herber, L. Alessio) pp. 123-133. IARC Scientific Publications.

Eriksson, J. 1990b. A field study on factors influencing Cd levels in soils and in grain of oats and winter wheat. Water Air Soil Pollut. 53: 69-81.

Eriksson, J. E. Factors influencing absorption and plant uptake of cd from agricultural soils. Dissertation. Swedish University of Agricultural Science, Department of Soil Science, Reports and Dissertations, 1990.

Feillet, P., and J. E. Dexter. 1996. Quality requirements of durum wheat for semolina milling and pasta pr oduction. Pages 95-131 in: Pasta and Noodle Technology. J. E. Kruger, R. R. Matsuo and J. W. Dick, eds. American Association of Cereal Chemists, St. Paul, MN.

Filius, A., T. Streck. and J. Richter. 1998. Cd sorption and desorption in limed topsoils as influenced by pH: Isotherms and simulated leaching. J. Environ. Qual. 27:12-18.

Flint-Garcia , S., J. Thornsberry, E. I. Buckler. 2003. Structure of linkage disequilibrium in plants. Annu Rev Plant Biol 54:357-374.

Gaut B. S., A. D. Long. 2003. The lowdown on linkage disequilibrium. Plant Cell 15: 1502-1506.

Gerritse, R., W. Van Driel, K. Smilde, and B. Van Luit. 1983. Uptake of heavy metals by crops in relation to their concentration in the solution. Plant and Soil 75: 393-404.

Grant, C., W. Buckley, L. Bailey, and F. Selles. 1998. Cadmium accumulation in crops. Can. J. Plant Sci. 78: 1-17.

Gunn, A., T. Gould, and W. Andersson. 1963. Cadmium induced interstitial cell tumours in rat and their prevention by zinc. J. Nat. Cancer. Inst. 31:745-759.

Gunn, S., T. Gould, and W. Andersson. 1961. Zinc protection against cadmium injury to rat

testis. Arch. Pathol. 71: 274-281.

Guo, Y., and H. Marschner. 1995. Uptake, distribution, and binding of cadmium and nickel in different plant species. J. Plant Nutr. 18: 2691-2706.

Gupta, P., S. Rustgi, and R. Mir. 2008. Array-based high-throughput DNA markers for crop improvement. Heredity 101: 5-18.

Haghiri, F. 1974. Plant uptake of cadmium as influenced by certain exchange capacity, organic matter, zinc, and soil temperature. Journal of Environmenttal Quality. 3: 180-182.

Hedrick, P.W. 1987. Gametic disequilibrium measures: Proceed with caution. Genetics117: 331-341.

Henderson, C. R. 1975. Best linear unbiased estimation and prediction under a selection model. Biometrics 31:423-447.

Herms, U., and G. Brummer. 1984. Einflussgrossen der Schwermetalloslichkeit und -bindung in Boden. Z. Pflanzenernaehr. Bodenkd. 147: 421-424.

Hill, W.G., Robertson, A.(1968). Linkage disequilibrium in finite populations.Theoretical and Applied Genetics38, 226-231.

Jannink, J. L., and B. Walsh. 2002. Association mapping in plant populations. In Quantitative Genetics, Genomics and Plant Breedinging, M. S. Kang, Ed., pp. 59-68, CAB International, Oxford, UK.

Jaccoud, D., K. Peng, D. Feinstein, and A. Kilian. 2001. Diversity arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Res. 29: e25.

Jalil, A., F. Selles, and J. Clarke. 1994a. Effect of Cd on growth and uptake of Cd and other elements by durum wheat. Journal of Plant Nutrition 17: 1839-1858.

Jalil, A., F. Selles, and J. Clarke. 1994b. Growth and cadmium accumulation in two durum wheat cultivars. Soil Sci. Plant Anal. 25: 2597-2611.

Jay, R. K. 2011. Candidate gene association mapping in spring wheat. PhD diss. Montana State University, USA.

John, M. 1973. Cadmium uptake by eight food crops as influenced by various soil levels of cadmium. Environ. Pollution 4: 7-15.

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.

King, L. 1988. Effect of selected soil properties on cadmium content of tobacco. Journal of Environmental Quality 17: 251-255.

Lande, R., and R. Thompson. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124: 743-756.

Lewontin, R. C. 1963. The interaction of selection and linkage. I. General considerations; heterotic models. Genetics 49: 49-67.

Li, M., M. P. Reilly, D. J. Rader, L. S. Wang. 2010. Correcting population stratification in genetic association studies using a phylogenetic approach. Bioinformatics (Oxford) 26:798

Li, Y. M., R.L. Chancy, A. A. Schneiter and J. F. Miller. 1995. Genotypic variation in kernel cadmium concentration in sunflower germplasm under varying soil conditions. Crop Sci. 35: 137-141.

Li, Y., R. Chaney, A. Schneiter, E. Elias, and C. Green. 1993. Genetic and soil factors related to Cd levels in sunflower kernels and durum wheat. Agron. Abs. pp. 277.

Loiselle, B.A., V. L. Sork, J. Nason, and C. Graham. 1995. Spatial genetic-structure of a tropical understory shrub, psychotria officinalis (Rubiaceae). American Journal of Botany 82:1420-1425.

Lorenz, S., R. Hamon, S. McGrath, P. Holm, and T. Christensen. 1994. Application of fertilizer cations affect cadmium and zinc concentrations in soil solutions and uptake by plants. Eur. J. Soil Sci. 45: 159-165.

McClean, A. 1976. Cadmium in different species and its availability in soils as influenced by organic matter and addition of lime, P, Cd and Zn. Can. J. Soil Sci. 56: 129-138.

McLaughlin, M. J., K.G. Tiller, and M. K. Smart. 1997. Speciation of cadmium in soil solutions of saline/sodic soils and relationships with cadmium concentrations in potato tubers (Solanum tuberosum L.). Aust J Soil Res 35: 183-198.

McLaughlin, M., C. Williams, A. McKay, R. Kirkham, J. Gunton, K. Jackson, R. Thompson, B. Dowling, D. Partington, M. Smart, and K. Tiller. 1994. Effect of cultivar on uptake of cadmium by potato tubers. Australian Journal of Agricultural Research 45: 1483-1495.

McLaughlin, M., N. Maier, K. Freeman, K. Tiller, C. Williams, and M. Smart. 1995. Effect of potassic and phosphatic fertilizer type, fertiizer Cd concentration and zinc rate on cadmium uptake by potatoes. Fertilizer Res. 40: 63-70.

Michelmore, R. 1995. Molecular approaches to manipulation of disease resistance genes. Annual Review of Phytopathology 33: 393-427.

Moreau, L., A. Charcosset, F. Hospital, and A. Gallais. 1998. Marker assisted selection

efficiency in populations of finite size. Genetics 148: 1353-1365.

National Agricultural Statistics Service (NASS) Retrieved December 4, 2012, from http://www.nass.usda.gov/Statistics_by_State/North_Dakota/index.asp

Nordberg, G. F. 1972. Cadmium metabolism and toxicity. Experimental studies on mice with special reference to the use of biological materials as indices of retention and the possible role of metallothionein in transport and detoxification of cadmium. Environm Physiol Biochem 2: 7-36.

Oliver, D., R. Hannam, K. Tiller, N. Wilhelm, R. Merry, and G. Cozens.1994. The effects of zinc fertilization on cadmium concentration in wheat grain. Journal of Environmental Quality 23: 705-711.

Parizek, J. 1957. The destructive effect of cadmium on testicular tissue and its prevention by zinc. Journal of Endocrinology 15: 56-63.

Penner, G., J. Clarke, L. Bezte, and D. Leisle. 1995. Identification of RAPD markers linked to a gene governing cadmium uptake in durum wheat. Genome 38: 543-547.

Price, A.L., N. J. Patterson, R. M. Plenge, M. E. Weinblatt, N. A. Shadick, and D. Reich. 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nature Genetics 38: 904-909.

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

Rafalski, A. 2002. Applications of single nucleotide polymorphisms in crop genetics. Curr Opin Plant Biol 5:94-100.

Sparrow, L., A. Salardini, and A. Bishop.1994. Field studies of cadmium in potatoes (Solanum tuberosum L.). III. Response of cv. Russet Burbank to sources of banded potassium. Aust. J. Agric. Res. 45: 243-249.

Street, J., B. Sabey, and W. Lindsay. 1978. The influence of pH, phosphorus, cadmium, sewage sludge and incubation time on the stability and plant uptake of cadmium. Journal of Environmental Quality 7: 286-290.

Stuber, C., M. Polacco, and M. Senior. 1999. Synergy of empirical breeding, markers-assisted selection, and genomics to increase crop yield potential. Crop Sci. 39: 1571-1583.

Tanksley, S. 1983. Molecular markers in plant breeding. Plant Mol. Biol. Rep. 1: 3-8.

Tanksley, S., N. Young, A. Paterson, and M. Bonierbale. 1989. RFLP mapping in plant breeding: New tools for an old science. Biotechnology 7: 257-263.

Tiller, K., D. Oliver, M. McLaughlin, M. Conyers, R. Merry, and R. Naidu.1997. In: Iskandar, I. K. Ed., Remediation of Soils Contaminated by Metals, Science and Technology Letters, Northwood. pp.225-250.

Van Berloo, R., and P. Stam. 1998. Marker assisted selection in autogamous RIL populations: a simulation study. Theor. Appl. Genet. 96: 147-154.

Wagner, G. 1993. Accumulation of cadmium in crop plants and its consequences to human health. Adv. Agron. 51: 173-212.

Weber, W., and G. Wricke. 1994. Genetic markers in plant breeding. Advances in Plant Breeding No. 16. Parey Scientific Publ., Berlin.

Weeden, N., G. Timmerman, and J. Lu, 1994. Identifying and mapping genes of economic significance. Euphytica .73: 191-198.

William, Y., S. Wang, J. B. Bryan, G. C. David, and A. John. 2005. Genome-wide association studies: theoretical and practical concerns. Nature Reviews Genetics 6: 109-118

Yang, G., M. Saghai-Maroof, C. Xu, Q. Zhang and R. Biya-shev. 1994. Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice. Molecular and General Genetics 245:187-194.

Yousef, G. and J. Juvik. 2001. Comparison of phenotypic and marker-assisted selection for quantitative traits in sweet corn. Crop Sci. 41: 645-655.

Yu, J. M., G. Pressoir, W. H. Briggs, I. Bi, M. Yamasaki, J.F. Doebley, M. D. McMullen, B. S. Gaut, D. M. Nielsen, J. B. Holland, S. Kresovich, and E. S. Buckler. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nature Genetics 38: 203-208.

Zhu, C., M. Gore, E.S. Buckler, and J. Yu. 2008. Status and prospects of association mapping in plants. Plant Genome 1: 5-20.

CHAPTER II. GENOTYPING AND IDENTIFICATION OF QTL GOVERNING CADMIUM CONTENT IN DURUM WHEAT (*TRITICUM TURGIDUM* L. VAR. *DURUM* DESF.)

Abstract

Cadmium (Cd) is a biotoxic heavy metal present in low concentrations in the agricultural soils worldwide. High Cd level in durum wheat has become a serious problem that might have an impact on the national pasta industry and the international export market due to the large amount of Cd accumulated in their products. The objective of durum wheat breeding is to identify lines that are low in Cadmium accumulation by using linkage mapping. To achieve this goal, 178 recombinant inbred lines (RILs) derived from a cross between Grenora x Haurani were used to identify QTL for Cadmium tolerance. A total of 1,132 polymorphic loci (1,111 SNP and 21 SSR loci) were used for linkage map analysis. Subsequently, 330 of 1,132 loci (29 percent) markers were successfully integrated into the linkage maps at a LOD score of 3.0. The linkage map had a total genetic distance of 720.2 cM with an average distance of 2.2 cM between adjacent markers distributed on 14 chromosomes. Quantitative trait loci analysis was conducted with composite interval mapping (CIM). A single putative QTL associated with Cd uptake was detected on chromosome 5B. This QTL increased Cd uptake by the presence of alleles from Grenora. The additive effect of the Grenora allele was 0.02 mg/kg. The variation in Cd accumulation explained by this QTL (r^2) was 54.3 percent.

Introduction

Cadmium (Cd) is a biotoxic heavy metal present in low concentrations (milligrams per kilogram range) in the agricultural soils worldwide. Cadmium has no known biological function and may cause chronic toxicity in human bodies (NRCC, 1979). Due to the high mobility of Cd

in the soil-plant system, it can easily enter the food chain (Grant *et al.,* 1998) and cause harmful outcomes for human health, including damaged kidneys and osteoporosis due to the lack of calcium (Roberts, 1999). To form complex components that are transformed to the kidney, Cd first accumulated in the liver via the blood cells, where it combines with proteins. The accumulation of Cd in the liver damages its functions causing loss of protein and sugar from the body. As a result, the liver's tissue is harmed (Lind *et al.,* 1997).

The large-scale application of fertilizers, manure, sewage sludge, and aerial deposition of smelters dust play a critical role in soils contamination by Cd. Norvel *et al.* (2000) reported that the concentration of Zn, soil pH, organic matter content, clay content, and fertilization practices play an important role in the bioavailability of the cadmium element in soils. Phosphorus fertilizers have been shown to either increase (Williams and David, 1976) or have little or no effect (Mortvedt *et al.,* 1981) on Cd concentrations in crops. Eriksson (1990) reported that nitrogen fertilizers tend to increase Cd accumulation in plants either through increased yield or in their effects on soi1 chemistry. Oliver *et al.* (1993) have reported that crop rotation and tillage can influence Cd concentration in grain crops. Due to differences in soil management practices, exposure, and formation the amount of Cd added from each source differs with location, but the Cd level in the soil tends to increase over time (Jones and Johnston, 1989; Jones *et al.,* 1992).

Durum wheat (*Triticum turgidum* L. var. *durum* Desf.) is an economically important crop used primarily for pasta production. North Dakota produces on average 60 percent of the durum wheat in the United States, which is worth more than $529 million per year (NASS, 2012). High Cd level in durum wheat has become a serious problem that might have an impact on the national pasta industry and the international export market due to the large amount of Cd accumulated in their products. The increase of Cd level in the soil is harmful to the consumer

and to the environment. As a result, the level of Cd in durum wheat cultivars that will be released in the future must not exceed the grievous level. The proposed level by the Food and Agriculture Organization (FAO)/ World Health Organization (WHO, 2010) is 70-84 μg per person per day.

Cereal products tend to be the largest source of Cd in human diets, e.g., durum wheat, and tend to accumulate high concentrations of Cd in edible portions. The national mean values determined around the world for wheat (*Triticum* spp.) vary between 0.009 and 0.04 mg/kg (WHO, 2010). The Codex Alimentarius Commission of the FAO/WHO (2010) has proposed a maximum level of 0.2 mg kg^{-1} Cd for wheat grains traded in international markets. Wheat and rice are the most important source of Cd. In Japan, wheat grain Cd concentration and genotype influence varied substantial by among the 273 wheat cultivars analyzed (Kubo *et al*., 2008). The genotypic influence was confirmed in a study performed in Canada by using 13 bread wheat and two durum wheat cultivars. This trial was tested over six different site years (Gao *et al.,* 2011). A higher Cd concentration was found in durum wheat than in bread wheat (Grant *et al.,* 2008). Gao *et al.* (2011) have reported the level of Cd in bread wheat grain to be 0.010-0.045 mg/kg Cd compared with 0.060–0.145 mg/kg Cd in durum wheat.

Cadmium level determination in plants is relatively expensive; therefore, the use of molecular markers linked to the desired gene is an alternative to phenotyping. In durum wheat, two RAPD (random amplified polymorphic DNA) markers linked to the Cd gene were identified by using bulked segregant analysis (Penner *et al.,* 1995). Genetically, the regulation of Cd accumulation in grains is controlled by multiple genes with combined effects on uptake, translocation, and sequestration (Tanhuanpaa *et al.,* 2007). A single gene controlling grain Cd concentration, with low Cd being dominant, was found in durum wheat (Clarke *et al.,* 1997).

QTL for Cd concentration was also identified in other species. In rice grain (Oryza sativa L.), six putative QTL controlling Cd uptake have been mapped on chromosomes 3, 6, and 8 (Ishikawa *et al.,* 2010) and chromosomes 2, 5, and 11 (Ueno *et al.,* 2009), respectively, demonstrating the complexity of this trait. In contrast, a major QTL that controls Cd translocation from roots to shoots has been mapped on rice chromosome 7 by Tezuka *et al.* (2010). This QTL explained 88 percent of the phenotypic variation. Additional QTL on chromosomes 4 and 11 are identified by Kashiwagi *et al.* (2009). It was also suggested that low Cd accumulation is a dominant trait (Tezuka *et al.*, 2010). Soric *et al.* (2009) identified QTL associated with Cd uptake on chromosome 2 of maize. In maize (Zea mays L.) Wiebe et al. (2010) suggested that different physiological mechanisms are responsible for the phenotypic variation in Cd accumulation in plants. In durum wheat, it was suggested that the low cadmium phenotype is the result of restricted root-to-shoot Cd translocation (Harris and Taylor, 2001). This limits the size of the shoot Cd pool for remobilization to the grain. It is observed that the sequestration of Cd could occur in root tissues thereby reducing its availability in shoot. In addition, these several ABC transporters are involved in sequestration of cadmium into the vacuole (Song *et al.*, 2003; Klein *et al.,* 2006). Stolt *et al.* (2003) suggested the possibility that ABC transporters of root cells limit Cd translocation to the shoot for subsequent accumulation to the grain. Since the low Cd phenotype is dominant, the presence of a functional transporter or chelator that sequesters Cd in roots would result in a low cadmium phenotype (Wiebe *et al.,* 2010).

The objectives of this study were to 1) generate a molecular marker linkage map using the Grenora x Haurani RILs and 2) identify QTL controlling Cd accumulation in durum wheat.

Materials and methods

Plant materials and field evaluation

To develop the RIL mapping population, two durum wheat cultivars were used that had wide phenotyping variations and natural allelic differences between parental lines in Cd accumulation level. Grenora (Elias and Manthey, 2007a) was developed by the North Dakota Agricultural Experiment Station and is a high Cd accumulator. Haurani is a low Cd accumulator and is a local landrace selection from Syria developed by the International Center for Agricultural Research in Dry Areas (ICARDA). A cross of Grenora × Haurani was made in the fall of 2008. In the summer of 2009, the F_2 seed was grown in 8-row plots (25 feet long) at Casselton, ND. Three hundred spikes were harvested from the population for advancement.

The single seed descent method was used to develop the subsequent generations. In each generation, spikes were threshed individually, and two seeds from each spike were planted in the greenhouse in a single clay pot with a diameter of 15.2 cm and thinned to one plant after seed germination and seedling establishment. The seed from the F_2 spikes were planted in the fall 2009 greenhouse as an F_3 generation. Single spikes from each line were later harvested and 2 seeds grown for another season in the spring 2010 greenhouse to obtain the $F_{3:4}$ generation.

In the greenhouse, plants were grown in Sunshine Mix growth medium, fertilized with the recommended rate of Osmacoat at the Feekes growth stage 1.3, hand watered until approximately growth stage 6, and watered by a drip irrigation system in the later stages. Marathon 1% G was applied at Feekes growth stage 1.3 to control aphids. Powdery mildew (*Erysiphe graminis*) was controlled by sulfur vapors. Greenhouse temperature was maintained between 23°C and 29°C, and supplemental light was used to provide 16 hours of daylight and 8 hours of darkness.

Three spikes from each $F_{3:4}$ plants were harvested; seed from each of the harvested $F_{3:4}$ plants was planted at Prosper, ND, in the summer of 2010, and seed from the second spike was planted at Langdon, ND, as $F_{4:5}$ recombinant inbred lines (RIL) for preliminary Cd uptake evaluation and advancement. The third spike was kept as a remnant. The $F_{4:5}$ RILs, the parents, and checks were grown as an augmented block design (Federer, 1961) due to the small amount of seeds at the two locations Langdon and Prosper, ND. Field experiments were initiated on two soil types, representative of soils in North Dakota. The soil type at Langdon was Svea (fine-loamy, mixed Pachic Udic Haploborolls) and Barns (fine-loamy, mixed Udic Haplborolls). The Prosper soil types were Perella (fine-silty, mixed, frigid Typic Haplaquoll) and Bearden (fine-silty, frigid, Aeric Calciaquoll).

Because of high FHB disease pressure in 2010 and seed quality, the Prosper location was not harvested. At Langdon, three balanced bulks of subsamples from each line were used. The first set consisted of ten spikes that were selected randomly from each row to be tested for Cd content. To avoid any contamination of samples for Cd analysis, the harvested spikes were cut by stainless steel knives containing chromium that is resistant to corrosion and threshed in a thresher that contained no galvanized metal. Spikes from each RIL were threshed separately, and the balance bulk was used for Cd evaluation. The second subsample of seed was sent to an off-season nursery in New Zealand for increase to conduct a yield trial in 2011. The third subsample was saved as a remnant source. Because of the inadaptability of Haurani, heavy FHB infection, and poor seed germination due to the lack of viability of seeds which affected the germination rate, the population size of the Grenora × Haurani was reduced from 300 to 167 RILs. Ten grams of the seed from off-season nursery were evaluated in preliminary yield trails (PYT) in the summer of 2011 as $F_{4:6}$ RILs in two locations (Langdon and Prosper), and a

subsample was grown in the greenhouse in the spring of 2011 for DNA extraction. During the growing season, agronomic data, including yield, test weight, height, heading date, lodging, plant aspect, and disease infection, were collected. Grain yield was obtained as g/plot and converted to Kg/ha. Test weight was measured with a half-pint size container, using formula: Test weight (lb/bu) = (seed weight of half pint (g) – 294.4) ×0.28224 and converted to Kg/m^3. Height was measured from the ground level to the tip of the spikes, excluding awns, and recorded in centimeters (cm). Heading date was expressed as days after seeding to which 50 percent of the spikes of a specific plot were completely out of the boot.

In the summer of 2011, the population was grown and evaluated as a (13 x 13) simple lattice design in two locations (Langdon and Prosper, ND) and two replicates. The checks Strongfield (Clarke *et al.*, 2005) and Alkabo (Elias and Manthey, 2007b), parents, and RILs were planted as single-row plots 2.75 m long with a 30 cm row spacing between rows. Seeding rate was 10 g of seeds per row which translates to approximately 2.5 million plants per hectare. Planting, field fertility, herbicide and fungicide application, agronomic data collection, and harvest were concurrent with the North Dakota State University Durum Wheat Breeding Project and are typical of North Dakota growing conditions. Fungicide applications consisting of Folicur 3.6F (alpha-[2-(4-chlorophenyl)-ethyl]-alpha (1, 1-dimethylethyl)-1H-1, 2, 4-triazole-1-ethanol) (Bayer Corporation, Kansas City, MO) at flag leaf stage, using the rate of 267.1 ml/ha, was used at both locations and the two years (2010 and 2011) to minimize the infection by Fusarium head blight caused by *Fusarium graminearum* Schwabe. A herbicide tank-mixture of Discover (Syngenta Crop Protection, Inc. Greensboro, NC, at rate 292.2 ml/ha), Buctril (Bayer Corporation, Kansas City, MO, at rate 1.17 l/ha), and MCPA ester (584.3 ml/ha) was applied at both locations for weed control.

Cadmium content estimation

Samples were sent for Cd analysis to Dr. Michael A. Rutzke, College of Agriculture and Life Sciences Nutrient Analysis Laboratory, Cornell University, NY. A representative flour sample of up to 0.5 g from each RIL was digested in 4ml of mix (40 percent of concentrated nitric acid and 60 percent of perchloric acid) and an extra 1 ml perchloric acid. The sample and acid were placed in a fluorocarbon vessel. The open vessel was heated on a hot plate unit. After cooling, the vessel contents were allowed to settle and then diluted to a 20 ml volume and analyzed with the appropriate SW-846 method (EPA Method No. 3051, 3050, 3052).

Experimental design and statistical analysis

Entries in the yield trial experiments were assigned to experimental units using an augmented block design in 2010 and a 13 x 13 simple lattice in 2011. In 2010, the augmented block design consisted of eight blocks with 22 RIL per block and the two parents replicated in each block. The phenotypic data were analyzed using the Statistical Analysis System (SAS Institute, Cary, NC, 2012). Homogeneity of error variances among environments was determined by the ratio of the largest error variance to the smallest. If the ratio was less than 10, then variances were considered homogeneous. The individual environment-adjusted genotypes were used for the combined analysis using randomized complete block design where years were considered as replicates. Locations were treated as a random effect, whereas the genotypes were considered as a fixed effect. Means were separated by Fisher's Protected LSD at the 5 percent level of significance. Spearman rank ccorrelations among traits were calculated in each individual location. The correlation coefficients from each location were tested for homogeneity and pooled when appropriate as described in Gomez and Gomez (1984).

Genetic analysis

DNA isolation

Three remnant seeds of the 169 RILs from the Grenora x Haurani cross were grown in the spring 2011 greenhouse for DNA extraction. Three fresh leaf tissues were harvested from three plants of the same line, bulked, and stored at -80^{o}C, then lyophilized for 48 hours and kept at room temperature. The DNA was extracted using the procedure developed by Triticarte Pty.Ltd. (Canberra, Australia; http://www.triticate.com.au) with a few modifications. The tissues were ground in a mixer mill at frequency 26 for 3 min to fine powder. The resulting tissue powder was transferred to a 2 ml polypropylene tube. The samples in each tube were treated with 900 µl of fresh buffer solution (0.15 g sodium disulfite, 0.6 g PVP-40 (K29-32) preheated to 65 ^{o}C, 0.6 g CTAB, 12.5 ml extraction buffer (0.35 M sorbitol, 0.1 M TrisHCl pH 8.0, 5 mM EDTA pH 8.0), 12.5 ml lysis buffer stock (0.2 M TrisHCl pH 8.0, 0.05 M EDTA pH 8.0, 2 M NaCl), and 5 ml 5 percent sarcosyl stock). The contents were mixed and incubated in a water bath at 65^{o}C for 1 hour. Approximately 900 µl of chloroform: isoamyl alcohol (24:1) was added to each tube. All the tubes were balanced and centrifuged (25 min, 11,000x, 25^{o}C). Around 600 µl of aqueous phase was removed and transferred from each tube to a fresh 1.5 ml tube, and 600 µl of isopropanol was added to each tube followed by mixing 5 min. The new tubes were balanced and centrifuged (15 min, 11,000x, 25^{o}C). The isopropanol was removed to leave a DNA pellet in bottom of tube. The DNA pellet was washed with 950 µl of 70 percent ethanol by using a centrifuge (15 min, 11,000x, 25^{o}C). Then, the ethanol was removed carefully. Once the washes were complete, the DNA was oven-dried at 37^{o}C for 15-30 min to remove any traces of ethanol and then resuspended in 200 µl of double distilled water and left at 14^{o}C to dissolve overnight.

NanoDrop 2000 (Thermo Fisher Scientific, West Palm Beach FL, USA) was used to measure the concentration of the DNA. The average DNA yield from the given procedure was 323.5 ng/µl. The samples were checked for integrity, purity and quality by running 1 µl of each sample on a 1 percent agarose gel after incubation in digestion buffer for 1 hour.

DNA bulks

The phenotypic data were grouped into two groups: high and low cadmium. The extracted DNA of every 10 lines was combined into one pool, taking into consideration the number of individuals in each group, to yield seven pools containing an equal amount of DNA up to approximately 75 ng/µl with a total volume of 100 µl. The total number of pools was 14 for the two groups in addition to two pools representing the low and high cadmium uptake parents. The extracted DNA samples were sent to the USDA-ARS Cereal Crops Genotyping Laboratory (Fargo, ND) for genotyping with microsatellite markers (SSR) using Genetic Analyzer 3130 (www.appliedbiosystems.com). The software GeneMapper v.3.7 was used to identify 255 polymorphic markers that segregate for low and high cadmium across all durum wheat chromosomes. Two SSR markers one from each arm were selected to screen 167 lines generated from the cross of Grenora x Haurani. Thirty SSR markers were used for the genotypic evaluation of lines. The genotypic data of the SSR markers were combined with SNPs data to construct the linkage map and to locate the QTL on chromosomes.

Genotyping

The DNA was extracted using the procedure developed by Triticarte Pty. Ltd. (Canberra, Australia; http://www.triticate.com.au) with few modifications. The extracted DNA samples were sent to the USDA-ARS Cereal Crops Genotyping Laboratory (Fargo, ND) for genotyping with high density single nucleotide polymorphism (SNP) markers with a custom-designed

Illumina iSelect BeadChip platform containing 9,000 SNPs and using the Illumina Infinium Assay protocol (Steemers *et al.,* 2006). In addition thirty SSR markers, two SSR markers from each chromosome, one on each arm, were selected to screen RIL generated from the cross of Grenora x Haurani.

Infinium assay protocol consists of three steps: genomic DNA hybridization to beads, single-base extension, and allele-specific detection methods. The bead chips were scanned, and the hybridization intensity values were obtained by using the BeadStation scanner. The hybridization intensity values were used as the basis for genotype calling. GenomeStudio software, developed by Illumina, was used to determine the genotype calls. A cluster is formed if samples exhibit the same genotype. Three clusters were identified by the calling algorithm, each cluster corresponding to one of the three genotypes AA, AB, and BB (Figure 2.1). Each genotype call was validated manually and edited to avoid errors related to cluster compression. The genotypic data were then exported from the software and used for further analysis.

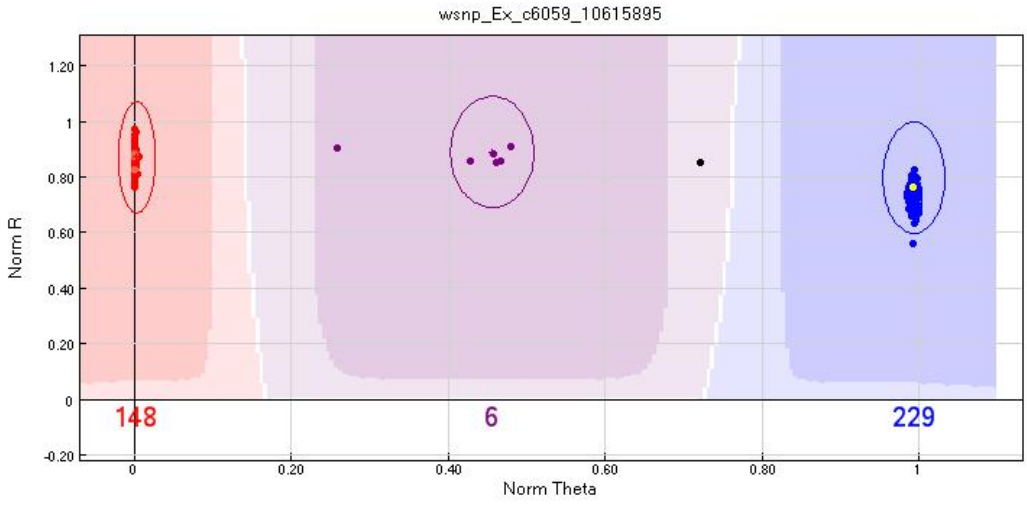


Figure 2.1. Genotype call output from GenomeStudio software. Three clusters representing AA, AB, and BB genotype classes (left to right).

Genetic linkage map construction

A map for the Grenora x Haurani population was obtained using 549 SNP and 21 polymorphic SSR markers. CarthaGene (Schiex and Gaspin, 1997) which is a maximum-likelihood multipoint genetic data mapping tool (http://www.inra.fr/bia/T/CarthaGene) was used. The command "*mrkmerges"* was used to combine the co-segregating markers (no recombination). Linkage groups were then identified by using the command "*group"* with a maximum of 50 cM distance and a LOD score of 3. With a keep and add threshold of LOD 3, *"Buildfw"* was used to construct the linkage groups. Further, the command *"flips"* was used to check for all possible permutations of the order, and the command *"polish"* was used to check the reliability of the map by removing one marker from the initial map and trying to insert it in all possible intervals. Kosambi (1944) map function was used to present map distances in centimorgans (cM).

QTL analysis

The data obtained from the linkage maps were imported for QTL identification to QGene 4.3.8 (Joehanes and Nelson, 2008). QTL analysis was conducted with composite interval mapping (CIM). This type of analysis considers the markers outside the interval as cofactors that can be helpful in estimating the effect of a QTL within the interval (Zeng, 1994). The default parameters in QGene were used to select cofactors. A walking distance of 2 cM was used for QTL analyses. To locate chromosomal regions associated with the trait, permutation tests (1,000 iterations) were run to determine the minimum LOD value needed before a marker-trait association would be considered significant ($\alpha_{0.01}$ and $\alpha_{0.05}$ experiment-wide error). The additive effect and percentage of phenotypic variance explained by each QTL were also estimated using QGene. To account for the unlinked markers (markers that do not belong to any linkage group),

QTL analysis was also performed using single marker regression that looks at each individual marker and performs one-way analysis of variance.

Marker validation

To verify the effectiveness and efficiency of the identified SNP marker in selecting for low Cd accumulation genotypes in durum wheat KASPar (Kompetitive Alelle Specific PCR), genotyping assay was used for the validation (LGC Ltd., Teddington, UK). KASP assay is a PCR-based genotyping assay that enables accurate bi-allelic discrimination of known SNPs at specific loci across a wide range of genomic DNA samples. The SNP marker was subjected to validation using other populations of recombinant inbred lines consisting of 178 RILs developed from a cross between a low Cd cultivar Strongfield and high Cd cultivar Alkabo.

Results and discussion

Descriptive statistics of phenotypic data

In 2010, despite of fungicide application, the Prosper location was not harvested because of a heavy infection with FHB that resulted in a very shriveled and low number of seed per spike. In 2011, the Prosper location again was lost because of heavy rainfall that flooded the whole durum wheat nursery. Therefore, only the 2010 and 2011 Langdon location data were used in the analysis. Spearman rank correlation was used to test for genotype by environment interaction whether it was due to the rank or magnitude between the data obtained in 2010 and 2011. A positive and significant correlation was observed between 2010 and 2011 for Cd content ($r=0.50$, $P\leq 0.05$), indicating RILs maintained their rank for both years. Mean values and descriptive statistics from the analyses of variance (ANOVA) of the parents and RILs for the phenotypic trait are presented in Table 2.1. The parental lines Grenora and Haurani differed significantly for Cd uptake. Variation in Cd concentration showed a skewed distribution toward

low concentrations. The mean parental values for Cd concentration were 0.0652 mg/kg for Grenora and 0.0275 mg/kg for Haurani, with the progenies ranging from 0.0123 to 0.1193 mg/kg (Table 2.1, Figure 2.1).

The top 15 durum wheat lines were selected based on Cd level in grains across 2010 and 2011 (Table 2.2). Cadmium level in these lines ranged from 0.014 to 0.026 mg/kg which is still numerically lower than the low Cd parent Haurani and the checks (Table 2.2). This indicates the possibility of developing low Cd uptake germplasm/cultivars from this population. Days to heading, plant height, test weight, and yield also are presented in Table 2.2. In general, most of the RIL's had yield lower than the parent Grenora but higher than the check Alkabo. Yield ranged from 3157.9 Kg/ha to 5610.4 Kg/ha (Table 2.2). This big difference is due to the fact that Haurani is not an adapted cultivar grown in North Dakota. However, some lines showed promising yields for future selections. The RIL D102605 had a Cd concentration lower than the parents and the checks and a higher yield than both parents and the checks. The test weight of this line was lower than the parents, but higher than the check Alkabo. This line will be tested further for quality traits because of its low Cd uptake and yield potential. In general, sound, clean, vitreous wheat with low moisture content tended to give the highest test weights (Matsuo and Dexter, 1980). Stresses that damage kernels (e.g., frost, weathering, drought, and disease) tend to reduce kernel size and test weight (Saadalla *et al.*, 1990). A plump kernel, free of any damage and fully mature, describes the sound wheat that yielded a high test weight. The acceptable test weight for durum wheat produced in North Dakota is 772.8 Kg/m^3. The average test weight of durum wheat grown in North Dakota in 2010 and 2011 was 756.1 Kg/m^3 and 777.9 Kg/m^3, respectively (NASS). Test weight ranged from 724.7 to 768.9 Kg/m^3. The low test weight of these lines is probably due to the low test weight contributed by the Haurani parent. The

highest value of test weight of the RIL D102696 (768.9 Kg/m^3) was very close to the Grenora value of 784.4 Kg/m^3 but higher than Alkabo (750.9 Kg/$m^{3)}$. Days to heading of the RILs ranged from 63 to 65 days from the planting date which was lower than the parents and the checks (Table 2.2).

Table 2.1. Overall means and descriptive statistics of Cadmium level for Grenora x Haurani recombinant inbred lines evaluated at Langdon, ND, across 2010 and 2011.

	Cd
	(mg/kg)
Parents	
Grenora	0.0652
Haurani	0.0275
Population statistics	
Mean	0.0496
Minimum	0.0123
Maximum	0.1193
Standard deviation	0.0253
Strongfield	0.0312
Alkabo	0.0915
‡$LSD_{1(0.05)}$	0.0315
‡$LSD_{2(0.05)}$	0.0253
†CV%	44.14

‡ $LSD_{1\ (0.05)}$ = Least significant difference comparing individual genotypes; $LSD_{2\ (0.05)}$ Comparing individual genotype with population mean.
†CV%= Coefficient of variation.

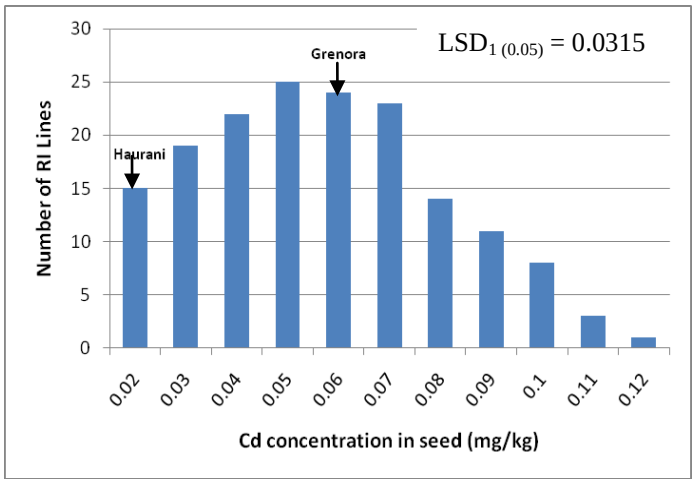


Figure 2.2. Phenotypic distribution of Cd concentration in the RILs mapping population derived from the cross Grenora (high Cd) x Haurani (low Cd) evaluated in 2010 and 2011. Arrow indicates the level of the two parental lines.

Table 2.2. Cadmium uptake level, test weight, yield, days to heading, and plant height of the top 15 durum wheat lines developed from a cross between Grenora x Haurani grown in Langdon, ND, in 2010-2011.

Line≠	Cadmium	Test weight	Yield	Days to heading	Plant height
	mg/kg	Kg/m^3	Kg/ha	day	cm
D102563	0.015	740.6	4085.2	63	104
D102574	0.017	740.6	4353.9	64	106
D102576	0.023	767.6	4219.5	64	105
D102582	0.019	761.2	4185.9	63	106
D102584	0.026	743.2	4488.3	64	108
D102601	0.023	750.9	3655.1	64	107
D102605	0.021	763.8	5610.4	65	102
D102606	0.021	758.6	4145.6	64	108
D102626	0.026	753.5	4521.9	65	104
D102641	0.022	743.2	4343.1	65	106
D102653	0.018	744.7	3486.5	65	107
D102655	0.025	724.7	3248.7	65	107
D102657	0.022	747.0	3157.9	63	110
D102696	0.014	768.9	3547.6	63	101
D102697	0.024	747.0	3346.1	63	109
[1]Grenora	0.061	784.4	4965.3	67	94
[2]Haurani	0.027	717.4	2936.2	66	95
StrongField	0.031	775.4	4091.9	68	95
Alkabo	0.091	750.9	3675.3	67	92
$LSD_{(0.05)}$	0.032	63	1955	6	15

[1] Grenora = high Cd, used as male parent. [2] Haurani = low Cd, used as female parent.

Phenotypic correlation

The correlation coefficient was calculated across locations for Cd uptake and all the phenotypic traits studied such as heading date, plant height, yield, and test weight (Table 2.3). There was no significant correlation between Cd uptake and all the agronomic traits, indicating Cd uptake is independent of yield, test weight, height and days to heading. Yield and test weight were significantly correlated ($r=0.51$, $p < 0.05$) (Table 2.3). No significant correlations were observed among test weight, days to heading, and height. Although non-significant, Cadmium level in general had a positive association with all agronomic traits, except with plant height (Table 2.3).

Table 2.3. Correlation coefficient among traits[†] for RILs grown in Langdon, ND, in 2010 and 2011.

	CD	TWT	YLD	DTH
TWT	0.13			
YLD	0.05	0.51*		
DTH	0.10	-0.04	-0.01	
HT	-0.15	-0.02	-0.04	-0.04

[†] DTH: days to heading (days); HT: height (cm); YLD: yield (Kg/ha); TWT: test weight (Kg/m^3); Cd: Cadmium level (mg/kg).
[*] Association between traits exists at $\alpha = 0.05$.

There were no correlation between plant height and Cd accumulation (Table 2.3). The top 15 low Cd lines were taller than the parents and the checks. The same results were seen for the top 15 high Cd lines. This suggests that Cd accumulation was not affected by plant height. Generally, roots are the major source of Cd uptake in crops rather than Cd absorbed from the atmosphere via the shoots (Smolders, 2001). In rice, as a model organism, Cd accumulation is regulated physiologically by Cd uptake by root, the translocation of xylem from root to shoot, and then the movement of phloem from shoot into grain (Clemens *et al.*, 2002; Hart *et al.*, 1998).

In soybean, Shute and Macfie (2006) reported that the application of Cd in a high dose (100 mg/kg) reduced plant height and dry weight by 40 percent and 34 percent of control, respectively. Although most of the correlations were not large, they still suggested a relatively weak association of Cd and other agronomic traits. Since the correlation coefficients were relatively small and showed a weak association, a larger scale of testing with more years, locations, and replications would be appropriate to give a positive conclusion about the association between Cd content and other agronomically important traits, especially the strength of the association. Since these correlations were weak, it may be possible to develop low Cd cultivars with a high yield and good agronomic traits.

Linkage map construction

A total of 1,132 polymorphic loci (1,111 SNP and 21 SSR loci) were used for linkage map analysis. 0330 of 1,132 loci (29 percent) markers were successfully integrated into the linkage analysis at a LOD 3.0 (Table 2.4). The linkage map had a total genetic distance of 720.2 cM with an average distance of 2.2 cM between adjacent markers distributed on 12 chromosomes. The average length per linkage group was 72 cM, but ranged from 6.9 cM (chromosome 3B-1) to 151.2 cM (chromosome 1A). The number of markers in each linkage group varied from 9 (chromosome 4B) to 63 (chromosome 1A) (Table 2.4). The average distance between linked markers for each linkage group varied from 0.7 cM (chromosome 3B-1) to 3 cM (chromosome 5B-1) (Table 2.4). SSR markers were used to identify each linkage group (Sourdille *et al.*, 2004).

QTL analyses

The linkage map developed with SNP and SSR markers was used to conduct QTL analyses. Table 2.5 provides a summary of the composite interval mapping (CIM) analyses using Grenora x Haurani RILs. A single putative QTL associated with Cd uptake was detected on Chromosome 5B (Table 2.5). Cadmium uptake QTL had its peak at 82.9 cM on Chromosome 5B with a LOD of 28.1 ($\alpha_{0.05} = 2.9$; $\alpha_{0.01} = 3.6$) (Figure 2.3). The QTL lies within the region of 76-85.5 cM. The only marker present in this region is *Ex_c1343_2570756.* This QTL increased Cd uptake because of the presence of alleles from Grenora. The additive effect of the Grenora allele was 0.02 mg/kg. The variation in Cd accumulation explained by this QTL (r^2) was 54.3 percent (Table 2.5). When a single marker analysis was performed, another five markers associated with the Cd accumulation in durum wheat were identified (Table 2.5). One (*Ex_c17754_26503892*) was located on Chromosome 5B at the same genetic distance of 82.9

Table 2.4. Distribution of molecular markers among 12 chromosomes established on a genetic linkage map using an RIL population derived from the cross Grenora (high Cd) x Haurani (low Cd).

Chromosome‡	No. of markers	Map length (cM)	Average distance (cM)	Marker type	
				SNP	SSR
1A	63	151.2	2.4	61	2
1B	55	145.1	2.6	53	2
3A	52	61.9	1.2	52	0
3B-1	10	6.9	0.7	10	0
3B-2	41	99.4	2.4	41	0
4B	9	25.1	2.8	9	0
5B-1	31	92.7	3.0	30	1
5B-2	44	76.1	1.7	44	0
6B	13	34.5	2.7	12	1
7B	12	27.3	2.3	10	2
Total	330	720.2	2.2	322	8

‡ The numbers followed by each letter represent different linkage groups for the same chromosome.

cM as the major QTL (*Ex_c1343_2570756*) found earlier in this study. The LOD score for both markers differed by 1.8 units. Each of these markers explained more than 50 percent of the phenotypic variation (Table 2.5).

Two putative QTL for yield (Figure 2.4) were detected on Chromosomes 1A and 3A. These QTL mapped at 11.5 cM and 89.2 cM, respectively (Table 2.6). The additive effect of the Grenora allele was 2.25 and 3.16 bu/a, respectively. The significant LOD scores for each linkage group for the $\alpha_{0.05}$ experiment-wise errors were 1.5 and 2.9. Each QTL for yield explained 4.2 to 7.8 percent of the phenotypic variation. Two putative QTL for plant height (Figure 2.4) were detected on chromosome 3A and 5B. These QTL mapped at 11.4 cM and 11.5, respectively (Table 2.6). The significant LOD scores for each linkage group for the $\alpha_{0.05}$ experiment-wise errors were 0.7 and 4.6. The additive effect of the Grenora allele was 1.33 cm, and the additive effect for the Haurani allele for the same trait was -0.98 cm. Each QTL explained 2.2 to 4.5 percent of the phenotypic variation. Several other QTL were also detected on Chromosome 3A,

but they were not significant when conducting the permutation test. In this study, no QTL associated with test weight was detected.

Table 2.5. LOD score and phenotypic variation of markers associated with Cd accumulation in durum wheat.

Locus name	Chromosome‡	LOD Score	r^2
Ex_c1343_2570756	5B	28.1	54.3
Ex_c17754_26503892	5B	26.3	51.6
Ku_c19334_28808006	n/a	25.4	50.8
Ex_c210_411604	n/a	25.4	50.8
JD_c37023_27225840	n/a	26.3	51.6
Ex_c29304_38355434	n/a	26.3	51.6

‡n/a = information about the location of SNP marker on chromosomes was not available.

The identified SNP marker (*Ex_c1343_2570756*) was validated in different genetic background populations (StrongField x Alkabo), using the KASPar genotyping assay. The population consists of 178 RILs. Twenty five of 178 lines showed false positives which are 14 percent of total lines. The SNP marker was able to identify 43 percent of the low Cd lines. The same data set was analyzed by using regression analysis. The R^2 was 64.7% and highly significant. This suggests the effectiveness of this marker in different populations with different genetic backgrounds and therefore can be used in future selections for low Cd accumulation in durum wheat.

The need to develop low Cd accumulating durum cultivars is a necessity to minimize the level transferred from the soil to the plant and further to the human diet. Because Cd phenotyping is laborious and expensive, the use of marker-assisted selection could be an alternative approach for selecting durum wheat lines with low Cd accumulation (Penner *et al.*, 1995) without compensating other grain quality traits such as test weight and yield. In this study, 15 lines were selected phenotypically based on cadmium level in grain.

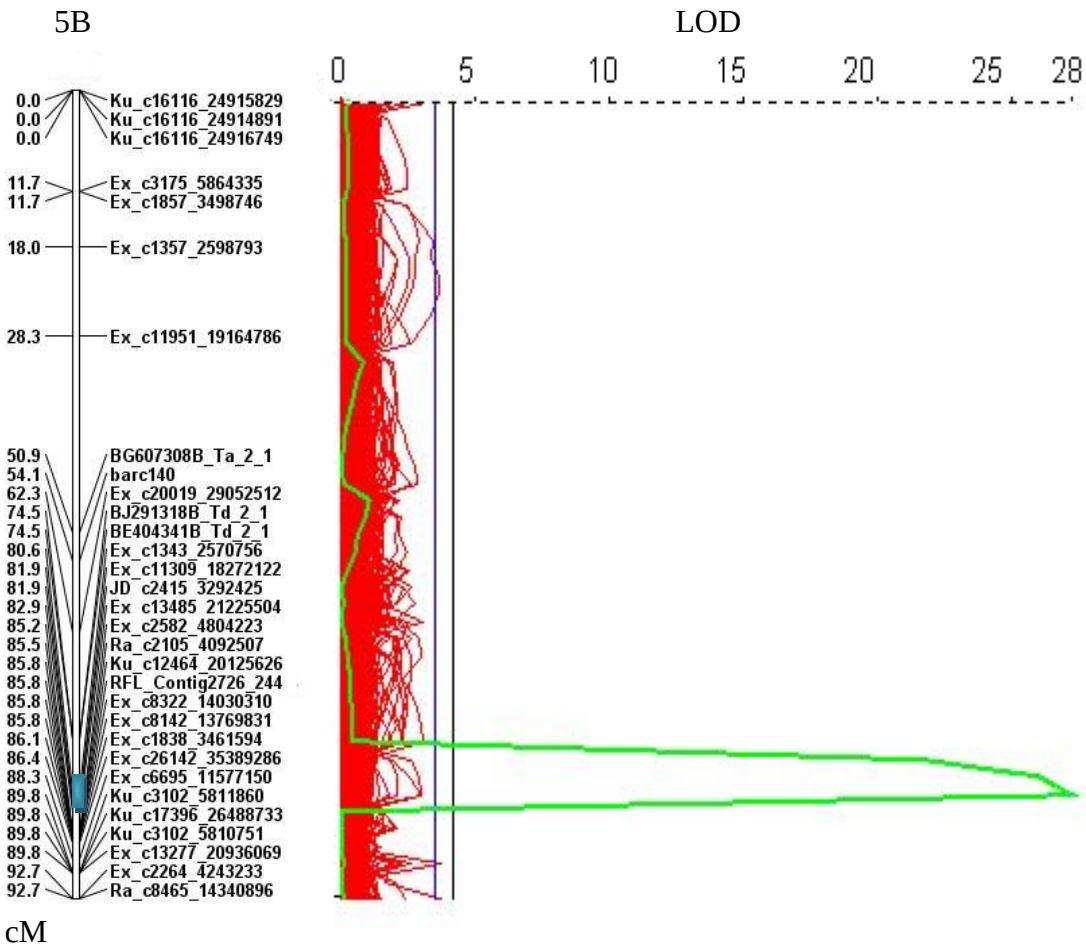


Figure 2.3. Major QTL for Cd accumulation on Chromosome 5B in the RILs population derived from the cross Grenora (high Cd) x Haurani (low Cd) was identified using composite interval mapping. Significant QTL for Cadmium was indicated by bars on Chromosome 5B.

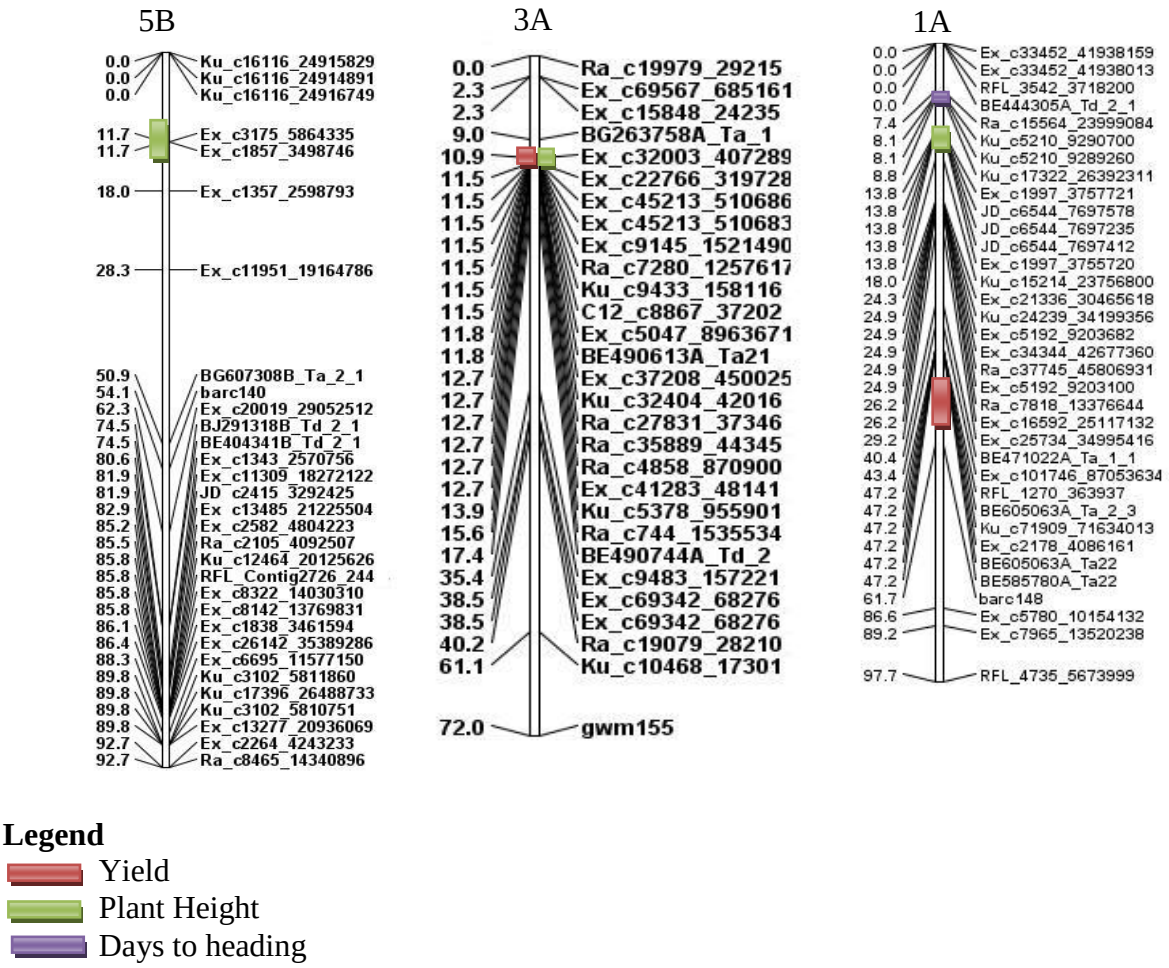


Figure 2.4. Genetic linkage map of durum wheat (*Triticum turgidum* L. var. *durum*) based on an RILs population derived the cross of Grenora (high Cd) x Haurani (low Cd).The marker name and map distances in centiMorgans (cM) are indicated on the left and right side of the chromosomes, respectively. Significant QTL for days to heading, yield, and plant heights were indicated by bars on each corresponding chromosome.

Table 2.6. Composite interval mapping analysis based on overall means from the Grenora x Haurani RILs for Cd, test weight, yield, plant height, and days to heading.

Trait‡	SNP	Chromosome	Position (cM)	LOD[a]	r^2† (%)	Source	A[b]
Cd	Ex_c1343_2570756	5B	82.9	28.1	54.3	Grenora	0.017
DTH	Ex_c33452_41938159	1A	0.0	2.3	6.1	Haurani	-0.281
YLD	Ex_c22766_31972812	3A	11.5	1.5	4.2	Grenora	2.25
	Ex_c7965_13520238	1A	89.2	2.9	7.8	Grenora	3.16
PHT	Ex_c3175_5864335	5B	11.4	1.6	4.5	Grenora	1.33
	Ex_c22766_31972812	3A	11.5	0.7	2.2	Haurani	-0.98

‡Cd = Cadmium, DTH = Days to Heading, YLD = Yield, PHT = Plant Height
†The percent of variation explained by the QTL.
a based on 1,000 permutations.
b additive effect. Positive or negative value indicates that the allele from Grenora or Haurani increases the phenotypic value, respectively.

Our goal was to identify a QTL that governs Cd accumulation in durum wheat. To achieve this, a linkage map of the Cd region in durum wheat is required. Knox *et al.,* (2009) and Wiebe *et al.,* (2010) localized *Cdu1* as a Mendelian factor to chromosome 5B in durum wheat. Another minor QTL has been mapped on the same chromosome near *Cdu1* (Wiebe *et al.,* 2010). Both grain cadmium content and root-to-shoot partitioning are associated with the *Cdu1* gene (Hart *et al.*, 2006). The QTL detected on Chromosome 5BL for Cd accumulation in this study was similar to the one detected on the same chromosome arm by Knox *et al.* (2009) and Wiebe *et al.* (2010). The other markers identified here could be either near this QTL or could be another minor QTL. The identified marker in this study appeared to be linked to Cd level and was co-segregating with the trait with low recombination rates. We are confident that these SNP markers are good for marker-assisted selection for Cd content in durum wheat populations. In a study conducted by Faris et al. (1996), several microsatellite markers (e.g., *Xfcp1*, *Xfcp2*, *Xfcp393*and *Xfcp394*) have been identified on Chromosome 5BL. These SSR markers are close to tan spot locus (*Tsn1).* The tan spot gene (*Tsn1)* lies within a gene-rich region of Chromosome 5BL between deletions 5BL-14 and 5BL-9, which accounts for about 1 percent of the physical

size of the long arm (Chu *et al.*, 2008). The markers that are associated with tan spot of wheat and low Cd level might be used in future breeding as an efficient selection method for low Cd in durum since they are all located on the long arm of Chromosome 5B and as long as they provide satisfactory polymorphism among lines of the populations under selection. The use of markers for selecting low Cd lines may help to reduce the environmental effect, thus indicating the true genetic basis of the materials undergoing selection.

In this study, the level of Cd for the top 15 RILs developed by the cross between Grenora and Haurani ranged from 0.014 to 0.026 mg/kg. The level of Cd for these lines was numerically lower than the low Cd parent (Hurani) and also lower than the Strongfield cultivar which exhibited 0.031mg/kg. In addition, all these lines had the marker associated with low Cd uptake. The high level of Cd accumulation in grains may be due to the translocation of Cd from leaves to stem to grain (Harris and Taylor, 2001; Chan and Hale, 2004). In this study, some RILs were extreme to the Cd level relative to those of either parental line. This transgressive segregation of lines, however, suggests that some additional minor genes may also influence Cd accumulation, indicating that Cd accumulation in durum might be a quantitative trait. The low Cd lines revealed in this study are potential candidates as sources of low Cd cultivars for food safety and human health in the breeding program at North Dakota State University. The five markers identified using single marker analysis could also be used as a selection tool in durum wheat breeding programs targeting reduced grain Cd levels. However, these markers should be validated in a larger germplasm pool to determine their effectiveness for selecting low Cd lines in durum wheat. A number of recombinant inbred lines showed potential for possible release, according to both Cd level data and their agronomic performance (heading date, plant height, and yield). There were 82 RILs that had a high yield, high test weight, and low Cd than Grenora or

Haurani, with an acceptable heading date and plant height in the environments tested (data not shown). Since RILs are already homozygous, no generation advancing is needed. These lines will be tested in advanced yield trials with four-row plots, more replicates, and more environments (year, location) to give a reliable estimate of their Cd level, yield potential, quality, and performance stability.

It is important to use durum wheat cultivars that absorb less Cd, suggesting the importance of incorporating a low Cd concentration trait in the durum wheat breeding program. The identified marker can be useful for marker-assisted selection and for better understanding of Cd translocation and accumulation in grain. Additional markers would facilitate fine mapping and positional cloning of this region and confirm further the presence of *Cdu1* in this region in addition to identifying the minor QTL that could be responsible for Cd accumulation in durum wheat.

Conclusion

In this study, a major QTL that explained 54 percent of the phenotypic variation was identified. This QTL is located on Chromosome 5B. The marker associated with this QTL is *Ex_c1343_2570756*. This marker will be used for future selections for Cd uptake in the breeding program. The use of this marker and other markers identified in different studies will be valuable and cost effective. The lines developed in the breeding program can be selected for low Cd in the early generations. This way, lines with high Cd can be eliminated without the need for phenotyping.

References

Chan, D.Y., and B. A. Hale. 2004. Differential accumulation of Cd in durum wheat cultivars: uptake and retranslocation as sources of variation. Journal of Experimental Botany 55: 2571-2579.

Chu, C.G., S.S. Xu, J.D. Faris, E. Nevo, and T.L. Friesen. 2008. Seedling resistance to tan spot and Stagonospora nodorum leaf blotch in wild emmer wheat (Triticum dicoccoides). Plant Dis. 92: 1229-1236.

Clarke, J. M., D. Leisle, and G.L. Kopytko. 1997. Inheritance of cadmium concentration in five durum wheat crosses. Crop Sci. 37: 1722-1726.

Clarke, J.M., T.N. McCaig, R.M. DePauw, R.E. Knox, F.R. Clarke, M.R. Fernandez, and N.P. Ames. 2005. Strongfield durum wheat. Canadian Journal of Plant Science 85: 651-654.

Clemens, S., M.G. Palmgren, and U. Kramer. 2002. A long way ahead: understanding and engineering plant metal accumulation. Trends in Plant Sci. 7: 309-315.

Elias, M. E., and F.A. Manthey. 2007a. Registration of "Grenora" Durum wheat. Journal of Crop Registrations 1: 8-9.

Elias, M. E., and F.A. Manthey. 2007b. Registration of 'Alkabo' durum wheat. Journal of Crop Registrations 1: 10-11.

EPA Method No. 3051, 3050, 3052. Test methods for evaluating solid waste, physical/chemical methods, 3rd ed; U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response. U.S. Government Printing Office: Washington, DC, 1986; SW-846.

Eriksson, J. 1990. A field study on factors influencing Cd levels in soils and in grain of oats and winter wheat. Water Air Soil Pollut. 53: 69-81.

FAO/WHO. 2010. Evaluation of Certain Food Additives and Contaminants (Seventy-third Report of the Joint FAO/WHO Expert Committee on Food Additives). WHO Technical Report Series, No. 960 (http://whqlibdoc.who.int/trs/WHO_TRS_960_eng.pdf)

Faris, J.D., J.A. Anderson, L.J. Francl, and J.G. Jordahl. 1996. Chromosomal location of a gene conditioning insensitivity in wheat to a necrosis-inducing culture filtrate from Pyrenophora tritici-repentis. Phytopathology 86: 459-463.

Federer, W.T. 1961. Augmented designs with one way elimination of heterogeneity. Biometrics 17: 447-473.

Gao, X., R.M. Mohr, D.L. McLaren, and C.A. Grant. 2011. Grain cadmium and zinc concentrations in wheat as affected by genotypic variation and potassium chloride fertilization. Field Crops Res. 122: 95-103.

Gomez, K.A., and A.A. Gomez, 1984. Statistical Procedures for Agricultural Research. John Wiley & Sons, Inc., New York.

Grant, C., W. Buckley, L. Bailey, and F. Selles. 1998. Cadmium accumulation in crops, Can. J.

Plant Sci .78: 1-17.

Grant, C.A., J.M. Clarke, S. Duguid, and R.L. Chaney. 2008. Selection and breeding of plant cultivars to minimize cadmium accumulation. Sci. Total Environ. 390: 301-310.

Harris, N.S., and G.J. Taylor. 2001. Remobilization of cadmium in maturing shoots of near isogenic lines of durum wheat that differ in grain cadmium accumulation. Journal of Experimental Botany 52: 1473-1481.

Hart, J.J., R.M. Welch, W.A. Norvell, L.A. Sullivan, L.V. Kochian. 1998. Characterization of cadmium binding, uptake, and translocation in intact seedlings of bread and durum wheat cultivars. Plant Physiol 116: 1413-1420.

Hart, J.J., R.M. Welch, W.A. Norvell, and L.V. Kochian.2006.Characterization of cadmium uptake, translocation and storage in near-isogenic lines of durum wheat that differ in grain cadmium accumulation. New Phytol. 172: 261-271.

Hart, J.J., R.M. Welch, W.A. Norvell, L.A. Sullivan, and L.V. Kochian.1998. Characterization of cadmium binding, uptake, and translocation in intact seedlings of bread and durum wheat cultivars. Plant Physiology 116: 1413-1420.

Ishikawa, S., T. Abe, M. Kuramata, M. Yamaguchi, T. Ando, T. Yamamoto, and M. Yano. 2010. A major quantitative trait locus for increasing cadmium specific concentration in rice grain is located on the short arm of chromosome. Journal of Experimental Botany 61: 923-934.

Joehanes, R., and Nelson, J. C. 2008. QGene 4.0, an extensible Java QTL-analysis platform. Bioinformatics 24: 2788-2789.

Jones, K., A. Jackson, and A. Johnston. 1992. Evidence for an increase in the Cd content of herbage since 1860's. Environ. Sci. Technol. 26: 834-836.

Jones, K., and A. Johnston. 1989. Cadmium in cereal grain and herbage from long-term experimental plots at Rothamsted, U.K. Environ. Pollut. 57: 199-216.

Kashiwagi, T., K. Shindoh, N. Hirotsu, and K. Ishimaru. 2009. Evidence for separate translocation pathways in determining cadmium accumulation in grain and aerial plant parts in rice. B.M.C. Plant Biol. 9:8.

Klein, M., B. Burla, and E. Martinoia. 2006. The multidrug resistance-associated protein (MRP/ABCC) subfamily of ATP-binding cassette transporters in plants. FEBS Lett.580: 1112-1122.

Knox, R.E., C.J. Pozniak, F.R. Clarke, J.M. Clarke, S. Houshmand, and A.K. Singh. 2009. Chromosomal location of the cadmium uptake gene (Cdu1) in durum wheat. In: Genome 52:741-747.

Kosambi, D.D. 1944.The estimation of map distances from recombination values. Ann.Eugen. 12: 172-175.

Kubo, K., Y. Watanabe, A. Oyanagi, S. Kaneko, M. Chono, H. Matsunaka, M. Seki, and M. Fujita. 2008. Cadmium concentration in grains of Japanese wheat cultivars: genotypic difference and relationship with agronomic characteristics. Plant Prod. Sci. 11: 243-249.

Lind, Y., j. Engman, L. Jorhem, W. Wicklund-Glynn. 1997. Cadmium accumulation in liver and kidney of mice exposed to the same weekly cadmium dose continuously or once a week. Food Chem. Toxicol. 35: 891-895.

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.

Mortvedt, J., D. Mays, and G. Osbom. 1981. Uptake by wheat of cadmium and other heavy metal contaminants in phosphate fertilizers. J. Environ. Qual. 10: 193-197.

National Agricultural Statistics Service (NASS) Retrieved December 4, 2012, from http://www.nass.usda.gov/Statistics_by_State/North_Dakota/index.asp National Research Council Canada (NRCC), Associate Committee on Scientific Criteria for Environmental Quality (1979) Effects of cadmium in the Canadian environment. l48 pp.

Norvell, W.A., J. Wu, D.G. Hopkins, and R.M. Welch. 2000. Association of cadmium in durum wheat grain with soil chloride and chelate-extractable soil cadmium. Soil Sci. Soc. Am. J. 64: 2162-2168.

Oliver, D., J. Schultz, K. Tiller, and R. Merry. 1993.The effect of crop rotations and tillage practices on cadmium concentration in wheat grain. Australian Journal of Agricultural Research 44: 1221-1234.

Penner, G.A., L.J. Bezte, D. Leisle, and J. Clarke. 1995. Identification of RAPD markers linked to a gene governing cadmium uptake in durum wheat. Genome 38: 543-547.

Roberts, J.R. 1999. Metal toxicity in children. In Training Manual on Pediatric Environmental Health: Putting It into Practice 1999 Jun. Emeryville, CA: Children's Environmental Health Network (http://www.cehn.org/cehn/trainingmanual/pdf/manual-full.pdf).

Saadalla, M.M., J.F. Shanahan, and J.S. Quick. 1990. Heat tolerance in winter wheat: I. Hardening and genetic effects on membrane thermostability. Crop Sci. 30: 1243-1247.

SAS Institute. 2012. SAS/STAT 9.3 User's Guide. SAS Institute, Cary, NC.

Schiex, T., and C. Gaspin. 1997. CARTAGENE: constructing and joining maximum likelihood genetic maps. Proceedings of the 5th International Conference on Intelligent Systems for Molecular Biology, Porto Carras, Halkidiki, Greece, pp. 258-267.

Shute, T., and S.M. Macfie. 2006. Cadmium and zinc accumulation in soybean: A threat to food safety. Sci. Total Environ. 371:63-73.

Smolders, E. 2001. Cadmium uptake by plants. International Journal of Occupational Medicine and Environmental Health 14: 177-183.

Song, W.Y., E.J. Sohn, E. Martinoia, Y.J. Lee, Y.Y. Yang, M. Jasinski, C. Forestier, I. Hwang, and Y. Lee. 2003. Engineering tolerance and accumulation of lead and cadmium in transgenic plants. Nat. Biotechnol. 21: 914-919.

Soric, R., Z. Loncaric, V. Kovacevic, I. Brkic, and D. Simic. 2009. A major gene for leaf cadmium accumulation in maize (Zea mays L). The Proceedings of the International Plant Nutrition Colloquium XVI. U.C. Davis, U.S.A. Available from: http://escholarship.org/uc/item/1q48v6cf.

Sourdille, P., S. Singh, T. Cadalen, G.L. Brown-Guedira, G. Gay, L. Qi, B.S. Gill, P. Dufour, and A. Murigneux. 2004. Microsatellite-based deletion bin system for the establishment of genetic-physical map relationships in wheat (*Triticum aestivum* L.) Functional and Integrative Genomics 4: 12-25.

Steemers, F.J., W. Chang, G. Lee, D.L. Barker, and R. Shen. 2006. Whole-Genome Genotyping with the single-base extension assay. Nat. Methods. 3: 31-33.

Stolt, J. P., F. E. C. Sneller, T. Bryngelsson, T. Lundborg, and H. Schat. 2003. Phytochelatin and cadmium accumulation in wheat. Environmental and Experimental Botany 49: 21-28.

Tanhuanpaa, P., R. Kalendar, A.H. Schulman, and E. Kiviharju. 2007. A major gene for grain cadmium accumulation in oat (Avena sativa L.). Genome 50: 588-594.

Tezuka, K., H. Miyadate, K. Katou, I. Kodama, S. Matsumoto, T. Kawamoto, S. Masaki, H. Satoh, M. Yamagucji, K. Sakurai, H. Takahashi, N. Satoh-Nagasawa, A. Watanabe, T. Fujimura, and H. Akagi. 2010. A single recessive gene controls cadmium translocation in the cadmium hyperaccumulating rice cultivar Cho-Ko-Kohu. Theor. Appl. Genet. 120: 1175-1182.

Ueno, D., I. Kono, K. Yokosho, T. Ando, M. Yano, and J. F. Ma. 2009. A major quantitative trait locus controlling cadmium translocation in rice (Oryza sativa). New Phytol. 182: 644-653.

WHO, regional office for Europe. 2010. Health risks of heavy metals from long-range transboundary air pollution. http://www.euro.who.int/__data/assets/pdf_file/0007/78649/E91044.pdf, accessed 7 November, 2011)

Wiebe, K., N.S. Harris, J.D. Faris, J.M. Clarke, R.E. Knox, G.J. Taylor, and C.J. Pozniak. 2010. Targeted mapping of *Cdu1*, a major locus regulating grain cadmium concentration in durum wheat (Triticum turgidum L. var durum).Theoretical and Applied Genetics 121: 1047-1058.

Williams, C., and D. David. 1976. The accumulation in soi1 of cadmium residues from phosphate fertilizers and their effect on the cadmium content of plants. Soil Sci. 12 l: 86-93.

Zeng, Z.B. 1994. Precision mapping of quantitative trait loci. Genetics 136: 1457-1468.

CHAPTER III. IDENTIFICATION OF QTL LINKED TO CD UPTAKE LOCUS USING ASSOCIATION MAPPING IN A DIVERSE COLLECTION OF DURUM WHEAT ADVANCED BREEDING LINES

Abstract

Cadmium is a toxic heavy metal that can retard plant development and can adversely affect the health of humans (*Homo sapiens*). The amount of Cd in grain depends on multiple factors including the uptake of Cd from soil and translocation through xylem and phloem to the grains. To discover markers associated with the amount of Cd accumulated in the grains, association mapping was used to analyze two durum wheat collections (*n=96*) consisting of advanced breeding lines from the North Dakota wheat breeding program grown in Langdon, ND in 2009 and 2010. Entries in the yield trial experiments were assigned to experimental units using simple lattice in both years. For the 2009 collection, one major QTL that explains 3 percent of phenotypic variation was identified on Chromosome 2B at 7.25 cM which could co-segregate with a height locus identified earlier. For the 2010 collection, one marker that explains 34 percent of phenotypic variation was identified on Chromosome 5B at 165.7cM and another at 178.3cM that explains 27 percent of phenotypic variation. These markers could be linked to the gene Cdu1 identified previously. Since the complexity of steps involved in the Cd accumulation in grains, identification of additional QTL using better coverage of markers and a large collection of genotypes would help MAS for Cd accumulation.

Introduction

Cadmium (Cd) is a heavy metal present in low concentrations in the soil worldwide (Lenntech, 2004). It is considered highly toxic and an important environmental pollutant (Liu *et al.*, 2007). Cadmium is also a nonessential element for plants and has the ability to form strong

complexes with biomolecules which can be potentially harmful to plants and animals even if present in small quantities. The main effects of cadmium on plants include reduced growth, inhibition of photosynthesis, changes in stomatal action, changes in respiration and nitrogen metabolism and oxidative stress (Weigel, 1985; Liu *et al.,* 2007; Ci *et al.,* 2012). In addition it is proposed that Cd has damaging effects on vital cell constituents, including oxidation and structural alterations of proteins, lipids, and DNA (Yin *et al.,* 2000). Cebeci *et al.* (2008) suggested that excess Cd in the cells can interfere with the transcription and signal transduction mechanisms.

Cadmium is also toxic to humans and can lead to kidney damage and is a suspected carcinogen (McLaughlin *et al.,* 1999). It can also cause genotoxic and cytotoxic effects, leading finally to the inhibition of cell proliferation and apoptosis (Xu *et al.,* 2012). Cadmium can accumulate in human bodies over time through ingestion of food containing Cd, leading to a risk of chronic toxicity with excessive intake (Grant *et al.,* 2008; Ueno *et al.,* 2009; Xu *et al.,* 2012).

The main source of Cd in plants is soil. Vast areas of agricultural soils are contaminated with Cd through the use of super phosphate fertilizers and sewage sludge and inputs from mining and smelting industries (McGrath *et al.,* 2001). Cadmium is taken up by the roots, then translocated into the aerial organs of the plant. With this, the increase of the Cd level in the soil is harmful to the crops, humans, and the environment. Cadmium accumulation in wheat grains is caused by Cd uptake by roots, translocation roots to shoot, and translocation within the shoot to the grains (Greger and Lofstedt, 2004). It has been found that Cd uptake from soil depends on the soil Cd concentration, soil pH, and the level of organic matter and zinc in the soil (Eriksson *et al.,* 1996; Eriksson and Söderström, 1996). The other factors of Cd accumulation are regulated by the plant itself. The amount of Cd in grain not only varies between different species, but also

varies widely among the cultivars within the same species (Khan *et al.,* 2007; Verma *et al.,* 2007; Wenzel and Jockwer, 1999; Chan and Hale 2004; Stolt *et al.,* 2006; Wangstrand *et al.,* 2007). Hart *et al.* (1998) estimated that the proportion of Cd translocated to shoots was 1.5 to 4.5 times higher in bread wheat than in a durum wheat cultivar. The uptake of Cd varies among wheat cultivars (Oliver *et al.,* 1995; Wenzel *et al.,* 1996; Li *et al.,* 1997; Ci *et al.,* 2012) and rice cultivars (Morishita *et al.,* 1987; Arao and Ishikawa, 2006; Uraguchi *et al.,* 2009).

Many factors are responsible for the large difference in Cd concentration in cultivars. For example, a large difference in Cd concentration was observed between the cultivar's xylem sap which is determined by factors such as xylem loading of Cd and transpiration rate (Ishikawa, 2010). It is also observed that this ratio is approximately equal to the ratio of grain Cd concentration among cultivars, suggesting that the importance of root-to-shoot Cd translocation via the xylem is an important factor to determine grain Cd concentration (Ishikawa, 2010). In addition, phloem has been suggested to be directly linked to Cd content in grains (Harris and Taylor, 2004, Chen *et al.,* 2007, Tanaka *et al.,* 2003).

Kashiwagi *et al.* (2009) identified QTL associated with phloem transport of Cd into rice grains. This region is suggested to harbor several candidate genes linked to Cd transport (Ishikawa, 2010). Hart *et al.* (1998) suggested that Cd accumulation in grains may occur via phloem in durum wheat. High Cd accumulation levels in durum wheat grains can be due to the elevated translocation from leaves and stems to maturing grain (Harris and Taylor, 2001; Chan and Hale, 2004) or directly from root to shoots (Greger and Lofstedt, 2004). Liu *et al.* (2007) and Kashiwagi *et al.* (2009) observed a positive correlation of Cd concentration in rice grain with Cd accumulation in the plant and the Cd distribution ratio from aerial parts to the grain. Briefly, the Cd accumulation in grain is affected by the three transport processes: (i) uptake of Cd by roots,

(ii) translocation to shoots via xylem, and (iii) translocation to seeds via phloem (Clemens *et al.,* 2002) which in turn can depend on the agronomic characteristics of the plant and many other factors.

The variation in Cd content in different cultivars varies because of the variation in Cd uptake and translocation to the grains. Cultivar selection is an important method to limit Cd uptake and accumulation. Developing markers for low Cd accumulation will facilitate marker assisted-selection (MAS) that will facilitate identification of cultivars that accumulate low Cd content (Jegadeesan *et al.,* 2010).

Association mapping (AM) is based on linkage disequilibrium (LD) and has been effective in identifying QTL. Association mapping identifies QTL in a collection of diverse individuals from natural populations and has the potential to improve trait and germplasm security in the future (Zhu *et al.,* 2008). On the contrary, the sampling is restricted to only the alleles, differing between the two parents in linkage mapping (Remington *et al.,* 2001; Kraakman *et al.,* 2004). In a population, changes in phenotype may occur due to a mutation in one individual, creating a new allele for a trait. This can be passed on to subsequent generations if the fitness of an individual is improved (Jay, 2011). In addition, loci surrounding the valuable locus are also passed on, creating a block of loci that are co-inherited non-randomly (Jay, 2011). These loci can be detected by the association mapping approach. Compared to bi-parental or linkage mapping, AM offers several advantages: 1) allelic variation sampling is increased, and 2) mapping resolution is increased (Abdurakhmonov and Abdukarimov, 2008; Myles *et al.,* 2009).

Association mapping studies were performed in different crop species, and many traits have been associated in wheat. For example, Breseghello and Sorrells (2006) reported association mapping of kernel size and milling quality of the U.S.A winter wheat collection

using SSRs. High molecular-weight glutenin (Ravel *et al.,* 2006) and blotch resistance (Tommasini *et al.,* 2007) were also reported, using different molecular markers such as SNPs, SSRs, and STS. Ghavami *et al.* (2011) identified QTL linked to FHB in durum Gurung *et al.* (2011) and Patel *et al.* (2013) identified markers linked to tan spot in wheat.

The need to develop low Cd accumulating durum cultivars is necessary to minimize the level transferred from the soil to the plant and further to the human diet. The objective of the durum wheat breeding program is to identify a QTL that governs Cd accumulation in seed. To achieve this goal, an association mapping study was implemented. The identification of QTL associated with important traits (e.g., Cd accumulation), using association mapping, is a common approach. In the last few years detecting associations via a mixed model (proposed by Yu *et al.,* 2006b) that controls for both population structure and relatedness is commonly used to control spurious associations (Mamidi *et al.,* 2011; Ghavami *et al.,* 2011; Iqbal *et al.,* 2011; Gurung *et al.,* 2011).

The objective of this study was to identify markers/QTL linked to Cd concentration using association mapping in a diverse collection of durum wheat cultivars and advanced breeding lines collected from the durum wheat breeding program at North Dakota State University, Fargo, USA.

Materials and methods

Plant materials and phenotype data

Two collections of durum wheat lines consisting of low and high cadmium line sources were selected from the durum wheat breeding program, representing a wide range of Cd accumulation level in the seeds. Each collection (n=96) of near inbred lines (NILs) were grown in Langdon, ND, in 2009 and 2010. The soil type at Langdon is Svea (fine-loamy, mixed Pachic

Udic Haploborolls) and Barns (fine-loamy, mixed Udic Haplborolls). The checks, Strongfield (Clarke *et al.*, 2006) and Alkabo (Elias and Manthey, 2007), and the durum wheat lines were planted as single-row plots 2.75 m long with 30 cm row spacing between rows. Seeding rate was 10 g of seed per row which translates to approximately 2.5 million plants per hectare. Planting, field fertility, herbicide and fungicide application, and harvest were concurrent with the North Dakota State University Durum Wheat Breeding Project which are typical of North Dakota growing conditions. Fungicide applications consisting of Folicur 3.6F (alpha-[2-(4-chlorophenyl)-ethyl]-alpha (1, 1-dimethylethyl)-1H-1, 2, 4-triazole-1-ethanol) (Bayer Corporation, Kansas City, MO) at flag leaf stage, using the rate of 267.1 ml/ha, was used for two years (2009 and 2010) to minimize the infection by Fusarium head blight caused by *Fusarium graminearum* Schwabe. An herbicide tank-mixture of Discover (Syngenta Crop Protection, Inc. Greensboro, NC, at rate 292.2 ml/ha), Buctril (Bayer Corporation, Kansas City, MO, at rate 1.17 l/ha), and MCPA ester (584.3 ml/ha) was applied for weed control. The data for each collection were screened for normality by using Kolmogorov-Smirnov test procedure in SAS 9.3

Cadmium level evaluation

Samples were sent for Cd analysis to Dr. Michael A. Rutzke, College of Agriculture and Life Sciences Nutrient Analysis Laboratory, Cornell University, Ithaca, NY. A representative flour sample of up to 0.5 g from each NIL was digested in 4 ml of mix (40 percent of concentrated nitric acid and 60 percent of perchloric acid) and an extra 1 ml perchloric acid. The sample and acid were placed in a fluorocarbon vessel. The open vessel is heated on a hot plate unit. After cooling, the vessel contents were allowed to settle, then diluted to 20 ml volume and analyzed with the appropriate SW-846 method (EPA Method No. 3051, 3050, 3052).

Experimental design and statistical analysis

Entries in the yield trial experiments were assigned to experimental units using simple lattice in 2009 and 2010. The phenotypic data gathered from each test year were analyzed separately, using the statistical software SAS v. 9.3 (SAS Institute, Inc. 2011). Means were separated by Fisher's Protected LSD at the 5 percent level of significance.

Genetic analysis

DNA isolation

The lines from both collections were grown in the greenhouse in the fall of 2010; three fresh leaf tissues were harvested and stored at -80^{o}C, then lyophilized for 48 hours and kept at room temperature. The dried leave tissues were used for DNA extraction by using the protocol developed by Triticarte Pty.Ltd. (Canberra, Australia; http://www.triticate.com.au) with few modifications. The tissues were ground in a mixer mill at frequency 26 for 3 min to fine powder. The resulting tissue powder was transferred to a 2-ml polypropylene tube. The samples in each tube were treated with 900 µl of Fresh buffer solution (0.15 g Sodium disulfite, 0.6 g PVP-40 (K29-32) preheated to 65 ^{o}C, 0.6 g CTAB, 12.5 ml extraction buffer (0.35 M sorbitol, 0.1 M TrisHCl pH 8.0, 5 mM EDTA, pH 8.0), 12.5 ml lysis buffer stock (0.2 M TrisHCl pH 8.0, 0.05 M EDTA pH 8.0, 2 M NaCl), and 5 ml 5 percent sarcosyl stock. The contents were mixed and incubated in a water bath at 65 ^{o}C for 1 hour. Approximately 900 µl of chloroform: isoamyl alcohol (24:1) was added to each tube. All the tubes were balanced and centrifuged (25 min, 11,000x, 25^{o} C). Around 600 µl of the aqueous phase was removed and transferred from each tube to a fresh 1.5-ml tube, and 600 µl of isopropanol was added to each tube and mixed for 5 min. The new tubes were balanced and centrifuged (15 min, 11,000x, 25^{o} C). The isopropanol was removed to leave a DNA pellet in bottom of tube. The DNA pellet was washed with 950 µl

of 70 percent ethanol by using a centrifuge (15 min, 11,000 x, 25^{o} C). Then the ethanol was removed carefully. Once the washes were complete, the DNA was oven-dried at 37 ^{o}C for 15-30 min to eliminate any traces of ethanol, then resuspended in 200 µl of double distilled water and left at 14 ^{o}C to dissolve overnight. NanoDrop 2000 (Thermo Fisher Scientific, West Palm Beach, FL, USA) was used to measure the concentration of the DNA. The samples were checked for integrity, purity, and concentration by running 1 µl of each sample on a 1 percent agarose gel, after incubation in a digestion buffer for 1 hour.

Genotyping

The extracted DNA samples were sent to the USDA-ARS Cereal Crops Laboratory (Fargo, ND) for genotyping with high-density, single nucleotide polymorphism (SNP) markers. A custom-designed Illumina iSelect BeadChip platform containing 9,000 SNPs and the Illumina Infinium Assay protocol (Steemers *et al.,* 2006) were used. Infinum assay protocol consists of three steps: genomic DNA hybridization to beads, single-base extension, and allele-specific detection methods. The bead chips were scanned and the hybridization intensity values were obtained by using the BeadStation scanner. The hybridization intensity values were used as the basis for genotype calling. GenomeStudio software, developed by Illumina, was used to determine the genotype calls. A cluster is formed if samples exhibit the same genotype. Three clusters were identified by the calling algorithm with each cluster corresponding to one of the three genotypes AA, AB, and BB (Figure 3.1). Each genotype call was validated manually and edited to avoid errors related to cluster compression. The genotypic data were exported from the software and used for further analysis.

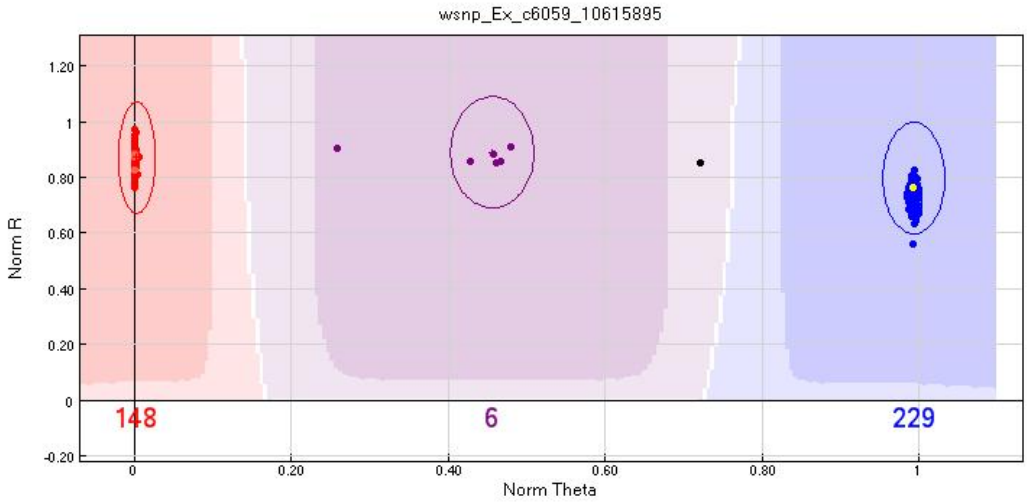


Figure 3.1. Genotype call output from GenomeStudio software. Three clusters representing AA, AB, and BB genotype classes (left to right).

Association mapping analysis

Missing data imputation

To minimize the problems caused by missing genotypic data, the default settings of FastPHASE 1.3 (Scheet and Stephens, 2006) were used to impute missing loci for the two data sets, using "likelihood" based imputation. The heterozygotes were considered missing. Markers with minor allele frequency (MAF) < 5 percent were removed for further analysis. The marker characteristics like polymorphism information content (PIC) were further estimated in PowerMarker 3.25 (Liu and Muse, 2005).

Population structure and kinship

Population structure can be defined as the differential relatedness among individuals of different subsets (e.g., breeding programs). Principal component analysis (PCA) was used to control for population structure and was performed, using PRINCOMP procedure in SAS 9.3 (SAS Institute, Inc. 2011) for the two data sets independently. We chose principal components that explain 50 percent of cumulative variation and number of principal components selected on

the MAP test (Shriner *et al.*, 2010) implemented in SAS/IML. A pairwise kinship coefficient matrix (K-matrix) that estimates the probability of recent co-ancestory between genotypes (Loiselle *et al.*, 1995) was determined, using SPAGeDi 1.2 (Hardy and Vekemans, 2002). Identity by state (IBS) matrix (Zhao *et al.*, 2007) that represents the proportion of shared alleles for all pairwise comparisons in each population was calculated, using DISTANCE procedure in SAS 9.3 (SAS Institute, Inc. 2011).

Linear regression models used and identification of marker-trait associations

Six linear regression models comprising both general linear models (GLM) and mixed linear models (MLM) were selected to determine *P*-values associated with tests for marker-trait associations. The MLM accounts for population structure and a pairwise kinship matrix, to control for Type I and Type II errors (Yu *et al.*, 2006). Descriptions for each of the six models (Naïve, IBD, IBS, PCA, PCA + IBS, and PCA + IBD) are provided in Table 3.1. All analyses were conducted using the MIXED procedure of SAS 9.3. The general linear regression model used follows the formula:

$$y = X\alpha + P\beta + K\nu + \varepsilon$$

where y is a vector for phenotypic observations, α is the fixed effects related to the SNP marker, β is a vector of the fixed effects related to the population structure, ν is a vector of the random effects related to the relatedness among the individuals, and ε is a vector of the residual effects. X is the genotypes of the SNP markers, P is the matrix of the principal components, and K is the kinship matrix. The variances of the random effects were estimated as Var (u) = 2KVg and Var (e) = IV_R, where K is a kinship matrix, I is an identity matrix with the off-diagonal elements recorded as 0, and diagonal elements is the reciprocal of the number of observations for which the phenotypic data were obtained, Vg is the genetic variance, and V_R is the residual variance.

Table 3.1. Statistical models tested for association mapping.

Model	Statistical model	Description
Naïve	$y = X\alpha + \varepsilon$	No control over population structure
IBD	$y = X\alpha + K\nu + \varepsilon$	Loiselle coefficient used as kinship matrix
IBS	$y = X\alpha + K\nu + \varepsilon$	Similarity used as kinship matrix
PCA	$y = X\alpha + P\beta + \varepsilon$	PC that explain 50% cumulative variance or PC identified by MAP test
PCA + IBS	$y = X\alpha + P\beta + K\nu + \varepsilon$	PC that explain 50% cumulative variance or PC identified by MAP test and Loiselle coefficient used as kinship matrix
PCA + IBD	$y = X\alpha + P\beta + K\nu + \varepsilon$	PC that explain 50% cumulative variance or PC identified by MAP test and similarity used as kinship matrix

For the selection of best model, mean square difference (MSD) was calculated as:

$$MSD = \frac{\sum_{i=1}^{n} (p_i - \frac{1}{n})^2}{n}$$

where i is the rank number, P_i is the probability of the i^{th} ranked p-value, and n is the number of markers (Mamidi *et al.*, 2011). The best model is defined as the one with the lowest MSD value and the least percent of observations with a p-value < 0.05. For the best model, MULTTEST procedure in SAS was used to estimate the positive false discovery rate (pFDR) for each marker to correct for multiple marker trait association. Markers are considered significant if in the best model they have a pFDR < 0.1. The multiple R^2_{LR} values for the significant loci were calculated using MIXED procedure in SAS as described in Sun *et al.* (2010). In addition to estimating the interaction effect, we used MIXED procedure in SAS, including an interaction term in the best model equation.

Linkage disequilibrium analysis

For each year separately, genome-wide LD analysis was performed making intra-chromosomal pairwise comparisons among SNP markers, using the squared correlation coefficient (R^2) as suggested by Hill and Robertson (1968). This was calculated using the CORR procedure of SAS 9.3. Linkage disequilibrium decay graphs were plotted with genetic distance

and R^2 for those pairwise comparisons that are on the same chromosome using nonlinear regression as described by Remington *et al.* (2001) and Pyhajarvi *et al.*, (2007); and implemented by the NLIN procedure of SAS 9.3. Separate LD analyses for each year were performed. The pattern and distribution of intrachromosomal LD was graphically depicted for each chromosome and the whole genome by plotting significant intra-chromosomal pairwise R^2 values against the genetic distance (cM) between markers.

Results and discussion

Phenotypic analysis for cadmium

For the 96 lines evaluated in 2009, the Cd level ranged from 0.024 to 0.108 mg/k*g* with an average of 0.066 mg/kg. For the 96 lines evaluated in 2010, the Cd level ranged from 0.015 to 0.120 mg/kg with an average of 0.051 mg/kg (Figure 3.2). Variation in Cd concentration showed a skewed distribution toward low concentrations. The p-value for the KS normality test was > 0.15 in 2009 and <0.01 in 2010. Mean values and descriptive statistics from the analyses of variance (ANOVA) of the checks and the experimental lines for the phenotypic trait are presented in Table 3.2. The checks Strongfield and Alkabo differed significantly for Cd content in both years.

Single nucleotide polymorphism marker analysis

In both years, a very low percent of alleles (~1.5 percent) was called as heterozygotes, and the rest of the genotype calls were homozygous as would be expected for a self-pollinated crop. This was converted to missing data and imputed. The PIC for the markers in 2009 is in the range of 0.0939 to 0.3750 with a mean of 0.2466. For 2010, this value ranges from 0.0939 to 0.3750 with a mean of 0.2748.

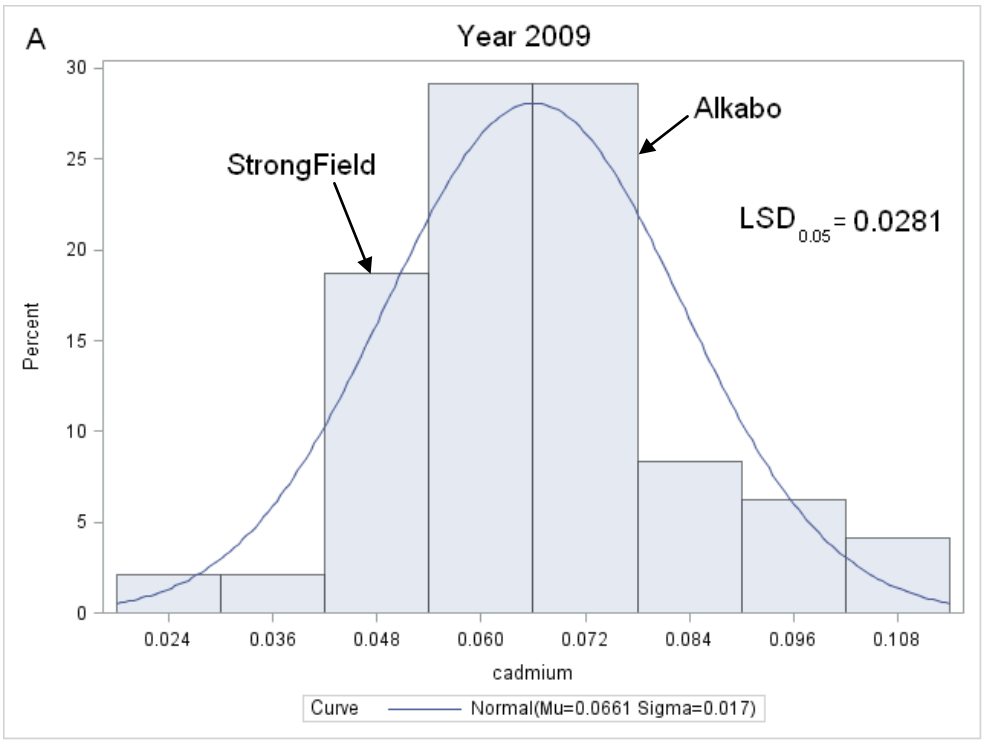
Association mapping analysis

Missing SNP genotypes can be estimated by using imputation analysis. This type of analysis relies on the expectation-maximization algorithms that compare each marker's locus and individual against the entire data from other individuals. The accuracy of missing data imputation depends on the amount of information gathered from the use of thousands of markers (Li *et al.*, 2009; Browning, 2008). Approximately 5 percent of the missing data is imputed to increase the power of AM (Marchini and Howie, 2010). Further 1,071 markers in 2009 and 872 markers in 2010 that have a MAF < 5 percent were removed. The number of principal components that are necessary to explain the percentage of variation in each year was varied. In 2009, 16 principal components were selected based on the minimum average partial (MAP) test and six of them explained 50 percent of the cumulative variation. In 2010, 16 principal components were selected based on the MAP test, and four principal components explain 50 percent of cumulative variation.

The individuals grouped into three distinct clusters in 2009 and four distinct clusters in 2010, using the first two principal components that explain 30.87 and 40.07 percent cumulative percent of variation in 2009 and 2010, respectively (Figure 3.3). In 2009, the similarity of individuals was in the range 0.45 to 0.97 with a mean of 0.694 and a median of 0.690; and in 2010, the range was between 0.42 and 0.99 with a mean of 0.655 and a median of 0.6238.

Single nucleotide polymorphism and cadmium uptake marker-trait associations

Based on our best model selection criteria, we found PC6+IBS model in 2009 and PC16+IBS in 2010 to perform better over the six models tested (Table 3.3, Figure 3.4). Three markers were found to be significant in 2009, and two makers were found to be significant in 2010 (Table 3.4). In 2009, two of the markers are at 7.25 cM on Chromosome 2B, and the



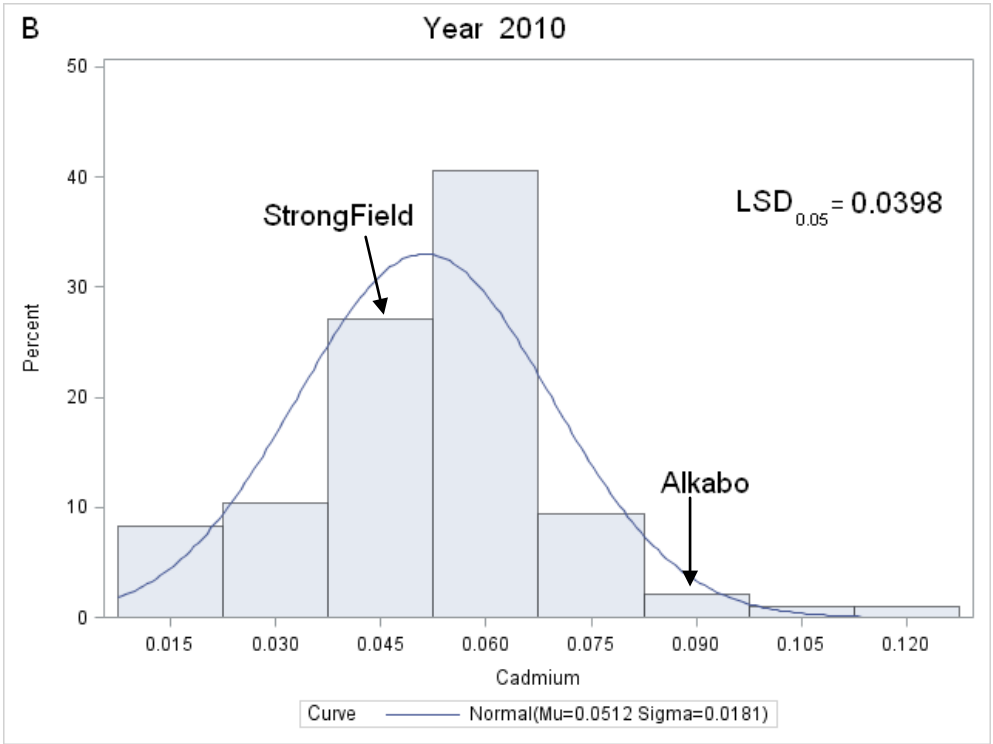


Figure 3.2. Phenotypic distribution of Cadmium accumulation level. (A) 2009 and (B) 2010.

Table 3.2. Overall means and descriptive statistics of Cadmium level for durum wheat lines evaluated at Langdon, ND, across 2009 and 2010.

	2009	2010
	(mg/kg)	
Checks		
Strongfield	0.0421	0.0444
Alkabo	0.0725	0.0856
Population statistics		
Mean	0.0661	0.0512
Minimum	0.0249	0.0159
Maximum	0.1082	0.1208
Standard deviation	0.0170	0.0181
‡ $LSD_{1(0.05)}$	0.0281	0.0398
‡ $LSD_{2(0.05)}$	0.0313	0.0325
‡ CV%	35.08	40.36

‡ $LSD_{1(0.05)}$ = Least significant difference comparing individual genotypes; $LSD_{2(0.05)}$ comparing individual genotype with population mean.
CV= Coefficient of variation.

location of the third marker was unknown (Table 3.4). This could be similar to the QTL linked to plant height in durum wheat reported by Pozniak *et al.,* (2012) and hexaploid wheat (Cui *et al.,* 2012; Griffiths *et al.,* 2011; Neumann *et al.,* 2011). The p-value is 0.0003 that can explain about three percent of Cd variation. These loci co-segregate, so the combined phenotypic variation explained is 3 percent. The other allele that is unmapped has the same p-value and R^2 and is in perfect LD with the other two significant markers. The allelic mean for each of these markers is 0.063 and 0.068 (Table 3.4). In 2010, two markers were found on chromosome 5B at 165.7 and 178.3 cM (Table 3.4). The p-values were 5.50E-06 and 0.0001, respectively; and the phenotypic variation explained by these markers was 33.74 and 26.96 percent. The allelic means for the

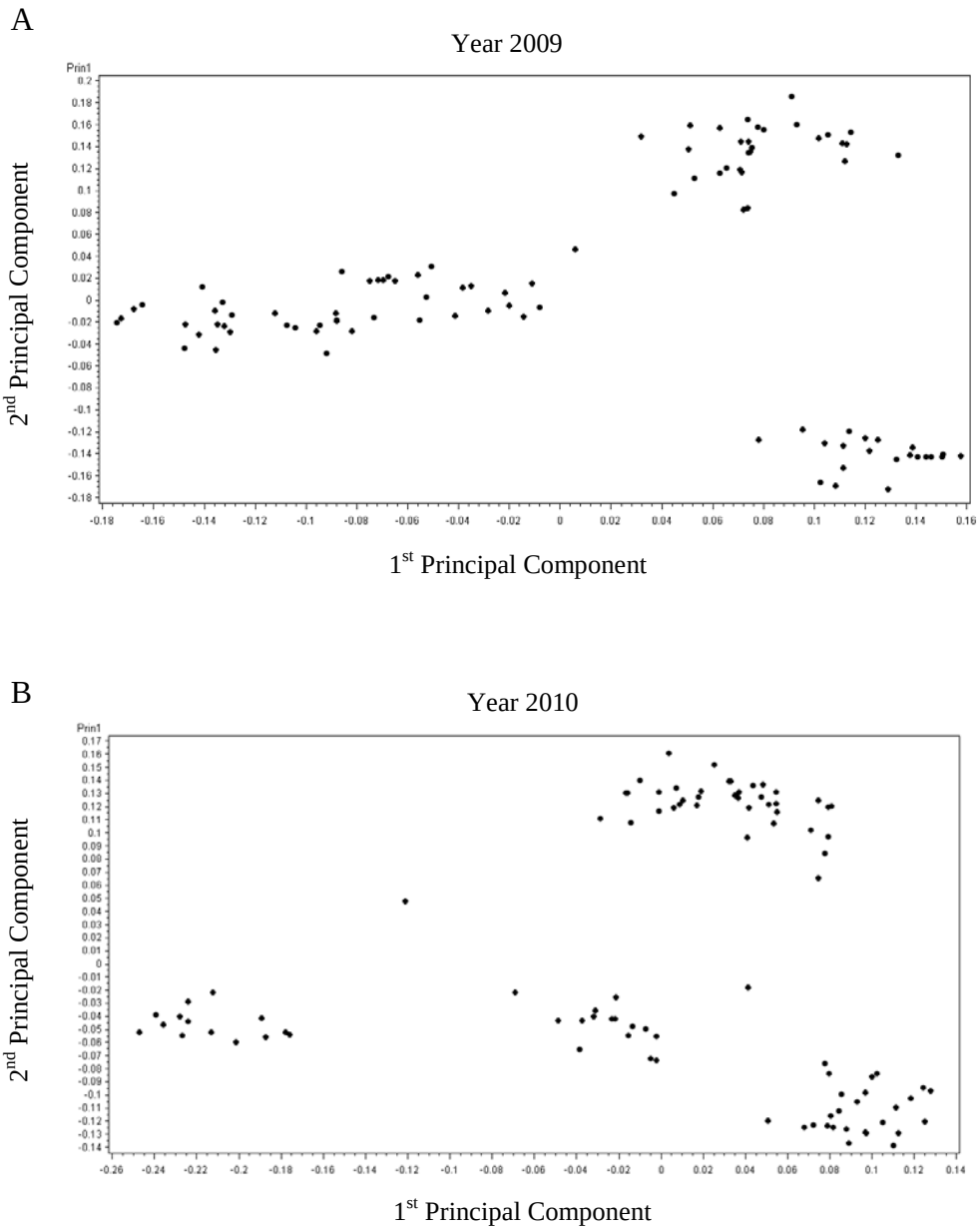


Figure 3.3. PCA analysis. Plots of PC1 on x-axis and PC2 on y-axis to visualize the population structure in the genotype collections of A) 2009 and B) 2010.

minor allele of the marker Ex-c17754-26503892 were 0.026 and 0.056 for the major allele. The minor and the major allelic means for the marker Ex-c20019-29052512 were 0.025 and 0.055, respectively (Table 3.4). This is similar to the single dominant gene identified earlier by Penner *et al.* (1995) and Clarke *et al.* (1997) and named as *Cdu1* by Knox *et al.* (2009). Even though multiple factors affect Cd accumulation, one major QTL has been identified that explains a higher percent of phenotypic variation. These QTL that we identified explain a high percent of phenotypic variation in these populations especially in 2010. This is similar to many of the Cd QTL identified earlier in other crops. For example, in maize, one major QTL affecting Cd accumulation was detected on chromosome 2 that has a LOD score of 32.5 and can explain 49.8 percent of the phenotypic variation (Soric, 2009). Similarly in rice, a major QTL grain Cd was identified on Chromosome 7 that explains 35.5 percent of all phenotypic variance (Ishikawa *et al.*, 2010). In radish, a major QTL which accounted for 48.64 percent of the total phenotypic variance in Cd accumulation was identified on LG 9 (Xu *et al.*, 2012). In soybean, a QTL that explains 57.3 percent of the phenotypic variation that are closely linked to *Cda1* was identified (Jegadeesan *et al.*, 2010).

Table 3.3. Model selection for the different mixed models in each year

2009		2010	
Model	MSD[†]	Model	MSD[†]
IBS	0.0012	IBS	0.0069
Naive	0.0510	Naive	0.0058
PC16	0.0009	PC16	0.0010
PC16+IBS	0.2372	PC16+IBS	0.0005
PC6	0.0017	PC4	0.0065
PC6+IBS	0.0003	PC4+IBS	0.0050

[†]MSD = mean square difference.

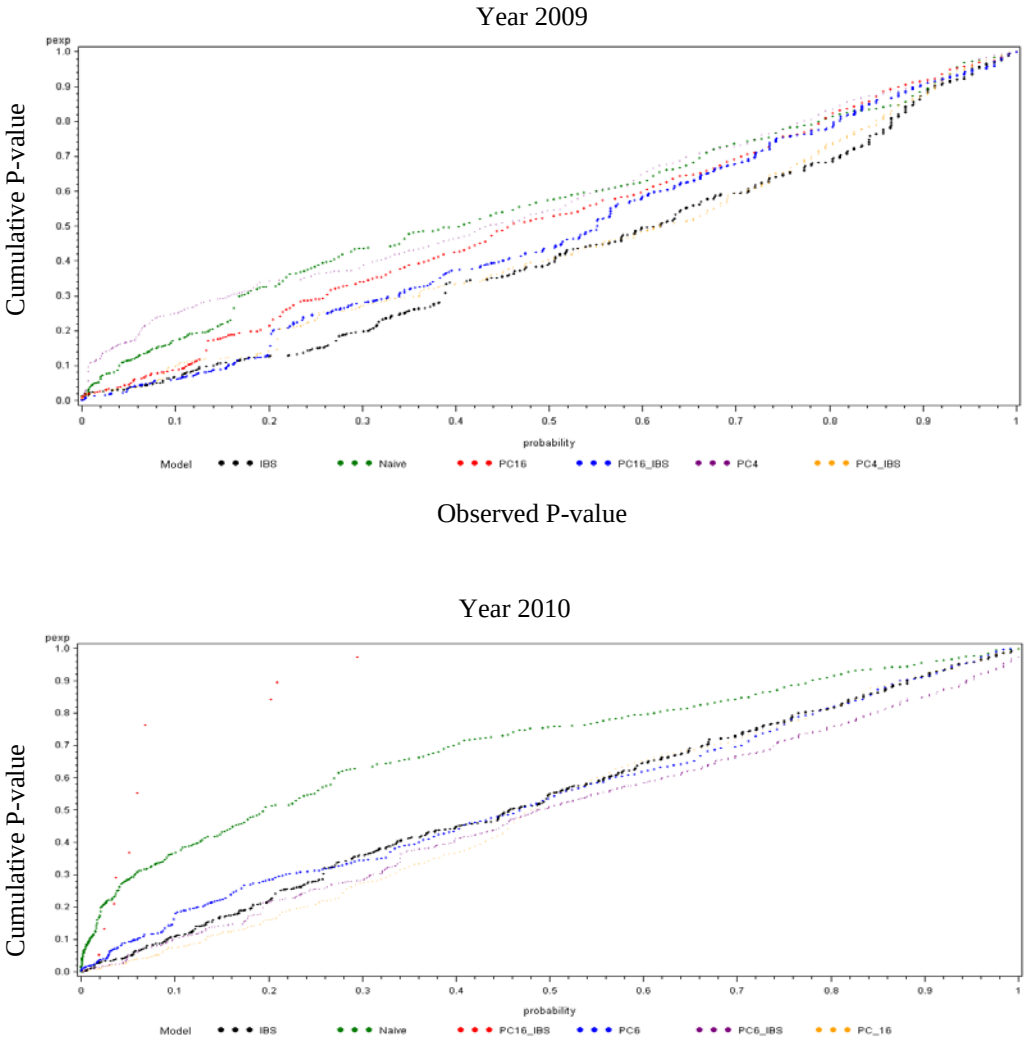


Figure 3.4. P-P plots of the association models in 2009 and 2010

Table 3.4. Test statistics for single nucleotide polymorphism (SNP) loci significantly associated with cadmium uptake in the 2009 and 2010 populations.

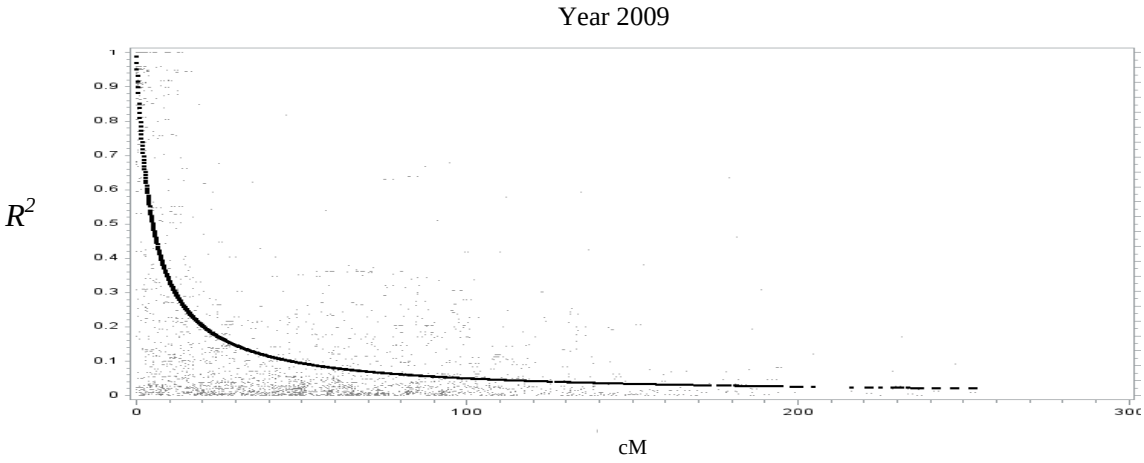
Year	SNP marker	Chromosome	Genetic position (cM)	–log10(p)	pFDR†	R^2 (%)	Minor allele frequency	Minor allele	Minor allele mean	Major allele	Major allele mean
2009	Ex-c1996-3754394	2B	7.25	3.52	0.0989	3.04	47.92	C	0.063	T	0.069
	Ra-rep-c106727-90434958	2B	7.25	3.52	0.0989	3.04	47.92	C	0.063	T	0.069
2010	Ex-c17754-26503892	5B	165.7	5.26	0.0047	33.74	12.50	C	0.026	A	0.056
	Ex-c20019-29052512	5B	178.3	3.99	0.0448	26.96	11.45	G	0.025	A	0.055

†pFDR = positive false discovery rate.

Linkage disequilibrium decay

The LD decay calculation is based on 687 markers that have been mapped on to linkage groups. For the 2009 population, LD decays to 0.5 at 6 cM and to 0.1 at 45cM. For the 2010 population, it decays to 0.5 at 7cM and 0.1 at 60 cM (Figure 3.5). With this finding, it is imperative that 410 and 352 equidistant markers are necessary to identify QTL based on the linkage map of about 2,470 cM. The average distance between the two markers is 5.43 cM (range 0.1 to 43.1 cM).

A



B

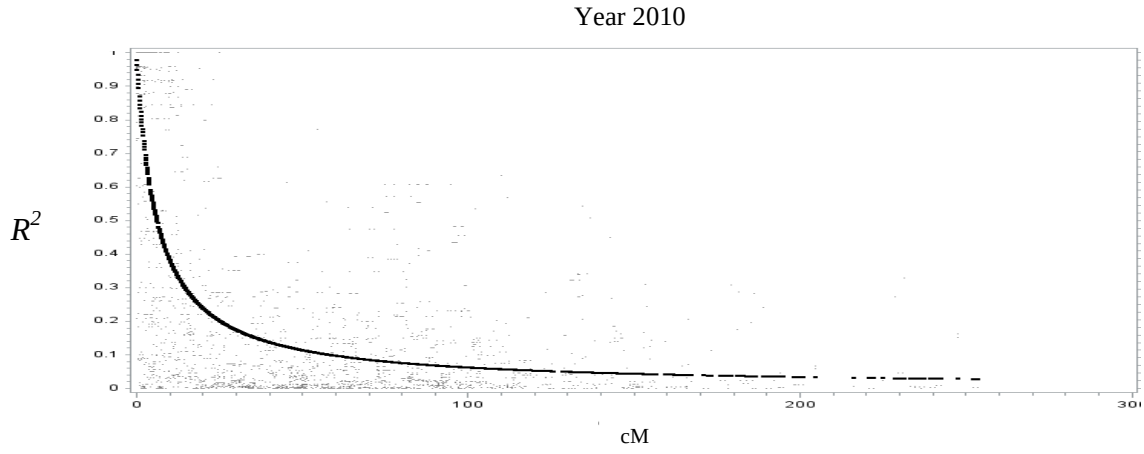


Figure 3.5. Genome-wide linkage Disequilibrium (LD) decay Plot for the two populations. Linkage disequilibrium, measured as R^2, between pairs of polymorphic marker loci is plotted against the genetic distance (cM). (A) 2005 population. (B) 2006 population.

Previously, Knox *et al.* (2009) and Wiebe *et al.* (2010) localized *Cdu1* as a Mendelian factor to Chromosome 5B in durum wheat. Another minor QTL has been mapped on the same chromosome near *Cdu1* (Wiebe *et al.*, 2010). Both grain cadmium content and shoot-to-root partitioning are associated with the *Cdu1* gene (Hart *et al.*, 2006). The QTL detected on Chromosome 5BL for Cd accumulation in this study was similar to the one detected on the same chromosome arm by Knox *et al.* (2009) and Wiebe *et al.* (2010). The other markers identified here could be either in the proximity of this QTL or could be another minor QTL. In the linkage mapping (Chapter II) and association mapping (Chapter III) studies, the same marker was identified on chromosome 5B. In the linkage mapping study the marker explained 51.6 percent of the phenotypic variation and 33.7 percent of the phenotypic variation in the association mapping study. The results suggest the use of this marker (*Ex_c17754_26503892*) in the selection process for low Cd accumulation lines in the breeding program. In addition, the marker (*Ex_c1343_2570756*) associated with QTL was also identified on Chromosome 5B that explained 54.3 percent of the phenotypic variation in Chapter II. Both markers show promising results. The marker (*Ex_c1343_2570756*) was validated in a different genetic background and was able to differentiate between low and high Cd lines. However, the second marker from the association mapping (*Ex_c17754_26503892*) needs to be validated before using it in the breeding program. The identified marker in this study appeared to be linked to the Cd level and was co-segregating with the trait with low recombination rates. The results indicate that these SNP markers are good for marker-assisted selection for Cd content in the durum wheat breeding program.

Multiple linear models and multiple PC were used to account for population structure (PC that control for 50 percent cumulative variation and the number of PCs identified by

Velicer's MAP test) and relatedness and a combination of these. This was necessary because for a population and trait, the model that better controls for false positives varies. For example, in barley CAP genotypes, only a kinship matrix was sufficient to control for population structure (Bradbury *et al.,* 2011). On the other hand, models with both PC/structure and kinship were necessary to control false positives (Mamidi *et al.,* 2011; Ghavami *et al.,* 2011; Iqbal *et al.,* 2011; Gurung *et al.,* 2011). The genotypes used in this study represent low and high Cd content. The low Cd NILs have the potential to be used in the breeding program to develop low Cd durum cultivars. These genotypes will be evaluated for their agronomic traits, quality, and disease resistance for advance and possible release. Use of association mapping for Cd accumulation is necessary because the identification of Cd concentration in seeds is costly.

In addition to the QTL identified here, there might be multiple other genes responsible for differential Cd content in durum wheat. The QTL/genes could be the same QTL/genes identified earlier for other traits. For example, rice Fe transporters OsIRT1 and OsIRT2 and Zn transporter OsZIP1 are responsible for Cd uptake by roots (Ramesh *et al.,* 2003; Nakanishi *et al.,* 2006). Verret *et al.* (2004) proposed that P-type ATPase, AtHMA4, plays a major role in the xylem loading of both Zn and Cd in *Arabidopsis thaliana*. In barley, Cd QTL relate to QTLs for DTH (Yano *et al.,* 2001; Ishikawa *et al.,* 2005; Ishikawa 2010). Liu *et al.* (2005) reported that Cd concentration is significantly correlated with grain yield and agronomic traits in brown rice. In *Arabidopsis halleri* sp. Halleri, a gene for heavy metal transporting, ATPase4 (HMA4) was found to be in the interval of a major QTL (Courbot *et al.,* 2007). In rice, the major QTL is identified and is not related to any QTL identified earlier for any agronomic trait or other essential elements (Ishikawa, 2010).

There are many mechanisms involved for Cd accumulation from the root to the grain, and most of the mechanisms behind these processes are still unknown (Greger and Lofstedt, 2004). It is necessary for a larger population with more diversity and a high marker density along with availability of the genome to help dissect the accumulation of Cd. The four markers identified might be used as selection tools in the durum wheat breeding programs targeting reduced grain Cd levels. However, these markers should be validated in a larger germplasm pool to determine their effectiveness for selecting low Cd lines in durum wheat. It is important to use durum wheat cultivars that absorb less Cd, suggesting the importance to incorporate the lower Cd concentration trait in the durum wheat breeding program. The identified markers can be useful for marker-assisted selection and for better understanding of Cd translocation and accumulation in grain. Additional markers would facilitate fine mapping and positional cloning of this region; confirm further the presence of *Cdu1* in this region in addition to identifying the minor QTL that could be responsible for Cd accumulation in durum wheat.

Conclusion

In this chapter, four markers (*Ex-c1996-3754394, Ra-rep-c106727-90434958, Ex-c17754-26503892* and *Ex-c20019-29052512*) associated with low Cd accumulation were identified on chromosomes 2B and 5B. The markers identified on chromosome 2B explained only 3 percent of the phenotypic variation. However, the two markers located on Chromosome 5B explained 33.74 and 26.96 percent of the phenotypic variation. One was also identified on the same chromosome in Chapter I. The use of this marker and other markers identified early will be a valuable asset in the breeding program. However, this marker needs to be validated in a different genetic background.

References

Abdurakhmonov, I., A. Abdukarimov. 2008. Application of association mapping to understanding the genetic diversity of plant germplasm resources. International Journal of Plant Genomics 2008: 1-18.

Arao, T., and S. Ishikawa. 2006. Genotypic differences in cadmium concentration and distribution of soybean and rice. Japan Agricultural Research Quarterly 40: 21-30.

Bradbury, P., T. Parker, M.T. Hamblin, J.L. Jannink. 2011. Assessment of power and false discovery rate in genome-wide association studies using the BarleyCAP germplasm. Crop Sci 51: 52-59.

Breseghello, F., and M. Sorrells. 2006. Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. Genetics 172: 1165-1177.

Browning, S.R. 2008. Missingdataimputationandhaplotypephase infer-ence for genome-wide association studies. Hum Genet 124: 439-450.

Cebeci, O., B. Kokturk, N. Ergen, L. Ozturk, I. Cakmak, and H. Budak. 2008. Differential expression of wheat transcriptomes in response to varying cadmium concentrations. Journal: Biologia Plantarum – Biol. Plant 52: 703-708.

Chan, D.Y., and B. A. Hale. 2004. Differential accumulation of Cd in durum wheat cultivars: uptake and retranslocation as sources of variation. Journal of Experimental Botany 55: 2571-2579.

Chen, F., J. Dong, F. Wang, F. Wu, G. Zhang, G. Li, Z. Chen, J. Chen, and K. Wei. 2007. Identification of barley genotypes with low grain Cd accumulation and its interaction with four microelements. Chemosphere 67: 2082-2088.

Ci, D., D. Jiang, S. Li, B. Wollenweber, T. Dai, and W. Cao. 2012. Identification of quantitative trait loci for cadmium tolerance and accumulation in wheat. Acta. Physiol. Plant. 34: 191-202.

Clarke, J.M., D. Leisle, G.L. Kopytko.1997. Inheritance of cadmium concentration in five durum wheat crosses. Crop Sci. 37: 1722-1726.

Clarke, J. M., T. N. McCaig, R. M. DePauw, R. E. Knox, F. R. Clarke, M. R. Fernandez, and N. P. Ames. 2006. Registration of 'Strongfield' durum wheat. Crop Sci. 46: 2306-2307.

Clemens, S., M.G. Palmgren, and U. Kramer. 2002. A long way ahead: understanding and engineering plant metal accumulation. Trends in Plant Sci. 7: 309-315.

Courbot, M., G. Willems, P. Motte, S. Arvidsson, N. Roosens, P. Saumitou-Laprade, and N.

Verbruggen. 2007. A major quantitative trait locus for cadmium tolerance in Arabidopsis halleri co-localizes with HMA4, a gene encoding a heavy metal ATPase. Plant Physiology 144: 1052-1065.

Cui, F., J. Li, A.M. Ding, C.H. Zhao, X.F. Li, D.S. Feng, X.Q. Wang, L. Wang, H.G. Wang. 2012. QTL detection of internode length and its component index in wheat using two related RIL populations. Cereal Research Communications 40: 373 – 384.

Shriner, D., A. Adeyemo, G. Chen, and C.N. Rotimi. 2010. Practical considerations for imputation of untyped markers in admixed populations. Genetic Epidemiology 34: 258-265.

Elias, E. M., and F. A. Manthey. 2007c. Registration of 'Alkabo' durum wheat. Journal of plant registration. 1: 10-11.

Eriksson, J., I. Oborn, G. Jansson, and A. Andersson. 1996. Factors influencing Cd-content in crops. Results from Swedish field investigations. Swed. J. Agric. Res. 26: 125-133.

Eriksson, J., and M. Soderstrom. 1996. Cadmium in soil and winter wheat grain in southern Sweden. I. Factors influencing Cd levels in soils and grains. Acta Agric. Scand. sec. B. Soil Sci. 46: 240-248.

Gebhardt, C., A. Ballvora, B. Walkemeier, P. Oberhagemann, and K. Shueler. 2004. Assessing genetic potential in germplasm collections of crop plants by marker-trait association: a case study for potatoes with quantitative variation of resistance to late blight and maturity type. Mol Breed 13:93-102.

Ghavami, F., E.M. Elias, S. Mamidi, O. Ansari, M. Sargolzaei, T. Adhikari, M. Mergoum, and S.F. Kianian. 2011. Mixed model association mapping for Fusarium head blight resistance in Tunisian derived durum wheat populations. G3 (Bethesda, Md) 1: 209-218.

Grant, C.A., J.M. Clarke, S. Duguid, and R.L. Chaney. 2008. Selection and breeding of plant cultivars to minimize cadmium accumulation. Sci. Total Environ., 390: 2-3.

Greger, M., and M. Lofstedt. 2004. Comparison of uptake and distribution of cadmium in different cultivars of bread and durum wheat. Crop Sci. 44: 501-507.

Griffiths, S., J. Simmonds, M. Leverington, Y. Wang, L. Fish, L. Sayers, L. Alibert, S. Orford, L. Wingen, and J. Snape. 2011. Meta-QTL analysis of the genetic control of crop height in elite European winter wheat germplasm. Mol. Breeding 29: 159-171.

Gurung, S., S. Mamidi, J.M. Bonman, E.W. Jackson, L.E. Del Rio, M. Acevedo, M. Mergoum, and T.B. Adhikari. 2011. Identification of novel genomic regions associated with resistance to Pyrenophora tritici-repentis races 1 and 5 in spring wheat landraces using association analysis. Theoretical and Applied Genetics 123: 1029-1041.

Hardy, O.J., and X. Vekemans. 2002. Spagedi: A Versatile computer program to analyze spatial genetic structure at the individual or population levels. Molecular Ecology 2: 618-620.

Harris, N.S., and G.J. Taylor. 2001. Remobilization of cadmium in maturing shoots of near isogenic lines of durum wheat that differ in grain cadmium accumulation. J. Exp. Bot. 52: 1473-1481.

Harris, N.S., and G.J. Taylor. 2004. Cadmium uptake and translocation in seedlings of near isogenic lines of durum wheat that differ in grain cadmium accumulation. B.M.C. Plant Biol. 4: 4.

Hart, J.J., R.M. Welch, W.A. Norvell, L.A. Sullivan, and L.V. Kochian.1998. Characterization of cadmium binding, uptake, and translocation in intact seedlings of bread and durum wheat cultivars. Plant Physiol. 116: 1413-1420.

Hart, J.J., R.M. Welch, W.A. Norvell, and L.V. Kochian.2006. Characterization of cadmium uptake, translocation and storage in near-isogenic lines of durum wheat that differ in grain cadmium accumulation. New Phytol. 172: 261-271.

Hellstrand, S., and L. Landner. 1998. Cadmium in fertilizers, soil, crops and foods - the Swedish situation: In: Cadmium Exposure in the Swedish Environment. KEMI report no 1/98 Hill WG, Robertson A (1968) Linkage disequilibrium in finite populations. Theoretical and Applied Genetics 38: 226-231.

Iqbal, M.J., S. Mamidi, R. Ashan, S.F. Kanian, C.J. Coyne, A.A. Hamama, S.S. Narina, and H.L. Bhardwaj. 2011. Population structure and linkage disequilibrium in Lupinus albus L. germplasm and its implication for association mapping. Theor. Appl. Genet. 3: 517-530.

Ishikawa, S., T. Abe, M. Kuramata, M. Yamaguchi, T. Ando, T. Yamamoto, and M. Yano. 2010. A major quantitative trait locus for increasing cadmium specific concentration in rice grain is located on the short arm of chromosome. Journal of Experimental Botany 61: 923-934.

Ishikawa, S., N. Ae, M. Sugiyama, M. Murakami, T. Arao. 2005. Genotypic variation in shoot cadmium concentration in rice and soybean in soils with different levels of cadmium contamination. Soil Science and Plant Nutrition 51: 101-108.

Jain, S.K. 1975. Population structure and the effects of breeding system. In: Frankel, O.H. & Hawkes, J.G. (eds.) Crop Genetic Resources for Today and Tomorrow. Cambridge Univ. Press. 15-36.

Jay, R.K. 2011. Candidate gene association mapping in spring wheat. Ph.D. diss. Montana State University, Montana, Billing, U.S.A.

Jegadeesan, S., K. Yu, V. Poysa, E. Gawalko, M.J. Morrison, C. Shi, E. Cober. 2010. Mapping

and validation of simple sequence repeat markers linked to a major gene controlling seed cadmium accumulation in soybean [Glycine max (L.) Merr]. Theor. Appl. Genet.121: 283-294

Kashiwagi, T., K. Shindoh, N. Hirotsu, and K. Ishimaru. 2009. Evidence for separate pathways in determining cadmium accumulation in grain and aerial plant parts in rice. BMC Plant Biol 9: 8.

Khan, S., Q. Cao, A.L. Hesham, Y. Xia, J.Z. He. 2007. Soil enzymatic activities and microbial community structure with different application rates of Cd and Pb. J. Environ. Sci. 19: 834-840.

Knox, R.E., C.J. Pozniak, F.R. Clarke, J.M. Clarke, S. Houshmand, and A.K. Singh. 2009. Chromosomal location of the cadmium uptake gene (*Cdu1*) in durum wheat. Genome 52: 741-747.

Kraakman, A.T.W., R.E. Niks, P.M. Van den Berg, P. Stam, and F.A.Van Eeuwijk. 2004. Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. Genetics 168: 435-446.

Lenntech, R., 2004. Water treatment and air purification. Published by Lenntech, Rotterdamseweg, Netherlands (www.excelwater.com/thp/filters/Water-Purification.htm).

Lewontin, R.C. 1964.The interaction of selection and linkage. I. General considerations, heterotic models. Genetics 49: 49-67.

Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754-1760.

Li, Y.M., R.L. Chaney, A.A. Schneiter, J.F. Miller, M.E. Elias, and J.J. Hammond. 1997. Screening for low grain cadmium phenotypes in sunflower, durum wheat and flax. Euphytica 94: 23-30.

Liu, J., M.L. Cheng, Q. Yang, K.R. Shan, J. Shen, Y. Zhou, X. Zhang, A.L. Dill, and M.P. Waalkes. 2007. Blood metallothionein transcript as a biomarker for metal sensitivity: low blood metallothionein transcripts in arsenicosis patients from Guizhou, China. Environ Health Perspect 115: 1101-1106.

Liu, K., and S. V. Muse. 2005. PowerMarker: Integrated analysis environment for genetic marker data. Bioinformatics 21:2128-2129.

Liu, J., Q. Zhu, Z. Zhang, J. Xu, J. Yang, and M.H. Wong. 2005. Variations in cadmium accumulation among rice cultivars and types and the selection of cultivars for reducing cadmium in the diet. Journal of the Science of Food and Agriculture 85: 147-153.

Loiselle, B.A., V.L. Sork, J. Nason, and C. Graham. 1995. Spatial genetic-structure of a tropical

understory shrub, psychotria officinalis (Rubiaceae). American Journal of Botany 82:1420-1425.

Macnair, M.R., V. Bert, S.B. Huitson, P. Saumitou-Laprade, and D. Petit. 1999. Zinc tolerance and hyperaccumulation are genetically independent characters. Proc. R. Soc. Lond. B. 266:2175-2179.

Malosetti, M., C.G. Van der Linden, B. Vosman, and F.A. Van Eeuwijk. 2007. A mixed-model approach to association mapping using pedigree information with an illustration of resistance to Phytophthora infestans in potato. Genetics 175: 879-889.

Mamidi, S., S. Chikara, R.J. Goos, D.L. Hyten, S.M. Moghaddam, P.B. Cregan, P.E. McClean. 2011. Genome-wide association analysis identifies candidate genes associated with iron deficiency chlorosis in soybean. Plant Genome 4: 154-164.

Marchini, J., and B. Howie. 2010. Genotype imputation for genome-wide association studies. Nat. Rev. Genet. 11: 499-511.

McGrath, S.P., F.J. Zhao, E. Lombi. 2001. Plant and rhizosphere process involved in phytoremediation of metal-contaminated soils. Plant Soil 232: 207-214.

McLaughlin, M.J., D.R. Parker, J.M. Clarke. 1999. Metals and micronutrients - food safety issues. Field Crops Res. 60: 143-163.

Morishita, T., N. Fumoto, T. Yoshizawa, and K. Kagawa. 1987. Varietal differences in cadmium levels of rice grains of Japonica, Indica, Javanica and hybrid varieties produced in the same plot of a field. Soil Sci. Plant Nutri. 33: 629-637.

Myles, S., J. Peiffer, P. Brown, E. Ersoz, Z. Zhang, D. Costich, and E. Buckler. 2009. Association mapping: critical considerations shift from genotyping to experimental design. The Plant Cell 21: 2194-2202.

Nakanishi, H., I. Ogawa, Y. Ishimaru, S. Mori, N.K. Nishizawa. 2006. Iron deficiency enhances cadmium uptake and translocation mediated by the Fe2+ transporters OsIRT1 and OsIRT2 in rice. Soil Science and Plant Nutrition 52: 464-469.

National Agricultural Statistics Service (NASS) Retrieved December 4, 2012, from http://www.nass.usda.gov/Statistics_by_State/North_Dakota/index.asp
Neumann K, Kobiljski B, Dencic S, Varshney RK, Borner A (2011) Genome-wide association mapping: a case study in bread wheat (Triticum aestivum L.). Mol. Breeding 27:37-58.

Oliver, D.P., J.W. Gartrell, K.G. Tiller, R. Correll, G.D. Cozens, B.L. Youngberg.1995. Differential responses of Australian wheat cultivars to cadmium concentration in wheat grain. Aust. J. Agric. Res. 46: 873-886.

Patel, J.S., S. Mamidi, J. M. Bonman, and T.B. Adhikari. 2013. Identification of QTL in Spring Wheat Associated with Resistance to a Novel Isolate of Pyrenophora tritici-repentis. Crop Sci. 53: 842-852.

Penner, G.A., L.J. Bezte, D. Leisle, and J. Clarke. 1995. Identification of RAPD markers linked to a gene governing cadmium uptake in durum wheat. Genome 38: 543-547.

Pozniak, C.J., J.M. Clarke, and F.R. Clarke. 2012. Potential for detection of marker-trait associations in durum wheat using unbalanced, historical phenotypic datasets. Mol. Breeding 30: 1537-1550.

Pyhäjärvi, T., M.J. Salmela, O. Savolainen. 2007. Colonization routes of Pinus sylvestris inferred from distribution of mitochondrial DNA variation. Tree Genetics & Genomes 4: 247-254.

Ramesh, S.A., R. Shin, D.J. Eide, D.P. Schachtman. 2003. Differential metal selectivity and gene expression of two zinc transporters from rice. Plant Physiology 133: 126-134.

Ravel, C., S. Praud, A. Murigneux, L. Linossier, M. Dardevet, F. Balfourier, P. Dufour, D. Brunel, and G. Charmet. 2006. Identification of Glu-B1-1 as a candidate gene for the quantity of high molecular- weight glutenin in bread wheat (Triticum aestivum L.) bymeans of an association study. Theoretical and Applied Genetics 112: 738-743.

Remington, D.L., J.M. Thornsberry, Y. Matsuoka, L.M. Wilson, S. R. Whitt, J. Doebley, S. Kresovich, M.M. Goodman, I.V. Buckler. 2001. Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proc. Natl. Acad. Sci. U.S.A. 98: 11479-11484.

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.

Simko, I., S. Costanzo, K.G. Haynes, B.J. Christ, and R.W. Jones. 2004. Linkage disequilibrium mapping of a Verticillium dahliae resistance quantitative trait locus in tetraploid potato (Solanum tuberosum) through a candidate gene approach. Theoretical and Applied Genetics 108: 217-224.

Soric, R., Z. Loncaric, V. Kovacevic, I. Brkic, and D. Simic. 2009. A major gene for leaf cadmium accumulation in maize (Zea mays L.). The Proceedings of the International Plant Nutrition Colloquium XVI, Department of Plant Sciences, UC Davis, UC Davis.

Steemers, F.J., W. Chang, G. Lee, D.L. Barker, R. Shen, K.L. Gunderson. 2006. Whole-Genome Genotyping with the single-base extension assay. Nat Methods 3: 31-33.

Stolt, P., H. Asp, S. Hultin. 2006. Genetic variation in wheat cadmium accumulation on soils with different cadmium concentrations. J. Agron. Crop Sci. 192: 201-208.

Sun, X.L., G.C. Wang, X.M. Cai, S. Jin, Y. Gao, Z.C. Chen. 2010. The tea weevil, Myllocerinus aurolineatus, is attracted to volatiles induced by conspecifics. J. Chem. Ecol. 36: 388-395.

Tanaka, K., S. Fujimaki, T. Fujiwara, T. Yoneyama, and H. Hayashi. 2003. Cadmium concentrations in the phloem sap of rice plant (Oryza sativa L.) treated with a nutrient solution containing cadmium. Soil Sci. Plant Nutr. 49: 311-313.

Tommasini, L., T. Schnurbusch, D. Fossati, F. Mascher, and B. Keller. 2007. Association mapping of Stagonospora nodorum blotch resistance in modern European winter wheat varieties. Theoretical and Applied Genetics 115: 697-708.

Ueno, D., I. Kono, K. Yokosho, T. Ando, M. Yano, and J.F. Ma. 2009. A major quantitative trait locus controlling cadmium translocation in rice (Oryza sativa). New Phytologist 182: 644-653.

Uraguchi, S., S. Mori, M. Kuramata, A. Kawasaki, T. Arao, and S. Ishikawa. 2009. Root-to-shoot Cd translocation via the xylem is the major process determining shoot and grain cadmium accumulation in rice. Journal of Experimental Botany 60: 2677-2688.

Verma, V.K., Y.P. Singh, J.P. Rai. 2007. Biogas production from plant biomass used for phytoremediation of industrial wastes. Bioresource Technology 98: 1664-1669.

Verret, F., A. Gravot, P. Auroy, N. Leonhardt, P. David, L. Nussaume, A. Vavasseur, P. Richaud. 2004. Over expression of AtHMA4 enhances root-to-shoot translocation of zinc and cadmium and plant metal tolerance. FEBS Letters 576: 306-312.

Wangstrand, H., J. Eriksson, I. born. 2007. Cadmium concentration in winter wheat as affected by nitrogen fertilization. Europ. J. Agron. 26: 209-214.

Weigel, H.J. 1985. The effect of Cd2+ on photosynthetic reactions of mesophyll protoplasts.. Plant. Physiol. 63: 192-200.

Wenzel, W.W., W.H. Blum, A. Brandstetter, F. Jockwer, A. Kochl, M. Oberforster, H.E. Oberlander, C. Riedler, K. Roth, I. Vladeva. 1996. Effects of soil properties and cultivar on cadmium accumulation in wheat grain. Z. Pflantzenernahr. Bo- denkd. 159: 609-614.

Wenzel, W.W., and F. Jockwer.1999. Accumulation of heavy metals in plants grown on mineralized soils of the Austrian Aips. Environ. Pollut. 104: 145-155.

Wiebe, K., N.S. Harris, J.D. Faris, J.M. Clarke, R.E. Knox, G.J. Taylor, and C.J. Pozniak. 2010. Targeted mapping of Cdu1, a major locus regulating grain cadmium concentration in durum wheat (Triticum turgidum L. var durum). Theoretical and Applied Genetics 121: 1047-1058.

Xu, L., L. Wang, Y. Gong, W. Dai, Y. Wang, X. Zhu, T. Wen, L. Liu. 2012. Genetic linkage

map construction and QTL mapping of cadmium accumulation in radish (Raphanus sativus L.).Theor. Appl. Genet. 125: 659-670.

Yano, M., S. Kojima, Y. Takahashi, H.X. Lin, and T. Sasaki. 2001. Genetic control of flowering time in rice, a short-day plant. Plant Physiology 127: 1425-1429.

Yin, X.J., T. Tani, Y. Kato and Y. Tsunoda. 2000. Development of rabbit parthenogenetic oocytes and nuclear-transferred oocytes receiving cultured cumulus cells. Theriogenology 54: 1469-1476.

Yu, J., G. Pressoir, W.H. Briggs, B.I. Vroh, M. Yamasaki. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38: 203-208.

Zhao, K., M.J. Aranzana, S. Kim, C. Lister, C. Shindo, C. Tang, C. Toomajian, H. Zheng, C. Dean, P. Marjoram, M. Nordborg. 2007. An Arabidopsis example of association mapping in structured samples. PLOS Genet 3: e4.

Zhu, C., G. Michael, S.B. Edward, and J. Yu. 2008. Status and prospects of association mapping in plants. The Plant Genome 1: 5-20.

Zook, E.G., F.E. Greene, and E.R. Morris. 1970. Nutrient composition of selected wheats and wheat products. VI. Distribution of manganese, copper, nickel, zinc, magnesium, lead, tin, cadmium, chromium, and selenium as determined by atomic absorption spectroscopy and colorimetry. Cereal Chem. 47: 720-731.

CHAPTER IV. RESEARCH HIGHLIGHTS

Key findings of this research included:

Biparental mapping study:

- To develop the RIL mapping population, two durum wheat cultivars were used in this study that show wide phenotyping variations and natural allelic differences between parental lines in Cd accumulation level.
- The single seed descent method was used to develop the subsequent generations.
- The extracted DNA samples were sent to the USDA-ARS Cereal Crops Genotyping Laboratory (Fargo, ND) for genotyping with high density single nucleotide polymorphism (SNP) markers.
- A map for the Grenora x Haurani population was obtained using 549 SNP and 21 polymorphic SSR markers.
- The data obtained from the linkage maps were imported for QTL identification to QGene 4.3.8
- QTL analysis was conducted with composite interval mapping (CIM).
- A Single marker (*Ex_c1343_2570756*) associated with Cd uptake was detected on Chromosome 5B and explained 54.3% of the phenotypic variation.
- The identified SNP marker was validated in different genetic background populations (Stronfield x Alkabo), using the KASPar genotyping assay
- Regression analysis were preformed, the R^2 was 64.7% and highly significant and effective in different populations with different genetic.

Association mapping study:

- Two collections of durum wheat lines consisting of low and high cadmium line sources were selected from the durum wheat breeding program, representing a wide range of Cd accumulation level in the seeds.
- The extracted DNA samples were sent to the USDA-ARS Cereal Crops Genotyping Laboratory (Fargo, ND) for genotyping with high density single nucleotide polymorphism (SNP) markers.
- Four markers (*Ex-c1996-3754394*, *Ra-rep-c106727-90434958*, *Ex-c17754-26503892* and *Ex-c20019-29052512*) identified were associated with low Cd accumulation on chromosomes 2B and 5B.

The markers identified in both studies could be a valuable asset as a selection tool in the durum wheat breeding programs targeting reduced grain Cd levels. However, these markers should be validated in a larger germplasm pool to determine their effectiveness for selecting low Cd lines in durum wheat.