WHOLE GENOME APPROACHES FOR CHARACTERIZING AND UTILIZING SYNTHETIC WHEAT

by

SANDRA MARGARITA DUNCKEL

B.S., ETH Zurich, 2009
M.S., ETH Zurich, 2011

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Interdepartmental Genetics
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2015

Abstract

The global population is estimated to reach 9.1 billion by 2050. Together with climate change, insuring food security for this population presents a significant challenge to agriculture. In this context, a large number of breeding objectives must be targeted. The focus of the work presented here is to explore genomic approaches for tapping exotic germplasm for valuable alleles to increased yield, disease resistance and abiotic stress tolerance.

The loss of genetic diversity in bread wheat (*Triticum aestivum* L.) due to bottlenecks during polyploidization, domestication and modern plant breeding can be compensated by introgressing novel exotic germplasm. Here, the potential of genomic selection (GS) for rapid introgression of synthetic derived wheat is evaluated in field trials. Overall, the GS models had moderate predictive ability. However, prediction accuracies were lower than expected likely due to complex and confounding physiological effects. As such, implementation of rapid cycle GS for introgression of exotic alleles is possible but might not perform very well with synthetic derived wheat.

Disease resistance is another important trait affecting grain yield. Stem rust (*Puccinia graminis* f. sp. *tritici*) has historically caused severe yield loss of wheat worldwide. In a quantitative trait loci (QTL) mapping study with a synthetic-derived mapping population, QTLs for resistance to stem rust races TRTTF and QTHJC were identified on chromosomes 1AS, 2BS, 6AS and 6AL. Some of these genes could be new resistance genes and useful for marker-assisted selection (MAS).

In addition to food insecurity through lack of sufficient source of calories, nutrient deficiency is considered the ‘hidden hunger’ and can lead to serious disorders in humans. Through biofortification, essential nutrients are increased in staple crops for improved quality of

food and human health. A high-throughput elemental profiling experiment was performed with the same synthetic derived mapping population to study the wheat ionome. Twenty-seven QTL for different elements in wheat shoots and two QTL in roots were identified. Four “hotspots” for nutrient accumulation in the shoots were located on chromosomes 5AL, 5BL, 6DL and 7AL.

Overall, exotic germplasm is a valuable source of favorable alleles, but improved breeding methodologies are needed to rapidly utilize this diversity.

WHOLE GENOME APPROACHES FOR CHARACTERIZING AND UTILIZING SYNTHETIC WHEAT

by

SANDRA MARGARITA DUNCKEL

B.S., ETH Zurich, 2009
M.S., ETH Zurich, 2011

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Interdepartmental Genetics
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2015

Approved by:

Major Professor
Jesse Poland

Copyright

Abstract

The global population is estimated to reach 9.1 billion by 2050. Together with climate change, insuring food security for this population presents a significant challenge to agriculture. In this context, a large number of breeding objectives must be targeted. The focus of the work presented here is to explore genomic approaches for tapping exotic germplasm for valuable alleles to increased yield, disease resistance and abiotic stress tolerance.

The loss of genetic diversity in bread wheat (*Triticum aestivum* L.) due to bottlenecks during polyploidization, domestication and modern plant breeding can be compensated by introgressing novel exotic germplasm. Here, the potential of genomic selection (GS) for rapid introgression of synthetic derived wheat is evaluated in field trials. Overall, the GS models had moderate predictive ability. However, prediction accuracies were lower than expected likely due to complex and confounding physiological effects. As such, implementation of rapid cycle GS for introgression of exotic alleles is possible but might not perform very well with synthetic derived wheat.

Disease resistance is another important trait affecting grain yield. Stem rust (*Puccinia graminis* f. sp. *tritici*) has historically caused severe yield loss of wheat worldwide. In a quantitative trait loci (QTL) mapping study with a synthetic-derived mapping population, QTLs for resistance to stem rust races TRTTF and QTHJC were identified on chromosomes 1AS, 2BS, 6AS and 6AL. Some of these genes could be new resistance genes and useful for marker-assisted selection (MAS).

In addition to food insecurity through lack of sufficient source of calories, nutrient deficiency is considered the ‘hidden hunger’ and can lead to serious disorders in humans. Through biofortification, essential nutrients are increased in staple crops for improved quality of

food and human health. A high-throughput elemental profiling experiment was performed with the same synthetic derived mapping population to study the wheat ionome. Twenty-seven QTL for different elements in wheat shoots and two QTL in roots were identified. Four “hotspots” for nutrient accumulation in the shoots were located on chromosomes 5AL, 5BL, 6DL and 7AL.

Overall, exotic germplasm is a valuable source of favorable alleles, but improved breeding methodologies are needed to rapidly utilize this diversity.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I acknowledge the valuable contributions of my advisors, Dr. Jesse Poland and Dr. David Bonnett, who have been great mentors throughout my entire Ph.D. In addition, I would like to thank my other committee members Dr. Allan Fritz and Dr. Bikram Gill for their service and valuable inputs. A special thanks goes to all Poland lab members and students at KSU who have become collaborators and friends. Last but not least, I would like to acknowledge Monsanto and the Monsanto Beachell-Borlaug International Scholarship Program for providing the funding for my Ph.D. and the opportunity to become a member of the MBBIS community.

Dedication

I dedicate this dissertation to my parents Béatrice and Claudio and my dearest brother Guido, who have supported me throughout this journey in many different ways.

Chapter 1 - Genetic mapping of race-specific stem rust resistance in synthetic hexaploid W7984 x Opata M85 mapping population

This chapter has been published as following journal article:
Dunckel, Sandra M., Eric L. Olson, Matthew N. Rouse, Robert L. Bowden, and Jesse A. Poland (2015). Genetic Mapping of Race-Specific Stem Rust Resistance in the Synthetic Hexaploid W7984 x Opata M85 Mapping Population. Crop Science. doi: 10.2135/cropsci2014.11.0755

Abstract

Stem rust (caused by *Puccinia graminis* f. sp. *tritici*) has historically caused severe yield losses of wheat (*Triticum aestivum* L.) worldwide and has been one of the most feared diseases of wheat and barley (*Hordeum vulgare* L.). Stem rust has been controlled successfully through the use of resistant varieties. However, stem rust lineage Ug99 and its derivatives are virulent to many widely deployed stem rust resistance genes including *Sr31*. Doubled haploid lines from the Synthetic W7984 x Opata M85 wheat reference population were screened for seedling resistance to *P. graminis* f. sp. *tritici* races TRTTF and QTHJC. The phenotypic data were adjusted to a 1 to 5 scale and genes for resistance to races TRTTF and QTHJC were localized using composite interval mapping (CIM). Major effect quantitative trait loci (QTLs) for resistance to stem rust races TRTTF and QTHJC were identified on chromosome arms 1AS, 2BS, 6AS, and 6AL. The gene for resistance to both races on 2BS could potentially be a new stem rust resistance gene. The QTLs for resistance on 1AS and 6AL might be other new genes or alleles while the QTL on 6AS is likely an *Sr8* allele. Future work will determine if the resistance loci on 1AS, 2BS, and 6AL are novel. As shown here, the well-studied Synthetic x Opata reference population is a valuable source of potentially novel resistance genes for stem rust that can be leveraged in resistance breeding programs.

Introduction

Wheat stem rust caused by *Puccinia graminis* Pers. f. sp. *tritici* Eriks. & E. Henn. (*Pgt*), has historically been a devastating disease of wheat and barley. Globally, stem rust has caused large losses of wheat yields in the 20th century in Europe and North America (Singh et al., 2006). Through eradication of the alternate host, barberry (*Berberis vulgaris* L.), and the deployment of varieties with genetic resistance, the yield and economic losses due to stem rust were reduced substantially (Singh et al., 2011). For decades, resistance to stem rust has relied on a handful of genes, including resistance gene *Sr31* (Singh et al., 2006).

In 1999, an isolate of *Pgt* virulent to *Sr31* was discovered in Uganda and named Ug99 (Pretorius et al., 2000). The original *Sr31*-virulent Ug99 isolate is designated as race TTKSK based on the North American nomenclature (Jin et al., 2008; Roelfs and Martens, 1987). Stem rust race TTKSK and variant races TTKST (virulent on *Sr24*) and TTTSK (virulent on *Sr36*) are virulent to most known resistance genes (Jin et al., 2008, 2009; Singh et al., 2006, 2008; Visser et al., 2010). Since its first discovery, the Ug99 race group has been detected in Kenya (2001), Ethiopia (2003), Sudan and Yemen (2006), Iran (2007), Tanzania (2009), and South Africa and Zimbabwe (2010) (Hale et al., 2012; Jin et al., 2009; Nazari et al., 2009; Pretorius et al., 2010, 2012). It is estimated that more than 66% of the global wheat growing area is environmentally conducive for the development of stem rust, and in much of this area, susceptible cultivars are grown (Pardey et al., 2013). Therefore, the discovery of new resistance genes with development of markers to facilitate marker-assisted breeding and strategic deployment in gene pyramids is critical.

The genetic diversity for resistance to stem rust in the hexaploid bread wheat gene pool is rather limited (Singh et al., 2011). A natural whole-genome hybridization of cultivated tetraploid wheat (*T. turgidum* L.) ($2n = 4x = 28$, AABB) and diploid wild species *Aegilops tauschii* Coss.

(2*n* = 2*x* = 14, DD) about 8000 yrs. ago gave rise to the allohexaploid species known as bread wheat (2*n* = 6*x* = 42, AABBDD) (Dvorak et al., 1998; Kihara, 1944; McFadden and Sears, 1946; Talbert et al., 1998). Cultivated bread wheat went through multiple genetic bottlenecks during its evolution and domestication process and the diversity within modern wheat varieties has been narrowed further through modern crop improvement (Marcussen et al., 2014; Warburton et al., 2006).

Genetic diversity in the hexaploid germplasm pool can be increased through the production of synthetic hexaploid wheat by crossing tetraploid *T. turgidum* wheat with diploid *Ae. tauschii*. The development of synthetic hexaploids was first demonstrated by McFadden and Sears (1946). The technique has been iteratively improved and several research programs like International Maize and Wheat Improvement Center's (CIMMYT's) wide cross program have developed hundreds of primary synthetic bread wheat lines to capture genetic diversity from wheat progenitors (Reif et al., 2005; Zhang et al., 2005). Even though the primary synthetics generally have poor agronomic performance, they are known for harboring genes for tolerance to a range of biotic and abiotic stresses (Arraiano et al., 2001; Mujeeb-Kazi et al., 2004).

Triticum turgidum, has been a good source of new stem rust resistance genes including *Sr2*, *Sr9d*, *Sr9e*, *Sr9g*, *Sr11*, *Sr12*, *Sr13*, *Sr14*, and *Sr17* (Singh et al., 2011). Genes conferring resistance to race TTKSK are *Sr2*, *Sr13*, and *Sr14* (McIntosh, 1988; Simons et al., 2011; Singh et al., 2006, 2011). *Sr2* confers slow rusting adult plant resistance and is linked with the pseudo-black chaff (PBC) phenotype (Singh and Rajaram, 2002). It confers partial resistance to race TTKSK when homozygous and under low to moderate disease pressure (Mago et al., 2010; Singh et al., 2006). The recessive resistance gene *Sr2* is the primary component of the highly effective "*Sr2* complex" of several minor genes (Hare and McIntosh, 1979). Resistance gene

Sr13 confers resistance to race TTKSK (Jin et al., 2007). The sources of this gene are the Ethiopian land race ST464 and the emmer wheat cultivar Khapli (Klindworth et al., 2007; Knott, 1962). Like *Sr13, Sr14* was introduced from emmer wheat Khapli (Knott, 1962; McIntosh, 1980).

Six stem rust resistance genes or alleles from *Ae. tauschii* have been described including *Sr33, Sr45*, *Sr46, SrTA1662, SrTA1017* , and *SrTA10187* (Kerber and Dyck, 1979; Olson et al., 2013a, 2013b; Rouse et al., 2011). All are effective against TTKSK. Resistance genes *Sr33* and *Sr45* were incorporated into synthetic wheat by Kerber and Dyck (1979). Both, *Sr33* and *Sr45* originated from *Ae. tauschii* accessions found in Iran (Olson et al., 2013b). Resistance to several *Pgt* races has recently been identified in 98 *Ae. tauschii* accessions (Rouse et al., 2011). However, Rouse et al. (2011) were not able to postulate the presence of *Sr33, Sr45, Sr46,* or new genes in the accessions because of the complexity of the stem rust phenotypes. Recently, Olson et al. (2013a, 2013b) transferred *SrTA1662, SrTA1017 ,* and *SrTA10187* from *Ae. tauschii* by direct crossing between diploid *Ae. tauschii* and hexaploid *T. aestivum*.

Two synthetic wheat reference populations have recently been reconstructed by Sorrells et al. (2011). One is a doubled haploid population (named SynOpDH) and one consists of recombinant inbred lines (named SynOpRIL). Both populations were developed by crossing synthetic hexaploid line W7984 and elite bread wheat cultivar Opata M85 (Sorrells et al., 2011). The SynOpDH and RIL mapping populations were developed from the same parents as the original mapping population in the late 1980s. That population was also known as M6 x Opata, Synthetic x Opata, and ITMI mapping population. Here, we evaluated the recreated synthetic wheat doubled haploid mapping population (SynOpDH) for resistance to several stem rust races at the seedling stage to identify potentially new stem rust resistance genes.

Materials and Methods

Mapping population

The SynOpDH mapping population was developed by Sorrells et al. (2011). The pedigree of the population is synthetic W7984 (Altar 84/*Ae. tauschii* (219) CIGM86.940)/Opata M85. The population consists of 215 lines.

Genotypic Data

Genome-wide marker data on the SynOpDH mapping population was generated using a two-enzyme genotyping-by- sequencing (GBS) approach by Poland et al. (2012). The publicly available map and marker data were used to perform the QTL analysis. The previously constructed map consisted of 1485 single nucleotide polymorphisms (SNPs). Complete information about data filtering, SNP calling and map construction can be found in Poland et al. (2012). The female parental alleles from Synthetic W7984 were coded as "A", the male parental alleles from Opata M85 as "B". We verified the original assignment of the GBS markers to chromosomes and chromosome arms in Poland et al. (2012) by aligning the tags to the recently published draft sequence of the wheat genome (The International Wheat Genome Sequencing Consortium, 2014). The linkage groups from GBS markers corresponding to chromosomes 1B, 1D, 2A, 5B, 6D and 7A were inverted to match the correct short and long chromosome arms, respectively. An updated map with corrected linkage group orientation and chromo- some arm assignments is available in Supplemental Table S3 online.

Pgt Inoculation and Evaluation of Seedling Infection Types

The parents of SynOpDH, Synthetic W7984 and Opata M85, and a subset of the mapping population were screened with 14 *Pgt* races of diverse geographical origin at the USDA-ARS Cereal Disease Laboratory, Saint Paul, MN. Foreign *Pgt* races were evaluated in a biosafety

level 3 containment facility at the University of Minnesota. Based on the observed segregation pattern, *Pgt* races TRTTF (from Yemen) and QTHJC (United States) were targeted for further mapping experiments.

Reactions to *Pgt* races TRTTF (isolate 06YEM34–1) and QTHJC (isolate 75ND717C) were evaluated in St. Paul, MN, following protocols described previously. The experimental design was a randomized complete block design (RCBD) with two replications of 10 seedlings per DH line for races TRTTF and QTHJC. Seedling infection types (ITs) to races TRTTF and QTHJC were evaluated in 127 and 100 doubled haploid lines, respectively. The stem rust susceptible wheat line LMPG-6 and the 20 lines of the North American *Pgt* differential set were included in the experiments as controls (Jin et al., 2008).

Urediniospores stored at –80°C were heat-shocked in a water bath at 45°C for 5 min and suspended in Soltrol 170 isoparaffin oil (Chevron Phillips Chemical Company LP, The Woodlands, TX). The suspension was sprayed onto 7- to 9-day- old seedlings and inoculated seedlings were placed in mist chambers overnight at 20±1°C at 100% humidity and then transferred to a greenhouse bench with a 16 h light/8 h dark cycle at 18±2°C. Infection types to both races were scored 14 d after inoculation using the scale of Stakman et al. (1962). Seedlings showing low and intermediate ITs up to 2+3- were considered resistant and seedlings showing high ITs of 3 to 4 were considered susceptible. The avirulence/virulence formula for TRTTF was 8a, 24, 31/5, 6, 7b, 9a, 9b, 9d, 9e, 9g, 10, 11, 17, 21, 30, 36, 38, McN, Tmp,$1RS^{Amigo}$ and for QTHJC 7b, 9a, 9e, 24, 30, 31, 36, 38, Tmp, $1RS^{Amigo}$/5, 6, 8a, 9b, 9d, 9g, 10, 11, 17, 21, McN (Jin et al., 2008; Olivera et al., 2012a).

The qualitative phenotypic data using the Stakman scale were converted to a 1 to 5 quantitative scale to enable analysis with QTL mapping algorithms that assume an ordered

phenotypic distribution (Table 1.1). Low ITs with hypersensitive flecking including; ;, 0;, ;1,;1+,;2-,;2,;2+, 2-;, and ;12, were classified as 1. ITs of 2- to 2 were classified as 2. Lines displaying 2+ were classified as 3. Lines scoring ITs ranging from 2+3- and 32+ to 3 were scored as 4. High susceptible ITs of 3+ or 4 were scored as 5. The Stakman scale was converted for each replicate and then averaged across replicates. SynOpDH lines with missing data or inconsistent phenotypes (indicating mixed seed source) between replicates were removed before further analysis. The repeatability of the experiments was tested with the Pearson correlation coefficient after conversion to the 1 to 5 scale.

Quantitative Trait Loci Analysis

Identification of stem rust resistance QTLs was performed in the R software environment (R Core Team, 2013) using the R-package R/qtl (Broman et al., 2003). The most significant markers were identified with stepwise regression separately for both *Pgt* races. Composite Interval Mapping (CIM) was implemented applying a Haley–Knott regression using forward selection of marker covariates and a window size of 10 cM for both stem rust races. Three marker covariates were used for CIM for races TRTTF and QTHJC. The map position and markers of QTLs identified by single interval mapping (SIM) and CIM were used in multiple quantitative trait loci mapping (MQM) (Arends et al., 2010) to confirm identified resistance loci and to refine their position. MQM was implemented in R/qtl to obtain estimated QTL effects. The genome-wide logarithm of the odds value (LOD) for declaring significant QTL for each race was determined by 1000 permutations. The parental alleles for Synthetic W7984 and Opata M85 were coded as –1 and 1, respectively, as described by Broman and Saunak (2009).

Results

The parents of the SynOpDH population showed a wide range of ITs when tested with 14 *Pgt* races (Table 1.2). Both parents were observed to be susceptible to races TTKSK, TTKST, and TTTSK. However, both parents showed resistance to several *Pgt* races including TRTTF and QTHJC (Table 1.2). The observed IT of the synthetic parent to *Pgt* race TRTTF was 22-, and ; to QTHJC. Elite parent Opata M85 showed ITs of 22+ and 2 to races TRTTF and QTHJC, respectively.

To test the repeatability of the experiments the Pearson correlation coefficient *r* was calculated for the DH lines. *r* values of 0.84 and 0.9 for QTHJC and TRTTF, respectively, indicated good repeatability of each experiment. The SynOpDH mapping population segregated for distinct resistance ITs to *Pgt* race TRTTF (Figure 1.1 a, Supplemental Table S1 online). Under the binary, resistant/susceptible IT designation, the SynOpDH population segregated 108:19 resistant/susceptible for race TRTTF (Table 1.3) not significantly different from a 7:1 ratio ($\chi^2 = 0.70$, $P = 0.40$), suggesting three resistance genes were segregating. Mapping revealed three QTLs for resistance to race TRTTF with a genome-wide LOD of 4.20 at a 5% error rate as determined by 1000 permutations. The three loci identified with CIM are located on 2BS, 6AS, and 6AL (Figure 1.2). The QTL on 2BS is located proximal (most significant GBS marker is synopGBS355 at 53.8 cM). The other two QTLs are located on 6AS (GBS marker synopGBS1019 at 0.8cM) and 6AL (GBS marker synopGBS85 at 130.6cM) (Table 1.4).

The segregation pattern of resistance to stem rust race QTHJC showed a relatively larger proportion of lines was resistant to race QTHJC compared to race TRTTF (Figure 1.1 b). The SynOpDH population segregated 92:8 resistant/susceptible, which is not significantly different from either 7:1 or 15:1 ratios ($\chi^2 = 1.85$, $P = 0.17$; $\chi^2 = 0.52$, $P = 0.47$), suggesting three or four

resistance genes were segregating. Quantitative trait locus mapping with CIM identified one QTL on 2BS (Figure 1.2) with a genomewide LOD of 3.70 at a 5% error rate. The marker with the highest LOD score on 2BS is located at 44.1cM (GBS marker synopGBS616). A second resistance QTL was located at 27.9cM on chromosome 1AS (most significant GBS marker synopGBS665) (Table 1.4).

The MQM is a useful tool to obtain estimates on phenotypic variance, QTL, and allele effects. The allelic state of the markers with the highest LOD score at each QTL was used to represent the allelic state of the QTL. For race TRTTF, a model with three genes and no interaction explained more than 60% of the phenotypic variance. The resistance genes on 2BS, 6AS, and 6AL explained 23, 13, and 26% of the estimated phenotypic variation, respectively (Table 1.5). The estimated allele effects show that the resistance to race TRTTF is conferred through alleles from both the synthetic and elite wheat parent (Figure 1.3). The estimated QTL effect for the resistance gene on 2BS is 0.53 and 0.4 for the gene on 6AS with the resistant allele coming from the synthetic parent for both of these genes (Table 1.5 and Figure 1.3). The resistant allele at the locus on 6AL was contributed by Opata and had an estimated effect of –0.57 (Table 1.5). Analysis of race TRTTF confirmed that resistant lines of phenotype classes 1 to 3 have either all three resistant alleles, or any combination of two out of the three identified resistance alleles.

For race QTHJC a model considering only one QTL on 2BS explains 26.7% of the estimated phenotypic variation. Including the QTL on 1AS, the estimated phenotypic variation increases to 35.5%. Searching for further QTLs by means of MQM did not result in any LOD peaks over the set threshold. However, the LOD profile suggests the presence of a third resistance QTL on chromosome arm 1DL at 154.6cM, barely below the threshold of 3.70 at an

error rate of 5%. The estimated effect for the QTL on 2BS is 0.520 and 0.318 for the 1AS QTL. Based on the estimated allele effects both resistance QTLs are contributed by the synthetic parent (Table 1.5 and Figure 1.3).

Discussion

This study identified three QTLs on chromosome arms 2BS, 6AS, and 6AL for resistance to *Pgt* race TRTTF and two QTLs on 1AS and 2BS for resistance to race QTHJC in the newly reconstructed synthetic hexaploid W7984 × Opata M85 wheat reference population. The estimated allele effects showed that resistance was conferred through alleles from both the synthetic hexaploid and the bread wheat parents. Based on the deduced chromosomal locations and the pedigree of the mapping population, only a few previously described resistance genes from durum wheat and bread wheat could be candidates for the QTLs identified here.

The resistance QTL on 2BS that was detected with race TRTTF was derived from the synthetic hexaploid parent and mapped to position 53.8 cM (Figure 1.2 and Figure 1.3). The resistance QTL for race QTHJC was initially mapped at 44.1 cM on 2BS. Applying MQM, the position of the gene was refined and newly positioned to 51.8 cM. Given the close proximity, it is possible that these effects for the two races are from the same resistance gene. Chromosome arm 2BS harbors at least seven numerically designated stem rust resistance genes, but many are on alien translocations (McIntosh et al., 2012). Resistance genes on 2BS from common wheat and durum include *Sr10, Sr19, Sr20*, and *Sr23*. Based on its position and presence in CIMMYT germplasm (McIntosh et al., 1995), *Sr10* is a potential candidate for the 2BS resistance QTL. However, both races TRTTF and QTHJC are virulent on *Sr10*, which, there- fore, can be ruled out. Canadian spring wheat variety Marquis is the source of *Sr19* and *Sr20* (McIntosh et al., 1995) and has been used in crosses made at CIMMYT (Smale, 1996). Virulence was reported to

be very common on *Sr19* and *Sr20* (McIntosh et al., 1995), but neither can be ruled out as candidates for the 2BS QTL. Stem rust resistance gene *Sr23* is completely associated with *Lr16* (McIntosh et al., 1995). However, *Lr16* was mapped at the distal end of chromosome 2BS (McCartney et al., 2005). Consequently, the mapped gene is likely not *Sr23* due to the different map position. Additional data are needed to determine the allelic relationship between the QTL mapped on 2BS and the numerically designated genes on 2BS.

The QTL for resistance to race TRTTF identified on 6AS was placed at 11cM distal following MQM. Based on its location, the QTL mapped to 6AS is probably conferred by *Sr8a* or *Sr8b* (Mcintosh, 1972; McIntosh et al., 1995). *Sr8a* was first characterized from bread wheat and was widely used in lines developed in Europe and Mexico (McIntosh et al., 1995). The location of *Sr8* is at the distal end of chromosome 6AS (GrainGenes2, 2013) and its location matches the resistance QTL. However, based on the allele effect of –0.402 for allele A the resistance gene was contributed by the synthetic parent. *Sr8b* is known to be present in durum wheat (Bhavani et al., 2008). Race TRTTF is avirulent to *Sr8a* whereas race QTHJC is virulent. The response of both races to *Sr8b* is unknown. Additional data are needed to determine the relationship between the QTL on 6AS and the two *Sr8* alleles.

The gene located on 6AL maps to the region of *Sr13,* which is known to confer resistance against TTKSK and its variants TTKST and TTTSK (Klindworth et al., 2007). However, both parents showed susceptible ITs to all three races (Table 1.1) indicating that this identified resistance locus is not *Sr13*. Furthermore, TRTTF is virulent on *Sr13* (Olivera et al., 2012b). Therefore, the gene mapped on 6AL is likely a new gene or a novel allele of *Sr13*.

For race QTHJC, segregation ratios of 7:1 and 15:1 suggest three or four resistance genes. However, only two resistance QTLs on chromosomes 1AS and 2BS were identified with

confidence. Based on the LOD profile, two additional QTL may be on chromosome arms 1DL and 7DS. Opata is known to carry the pleiotropic gene designated *Lr34*/*Yr18*/*Sr57*, which is an adult plant resistance gene on 7DS. It is possible that *Sr57* has a detectable effect at the seedling stage, especially in combination with other genes. One resistance gene has been reported on 1AS that originated from the 1AL.1RS wheat–rye translocation (Lein, 1975; McIntosh et al., 1998). However, this mapping population does not harbor this translocation, and, therefore, the resistance QTL identified here might be conferred by a new resistance gene. It is possible that the allele effect of each susceptible allele was not properly estimated due to some other gene(s) conferring resistance. This might lead to confounding estimations of the number of resistance genes. Further experiments will be needed to determine the mechanism of resistance to race QTHJC before proceeding with fine mapping and marker development for marker-assisted selection.

Through screening the SynOpDH reference mapping population for stem rust resistance, we have identified multiple resistance loci to the highly virulent *Pgt* race TRTTF and United States race QTHJC. Based on the virulence patterns and locations of known resistance genes, we conclude that the QTL on 6AL and 1AS could be new genes or new alleles of known genes. Additional data are needed to determine the relationship between the QTL on 2BS and 6AS and known genes on these chromosome arms. With the continued identification and marker tagging of effective stem rust resistance genes, the tools available to breeders for developing resistant breeding material and new varieties will further improve.

Supplemental information Available

Supplemental information is available with the online version of this manuscript.

Acknowledgment

The Graduate Research Assistantship of S. Dunckel is supported through the Monsanto Beachell-Borlaug International Scholars Program. This work was done under the auspices of the Wheat Genetics Resource Center (WGRC) Industry/University Collaborative Research Center (I/UCRC) supported by NSF grant contract (IIP-1338897) and industry partners. This work was funded by the U.S. Agency for International Development (USAID Cooperative Agreement no. AID-OAA-A-13-0005), the United States Department of Agriculture-Agricultural Research Service (Appropriations no. 5430-21000-006-00D and no. 3640-21220-021-00) and The Bill & Melinda Gates Foundation through the Durable Rust Resistance in Wheat project to Cornell University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. USDA is an equal opportunity provider and employer. This work represents contribution number 15-415-J from the Kansas Agriculture Experiment Station (KAES). Conflict of interest: The authors declare that they have no conflict of interest.

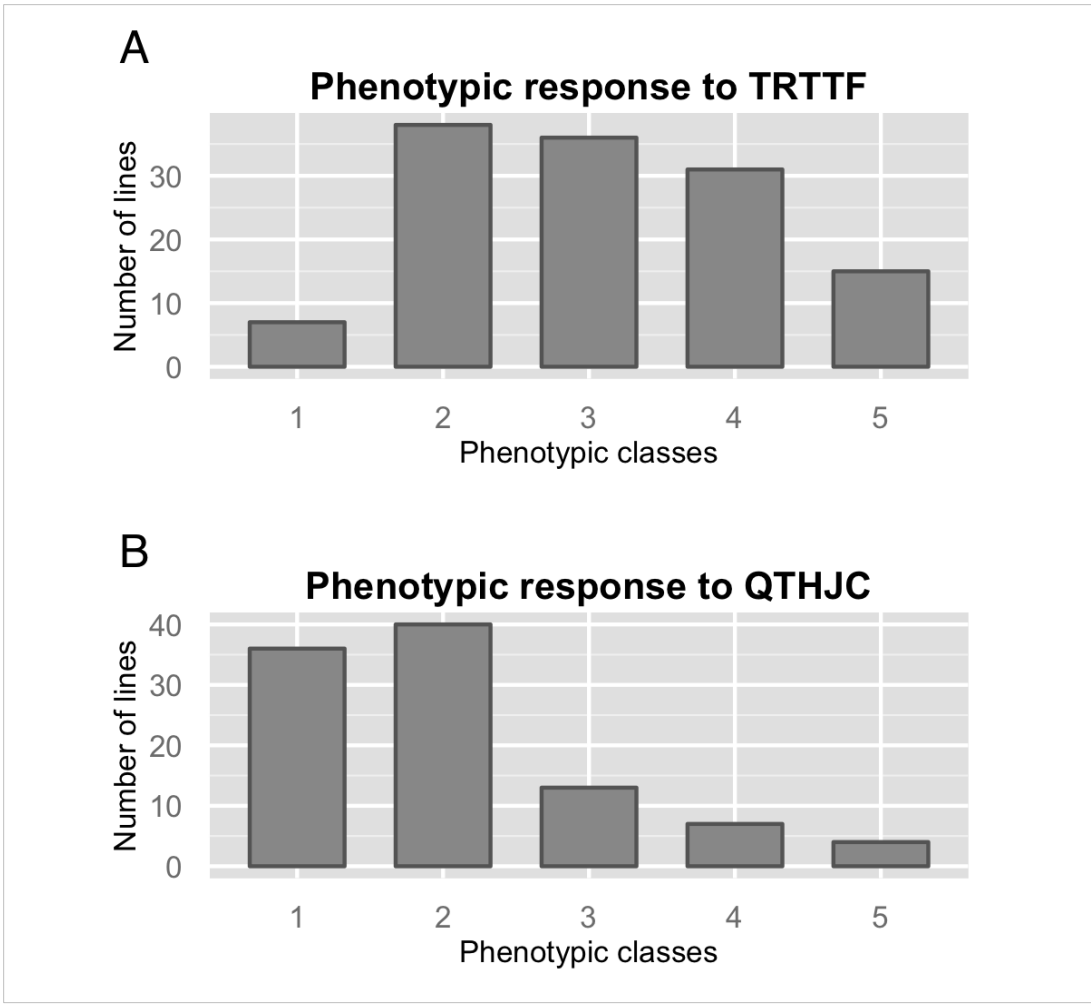

Figure 1.1 Phenotypic distribution of seedling infection types (ITs)

Phenotypic distribution of seedling infection types (Its) to stem rust race TRTTF after conversion of the Stakman scale to a 1 – 5 scale as described earlier. Lines falling into categories 1 to 3 are considered resistant and lines in categories 4 and 5 are susceptible to stem rust race TRTTF (A) and QTHJC (B).

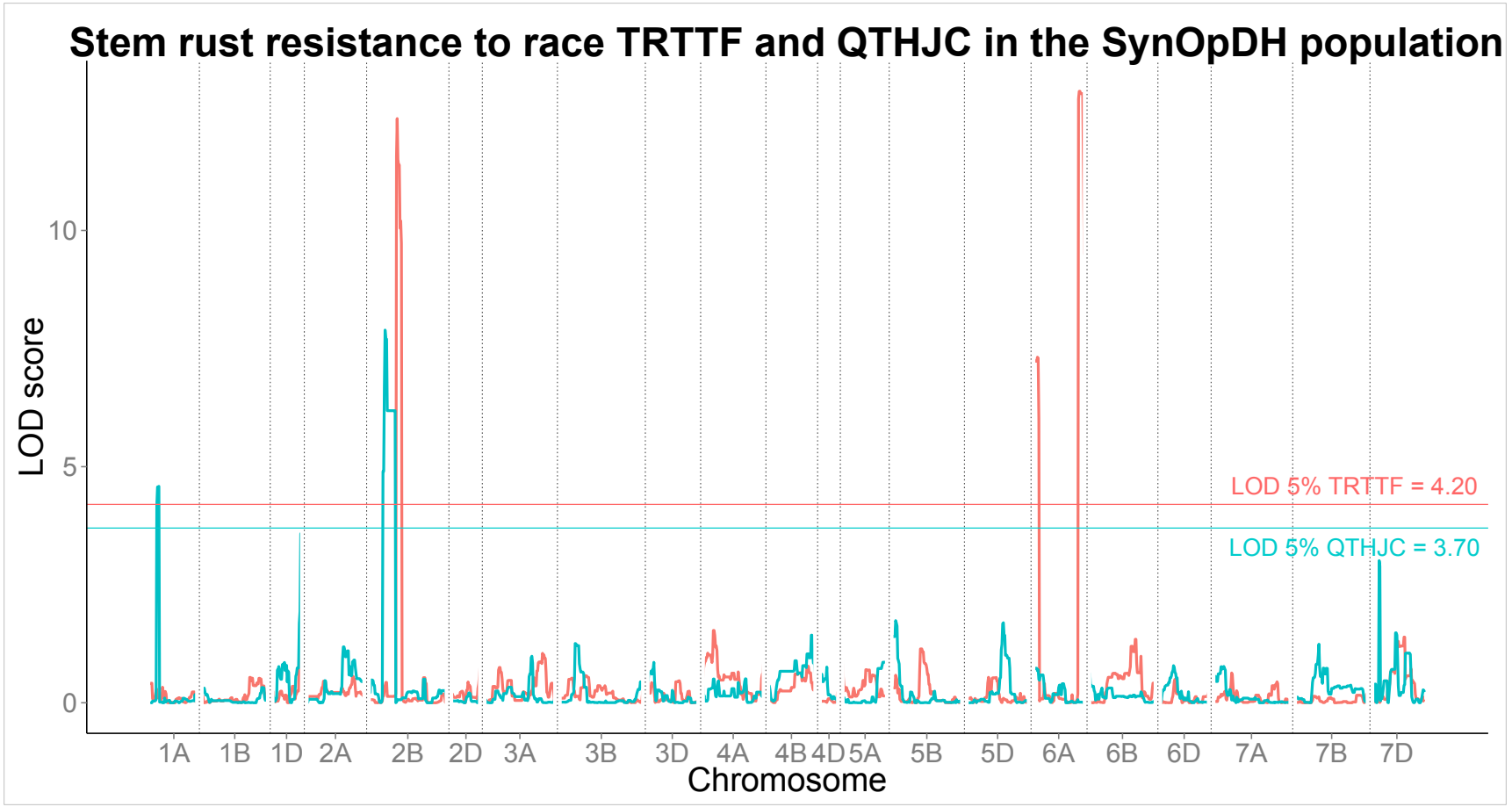

Figure 1.2 Logarithm of the odds (LOD) profile

The logarithm of the odds (LOD) profile for stem rust resistance to both races TRTTF and QTHJC shows the identified resistance genes for race TRTTF in orange on chromosomes 2BS, 6AS and 6AL (LOD 5% = 4.20), and for stem rust race QTHJC in turquoise on 2BS (LOD 5% = 3.70).

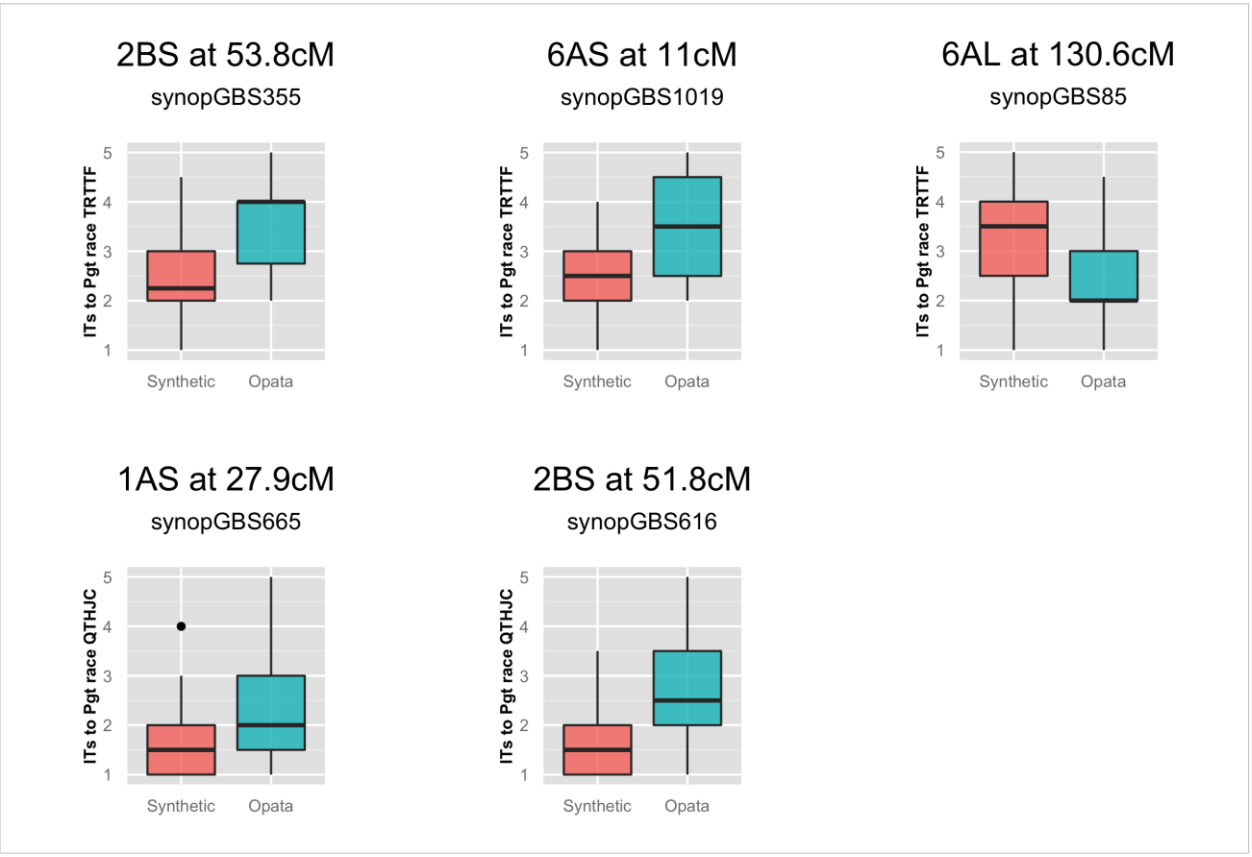

Figure 1.3 Estimated allele effect at each QTL

Estimated allele effects at each identified resistance QTL for both stem rust races TRTTF (above) and QTHJC (below). Allele A is the allele contributed by the synthetic parent, allele B by the elite parent Opata M85.

Table 1.1 Assessment of the resistance of the parents (Synthetic W7984 and Opata M85) of the mapping population with 14 different *Pgt* races from diverse origin

Race	Isolate	Origin	Synthetic W7984	Opata M85
TTKSK	04KEN156/04	Kenya	3	33+
TTKST	06KEN 19-V-3	Kenya	33+	3+
TTTSK	07KEN 24-4	Kenya	3+	3+
TRTTF	06YEM34-1	Yemen	22-	22+
TTTTF	01MN84A-1-2	United States	3	4
TPMKC	74MN1409	United States	2+3-	2+3-
RKQQC	99KS76A-1	United States	;1	0;
RCRSC	77ND82A	United States	2=;1	;1
QTHJC	75ND717C	United States	;	2
QFCSC	06ND76C	United States	0;	0;
MCCFC	59KS19	United States	;1-	;1-
QCCSM	75WA1652-A	United States	0;	0;/;1+
QCCJB	01SD80-A	United States	0;	;1
SCCSC	09ID73-2	United States	;13-	0

Table 1.2 Conversion of Stakman infection types to a 1 – 5 scale for mapping purposes

Stakman Infection Types	Conversion
;, ;0, ;1, ;1+, ;2-, ;2, ;2+, 2-;, ;12	Class 1
2- to 2	Class 2
2+	Class 3
2+3- and 32+ to 3	Class 4
3+ to 4	Class 5

Table 1.3 Number, segregation ratio, χ^2 and corresponding p-values of SynOpDH lines showing resistance and susceptibility to *Pgt* races TRTTF and QTHJC

	Resistant	Susceptible	Segregation ratio	χ^2	p-value
TRTTF	108	19	7:1[a]	0.70	0.40
QTHJC	92	8	7:1	1.85	0.17
			15:1[b]	0.52	0.47

[a] Segregation ratio for 3 gene model
[b] Segregation ratio for 4 gene model

Table 1.4 GBS marker and QTL position for identified resistance QTL to races TRTTF and QTHJC

Race	Chromosome	GBS Marker	Position
TRTTF	2BS	synopGBS355	53.8cM
	6AS	synopGBS1019	11.0cM
	6AL	synopGBS85	130.6cM
QTHJC	1AS	synopGBS665	27.9cM
	2BS	synopGBS616	51.8cM

Table 1.5 Estimated phenotypic variance explained by each gene, estimated QTL and allele effects for detected resistance QTL to races TRTTF and QTHJC

Race	Chromosome	Phenotypic variance	QTL effect	SE	Effect allele A	Effect allele B	Resistance from parent
TRTTF	2BS	23.1 %	0.53	0.06	-0.53	0.53	Synthetic
	6AS	13.0 %	0.40	0.06	-0.40	0.40	Synthetic
	6AL	26.4 %	-0.57	0.06	0.57	-0.57	Opata M85
QTHJC	1AS	9.7 %	0.32	0.08	-0.32	0.32	Synthetic
	2BS	24.4 %	0.52	0.09	-0.52	0.52	Synthetic

References

Arends, D., P. Prins, R.C. Jansen, and K.W. Broman. 2010. R/ qtl: High-throughput multiple QTL mapping. Bioinformatics 26:2990–2992. doi:10.1093/bioinformatics/btq565

Arraiano, L.S., A.J. Worland, C. Ellerbrook, and J.K.M. Brown. 2001. Chromosomal location of a gene for resistance to septoria tritici blotch (*Mycosphaerella graminicola*). In: The hexaploid wheat 'Synthetic 6x'. Theor. Appl. Genet. 103:758–764. doi:10.1007/ s001220100668

Bhavani, S., U.K. Bansal, R.A. Hare, and H.S. Bariana. 2008. Genetic mapping of stem rust resistance in durum wheat cultivar 'Arrivato'. Int. J. Plant Breed. 2:23–26.

Broman, K.W., and S. Saunak. 2009. A guide to QTL mapping with R/qtl. Springer, New York.

Broman, K.W., H. Wu, Ś Sen, and G.A. C. 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889–890. doi:10.1093/bioinformatics/btg112

Dvorak, J., M.C. Luo, Z.L. Yang, and H.B. Zhang. 1998. The structure of the *Aegilops tauschii* genepool and the evolution of hexaploid wheat. Theor. Appl. Genet. 97:657–670. doi:10.1007/ s001220050942

GrainGenes2.2013. Map of Sr8. USDA-ARS. http://wheat.pw.usda.gov/cgi-bin/graingenes/report.cgi?class=locus;name=Sr8 (accessed 12 Sept. 2013).

Hale, I.L., I. Mamuya, and D. Singh. 2012. *Sr31*-virulent races (TTKSK, TTKST, and TTTSK) of the wheat stem rust pathogen *Puccinia graminis* f. sp. *tritici* are present in Tanzania. Plant Dis. 97:557–557. doi:10.1094/PDIS-06-12-0604-PDN

Hare, R.A., and R.A. McIntosh. 1979. Genetic and cytogenetic studies of durable, adult-plant resistances in Hope and related cultivars to rusts. Planzenzüchtung 83:350–367.

Jin, Y., R.P. Singh, R.W. Ward, R. Wanyera, M. Kinyua, P. Njau, T. Fetch, Z.A. Pretorius, and A. Yahyoui. 2007. Characterization of seedling infection types and adult plant infection responses of monogenic *Sr* gene lines to race TTKS of *Puccinia graminis* f. sp. *tritici*. Plant Dis. 91:1096–1099. doi:10.1094/PDIS-91-9-1096

Jin, Y., L.J. Szabo, M.N. Rouse, T. Fetch, Z.A. Pretorious, R. Wanyera, and P. Njau. 2009. Detection of virulence to resistant gene *Sr36* within the TTKS race lineage of *Puccinia graminis* f. sp. *Tritici*. Plant Dis. 93:367–370. doi:10.1094/PDIS-93-4-0367.

Jin, Y., L.J. Szabo, Z.A. Pretorius, R.P. Singh, R. Ward, and T. Fetch. 2008. Detection of virulence to resistance gene *Sr*24 within race TTKS of *Puccinia graminis* f. sp. *tritici*. Plant Dis. 92:923–926. doi:10.1094/PDIS-92-6-0923

Kerber, E.R., and P.L. Dyck. 1979. Resistance to stem and leaf rust of wheat in *Aegilops squarrosa* and transfer of a gene for stem rust resistance to hexaploid wheat. In: S. Ramanjuam, editor, Proceedings of the 5th International Wheat Genetics Symposium, New

Delhi, India. February 1978. Indian Soc. of Genetics and Plant Breeding, Indian Agric. Res. Inst., New Dehli, India.

Kihara, H. 1944. Discovery of the DD-analyser, one of the ancestors of *Triticum vulgare*. Agric. Hortic. 19:889–890.

Klindworth, D.L., J.D. Miller, Y. Jin, and S.S. Xu. 2007. Chromosomal locations of genes for stem rust resistance in monogenic lines derived from tetraploid wheat accession ST464. Crop Sci. 47:1441–1450. doi:10.2135/cropsci2006.05.0345

Knott, D.R. 1962. The inheritance of rust resistance IX. The inheritance of resistance races 15B and 56 of stem rust in the wheat variety Khapstein. Can. J. Plant Sci. 42:415–419. doi:10.4141/ cjps62-068

Lein, A. 1975. Introgression of a rye chromosome to wheat strains by Georg Riebesel-Salzmunde after 1926. In: F. Dorofeev and T.Y. Zarubaiko, editors, Proceedings EUCARPIA Symposium Triticale: Studies and breeding. Vsesoyuzny, Inst. Rastenievodstva, Leningrad. p. 158–168.

Mago, R., G. Brown-Guedira, S. Dreisigacker, J. Breen, Y. Jin, R. Singh et al. 2010. An accurate DNA marker assay for stem rust resistance gene *Sr*2 in wheat. Theor. Appl. Genet. 122:1–10. doi: 10.1007/s00122-010-1482-7.

Marcussen, T.A., S.R. Sandve, L. Heier, M. Spannagl, M. Pfeifer, The International Wheat Genome Sequencing Consortium, K.S. Jakobsen, B.B. Wulff, B. Steuernagel, K.F.X. Mayer, and O.-A. Olsen. 2014. Ancient hybridizations among the ancestral genomes of bread wheat. Science (Washington, DC) 345:6194. doi:10.1126/science.1250092

McCartney, C.A., D.J. Somers, B.D. McCallum, J. Thomas, D.G. Humphreys, J.G. Menzies, and P.D. Brown. 2005. Microsatellite tagging of the leaf rust resistance gene *Lr*16 on wheat chromosome 2BS. Mol. Breed. 15:329–337. doi:10.1007/s11032- 004-5948-7

McFadden, E.S., and E.R. Sears. 1946. The origin of *Triticum spelta* and its free-threshing hexaploid relatives. J. Hered. 37:81–89.

Mcintosh, R. 1972. Cytogenetical studies in wheat VI. Chromosome location and linkage studies involving *Sr*13 and *Sr*8 for reaction to *Puccinia graminis* f. sp. *tritici*. Aust. J. Biol. Sci. 25:765–774.

McIntosh, R.A. 1980. Chromosome location and linkage studies involving the wheat stem rust resistance gene *Sr*14. Cereal Res. Commun. 8:315–320.

McIntosh, R.A. 1988. The role of specific genes in breeding for durable stem rust resistance in wheat and triticale. In: N.W. Simmonds and S. Rajaram, editors, Breeding strategies for resistance to the rust of wheat. CIMMYT, Mexico, DF. p. 1–9.

McIntosh, R.A., J. Dubcovsky, W.J. Rogers, C. Morris, R. Appels, and X.C. Xia. 2012. Catalogue of gene symbols for wheat: 2012 supplement. Natl. Inst. of Genetics, Japan.

http://www.shigen. nig.ac.jp/wheat/komugi/genes/macgene/supplement2012.pdf (accessed 27 Jan. 2014).

McIntosh, R.A., G.E. Hart, K.M. Devos, M.D. Gale, and W.J. Rogers. 1998. Catalogue of gene symbols for wheat. In: A.E. Slinkard, editor, Proceedings of the 9th International Wheat Genetics Symposium. Vol. 5, Saskatoon, SK, Canada. 2–7 Aug. 1998. Univ. Extension Press, Univ. of Saskatchewan, Saskatoon. p. 1–236.

McIntosh, R.A., C.R. Wellings, and R.F. Park. 1995. Wheat Rusts: An Atlas of Resistance Genes. CSIRO Publ., Clayton, Victoria, Australia.

Mujeeb-Kazi, A., R. Delgado, A. Cortes, S. Cano, V. Rosas, and J. Sanchez. 2004. Progress in exploiting *Aegilops tauschii* for wheat improvement. Annual Wheat News Letter 50:79–88.

Nazari, K., M. Mafi, A. Yahyaoui, R.P. Singh, and R.F. Park. 2009. Detection of wheat stem rust (*Puccinia graminis* f. sp. *tritici*) race TTKSK (Ug99) in Iran. Plant Dis. 93:317. doi:10.1094/PDIS- 93-3-0317B

Olivera, P.D., A. Badebo, S.S. Xu, D.L. Klindworth, and Y. Jin. 2012a. Resistance to race TTKSK of *Puccinia graminis* f. sp. *tritici* in Emmer Wheat. Crop Sci. 52:2234–2242. doi:10.2135/crop- sci2011.12.0645

Olivera, P.D., Y. Jin, M. Rouse, A. Badebo, T. Fetch, R.P. Singh, and A. Yahyaoui. 2012b. Races of *Puccinia graminis* f.sp. *tritici* with combined virulence to *Sr13* and *Sr9e* in a field stem rust screening nursery in Ethiopia. Plant Dis. 96:623–628. doi:10.1094/ PDIS-09-11-0793

Olson, E.L., M.N. Rouse, M.O. Pumphrey, R.L. Bowden, B.S. Gill, and J.A. Poland. 2013a. Introgression of stem rust resistance genes *Sr*TA10187 and *Sr*TA10171 from *Aegilops tauschii* to wheat. Theor. Appl. Genet. 126:2477–2484. doi:10.1007/ s00122-013-2148-z

Olson, E.L., M.N. Rouse, M.O. Pumphrey, R.L. Bowden, B.S. Gill, and J.A. Poland. 2013b. Simultaneous transfer, introgression, and genomic localization of genes for resistance to stem rust race TTKSK (Ug99) from *Aegilops tauschii* to wheat. Theor. Appl. Genet. 126:1179–1188. doi:10.1007/s00122-013-2045-5

Pardey, P.G., J.M. Beddow, D.J. Kriticos, T.M. Hurley, R.F. Park, E. Duveiller, R.W. Sutherst, J.J. Burdon, and D. Hodson. 2013. Right-sizing stem-rust research. Science (Washington, DC) 340:147–148 10.1126/science.122970. doi:10.1126/science.122970

Poland, J.A., P.J. Brown, M.E. Sorrells, and J.-L. Jannink. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS ONE 7:e32253. doi:10.1371/journal.pone.0032253

Pretorius, Z.A., C.M. Bender, B. Visser, and T. Terefe. 2010. First report of a *Puccinia graminis* f. sp. *tritici* race virulent to the *Sr*24 and *Sr*31 wheat stem rust resistance genes in South Africa. Plant Dis. 94:784–784. doi:10.1094/PDIS-94-6-0784C

Pretorius, Z.A., R.P. Singh, W.W. Wagoire, and T.S. Payne. 2000. Detection of virulence to wheat stem rust resistance gene *Sr*31 in *Puccinia graminis*. f. sp. *tritici* in Uganda. Plant Dis. 84:203–203. doi:10.1094/PDIS.2000.84.2.203B

Pretorius, Z.A., L.J. Szabo, W.H.P. Boshoff, L. Herselman, and B. Visser. 2012. First report of a new TTKSF race of wheat stem rust (*Puccinia graminis* f. sp. *tritici*) in South Africa and Zimbabwe. Plant Dis. 96:590–590. doi:10.1094/PDIS-12-11-1027-PDN

R Core Team. 2013. R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.

Reif, J.C., P. Zhang, S. Dreisigacker, M.L. Warburton, M. van Ginkel, D. Hoisington, M. Bohn, and A.E. Melchinger. 2005. Wheat genetic diversity trends during domestication and breeding. Theor. Appl. Genet. 110:859–864. doi:10.1007/s00122-004-1881-8

Roelfs, A.P., and J.W. Martens. 1987. An international system of nomenclature for *Puccinia graminis* f. sp. *tritici*. Phytopathology 78:526–533. doi:10.1094/Phyto-78-526

Rouse, M.N., E.L. Olson, B.S. Gill, M.O. Pumphrey, and Y. Jin. 2011. Stem rust resistance in *Aegilops tauschii* germplasm. Crop Sci. 51:2074–2078. doi:10.2135/cropsci2010.12.0719

Simons, K., Z. Abate, S. Chao, W. Zhang, M. Rouse, Y. Jin, E. Elias, and J. Dubcovsky. 2011. Genetic mapping of stem rust resistance gene *Sr*13 in tetraploid wheat (*Triticum turgidum* ssp. *durum* L.). Theor. Appl. Genet. 122:649–658. doi:10.1007/ s00122-010-1444-0

Singh, R.P., D.P. Hodson, J. Huerta-Espino, Y. Jin, S. Bhavani, P. Njau et al. 2011. The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production. In: N.K. VanAlfen, et al., editors, Annual review of phytopathology. Vol. 49. Annu. Rev. Phytopathol., Palo Alto, CA. p. 465–481.

Singh, R.P., D.P. Hodson, J. Huerta-Espino, Y. Jin, P. Njau, R. Wanyera, S.A. Herrera-Foessel, and R.W. Ward. 2008. Will stem rust destroy the world's wheat crop? Advances in Agronomy, Vol. 98. Elsevier Academic Press, San Diego, CA. p. 271–309.

Singh, R.P., D.P. Hodson, Y. Jin, J. Huerta-Espino, M.G. Kinyua, R. Wanyera, P. Njau, and R.W. Ward. 2006. Current status, likely migration and strategies to mitigate the threat to wheat production from race Ug99 (TTKS) of stem rust pathogen. CAB Reviews: Perspectives in Agriculture, Veterinary Science. Nutrition and Natural Resources 1:13.

Singh, R.P., and S. Rajaram. 2002. Breeding for disease resistance in wheat. FAO Plant Production and Protection Series 30. FAO, Rome, Italy. doi: ISBN 92-5-104809-6.

Smale, M. 1996. Understanding global trends in the use of wheat diversity and international flows of wheat genetic resources. CIMMYT, Mexico, DF.

Sorrells, M.E., J.P. Gustafson, D. Somers, S. Chao, D. Benscher, G. Guedira-Brown et al. 2011. Reconstruction of the synthetic W7984 x Opata M85 wheat reference population. Genome 54:875–882. doi:10.1139/g11-054

Stakman, E.C., D.M. Stewart, and W.Q. Loegering. 1962. Identification of physiological races of *Puccinia graminis* var. *tritici*. USDA Publ. E617. USDA, Washington, DC.

Talbert, L.E., L.Y. Smith, and N.K. Blake. 1998. More than one origin of hexaploid wheat is indicated by sequence comparison of low-copy DNA. Genome 41:402–407. doi:10.1139/g98-037

The International Wheat Genome Sequencing Consortium. 2014. A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science (Washington, DC) 645:286–300 10.1126/science.1251788.

Visser, B., L. Herselman, R.F. Park, H. Karaoglu, C.M. Bender, and Z.A. Pretorius. 2010. Characterization of two new *Puccinia graminis* f. sp. *tritici* races within the Ug99 lineage in South Africa. Euphytica 179:119–127. doi:10.1007/s10681-010-0269-x

Warburton, M.L., J. Crossa, J. Franco, M. Kazi, R. Trethowan, S. Rajaram et al. 2006. Bringing wild relatives back into the family: Recovering genetic diversity in CIMMYT improved wheat germplasm. Euphytica 149:289–301. doi:10.1007/ s10681-005-9077-0

Zhang, P., S. Dreisigacker, A.E. Melchinger, J.C. Reif, A.M. Kazi, M. Ginkel, D. Hoisington, and M.L. Warburton. 2005. Quantifying novel sequence variation and selective advantage in synthetic hexaploid wheats and their backcross-derived lines using SSR markers. Mol. Breed. 15:1–10. doi:10.1007/s11032-004-1167-5

Chapter 2 - Mapping the wheat ionome

Abstract

Nutrient deficiency can lead to serious disorders in humans. Through biofortification, essential nutrients are increased in staple crops for improved quality of food and human health. Within HarvestPlus, a major focus is biofortifying bread wheat. Biofortification requires inexpensive high-throughput elemental profiling to study a plant's ionome. Ionomics is a very useful tool to obtain a snapshot of the functional status of any plant tissue. Here, we describe a hydroponics experiment followed by elemental profiling for bread wheat (*Triticum aestivum* L.) on shoots and roots. The ionomics data were used for quantitative trait loci (QTL) mapping. We identified 27 QTL for different elements in the shoots and two QTL in the roots. Four "hotspots" for nutrient accumulation in the shoots were identified on chromosomes 5AL, 5BL, 6DL and 7AL. The hotspot on 5AL controls Cd, Cu, Mn, S and Zn, while the other three hotspots control two elements each (Se and P, Ca and Sr, P and Zn). Furthermore, we identified a hotspot for Ni and Na uptake on chromosome 2BS in the roots. None of the hotspots showed antagonistic pleiotropic effects. Furthermore, we identified one possible novel QTL for Cd concentration in wheat shoots on chromosome 4BS. This comprehensive wheat ionomics study provides more supporting evidence of the usefulness of ionomics in wheat, confirms previously described QTL and identifies new QTL. The described QTL, particularly the hotspots, could be used for marker –assisted selection (MAS) in plant breeding for a wide range of applications.

Key words

quantitative trait loci, genotyping-by-sequencing, wheat ionome, ionomics, crop improvement, biofortification

Introduction

Deficiency in one or more essential mineral elements can lead to serious deficiency disorders in humans. Two billion people suffer from malnutrition or "hidden hunger" globally. Iron, vitamin A, iodine and zinc deficiencies are considered among the most significant. Iron and zinc deficiencies affect physical growth, development and cognitive functions. Over 30% of children in developing countries are stunted as a consequence of malnutrition and an estimated 190 million are affected by preventable early childhood blindness due to vitamin A deficiency (Gainhealth, 2015). Biofortification is the process of increasing essential nutrients in staple crops for improved quality of food and human health. The quantity and/or quality of food are improved during crop production by biological means such as plant breeding, biotechnology, or agronomy (WHO, 2015). An example of biofortification through biotechnology is vitamin A fortified golden rice (Burkhardt et al., 1997, Potrykus, 2001, Goldenrice, 2015). However, the deployment of golden rice has not been as successful because of controversy surrounding genetically modified organisms (GMOs).

Wheat is the major cereal crop consumed in many regions of the world and provides over 20% of all calories consumed globally (FAO, 2015, Shiferaw et al., 2013). Many regions of the world, notably in South Asia, where wheat is a major staple, also have high incidence and risk of malnutrition and micronutrient deficiency. Improving the quality of wheat through biofortification would benefit human health. The HarvestPlus program is currently supporting the biofortification several crops such as wheat (*Triticum aestivum* L.), rice (*Oryza sativa* L.), maize (*Zea mays* L.) and pearl millet (*Pennisetum glaucum* L.) to help reduce the micronutrient malnutrition in Africa and Asia (HarvestPlus, 2015). Within HarvestPlus, a major focus is biofortifying bread wheat for zinc (Zn) and iron (Fe) (Velu et al., 2011). In this context, ionomics

represents a powerful tool to identify potential genes and quantitative trait loci (QTL) responsible for nutrient uptake, transport and accumulation in plants.

Ionomics is defined as the "quantitative and simultaneous measurement of the elemental composition of living organisms and changes in this composition in response to physiological stimuli, developmental state and genetic modification" (Salt et al., 2008). Ionomics is one of the elements of functional genomics, which is composed of proteomics, metabolomics, transcriptomics and ionomics (Salt, 2004). However, the four "omics" are interconnected and are the sum of all expressed genes, proteins and metabolites in an organism. The aim of ionomics is to provide a snapshot of the status of mineral in various tissues and cells of any complex biological organism under different conditions. The different conditions can be based on genetics, developmental stage of the organism or biotic and abiotic (stress) factors (Salt et al., 2008). Studying the ionome of whole plants provides information about the mineral nutrition status and productivity of a plant, which ultimately influence the concentration of nutrients available for consumption.

The development of inductively coupled plasma technologies (ICP) and ability to simultaneously measure and analyze important elements in plants has enabled the high-throughput study of the ionome. Briefly, an aqueous sample is transformed into an aerosol by a nebulizer in the instrument. The aerosol is then brought into the plasma by an argon gas stream (carrier gas). Once in the plasma, all atoms in the same sample are ionized into singly charged positive ions. At this stage, the ionized atoms are detected by ICP (through optical emission spectrometry or mass spectrometry) (Salt et al., 2008).

Ionomics studies to date were performed in *Arabidposis thaliana* (L.) Heynh. on shoots (Baxter et al., 2009, Baxter et al., 2008, Lahner et al., 2003) and seeds (Vreugdenhil et al., 2004),

soybean seeds (*Glycine max* L. Merr.) (Ziegler et al., 2013), *Lotus japonicus* (Regel) K. Larsen shoots (Chen et al., 2009), rice and pearl millet grain. One study using X-ray fluorescence spectrometry measuring the concentration of Zn, Fe and Se was conducted on whole kernels of wheat (Paltridge et al., 2012).

Here, we analyze the wheat ionome of shoots and roots in the newly reconstructed biparental hexaploid SynOpDH mapping population (Sorrells et al., 2011) under stable and controlled conditions in a hydroponic growth chamber experiment by surveying all mineral elements of biological relevance in wheat shoots and roots.

Materials and Methods

Mapping population

The Synthetic W7984 x Opata M85 (Altar84/*Aegilops tauschii* (219) CIGM86.940) / Opata M85) biparental mapping population developed in the late 1980s served as a resource for multiple QTL mapping studies in the wheat community (Sorrells et al., 2011). Two synthetic wheat reference populations with the same pedigree were recently reconstructed by Sorrells et al., (2011). One population consists of double haploids (named SynOpDH) and the other of recombinant inbred lines (named SynOpRIL). 154 SynOpDH were included in the experiment.

Genotypic data

The SynOpDH mapping population was genotyped using a two-enzyme genotyping-by-sequencing (GBS) approach by Poland et al., (2012). The previously constructed map consisted of 1,485 single nucleotide polymorphisms (SNPs). Detailed information about data filtering, SNP calling and map construction can be found in Poland et al. (2012). The female parental alleles from Synthetic W7984 were coded as “A”, the male parental alleles from OpataM85 as “B”. Dunckel et al. (2015) verified the original assignment of the GBS markers to chromosomes

and chromosome arms in Poland et al., (2012) by aligning the tags to the recently published draft sequence of the wheat genome (The International Wheat Genome Consortium, 2014).

Experimental design

The experiment was performed in a growth chamber as a randomized complete block design (RCBD) with three replicates of each line spread across four tubs. The data was analyzed using JMP Pro 11 Statistical Software (JMP®, 1989-2015). The model used to calculate BLUEs was $y_{ijk} = \mu + g_i + t_j + e_{ijk}$ where y_{ijk} is the analyzed element, g_i is the fixed effect for each genotype, t_j is the fixed effect of the jth tub, and e_{ijk} is the random error with N(0, σ_e^2). A random effects model was applied to estimate variance components for heritability calculations of each element. Heritability on an entry means basis was calculated as $H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_e^2}{t}}$ where σ_g^2 is the genotypic variance, *t* is the number of tubs, *r* is the number of replicates, and σ_e^2 is the error variance (Bernardo, 2010).

The plants were grown hydroponically under optimal conditions with a 12h day/night setting at 25/20°C. The nutrient solution (Table 2.1) was developed by Cobb et al. (2015) and was designed to deliver adequate amounts of micro- and macronutrients as well as sub-toxic amounts of heavy metals/metalloids of interest. The pH was held constant at pH 6.0 by adjusting it every second day. The plants were grown for 3 weeks before harvesting the seedling and separating roots and shoots the same day. Roots and shoots were rinsed with deionized water before harvesting and drying to avoid contamination from the nutrient solution. The tissue was dried in a 60°C oven for seven days. Thereafter, the dried tissue was digested in batches using a 60/40 perchloric/nitric acid solution, diluted and analyzed by Inductively Coupled Plasma -

Optical Emission Spectrometry (ICP-OES) (Fassel, 1974, Hou and Jones, 2000, Salt et al., 2008).

All measurements were taken separately for root and shoot tissue. The ICP-OES takes three independent measurements of a small sample and reports the mean. The precision of the measurements is defined through the average percent relative standard deviation (%RSD, = associated variance of the three samples) and indicates how consistent any element in the analysis is measured by ICP-OES over repeated measurements (Salt et al., 2008). Precision is essential as our goal is to detect differences in the ionome due to genotype and not experimental errors. We excluded measurements with %RSD higher than 15%. The accuracy of the analysis is defined through the lower limit of detection of every element. The average %RSD and lower limit of detection of all elements are shown in Table 2.2. Adjusting for the dilution volume and biomass of the analyzed sample normalizes raw ICP-OES data. The final data is expressed in terms of μg of element per gram dry weight tissue. Elements analyzed are aluminum (Al), arsenic (As), barium (Ba), boron (B), cadmium (Cd), calcium (Ca), chromium (Cr), cobalt (Co), copper (Cu), iron (Fe), lead (Pb), magnesium (Mg), manganese (Mn), molybdenum (Mo), nickel (Ni), phosphorus (P), potassium (K), selenium (Se), silicon (Si), sodium (Na), strontium (Sr), sulfur (S), zinc (Zn).

QTL analysis

Quantitative trait loci (QTL) mapping was performed separately for ion concentrations in roots and shoots in the R software environment (R Core Team, 2014) using R-package R/qtl (Broman et al., 2003). The same methods were applied as described in Dunckel et al., (2015). Briefly, QTL were mapped using Single Interval Mapping (SIM) and Composite Interval Mapping (CIM). The most significant markers were identified through stepwise regression. CIM

was implemented applying a Haley-Knott regression using forward selection of marker covariates and a window size of 10cM for all traits. Multiple QTL Mapping (MQM) was used to confirm identified QTL, refine their position and obtain estimated QTL effects phenotypic variance components (Arends et al., 2010). Additional QTL not identified through SIM and CIM were mapped through MQM. The allelic state of the markers with the highest LOD score at each QTL was used to represent the allelic state of the QTL. The genome-wide logarithm of the odds value (LOD) for declaring a QTL was determined by 1,000 permutations. The parental alleles for Synthetic W7984 and Opata M85 were coded as -1 and 1 respectively (Broman and Saunak, 2009).

Results

To test the repeatability of the experiment, we calculated heritability on an entry mean basis separately for each element and for roots and shoots (Table 2.3). Dry weight biomass (dw_mg) of roots and shoots was included in the analysis and subsequent QTL mapping. There was contrasting difference in heritability among the elements and tissues. For example, the heritability is high for Cd (0.75), P (0.73) and Mn (0.67) measured in the shoots and lower for roots (0.25, 0.34 and 0.10 respectively). The phenotypic distribution follows a normal distribution for elements measured from the shoots and is slightly skewed to the left for the root data (Figure 2.7 – Figure 2.14). We observed highly significant correlations of several elements indicating clusters and connected networks of traits (Figure 2.1 and Figure 2.2).

We identified 27 QTL for different elements in the shoots and two QTL in the roots. QTL mapped initially through SIM and CIM were confirmed and their position refined applying MQM. Furthermore, several additional QTL were identified through MQM. A summary of all QTL including LOD, estimated phenotypic variance, and allele effects is available in Table 2.5

and Table 2.6. The LOD profiles in Figure 2.3 – Figure 2.6 are based on LOD scores obtained through CIM.

We identified one QTL each for Ni and Na in the roots (Figure 2.3). The QTL for Na and Ni mapped to GBS marker synopGBS894 on chromosome 2BS at 78.7cM. MQM was used to obtain estimates on phenotypic variance, allele effects and refined position of the QTL. The QTL for Na explained 10.74% of the phenotypic variance and the QTL for Ni 9.24% respectively (Table 2.6). The estimated allele effects show that the allele conferred by the synthetic parent (allele A) increases the accumulation of Na and Ni in the roots.

We were able to map 27 QTL for As, Ba, Cd, Ca, Cu, Mn, Mo, P, Se, Sr, S and Zn in the shoots (Table 2.6, Figure 2.4 – Figure 2.6). Two QTL for biomass (dw_mg) were identified on chromosomes 1AL and 2BL explaining 17.89% of the estimated phenotypic variance. We were interested in mapping this trait to assure no QTL for any other element was confounded with biomass. No other QTL were mapped at the same marker positions.

Some QTL for different elements were identified in close proximity on the same chromosome (Table 2.6). QTL for P and Zn were mapped on chromosome 7AL at 85.5cM and 88.1cM explaining 7.13% and 8.46% of the estimated phenotypic variance, respectively. Chromosome 6DL harbors QTL for Ca and Sr at 120.0cM and 120.3cM, which contribute 16.50% and 13.09% of the estimated phenotypic variance, respectively. Elements P, Mo and Se have each a QTL on 5BL at 92.0cM, 84.8cM and 88.3cM. The QTL mapped for P explains 16.80 % of the estimated variance, while the QTL for Mo and Se 10.28 % and 12.65 % respectively. Furthermore, four QTL for Cd, Cu, Mn, and Zn mapped to GBS marker synopGBS429 on chromosome 5AL at 127.9cM. The QTL explained 9.84 %, 17.56 %, 27.1%, and 14.13% of the estimated phenotypic variance for Cd, Cu, Mn, and Zn, respectively. A QTL

for S was mapped in very close proximity on the same chromosome at 125.4cM explaining 8.90% of the estimated phenotypic variance.

We mapped seven QTL for P in the shoots. Two were identified by SIM and CIM and five QTL above the significance threshold were added by MQM (Table 2.6). The model with 7 QTL was the best fit and explains 66.62 % of the estimated phenotypic variance. The estimated allele effects show that alleles conferred by the synesthetic parent increased the P concentration in shoot tissue for QTL on chromosomes 1BL, 3DL, 4AS and 7AL, while having a negative effect on P concentration at the other QTL.

We mapped three QTL for Ca on chromosomes 6AL, 6DL and 7BS explaining 35.72% of the estimated phenotypic variance. The alleles conferred by the synthetic parent increased the Ca accumulation at two QTL, while reducing it at the QTL on chromosome 6AL. Two QTL were mapped for Ba, Cd, Cu, Mn and Zn. The QTL for Ba on chromosomes 4AS and 5BS explain 24.18 % of the estimated phenotypic variance. The two QTL for Cd mapped on chromosomes 4BS and 5AL and explain 42.78% of the estimated phenotypic variance, while the QTL on chromosomes 5AL and 7DS for Mn explain 25.89%. Furthermore, two QTL for Zn were identified on chromosomes 5AL and 7AL explaining 22.76% of the estimated variance, and two QTL for Cu on 5AL and 5BS explaining 25.4%. The estimated allelic effects indicate that the synthetic parent decreased the accumulation of Mn at both QTL and increased it for the QTL mapped for Zn on chromosome 7AL.

Discussion

We mapped 27 QTL in wheat shoots for As, Ba, Cd, Ca, Cu, Mn, Mo, P, Se, Sr, S and Zn and two QTL in wheat roots for Ni and Na in the newly reconstructed synthetic hexaploid

W7984 × Opata M85 DH wheat reference population. The estimated allele effects show that both parents contributed alleles affecting the mineral nutrient concentration in both tissues.

Pleotropic loci

We identified four loci with co-localizing QTL, or hotspots, in shoots and one hotspot in roots. The elements in each hotspot were highly correlated (Table 2.4, Figure 2.1 and Figure 2.2). None of them showed antagonistic pleiotropy, meaning that the estimated allele effects were always positively correlated. We did not identify any co-localizing where a given allele hotpot increased the concentration of one element but decreased concentration of another.

Hotspot for Cd, Cu, Mn, S and Zn

A QTL identified on chromosome 5AL at 127.9cM is a "hotpot" controlling the concentration of the four elements Cd, Cu, Mn, and Zn in wheat shoots. Our correlation analysis shows significant positive correlations of all elements (Figure 2.1 and Figure 2.2). Furthermore, a QTL for S was mapped in very close proximity to 125.4cM and is likely the same QTL. This hotspot is linked to increased Cd concentration in shoots. Cd is a toxic heavy metal and thought to share the nutrient uptake pathways with Zn. Increasing the Zn concentration through biofortification could possibly increase the Cd concentration in wheat as well (Palmgren et al., 2008). Furthermore, proteins involved in Cd uptake and translocation are related to Fe, Zn and Mn transport (Nakanishi et al., 2006; Sasaki et al., 2012). Guttieri et al. (2015) showed that the concentration of Cd in grain could be predicted by the Cd concentration in plant tissue at anthesis. This was not true for Zn. Furthermore, they showed that grain Cd and Zn are not highly correlated and concluded that breeders should be able to breed wheat with low Cd concentration without affecting Zn concentration in the grain.

Cadmium

Cd is a heavy metal, toxic and a naturally occurring non-essential element. The Cd concentration in some soils has increased through agricultural use. Roots take up Cd from contaminated soils and transport it to other parts of wheat, mainly the shoots and grain. This leads to higher Cd concentrations in the grain and Cd contamination of food (Clemens et al., 2013, Grant et al., 1998, McLaughlin et al., 1999). Harris et al. (2013) studied the changes in whole-plant Cd accumulation and allocation to different tissues during grain filling in durum wheat (*Triticum turgidum* L.). No difference of Cd uptake between high- and low-Cd lines was found. However, there were differences in root-to-shoot translocation indicating genetic variation underlying the Cd accumulation in the grain. Furthermore, they found a direct pathway from root-to-grain via xylem-to-phloem.

Guttieri et al. (2015) reported QTL for Cd accumulation in bread wheat on chromosome 5AL at 88.70 – 89.95cM. We identified the hotspot at 127.9cM and conclude this is likely not the same QTL. Furthermore, we identified a second QTL for Cd concentration in shoots on chromosome 4BS at 32.9cM explaining 33.19% of the estimated phenotypic variance. This QTL has not been reported in other studies and could be novel. The allele contributed by the synthetic parent decreased the Cd concentration for the QTL on chromosome 5AL. However, the synthetic allele at this QTL is also decreased concentrations of desirable elements Cu, Mn, S and Zn and would likely not be useful for breeding. The opposite was observed for the QTL on 4BS, the allele conferred by the synthetic parent increased Cd concentration in the shoots and could be tracked with markers to avoid during selection.

Zinc

One QTL for Zn was mapped to the hotspot described above and a second QTL on chromosome 7AL at 88.1cM. These QTL are interesting relative to the micronutrient objectives

of HarvestPlus. The synthetic parent contributed the allele for increased Zn concentration for the QTL on chromosome 7AL and decreased Zn concentration on 5AL. Peleg et al. (2009) reported QTL for Zn concentration in the grain of durum wheat on chromosomes 5A at 25.8 ± 22.0 cM and 7AL at 65.8 ± 4.6 cM. Two QTL for increased Zn concentration in the grain were identified on chromosome 7A in a *T. boeoeticum*/*T. monococcum* RIL mapping population (Tiwari et al, 2009). One QTL was mapped at 72.6 cM and one at 153.8 cM. Aligning both maps showed that the previously described QTL at 72.6cM and the QTL identified here are probably not the same. Other studies identified multiple QTL for Zn accumulation in shoots and grain of bread wheat on chromosomes 2A, 3A, 4A, 4B, 4D, 5A, 6A, 6B, 7A, 7B and 7D but none matched the QTL we identified (Bálint et al., 2007, Genc et al., 2009, Xu et al., 2012). Therefore, these might be novel QTL for Zn accumulation in wheat shoots. The QTL on 7AL contributed by the synthetic parent could be particularly interesting for marker-assisted selection in the context of biofortification and breeding for wheat with increased Zn concentration.

Copper, Manganese and Sulfur

QTL for copper accumulation were previously reported on chromosomes 1AL, 1BL, 2A, 2DS, 3B, 4AS, 4B, 5A, 5BL, 5DL, 6A, 6B, 7A, 7B and 7DS (Bálint et al., 2007, Peleg et al. 2009). Here, we only report one QTL for Cu on 5AL at 127.9cM. The QTL identified on 5A was previously mapped at 12.6 ± 11.2 cM does not match the location of the QTL reported here. QTL for Mn in grain and shoots were previously identified on chromosomes 2B, 3BL and 7B (Bálint et al., 2007, Peleg et al. 2009). In our study, we identified one QTL for Mn on chromosomes 5AL and 7DS. Both QTL might be novel. Peleg et al. (2009) reported QTL for macronutrient S on chromosomes 1A, 2A, 3A, 4A, 5A, 5B, 6B, 7A and 7B. They reported a QTL on

chromosome 5A at 95.3 ± 11.5cM, which is located in close proximity to the QTL on 5AL we reported here.

The usefulness of the hotspot on chromosome 5AL for MAS is questionable, because the alleles contributed by the synthetic parent decrease the concentration of all desired elements. However, this hotspot controls the concentration of multiple elements in wheat shoots and we conclude that it might be a functional gene worth studying.

Hotspot for P and Zn

The QTL on chromosome 7AL at 85.5 cM for P and 88.1 cM Zn for concentration are most likely the same QTL and represent the second hotspot. This QTL controls the concentration of P and Zn in the shoots. The estimated allelic effect shows that the synthetic parent contributed alleles for increased P and Zn. This hotspot could be of particular interested. We are interested in increased phosphorous uptake (PUP), phosphorous use efficiency (PUE) and Zn in wheat. Increased PUP could reduce As uptake, increased PUE reduce the P fertilizer requirement, and increased Zn is desirable for biofortification to increasing human health.

Phosphorous

Modern agriculture depends on phosphorus derived from phosphate rock. This is a non-renewable resource and may be depleted globally in only 50 to 100 years (Cordell et al., 2009). Therefore, it is important to breed crops with high PUP and PUE. We identified seven QTL in shoots for P on chromosomes 1BL, 3DL, 4AS, 5A, 5BL and 7AL (Table 2.6, Figure 2.6). The synthetic parent contributed the QTL on 1BL, 3DL, 4AS and 7AL. Weidong et al., (2001) studied the PUE in the same mapping population in a hydroponic experiment. They mapped QTL related to PUP under high and low P conditions in shoots and whole plant tissue (roots and shoots) and reported five QTL on chromosomes 1B, 2D, 3B, 5A and 6D in the shoots.

Furthermore, they identified four QTL on chromosomes 2B, 2D, 5A and 7A from whole plant tissue. The QTL they identified on 1B is located around 40cM and Opata M85 conferred the positive allele effect. We mapped our QTL at 62.2cM but identified the synthetic parent as the donor of the allele with positive effect. The QTL they reported on chromosome 5A in shoots is located between 30 – 45cM and the QTL form the whole plant at 45cM. These QTL might be the same, mainly because of their close proximity and because the synthetic parent contributed them both. We mapped our QTL on chromosome 5AL at 87.7cM and identified Opata M85 as the donor parent. Based on the estimated allele effects they are probably different QTL. However, the QTL Weidong et al., (2001) identified on chromosome 7A could be the same QTL we identified on 7AL at 85.5cM. In both cases, the synthetic parent conferred the allele. We did not identify any QTL on chromosomes 2D, 3B or 6D but mapped four more QTL on chromosomes 3DL, 4AS and 5BL.

In a more recent study Su et al., (2009) mapped seven reproducible QTL regulating PUP and six regulating PUE under high and low P conditions across multiple field experiments in a DH mapping population developed from two Chinese winter wheat varieties. Furthermore, they mapped QTL for other important agronomic traits and found that some were highly correlated with QTL for PUP and PUE. This suggests that by improving PUP and PUE, other important agronomic traits can be improved. It has to be noted that increasing PUP and PUE can be difficult due to their negative correlation. However, Su et al. (2009) found a few QTL positively linked for PUP and PUE that could be used in plant breeding through MAS. One of them was mapped on chromosomes 5A. Comparing the position of their QTL on 5A with ours shows that these are most likely the same QTL (GrainGenes database, Carollo et al. (2005)). Su et al., (2009) identified this QTL as one of their candidates for PUP and PUE and linked the QTL to

tiller number and shoot dry weight as well. They mapped several QTL we were not able to map, however, the QTL we report on 3DL, 4AS and 5BL might be novel.

Arsenic

Arsenic is a toxic heavy metal and the accumulation in wheat is one way of human exposure to As from the environment. The concern is mostly for wheat growing on As contaminated soil (Zhao et al., 2010). As is taken up by plants as arsenate across the plasma membrane via the phosphate (Pi) transport system (Dixon, 1996). P and As compete for plant resources. Pigna et al. (2010) showed that increased levels of P in the plant tissue reduced the toxicity of As by inhibiting the accumulation of As in wheat shoots and grain. Furthermore, Zhu et al. (2006) found that the As uptake varies across lines in a biparental wheat population, and even decreased over time for some lines. This has implications on the cropping system but also on plant breeding. Applying MAS would enable plant breeders to identify and select lines with QTL increasing PUP and PUE and reducing the As accumulation in wheat.

To reduce As accumulation in wheat, lines with low As uptake and/or accumulation and the identification of alleles reducing As uptake are of interest. We mapped one QTL for As on chromosome 3B at 96.6cM explaining 13.27 % of the estimated phenotypic variance. The allele reducing the As concentration was contributed by the synthetic parent. A recent study in maize showed that As accumulation in leaves, stems and kernels is controlled by different QTL and molecular mechanisms (Ding et al., 2011). Lines with low As concentration in the kernels and high As concentration in the stems and leaves could be useful for phytoremediation of As contaminated soils. However, we did not map any QTL for As in wheat roots and are not able to make the same observation in wheat.

Hotspot for Ca and Sr

Another hotspot was identified on chromosome 6DL. We mapped a QTL for macronutrient Ca at 120.0cM and one for Sr at 120.3 cM and conclude that this is the same QTL as well. Furthermore, we identified two more QTL for Ca on chromosomes 6AL and 7BS. Other QTL for Ca have been reported on chromosomes 1A, 2B, 4A, 4B, 5B, 6B and 7B (Peleg et al., 2009). We located the QTL on 7BS at 10.8cM, while Peleg at al. (2009) reported their QTL at 23.0 ± 7.8cM.

Hotspot for Mo, Se and P

The QTL mapped for Mo, Se, and P on chromosome 5BL are located at 84.8cM, 88.3cM and 92.0cM, respectively, and could be the same QTL. Se is essential for animals and humans, and can lead to deficiency related disease but has adverse health effects if the Se intake is chronically too high. However, the synthetic parent contributed the allele reducing the Mo, Se and P concentration in wheat shoots. Therefore, this QTL will probably not find any application in MAS.

Conclusion

With this study, we provide new information and insight into the wheat ionome. We have measured 23 elements in two different plant tissues in high-throughput fashion and mapped multiple QTL. Several hotspots controlling the concentration of multiple elements were identified here. Here, we analyzed the elemental composition of shoot and root tissue, which allows to draw some conclusions regarding nutrient translocation. However, experiments that are more specific might be needed, especially to assess nutrient sequestration to the wheat grain. While some of these hotspots might find their application in plant breeding through MAS, others may be more useful to study their functionality. Some hotspots might be functional genes, however this will have to be confirmed. The next step is to better understand the overlap between

the different networks, identify possible markers for MAS and develop gene hypotheses based on comparative genetics.

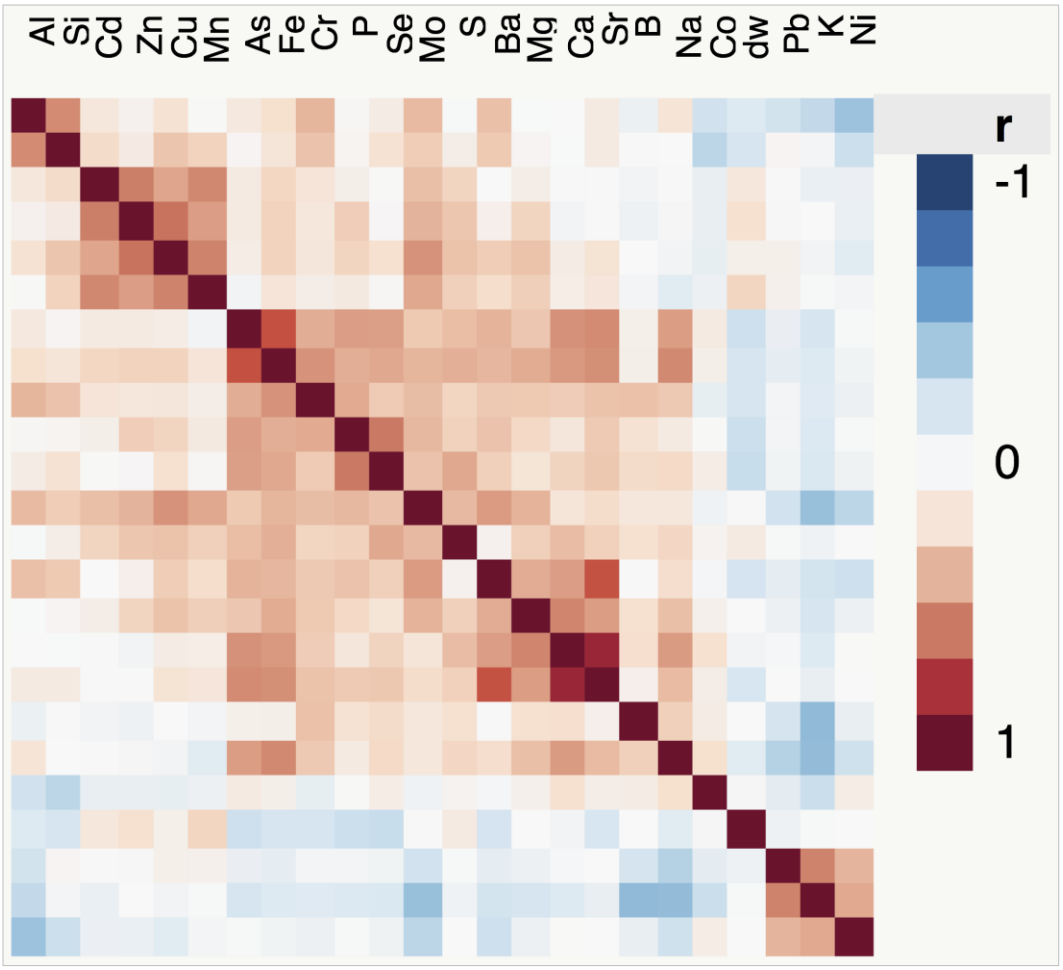

Figure 2.1 Correlation analysis shoots

Red indicates strong positive correlation of elements, white no correlation and blue strong negative correlation. Ni, K, Pb are positively correlated with each other while not or negatively with all other elements. Cd, Cu, Mn, Zn showed positive correlations indicating a cluster of elements. Several other traits are positively correlated (i.e. Ca and Sr, P and Se, and As and Fe).

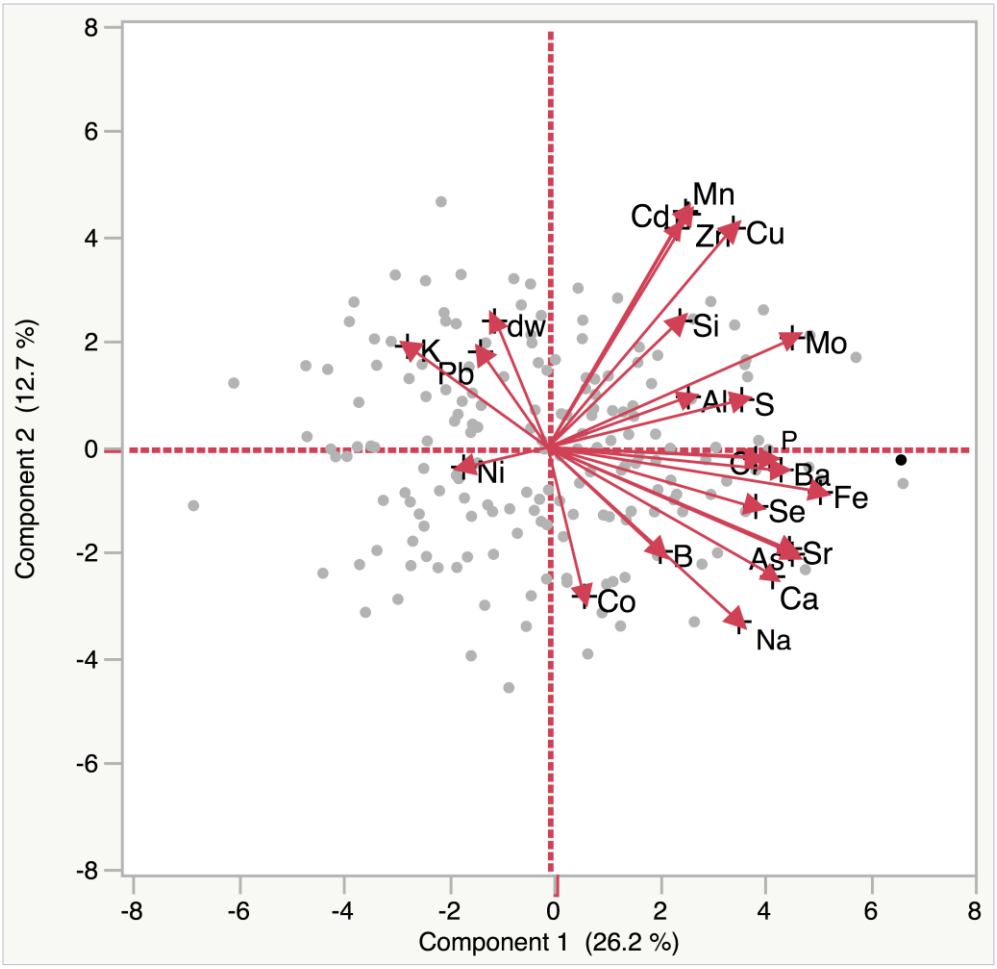

Figure 2.2 PCA biplot based on the correlations of elements in shoots

The PCA biplot is based on the correlations of all elements in the shoots. The two major principal components accounted for 38.9 % of the variation in the data set. PC1 explained 26.2% and was positively loaded for most elements while PC2 explained 12.7 % of the variation. Several correlated elements such as for example Cd, Cu, Mn, Zn and Si cluster together.

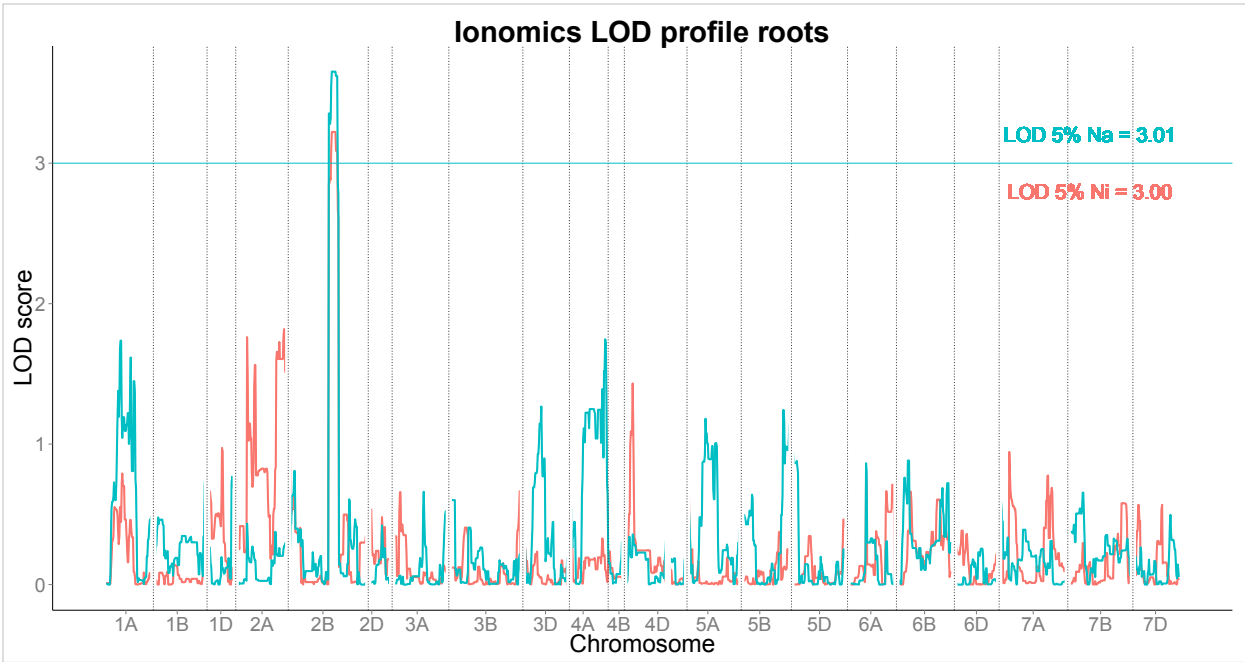

Figure 2.3 Logarithm of the odds (LOD) profile roots

The logarithm of the odds (LOD) profile for the root tissue shows one QTL for Ni (orange) and Na (turquoise) on chromosomes 2BS (LOD 5% =3.00 and 3.01 respectively).

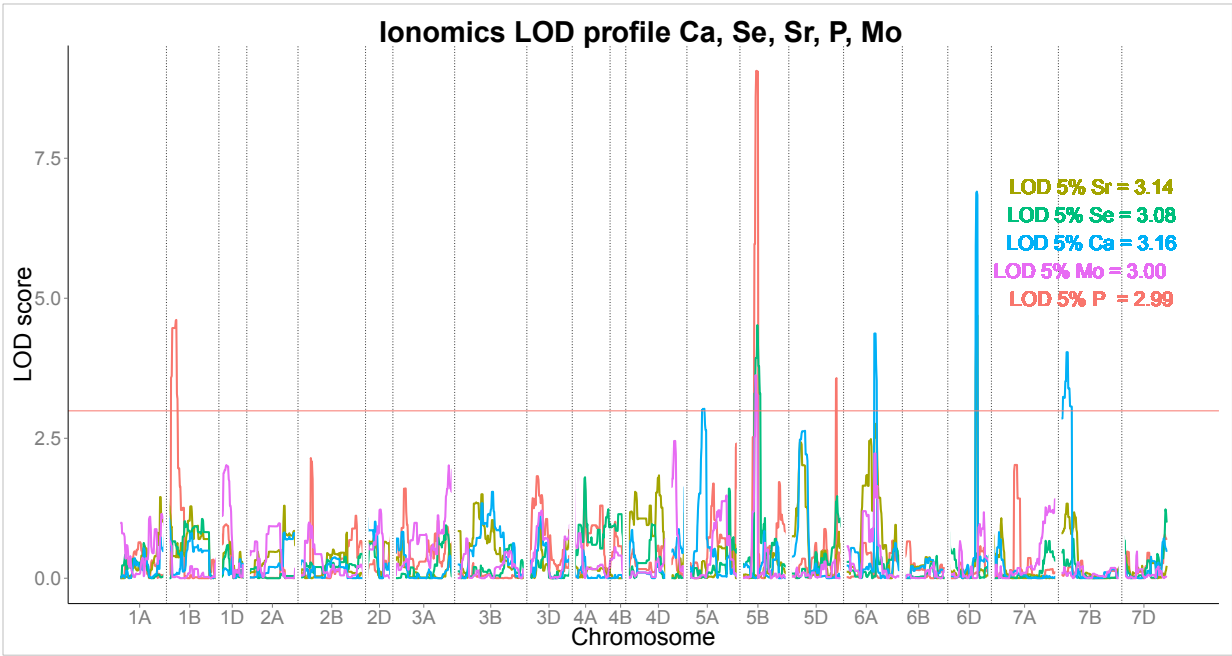

Figure 2.4 Logarithm of the odds (LOD) profile for Ca, Se, Sr, Mo and P in the shoots

The LOD profile for the shoot tissue shows two QTL for Ca (blue) on chromosomes 6AL, 6DL and 7BS (LOD 5% = 3.16), one for Se (green) on chromosome 5BL (LOD 5% = 3.08), one QTL for Sr (olive) on chromosome 6DL (LOD 5% = 3.14), three for P (red) on 1BL, 5BL and 5DL (LOD 5% =2.99) and one QTL for Mo on 5BL (LOD 5% = 3.00). Furthermore, MQM identified four more QTL for P on chromosomes 4AS, 3DL, 5A and 7AL.

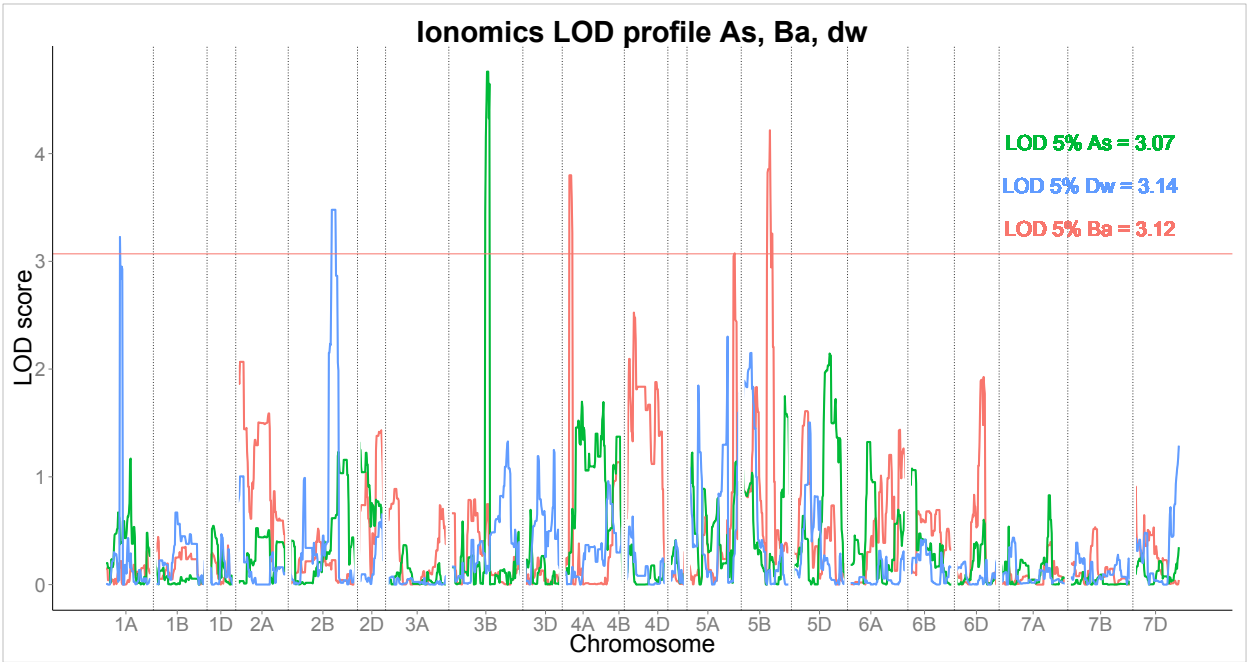

Figure 2.5 Logarithm of the odds (LOD) profile for As, Ba and biomass (dw) in the shoots

The LOD profile for the shoot tissue shows one QTL for As (green) on chromosome 3B (LOD 5% = 3.07), two for biomass (blue) on chromosomes 1AL and 2BL (LOD 5% = 3.14), two QTL for Ba (red) on chromosomes 4AS and 5BS (LOD 5% = 3.12).

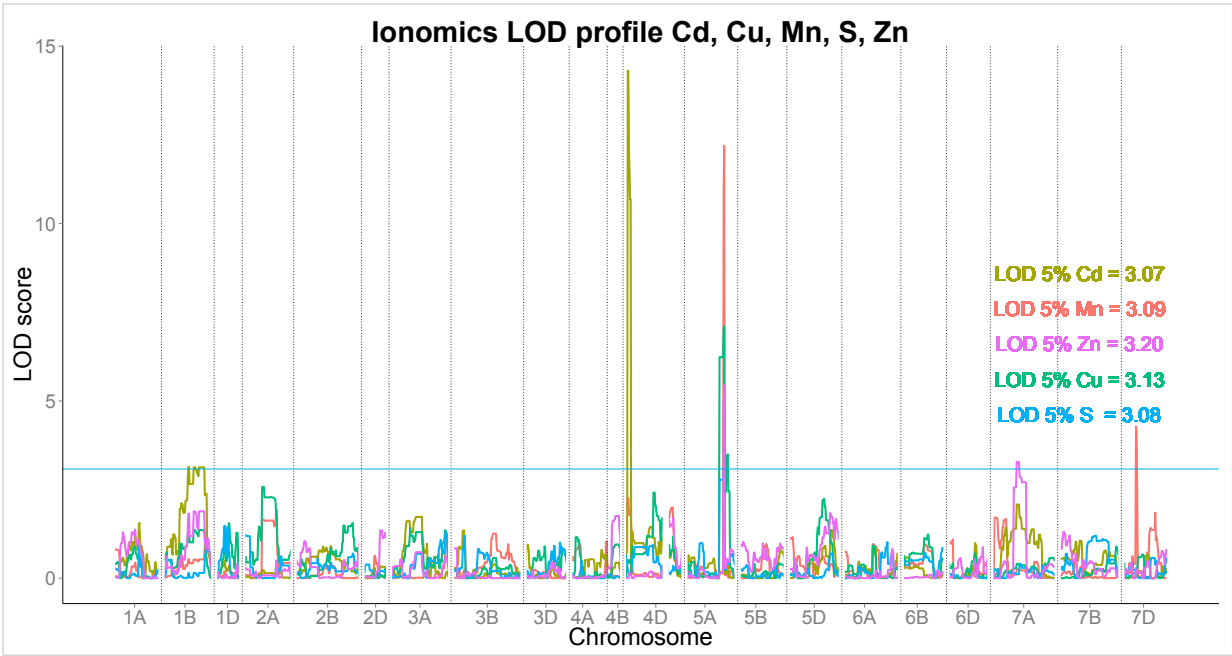

Figure 2.6 Logarithm of the odds (LOD) profile for Cd, Cu, Mn, S and Zn in the shoots

The LOD profile shows one QTL hotspot for Cd (olive), Cu (green), Mn (red), and Zn (pink) on chromosome 5AL at 127.9cM. A QTL for S (blue) was mapped in very close proximity at 125.4cM. The LOD significance thresholds at $\alpha = 0.05$ are 3.15, 3.08 and 3.20 for Cu, S and Zn respectively, and 3.09 and 3.07 for Mn and Cd. Another QTL for Zn was identified on chromosome 7AL, one QTL for Cd was mapped on chromosome 4BS and one for Mn on 7D.

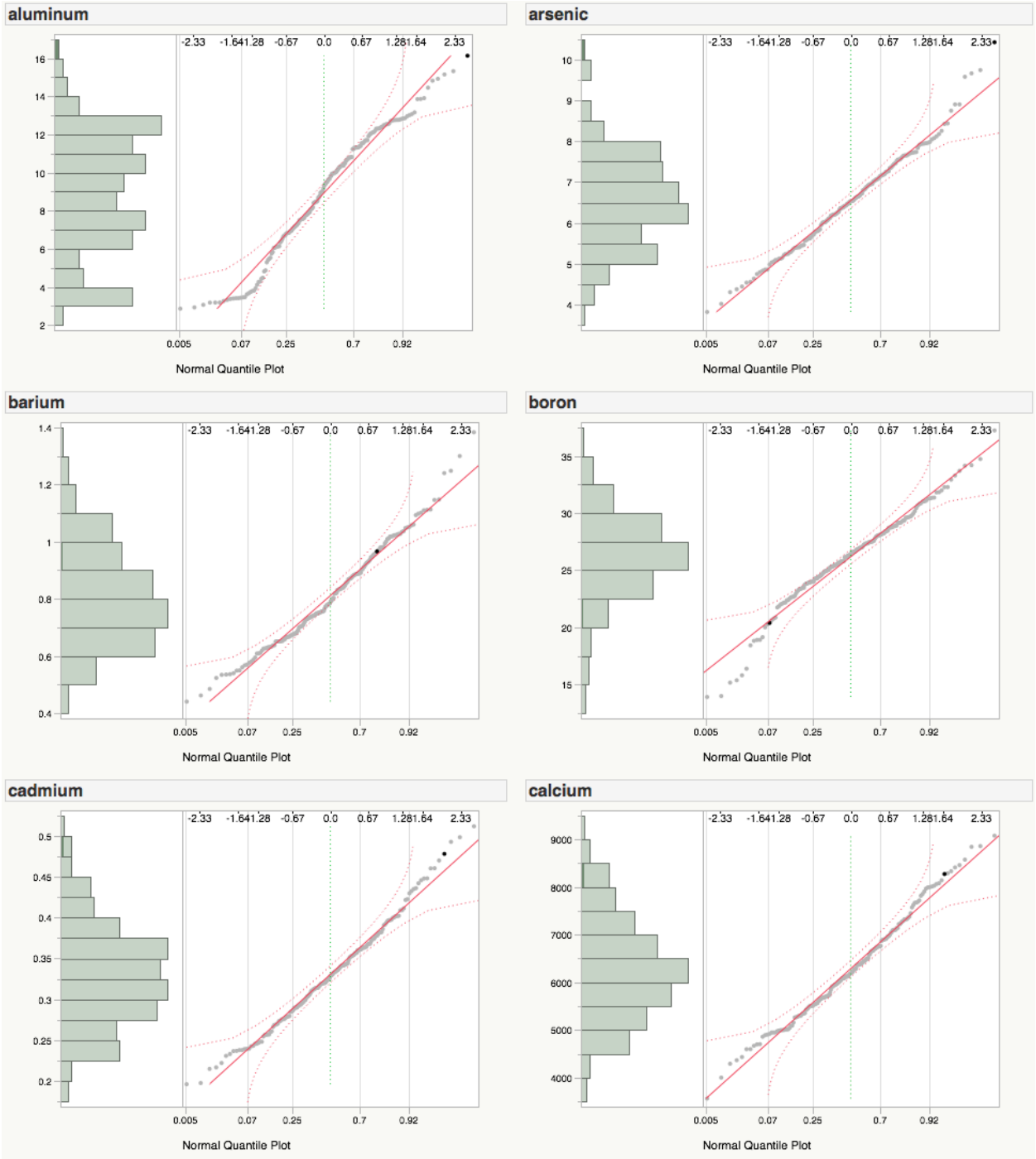

Figure 2.7 Phenotypic distributions of BLUEs of Al, As, Ba, B, Cd and Ca shoot tissue

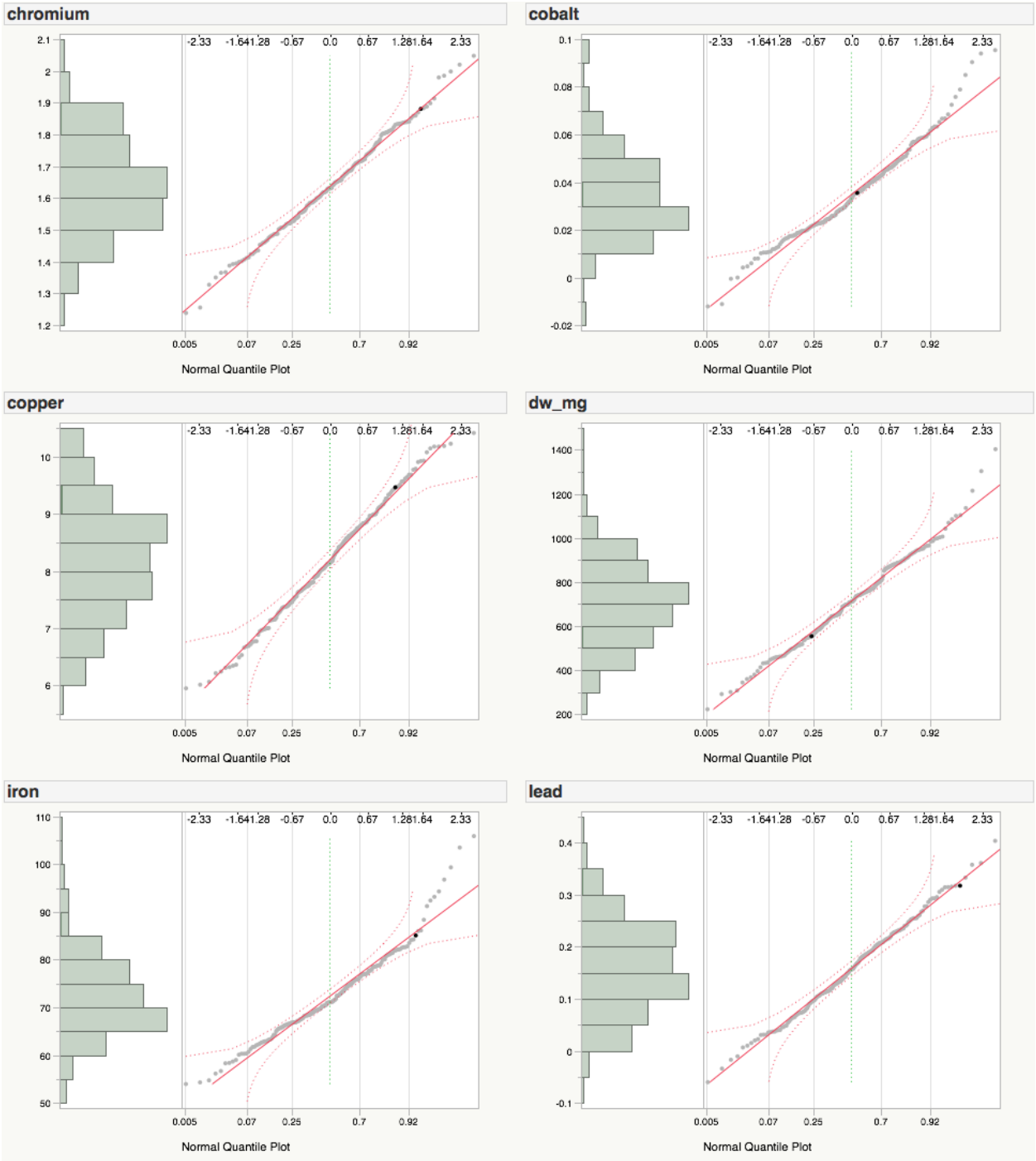

Figure 2.8 Phenotypic distributions of BLUES of Cr, Co, Cu, Fe, Pb and dry weight (dw_mg) shoot tissue

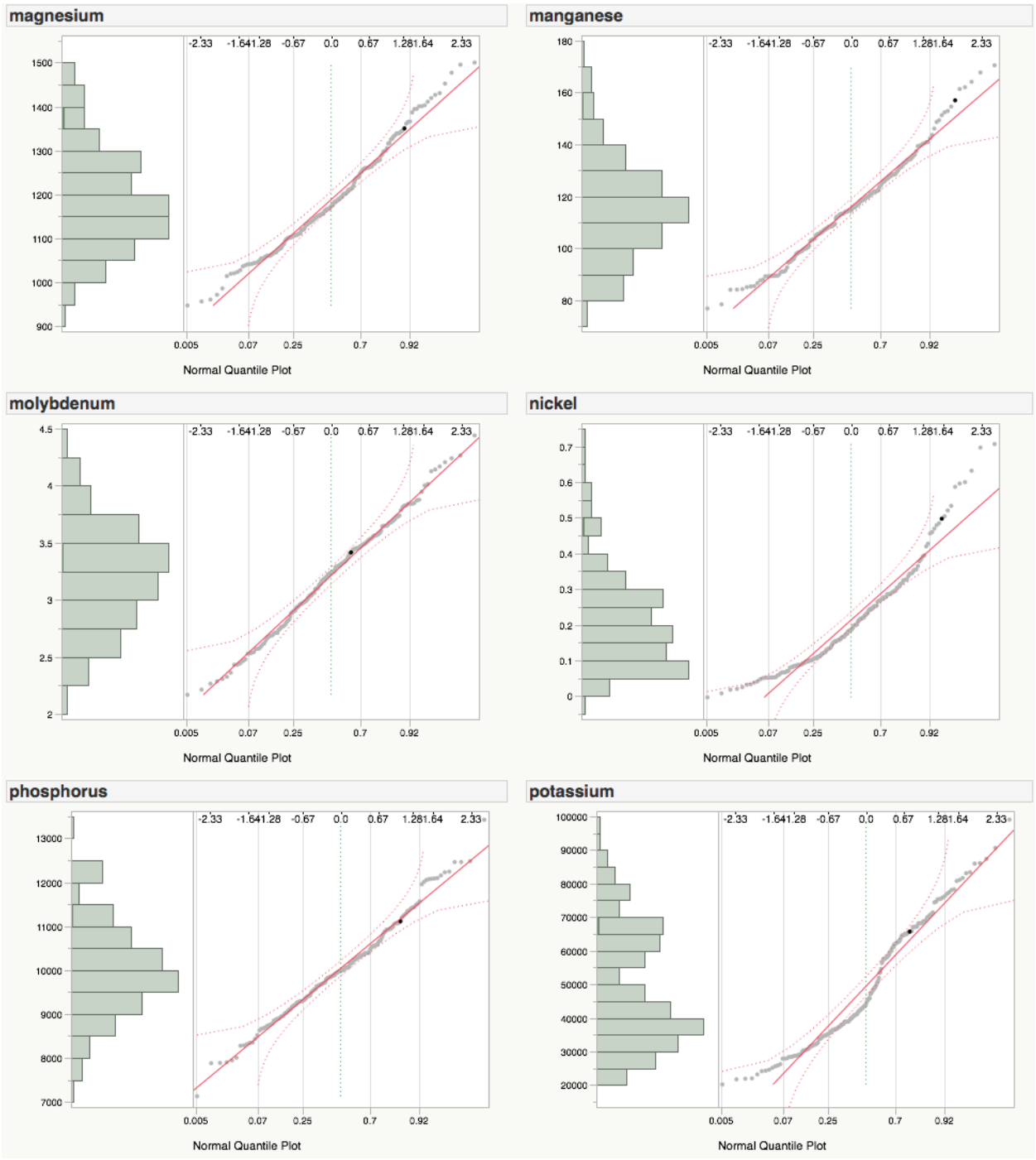

Figure 2.9 Phenotypic distributions of BLUEs of Mg, Mn, Mo, Ni, P and K shoot tissue

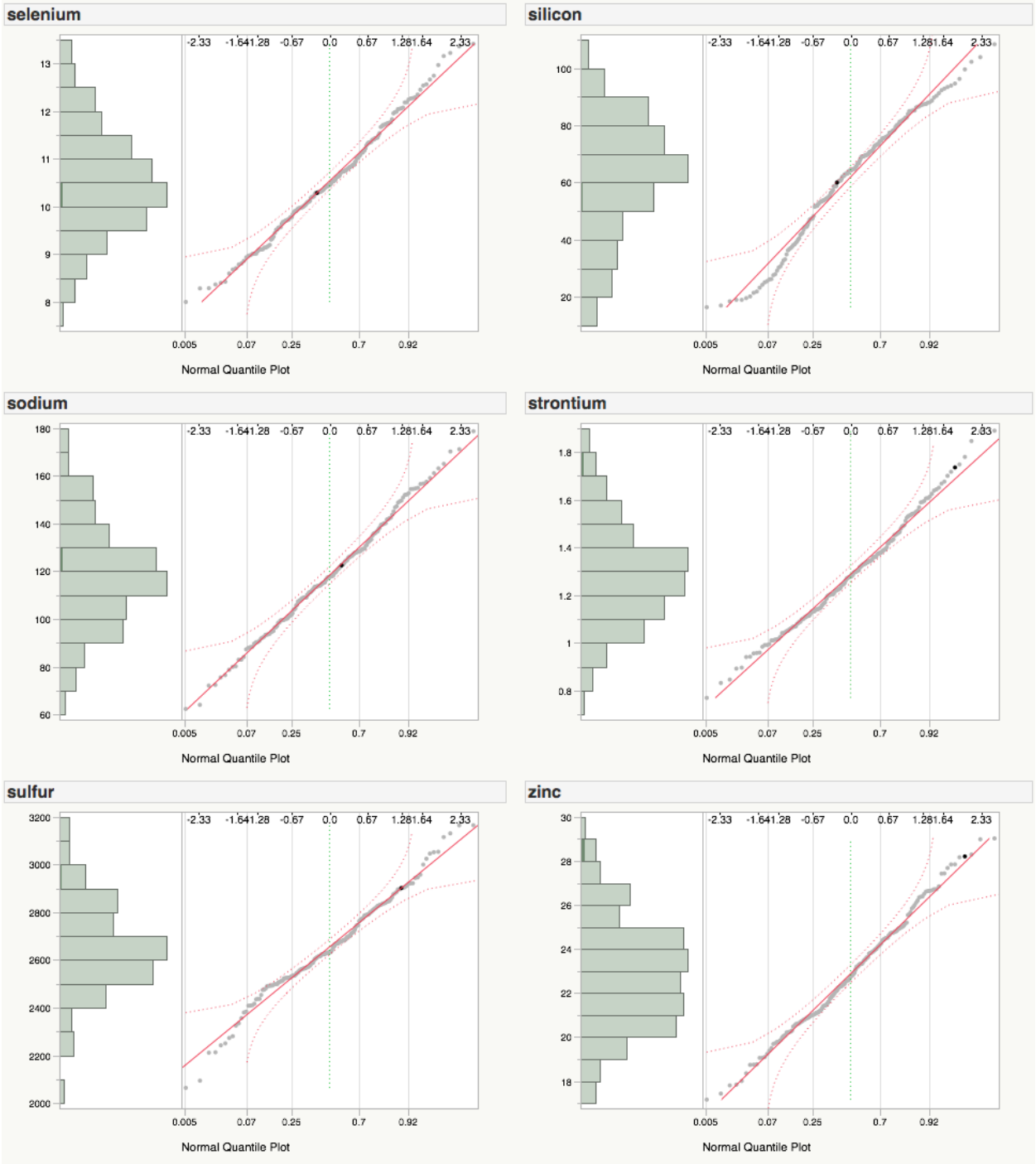

Figure 2.10 Phenotypic distributions of BLUEs of Se, Si, Na, Sr, S and Zn shoot tissue

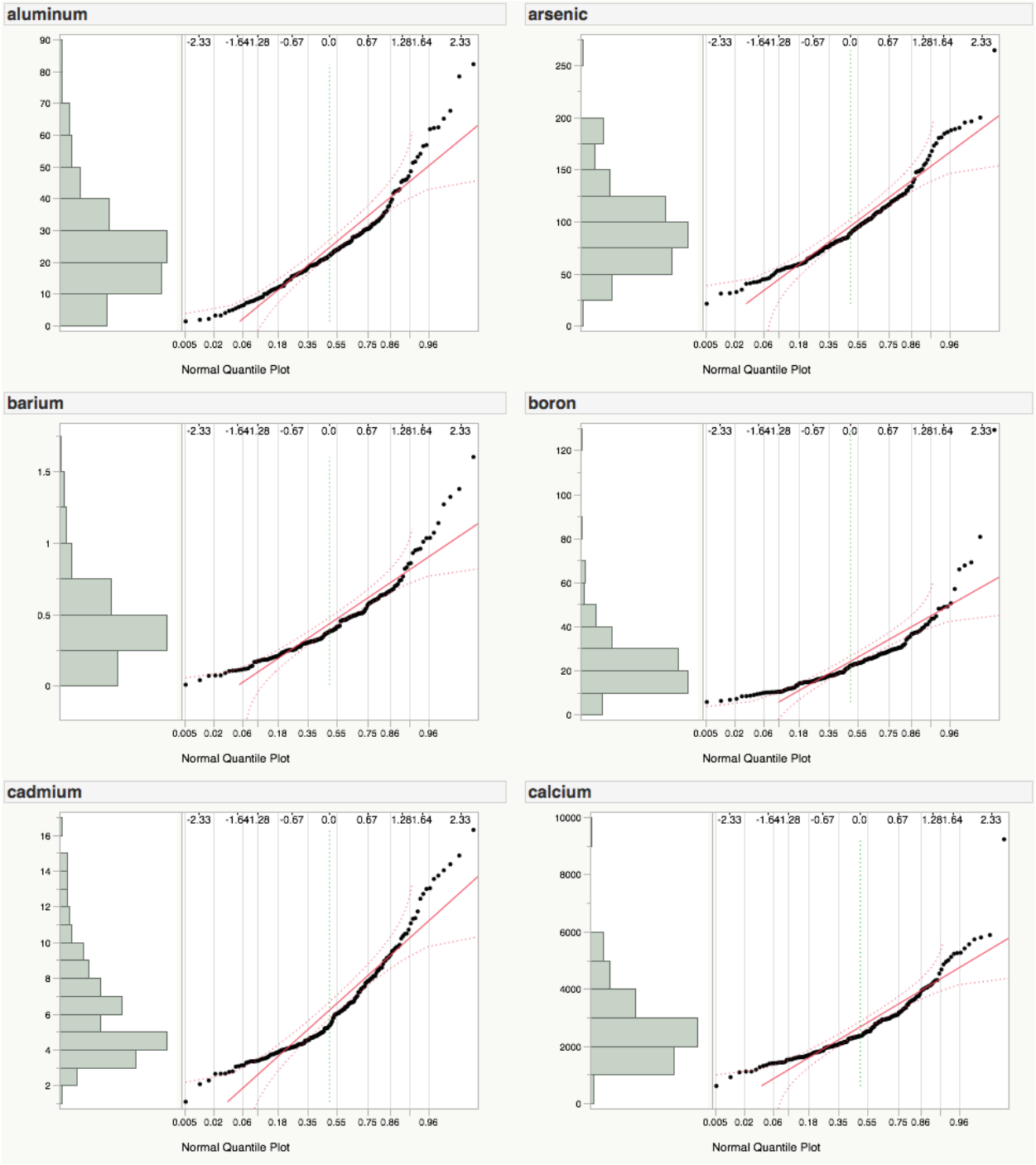

Figure 2.11 Phenotypic distributions of BLUEs of Al, As, Ba, B, Cd and Ca root tissue

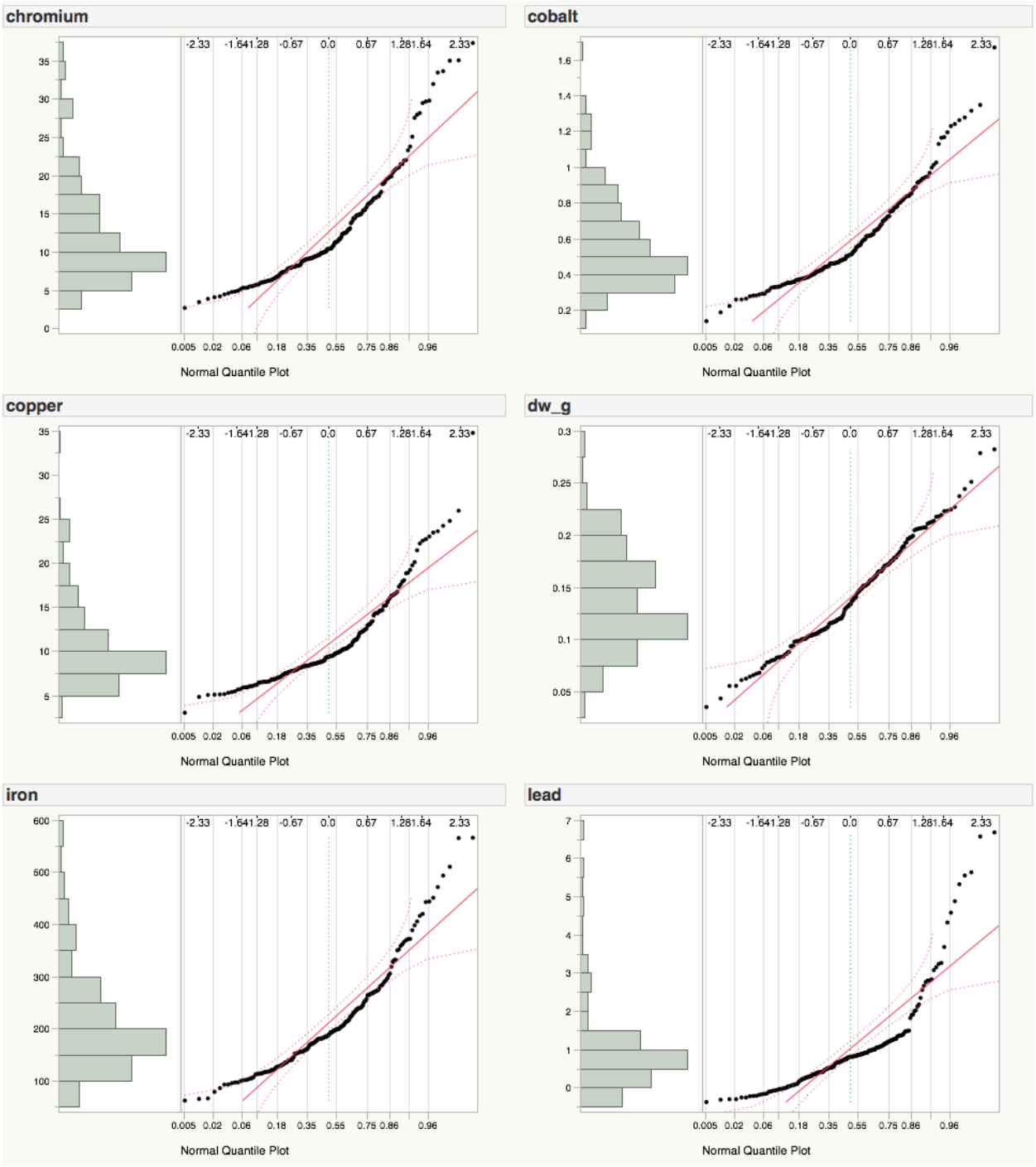

Figure 2.12 Phenotypic distributions of BLUES of Cr, Co, Cu, Fe, Pb and dry weight (dw_g) root tissue

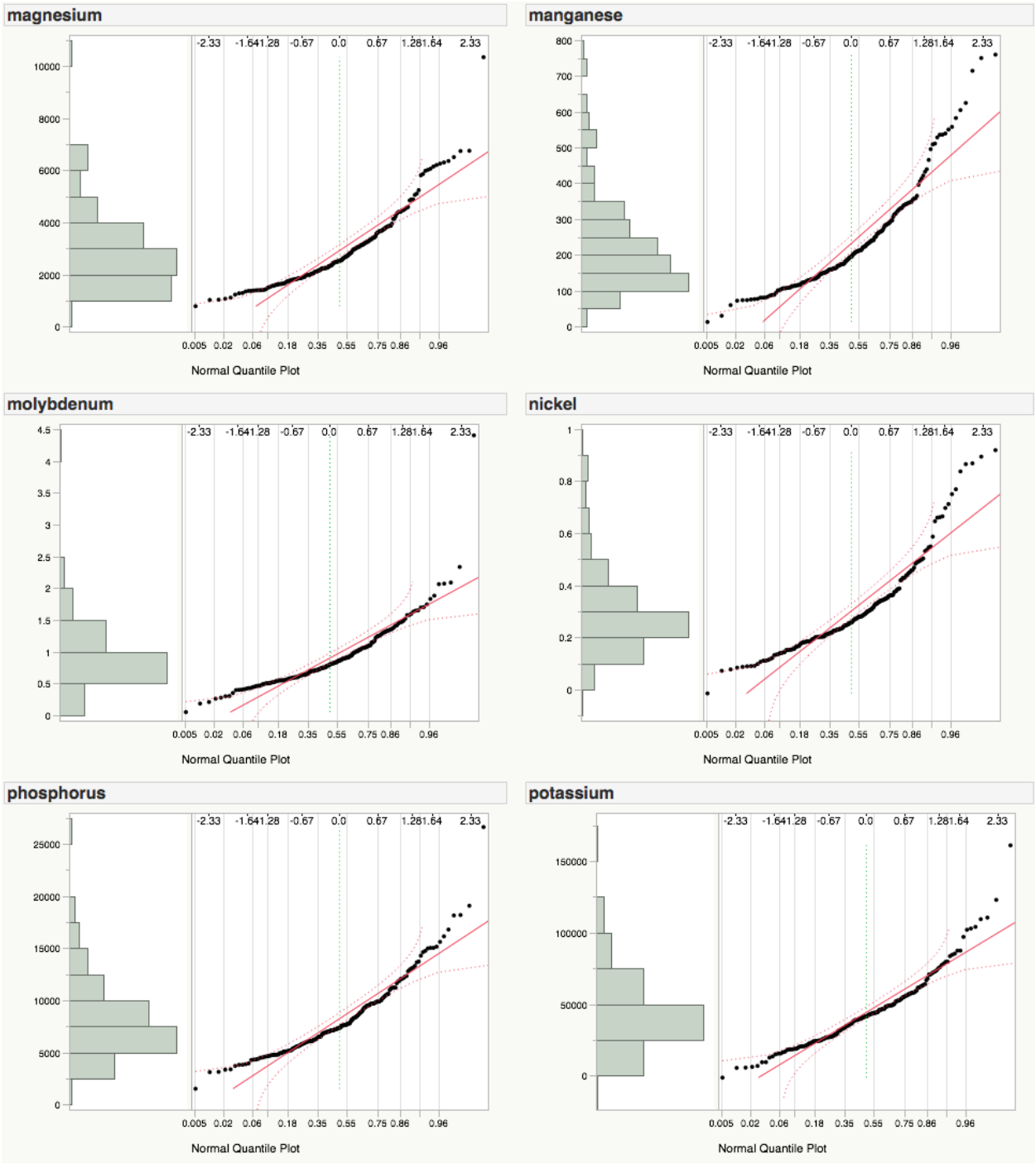

Figure 2.13 Phenotypic distributions of BLUEs of Mg, Mn, Mo, Ni, P and K root tissue

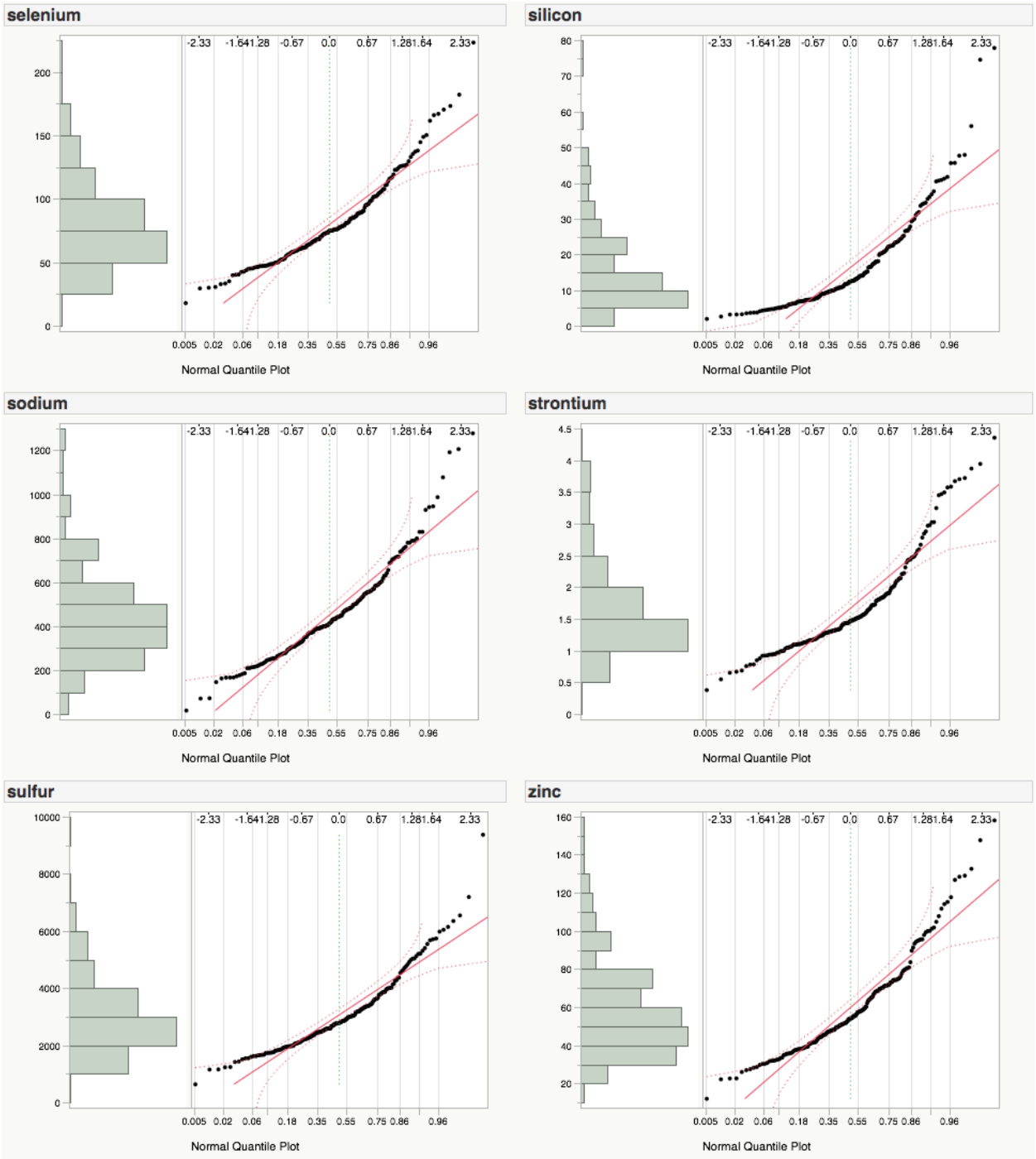

Figure 2.14 Phenotypic distributions of BLUEs of Se, Si, Na, Sr, S and Zn root tissue

Table 2.1 Nutrient solution wheat hydroponics

Element		[µM]	[µM_adj]
Ammonia	NO3	8981.8	8981.8
Ammonium	NH4	500.0	500.0
Arsenic	As	4.0	4.0
Boron	B	12.5	12.5
Buffer (K)	MES	1000.0	1000.0
Buffer (Zn)	HEDTA	77.0	77.0
Cadmium	Cd	0.1	0.1
Calcium	Ca	4000.0	4000.0
Chlorine	Cl	50.0	50.0
Chromium	Cr	0.1	0.1
Cobalt	Co	0.1	0.1
Copper	Cu	1.0	1.0
Iodine	I	1.0	1.0
Iron	Fe	77.0	77.0
Lead	Pb	0.1	0.1
Lithium	Li	0.1	0.1
Magnesium	Mg	250.0	250.0
Manganese	Mn	2.0	2.0
Molybdenum	Mo	0.1	0.1
Nickel	Ni	0.1	0.1
Phosphorus	P	250.0	250.0
Potassium	K*	587.0	587.0
Rubidium	Rb	0.1	0.1
Selenium	Se	2.0	2.0
Silicon	Si	5.0	5.0
Sodium	Na***	312.2	612.2
Strontium	Sr	0.1	0.1
Sulfur	S**	250.5	2250.5
Zinc	Zn	2.0	2.0

To hold the solution at pH 6.0 more K (*) was added as a result of pH adjustment of the MES buffer. MES contributes another ~2000 µM sulfur (**), and another 300 µM sodium (***) is added by the NaOH in the iron solution.

Table 2.2 Average % RSD and lower limit of detection (LOD)

Element		Avg. % RSD	LOD
Aluminum	Al	3.57	-
Arsenic	As	2.85	0.0383
Barium	Ba	7.21	-
Boron	B	6.01	0.1463
Cadmium	Cd	2.71	0.0044
Calcium	Ca	3.03	0.2913
Chromium	Cr	121.35	0.0219
Cobalt	Co	3.14	0.0028
Copper	Cu	2.91	0.0053
Iron	Fe	2.54	0.2399
Lead	Pb	2.41	0.0077
Magnesium	Mg	3.79	1.2190
Manganese	Mn	3.44	0.0009
Molybdenum	Mo	2.15	0.0034
Nickel	Ni	7.76	0.0028
Phosphorus	P	659.20	2.7940
Potassium	K	2.23	6.8640
Selenium	Se	2.61	0.0142
Silicon	Si	2.92	0.0886
Sodium	Na	659.70	-
Strontium	Sr	4.73	0.0008
Sulfur	S	660.00	0.5086
Zinc	Zn	2.99	0.0651

Table 2.3 Mean, standard deviation (Std. Dev.), coefficient of variation (CV), and heritability on entry mean basis for all elements

Element	Shoots				Roots			
	Mean	Std. Dev.	CV	H^2	Mean	Std. Dev.	CV	H^2
Al	9.26	2.20	23.72	0.36	24.00	13.91	57.94	0.25
As	6.57	1.13	17.14	0.45	95.83	40.08	41.82	0.30
Ba	0.81	0.17	21.68	0.62	0.44	0.27	60.89	0.33
B	26.24	4.41	16.79	0.32	24.15	14.47	59.94	0.43
Cd	0.33	0.06	18.73	0.75	6.21	2.83	45.58	0.25
Ca	6277.61	1063.79	16.95	0.65	2667.98	1177.41	44.13	0.35
Cr	1.64	0.15	8.95	0.03	12.84	7.04	54.81	0.31
Co	0.03	0.02	54.26	0.17	0.59	0.26	43.79	0.23
Cu	8.20	1.02	12.40	0.66	10.79	4.90	45.42	0.32
dw	711.05	200.87	28.25	0.73	139.86	47.84	34.21	0.50
Fe	72.14	8.70	12.06	0.15	212.86	98.90	46.46	0.38
Pb	0.16	0.09	56.01	0.20	1.02	1.18	116.40	0.35
Mg	1181.33	118.26	10.01	0.57	2914.70	1440.42	49.42	0.17
Mn	115.23	18.73	16.26	0.67	232.10	139.38	60.05	0.10
Mo	3.20	0.42	13.15	0.66	0.90	0.48	52.97	0.39
Ni	0.21	0.14	65.57	0.44	0.30	0.17	56.87	0.01
P	10054.50	1054.39	10.49	0.73	8229.54	3554.57	43.19	0.34
K	49234.03	15992.60	32.48	0.58	46599.15	23050.91	49.47	0.26
Se	10.56	1.14	10.84	0.62	80.11	32.89	41.06	0.38
Si	61.76	21.55	34.90	0.75	15.63	12.61	80.65	0.65
Na	119.70	25.33	21.16	0.60	455.58	213.51	46.86	0.37
Sr	1.30	0.21	16.42	0.45	1.68	0.74	44.26	0.43
S	2651.21	193.97	7.32	0.38	3057.46	1296.68	42.41	0.34
Zn	22.87	2.49	10.89	0.54	59.89	25.46	42.51	0.28

Table 2.4 Pairwise correlations between all elements in the shoots

Element	Al	As	Ba	B	Cd	Ca	Cr	Co	Cu	dw	Fe	Pb	Mg	Mn	Mo	Ni	P	K	Se	Si	Na	Sr	S	Zn
Al	1.00	0.05	0.156*	-0.44***	0.05	-0.01	0.10	-0.24***	0.09	-0.21**	0.13	-0.03	-0.06	0.05	0.06	-0.19*	-0.05	0.19**	0.08	0.60***	0.21**	0.07	-0.06	0.09
As		1.00	0.36***	0.10	0.14	0.52***	0.39***	0.21**	0.12	-0.22**	0.66***	-0.01	0.31***	-0.03	0.26***	0.11	0.47***	-0.12	0.45***	0.02	0.36***	0.54***	0.37***	0.14*
Ba			1.00	-0.08	-0.06	0.51***	0.19**	0.08	0.24***	-0.20**	0.36***	-0.10	0.46***	0.26***	0.43***	-0.15*	0.33***	-0.15*	0.28***	0.25***	0.24**	0.65***	0.09	0.10
B				1.00	-0.03	0.18*	0.36***	0.21**	-0.01	0.09	0.05	0.03	0.12	-0.03	0.02	0.23**	0.16*	-0.27***	0.10	-0.27***	-0.07	0.13	0.19*	-0.13
Cd					1.00	0.04	0.15*	-0.05	0.43***	0.21**	0.22**	0.09	0.10	0.55***	0.29***	0.07	0.11	0.01	0.02	0.13	-0.06	0.01	0.28***	0.55***
Ca						1.00	0.32***	0.23**	0.15*	-0.03	0.51***	-0.02	0.57***	0.14*	0.22**	0.02	0.18*	-0.19*	0.27***	0.04	0.44***	0.85***	0.37***	-0.03
Cr							1.00	-0.01	0.12	-0.15*	0.45***	0.16*	0.29***	0.14*	0.20**	0.24***	0.43***	0.05	0.25***	0.11	0.13	0.37***	0.27***	0.12
Co								1.00	-0.04	-0.08	0.19**	-0.09	0.15*	-0.04	0.03	0.13*	0.06	-0.31***	0.12*	-0.26***	0.13*	0.17*	0.10	-0.09
Cu									1.00	0.13	0.25***	0.15*	0.32***	0.57***	0.48***	-0.03	0.26***	0.06	0.21**	0.24***	-0.06	0.21**	0.37***	0.61***
dw										1.00	-0.19*	-0.07	0.01	0.27***	0.07	0.04	-0.24***	-0.01	-0.27***	-0.21**	-0.17*	-0.18*	0.15*	0.22*
Fe											1.00	-0.08	0.44***	0.20**	0.36***	0.01	0.41***	-0.13	0.44***	0.16*	0.49***	0.52***	0.44***	0.27***
Pb												1.00	-0.14	0.04	-0.15*	0.45***	-0.02	0.54***	-0.07	-0.01	-0.41***	0.11	0.03	-0.01
Mg													1.00	0.29***	0.45***	-0.14*	0.25***	-0.28***	0.20**	0.11	0.38***	0.45***	0.26***	0.28***
Mn														1.00	0.48**	-0.02	0.17*	-0.01	0.06	0.25***	-0.12	0.12**	0.28***	0.47***
Mo															1.00	-0.18*	0.38***	-0.34***	0.33***	0.23**	0.17*	0.19**	0.34***	0.40***
Ni																1.00	0.04	0.35***	-0.07	-0.29***	-0.40***	0.12	0.07	-0.09
P																	1.00	-0.13	0.59***	0.07	0.12*	0.30***	0.26***	0.30***
K																		1.00	-0.13	0.02	-0.44***	0.01	-0.02	0.05
Se																			1.00	0.26***	0.26***	0.28***	0.36***	0.08
Si																				1.00	0.26**	0.10	0.07	0.16*
Na																					1.00	0.24**	0.20**	0.05
Sr																						1.00	0.32***	0.00
S																							1.00	0.32***
Zn																								1.00

Asterisks indicate significance at * $p < 0.05$, ** $p < 0.001$, *** $p < 0.0001$

Table 2.5 Summary of identified QTL for elements measured in the roots

Element	Chr	GBS Marker	Pos (cM)	LOD	LOD α=0.05	Pheno variance	QTL effect	SE	Effect A	Effect B
Nickel	2BS	synopGBS894	78.7	3.22	3.00	9.24 %	-0.05	0.013	0.05	-0.05
Sodium	2BS	synopGBS894	78.7	3.65	3.01	10.74 %	-69.58	16.60	69.58	-69.58

GBS markers, QTL positions Pos (cM), LOD, LOD at the 5% significance threshold score, estimated phenotypic variance explained by QTL, estimated QTL effect and allele effects (effect A and effect B) for all QTL. QTL are sorted based on chromosome and location. The LODs reported here are LODs calculated through MQM.

Table 2.6 Summary of identified QTL for elements measured in the shoots

Element	Chr	GBS Marker	Pos (cM)	LOD	LOD α=0.05	Pheno variance	QTL effect	SE	Effect A	Effect B
dw_mg	1AL	synopGBS1113	88.2	3.24	3.14	8.39	-58.14	14.84	58.14	-58.14
Phosphorus	1BL	synopGBS151	62.2	3.16	2.99	4.66	-235.89	62.05	235.89	-235.89
dw_mg	2BL	synopGBS565	78.7	3.64	3.14	9.50	61.56	14.77	-61.56	61.56
Arsenic	3B	synopGBS697	92.6	4.76	3.07	13.27	0.41	0.09	-0.41	0.42
Phosphorus	3DL	synopGBS963	71.9	9.29	2.99	15.06	-943.00	137.92	934.00	-943.00
Phosphorus	3DL	synopGBS416	86.9	8.51	2.99	8.51	914.60	140.60	-914.60	914.60
Barium	4AS	synopGBS923	0.9	4.87	3.12	12.35	0.07	0.01	-0.07	0.07
Phosphorus	4AS	synopGBS620	33.8	5.00	2.99	7.56	-299.63	61.85	299.63	-299.63
Cadmium	4BS	synopGBS885	32.9	14.80	3.07	33.19	-0.04	0.01	0.04	-0.04
Phosphorus	5A	synopGBS836	87.7	4.61	2.99	6.90	290.21	62.51	-290.21	290.21
Sulfur	5AL	synopGBS147	125.4	3.09	3.08	8.90	56.63	14.75	-56.63	56.63
Cadmium	5AL	synopGBS429	127.9	4.99	3.07	9.59	0.02	0.01	-0.02	0.02
Copper	5AL	synopGBS429	127.9	7.10	3.13	17.52	0.44	0.07	-0.44	0.44
Manganese	5AL	synopGBS429	127.9	12.40	3.09	17.57	10.02	1.20	-10.02	10.02
Zinc	5AL	synopGBS429	127.9	5.44	3.20	14.32	0.99	0.19	-0.99	0.99
Molybdenum	5BL	synopGBS1229	84.8	3.63	3.00	10.28	0.13	0.03	-0.13	0.13
Selenium	5BL	synopGBS835	88.3	4.52	3.08	12.65	0.39	0.08	-0.39	0.39
Phosphorus	5BL	synopGBS1018	92.0	10.20	2.99	16.80	437.30	60.62	-437.30	437.30
Copper	5BS	synopGBS72	0.8	3.80	3.13	7.88	0.29	0.07	-0.29	0.29
Barium	5BS	synopGBS1229	84.8	4.67	3.12	11.83	0.07	0.01	-0.07	0.07
Calcium	6AL	synopGBS1401	95.6	4.50	3.16	10.17	353.47	71.89	-353.47	353.47
Calcium	6DL	synopGBS1323	120.0	7.02	3.16	16.50	-445.83	70.96	445.83	-445.83
Strontium	6DL	synopGBS322	120.3	4.69	3.14	13.09	-0.08	0.02	0.08	-0.08
Phosphorus	7AL	synopGBS452	85.5	4.72	2.99	7.13	-296.55	64.17	296.55	-296.55
Zinc	7AL	synopGBS1053	88.1	3.32	3.20	8.46	-0.75	0.19	0.75	-0.75
Calcium	7BS	synopGBS1380	10.8	4.04	3.17	9.05	-326.96	70.51	326.96	-326.96
Manganese	7DS	synopGBS761	72.5	4.25	3.09	8.32	5.63	1.24	-5.63	5.63

GBS markers, QTL positions Pos (cM), LOD, LOD at the 5% significance threshold score, estimated phenotypic variance explained by QTL, estimated QTL effect and allele effects (effect A and effect B) for all QTL. QTL are sorted based on chromosome and location. The LODs reported here LODs calculated through MQM.

References

Arends, D., P. Prins, R.C. Jansen and K.W. Broman. 2010. R/qtl: high-throughput multiple QTL mapping. Bioinformatics 26: 2990-2992. doi:10.1093/bioinformatics/btq565.

Bálint, A.F., M.S. Röder, R. Hell, G. Galiba and A. Börner. 2007. Mapping of QTLs affecting copper tolerance and the Cu, Fe, Mn and Zn contents in the shoots of wheat seedlings. Biol Plant 51: 129-134. doi:10.1007/s10535-007-0025-9.

Baxter, I., P.S. Hosmani, A. Rus, B. Lahner, J.O. Borevitz, B. Muthukumar, et al. 2009. Root Suberin Forms an Extracellular Barrier That Affects Water Relations and Mineral Nutrition in Arabidopsis. PLoS Genet 5: e1000492. doi:10.1371/journal.pgen.1000492.

Baxter, I., B. Muthukumar, H.C. Park, P. Buchner, B. Lahner, J. Danku, et al. 2008. Variation in Molybdenum Content Across Broadly Distributed Populations of *Arabidopsis thaliana* Controlled by a Mitochondrial Molybdenum Transporter *MOT1*. PLoS Genet 4: e1000004. doi:10.1371/journal.pgen.1000004.

Bernardo, R. 2010. Breeding for Quantitative Traits in Plants. 2nd ed. Stemma Press, Woodbury, Minnesota.

Broman K.W, Wu H., Sen Ś and C. G.A. 2003. R/ qtl: QTL mapping in experimental crosses. Bioinformatics 19: 889-890.

Broman, K.W. and S. Saunak. 2009. A Guide to QTL Mapping with R/qtl. Springer, New York.

Burkardt, P.K., P. Beyer, J. Wunn, A. Kloti, G.A. Armstrong, M. Schledz, J. vonLintig, I. Potrykus 1997. Transgenic rice (*Ozyra sativa*) endosperm expressing daffodil (*Narcissus pseudonarcissus*) phytoene synthase accumultes phytoene, a key intermediate of provitamine A biosynthesis. Plant Journal 11: 1071-1078.

Carollo, V., D.E. Matthews, G.R. Lazo, T.K. Blake, D.D. Hummel, N. Lui, et al. 2005. GrainGenes 2.0. an improved resource for the small-grains community. Plant Physiol 139: 643-651. doi:10.1104/pp.105.064485.

Chen, Z., T. Watanabe, T. Shinano, K. Okazaki and O. Mitsuru. 2009. Rapid characterization of plant mutants with an altered ion-profile: a case study using *Lotus japonicus*. The New phytologist 181: 795-801. doi:10.1111/j.1469-8137.2008.02730.x.

Clemens, S., M.G.M. Aarts, S. Thomine and N. Verbruggen. 2013. Plant science: the key to preventing slow cadmium poisoning. Trends in Plant Science 18: 92-99. doi:http://dx.doi.org/10.1016/j.tplants.2012.08.003.

Cobb, J.N., M. Rutzke, S. Giri, E. Craft, J. Shaff, P. Korneliev, et al. 2015. Ionomic Profiling in a Panel of Diverse Oryza sativa Accessions Is Associated with Natural Genetic Variation As Revealed by Genome-wide Association Mapping. In preparation.

Cordell, D., J.-O. Drangert and S. White. 2009. The story of phosphorus: Global food security and food for thought. Global Environmental Change 19: 292-305. doi:http://dx.doi.org/10.1016/j.gloenvcha.2008.10.009.

Ding, D., W. Li, G. Song, H. Qi, J. Liu and J. Tang. 2011. Identification of QTLs for Arsenic Accumulation in Maize (*Zea mays* L.) Using a RIL Population. PLoS ONE 6: e25646. doi:10.1371/journal.pone.0025646.

Dixon, H.B.F. 1996. The Biochemical Action of Arsonic Acids Especially As Phosphate Analogues. Advances in Inorganic Chemistry 44: 191-227. doi:10.1016/S0898-8838(08)60131-2.

Dunckel, S.M., E.L. Olson, M. Rouse, R. Bowden and J. Poland. 2015. Genetic Mapping of Race-Specific Stem Rust Resistance in the Synthetic Hexaploid W7984 x Opata M85 Mapping Population. Crop Sci. 55: 1-9. doi:10.2135/cropsci2014.11.0755.

FAO. 2015. FAOSTAT, Statistics Division FAO. Retrieved 08/0/2015 from http://faostat3.fao.org/home/E.

Fassel, V.A. and R.N. Kniseley. 1974. Inductively Coupled Plasma-Optical Emission Spectroscopy. Analytical Chemistry 46: 1110A-1120A. doi:10.1021/ac60349a722.

Gainhealth. 2015. Fast Facts About Malnutrition.

Genc, Y., A.P. Verbyla, A.A. Torun, I. Cakmak, K. Willsmore, H. Wallwork, et al. 2009. Quantitative trait loci analysis of zinc efficiency and grain zinc concentration in wheat using whole genome average interval mapping. Plant Soil 314: 49-66. doi:10.1007/s11104-008-9704-3.

Goldenrice. 2015. Rewiring the grain. Retreived 08/08/2015 from http://www.gainhealth.org/knowledge-centre/fast-facts-malnutrition/

Grant, C.A., W.T. Buckley, L.D. Bailey and F. Selles. 1998. Cadmium accumulation in crops. Canadian Journal of Plant Science 78: 1-17. doi:10.4141/P96-100.

Guttieri, M.J., P.S. Baenziger, K. Frels, B. Carver, B. Arnall, S. Wang, et al. 2015. Prospects for Selecting Wheat with Increased Zinc and Decreased Cadmium Concentration in Grain. Crop Science 55: 1712-1728. doi:10.2135/cropsci2014.08.0559.

Harris, N.S. and G.J. Taylor. 2013. Cadmium uptake and partitioning in durum wheat during grain filling. BMC Plant Biol 13: 103. doi:10.1186/1471-2229-13-103.

HarvestPlus. 2015. HarvestPlus. Retrieved 08/08/2015 from http://www.harvestplus.org/content/crops

Hou, X. and B.T. Jones. 2000. Inductively Coupled Plasma/Optical Emission Spectrometry. In: R. A. Meyers, editor Encyclopedia of Analytical Chemistry. John Wiley & Sons Ltd, Chichester. p. 9468-9485.

JMP®, V. 1989-2015. SAS Institute Inc, Cary, NC.

Lahner, B., J. Gong, M. Mahmoudian, E.L. Smith, K.B. Abid, E.E. Rogers, et al. 2003. Genomic scale profiling of nutrient and trace elements in *Arabidopsis thaliana*. Nat Biotech 21: 1215-1221.

McLaughlin, M.J., D.R. Parker and J.M. Clarke. 1999. Metals and micronutrients – food safety issues. Field Crops Research 60: 143-163. doi:http://dx.doi.org/10.1016/S0378-4290(98)00137-3.

Nakanishi, H., I. Ogawa, Y. Ishimaru, S. Mori and N.K. Nishizawa. 2006. Iron deficiency enhances cadmium uptake and translocation mediated by the Fe2+ transporters OsIRT1 and OsIRT2 in rice. Soil Science & Plant Nutrition 52: 464-469. doi:10.1111/j.1747-0765.2006.00055.x.

Palmgren, M.G., S. Clemens, L.E. Williams, U. Krämer, S. Borg, J.K. Schjørring, et al. 2008. Zinc biofortification of cereals: problems and solutions. Trends in Plant Science 13: 464-473. doi:http://dx.doi.org/10.1016/j.tplants.2008.06.005.

Paltridge, N., P. Milham, J.I. Ortiz-Monasterio, G. Velu, Z. Yasmin, L. Palmer, et al. 2012. Energy-dispersive X-ray fluorescence spectrometry as a tool for zinc, iron and selenium analysis in whole grain wheat. Plant Soil 361: 261-269. doi:10.1007/s11104-012-1423-0.

Peleg, Z., I. Cakmak, L. Ozturk, A. Yazici, Y. Jun, H. Budak, et al. 2009. Quantitative trait loci conferring grain mineral nutrient concentrations in durum wheat x wild emmer wheat RIL population. Theor Appl Genet 119: 353-369. doi:10.1007/s00122-009-1044-z.

Pigna, M., V. Cozzolino, A. Giandonato Caporale, M.L. Mora, V. Di Meo, A.A. Jara, et al. 2010. Effects of phosphorus fertilization on arsenic uptake by wheat grown in polluted soils. Journal of Soil Science and Plant Nutrition 10: 428-442.

Potrykus, I. 2001. Golden Rice and beyond. Plant Physiol 125: 1157-1161.

Poland, J.A. and T.W. Rife. 2012. Genotyping-by-Sequencing for Plant Breeding and Genetics. Plant Gen. 5: 92-102. doi:10.3835/plantgenome2012.05.0005.

R Core Team. 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Salt, D.E. 2004. Update on Plant Ionomics. Plant Physiology 136: 2451-2456. doi:10.1104/pp.104.047753.

Salt, D.E., I. Baxter and B. Lahner. 2008. Ionomics and the Study of the Plant Ionome. Annual Review of Plant Biology 59: 709-733. doi:doi:10.1146/annurev.arplant.59.032607.092942.

Sasaki, A., N. Yamaji, K. Yokosho and J.F. Ma. 2012. Nramp5 is a major transporter responsible for manganese and cadmium uptake in rice. The Plant cell 24: 2155-2167. doi:10.1105/tpc.112.096925.

Shiferaw, B., M. Smale, H.-J. Braun, E. Duveiller, M. Reynolds and G. Muricho. 2013. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. Food Sec. 5: 291-317. doi:10.1007/s12571-013-0263-y.

Sorrells, M., J. Gustafson, D. Somers, S. Chao, D. Benscher, G. Guedira-Brown, E. Huttner, A. Kilian, P.E. McGuire, K. Ross, J. Tanaka, P. Wenzel, K. Williams and C.O. Qualset. 2011. Reconstruction of the Synthetic W7984xOpata M85 wheat reference population. Genome 54: 875 - 882.

Su, J.-Y., Q. Zheng, H.-W. Li, B. Li, R.-L. Jing, Y.-P. Tong, et al. 2009. Detection of QTLs for phosphorus use efficiency in relation to agronomic performance of wheat grown under phosphorus sufficient and limited conditions. Plant Science 176: 824-836. doi:http://dx.doi.org/10.1016/j.plantsci.2009.03.006.

Tiwari, V.K., N. Rawat, P. Chhuneja, K. Neelam, R. Aggarwal, G.S. Randhawa, et al. 2009. Mapping of quantitative trait Loci for grain iron and zinc concentration in diploid A genome wheat. J Hered 100: 771-776. doi:10.1093/jhered/esp030.

Velu, G., I. Ortiz-Monasterio, R.P. Singh and T. Payne. 2011. Variation for Grain Micronutrients Concentration in Wheat Core-collection Accessions of Diverse Origin. Asian Journal of Crop Science 3: 43-48. doi:10.3923/ajcs.2011.43.48

Vreugdenhil, D., M.G.M. Aarts, M. Koornneef, H. Nelissen and W.H.O. Ernst. 2004. Natural variation and QTL analysis for cationic mineral content in seeds of *Arabidopsis thaliana*. Plant, Cell & Environment 27: 828-839. doi:10.1111/j.1365-3040.2004.01189.x.

Weidong, C., J. Jizeng and J. Jiyun. 2001. Identification and interaction analysis of QTL for phosphorus use efficiency in wheat seedlings. In: W. J. Horst, M. K. Schenk, A. Bürkert,

N. Claassen, H. Flessa, W. B. Frommer, H. Goldbach, H. W. Olfs, V. Römheld, B. Sattelmacher, U. Schmidhalter, S. Schubert, N. v. Wirén and L. Wittenmayer, editors, Plant Nutrition. Springer Netherlands. p. 76-77.

WHO. 2015. Biofortification of staple crops. Retrieved 08/08/2015 from http://www.who.int/elena/titles/biofortification/en/

Xu, Y., D. An, D. Liu, A. Zhang, H. Xu and B. Li. 2012. Molecular mapping of QTLs for grain zinc, iron and protein concentration of wheat across two environments. Field Crops Research 138: 57-62. doi:http://dx.doi.org/10.1016/j.fcr.2012.09.017.

Zhao, F.J., J.L. Stroud, T. Eagling, S.J. Dunham, S.P. McGrath and P.R. Shewry. 2010. Accumulation, distribution, and speciation of arsenic in wheat grain. Environmental Science & Technology 44: 5464-5468. doi:10.1021/es100765g.

Ziegler, G., A. Terauchi, A. Becker, P. Armstrong, K. Hudson and I. Baxter. 2013. Ionomic Screening of Field-Grown Soybean Identifies Mutants with Altered Seed Elemental Composition. The Plant Genome 6. doi:10.3835/plantgenome2012.07.0012.

Chapter 3 - Genomic selection for increased yield in synthetic derived wheat

Abstract

The loss of genetic diversity in bread wheat (*Triticum aestivum* L.) due to bottlenecks from polyploidy, domestication and modern plant breeding can be compensated by introgressing novel exotic germplasm. A successful approach to capture genetic diversity is the production of primary synthetic bread wheat, which are contemporary reconstitutions of the ancestral genomes of wheat from diverse wild relatives. To this end, wheat breeding and genetics programs around the world have developed many primary synthetics. However, this diverse germplasm has many undesirable characters, making direct use in breeding programs difficult. To increase the speed of introgression of exotic germplasm, genomic selection approaches could be applied to enable rapid cycles of selection. To test this approach, selected lines from double haploid (DH) and recombinant inbred line (RIL) populations between six different primary synthetics and the elite cultivar Opata M85 were evaluated for grain yield and other important agronomic traits. Field trials were conducted at CIMMYT (International Center for Maize and Wheat Improvement) over two years in irrigated, heat, and drought-stressed environments. Several synthetic derived lines outperformed the elite parent Opata M85 in all environments indicating that the primary synthetics contribute alleles increasing yield. Whole genome profiles were generated using genotyping-by-sequencing (GBS) to generate whole-genome prediction models in elite by synthetic populations. Five different whole-genome prediction models that can be applied for genomic selection (GS) were evaluated for prediction accuracy using cross-validation. Overall, the prediction models had moderate predictive ability. However, the prediction accuracies were

slightly lower than expected based on the heritability of the traits. As such, rapid cycle GS for introgression of exotic alleles might not perform as well as expected with synthetic derived wheat due to complex and confounding physiological effects.

Introduction

Domestication and modern plant breeding led to the reduction of genetic diversity in cultivated crops, including in bread wheat (*Triticum aestivum* L.) (Reif *et al.* 2005). Cultivated bread wheat (2n = 6x = 42, AABBDD) arose from a natural whole-genome hybridization of cultivated tetraploid wheat (*Triticum turgidum* L.) (2n = 4x = 28, AABB) and diploid wild species *Aegilops tauschii* Coss. (2n = 2x = 14, DD) about 8,000 years ago (Kihara 1944; McFadden and Sears 1946; Dvorak *et al.* 1998; Talbert *et al.* 1998; Marcussen *et al.* 2014). This speciation event resulted in the first genetic bottleneck and was followed by multiple genetic bottlenecks during its domestication process. The diversity of modern bread wheat varieties has been further narrowed through strong selection in breeding programs. It is well recognized that maintaining genetic diversity is crucial for sustaining gains through plant breeding and wild germplasm is a valuable source of novel genes for disease resistance, tolerance to abiotic stresses and increased yield (Mujeeb-Kazi *et al.* 2004, Reif *et al.*, 2005).

A successful method to compensate the loss of genetic diversity in bread wheat is the production of synthetic hexaploid wheat as first described by McFadden and Sears (1946). Since then several breeding programs have improved the technique. The CIMMYT wide-crossing program has developed over a thousand new primary synthetics from more than 600 different *Aegilops tauschii* accessions (Zhang *et al.* 2005). The value of landraces and synthetic derived wheat to improve genetic diversity has recently been demonstrated in wheat breeding programs

at CIMMYT as the use of synthetics resulted in a slight increase in genetic diversity (Warburton *et al.* 2006).

A synthetic derived mapping population from the cross between Synthetic W7984 (Altar 84 / *Aegilops tauschii* (219) CIGM86.940) and elite wheat cultivar 'Opata M85' was developed in the late 1980s and widely used in the wheat community. This original population is known as 'ITMI' or 'M6' mapping population. Recently, Sorrells *et al.* (2011) reconstructed two synthetic wheat reference populations with the same pedigree as the original ITMI mapping population. One population consists of doubled haploids (named SynOpDH) and one of recombinant inbred lines (named SynOpRIL) (Sorrells *et al.* 2011). Even though primary synthetics carry favorable genes associated with tolerance to a range of biotic and abiotic stresses, they often harbor unfavorable alleles associated with poor agronomic performance and low yield and need to be backcrossed to an elite cultivar or breeding line (Arraiano *et al.* 2001; Mujeeb-Kazi *et al.* 2004). Once backcrossed and intensively selected, the newly created plant material can start to be incorporated into the elite wheat breeding program for cultivar development. This process is very time-consuming, limiting the use of new genetic diversity in the breeding program. A previously proposed approach to introgress exotic germplasm more rapidly is through advanced backcross quantitative trait loci (QTL) mapping (Narasimhamoorthy *et al.*, 2006). However, number of identified QTL for yield and yield-related traits was low. Here we explore the potential for applying genomic selection for introgression of primary synthetics into the elite wheat breeding program.

Important agronomic traits, such as yield, are complex quantitative traits controlled by many loci of small effect. Traditionally, QTL mapping studies have been performed to identify loci underlying these traits. However, QTL mapping has failed to identify all loci controlling

them because the contribution of small effect loci to the total genetic variance cannot be detected applying traditional QTL mapping by linkage analysis (Meuwissen *et al.* 2001). Only loci with relatively large effect are identified by QTL mapping. In contrast, genomic selection (GS) can be used to predict complex quantitative traits in animal and plant breeding by omitting significance testing and modeling all marker effects. With all marker-marker effects simultaneously estimated across the entire genome, genomic estimated breeding values (GEBVs) can be calculated (Meuwissen *et al.* 2001; Heffner *et al.* 2009). Parents for the next cycle can then be chosen based on their GEBVs prior to phenotyping (Meuwissen *et al.* 2001; Heffner *et al.* 2009; Meuwissen 2009). With rapid selection on seedling plants, applying GS can significantly reduce selection cycle time (Heffner *et al.* 2010). Over the last several years, GS has been shown to increase the breeding efficiency in several crops using different types of populations (Heffner *et al.* 2011; Würschum *et al.* 2013; Rutkoski *et al.* 2014; Spindel *et al.* 2015; Zhang *et al.* 2015).

The underlying assumption for GS to work is extensive linkage disequilibrium (LD) between markers and QTL, at least one marker is assumed to be in LD with each QTL affecting a trait. This requires large numbers of genetic markers such as single nucleotide polymorphisms (SNPs). The rapid development and low cost of Next Generation Sequencing (NGS) technologies make it possible to include genotypic data into the equation and increase the information available to make more targeted selections. With inexpensive, whole-genome profiling, GS can be implemented as an effective approach to increase population size when traits are expensive to phenotype and reduce time to selection when traits are difficult or time consuming to measure (Poland *et al.* 2012b).

Abundant genetic diversity for agronomically important traits is found in wild relatives of most cultivated crops (Zamir 2001). One of the main limitations to use exotic germplasm in modern crops is the time it takes to introgress favorable alleles into elite material. Bernardo (2009) estimated duration of 10 – 20 years for successful introgression of exotic germplasm into elite maize. These numbers are not encouraging and most breeders will not exploit the use of exotic germplasm. Several breeding strategies such as selection in F_2, BC_1 and BC_2, recurrent selection of F_2, advanced backcross QTL mapping, exotic libraries etc. have been proposed for the introgression of exotic germplasm in different crops such as wheat, corn, sorghum, cowpea and maize (Crossa and Gardner 1987; Ehlers and Foster 1993; Zamir 2001; Narasimhamoorthy *et al.*, 2006; Feuillet *et al.* 2008; Ochanda *et al.* 2009). The most popular and recommended breeding strategy for an elite x exotic cross is to start phenotypic selection after the first or second backcross. In a simulation study in maize, Bernardo (2009) tested different breeding strategies to determine the most adequate for rapid introgression and improvement of a quantitative trait in an elite x exotic cross applying genome-wide selection with comparison of F_2, BC_1 and BC_2 populations. The most successful strategy for rapid introgression was when selection started in the F_2 followed by 7 – 8 cycles of GS. In this scenario, favorable alleles from the elite parent were increased while maintaining favorable alleles from the exotic parent. This was true even when the elite parent contributed many more favorable alleles than the exotic parent (Bernardo 2009). Furthermore, gains from 7 – 8 cycles of GS were larger than gains from two cycles of testcross phenotypic selection. Most importantly, the time to successfully introgress exotic germplasm applying GS could be reduced to three years following development of cycle 0 and the ability to advance three generations per year. This is a remarkable reduction

from the usual 10 – 20 years. Genome-wide selection has the potential to speed up the pre-breeding process and the introgression of exotic alleles into elite material.

Our goal is the development of a fast cycling biparental GS scheme for pre-breeding to rapidly introgress favorable alleles from primary synthetics into the elite bread wheat-breeding program. The proposed breeding scheme consists of two different genomic selection cycles GS1 and GS2 (Figure 3.1). GS1 represents the first cycle of our breeding scheme. Six synthetic derived populations are our base populations consisting of lines described in Table 3.1. SynOpDHs and SynOpRILs were added to increase the population size for the first cycle of GS. In GS2, we test newly derived synthetic material developed by crossing the top two lines of each population with elite CIMMYT cultivars. The goal of this breeding scheme is the development of a rapid cycle biparental GS to move synthetic derived lines faster into the elite wheat breeding program. In this study, we focus mainly on GS1 to test two hypotheses; (i) exotic alleles do contribute yield-promoting alleles and (ii) GS can be applied to exotic germplasm of diverse genetic background.

Materials and Methods

Plant material

All field trials were grown at CIMMYT's Norman E. Borlaug Research Station in Ciudad Obregon, Mexico. Synthetic derived spring wheat lines from 6 different biparental populations were used for this study. Table 3.1 includes the pedigrees and population sizes of the material in yield trials during seasons 2012/2013 and 2013/2014. All lines in population Synthetic 6 in Table 3.1 have the same pedigree and are part of the original Synthetic W7984 x Opata M85 population (also known as 'ITMI' or 'M6' population), and the new Synthetic W7984 x Opata M85 double haploid (SynOpDH) and recombinant inbred lines (SynOpRIL) mapping

populations (Sorrells *et al.* 2011). The other five populations have different synthetic parents but share the elite parent Opata M85. The plant material is very diverse and the range for days to flowering (DAYSFL) and plant height (PTHT) is large. All lines were grown in hill plots during season 2011/2012 and selected only for DAYSFL and PTHT to obtain meaningful yield trials by reducing the range of those two traits. A total of 429 lines were selected with reasonable height and maturity to include in yield trials during subsequent years.

Experimental Design

The conditions at CIMMYT's station in Cd. Obregon, Mexico are favorable to generate heat and drought stress. The soils are fertile but there is only very little precipitation during the off-season and no precipitation during the season. This allows controlling the irrigation precisely and simulating different growing conditions.

Irrigation

The irrigated trials were planted in early December and watered optimally with ~600mm per season to avoid drought stress at any stage of the experiment. The trials were harvested every year early May. The experimental design for the irrigated trials was a row-column spatially analyzable design with two replicates (Williams et al. 2006). In this design, the repeated checks and entries are randomly distributed in each replicate to prevent clustering of checks. The elite parent Opata M85 and other elite CIMMYT lines were included as repeated checks. The irrigated trials were planted in two-row plots of dimensions 0.8m x 3.0m (2.4m2).

Drought

Drought stress trials were planted and harvested at the same time as the irrigated trials. Over the course of the season the drought trials were only irrigated with half the water accumulating ~300mm per season. To avoid plot-edge effects the drought trials were planted in larger six-row

plots of 1.2m x 3.0m (3.6m2). Due to larger plot size and limited field space the drought trials were planted as augmented designs (Federer 1956). The included repeated checks were composed of different elite CIMMYT lines including OpataM85.

Heat

In contrast, the heat stress trials were planted in late February to expose the wheat to post-anthesis heat stress. However, to avoid confounding effects of drought stress, the trials were watered optimally with ~600mm per season. The heat trials were harvested late June and plants exposed to temperatures ranging from 35 – 40°C. In comparison, the irrigated and drought trials were exposed to maximal daytime temperatures of 25 – 30°C. The nighttime temperature was on average around 22°C. The experimental design and plot size for the heat trials were the same as for the irrigated trials.

Data collection and statistical analysis

We collected data on several important agronomic traits using the Field Book app (Rife and Poland 2014). Traits collected were days to heading date (DTHD), days to flowering (DASFL), days to maturity (DAYSMT), plant height (PTHT), grain weight per plot (GRWT) and grain yield per hectare (GRYLD). DTHD, DAYSFL and DAYSMT were recorded when more than 50% of the plants headed, flowered or reached maturity, respectively. PTHT was measured with a ruler and the average of three measurements recorded. GRWT refers to the weight of the harvested grain per plot. GRYLD was calculated based on GRWT per plot area (m^2) and extrapolated to t/ha. The data analysis was carried out using the statistical software R (R Core Team 2014). To calculate Best Linear Unbiased Estimators (BLUEs) we used ASReml for R for mixed model analysis (Gilmour *et al.* 2009).

BLUEs were calculated for all traits across both years. The model used to calculate BLUEs for the irrigated and heat trials was:

$$y_{ijk} = \mu + g_i + m_j + (gm)_{ij} + (mr)_{jk} + e_{ijk} \quad [1]$$

where y_{ijk} is the phenotypic trait analyzed, g_i is a fixed effect for each genotype, m_j is the random effect of the jth year, $(gm)_{ij}$ is the random interaction effect of the ith genotype with the jth year assumed N(0, Iσ^2_{gm}), $(mr)_{jk}$ is the kth replicate nested within jth year and e_{ijk} is the random error with N(0, Iσ^2_e). For the augmented drought trial only a few checks were repeated multiple times and the model to calculate BLUEs adjusted to

$$y_{ijk} = \mu + g_i + m_j + (gm)_{ij} + c_k + e_{ijk} \quad [2]$$

where g_i is the fixed effect for each experimental line (replicated only once), $(gm)_{ij}$ is the interaction of the ith entry with the jth year and c_k is the fixed effect of each replicated check. Broad sense heritability (H^2) was calculated on a entry means basis according to Bernardo, (2010) as

$$H^2 = \frac{\sigma^2_g}{\sigma^2_g + \frac{\sigma^2_{gy}}{y} + \frac{\sigma^2_e}{ry}} \quad [3]$$

where σ^2_g is the genotypic variance, σ^2_{gy} is the genotype by year interaction variance, y is the number of years, r is the number of replicates, and σ^2_e is the error variance (Bernardo 2010).

Genotyping, imputation and quality control

We genotyped all lines with genotyping-by-sequencing (Elshire *et al.* 2011) using the two-enzyme approach (Poland *et al.* 2012a). Briefly, 44,421 single-nucleotide polymorphisms (SNPs) with up to 70% missing data were retained in the data set and the tags aligned to the recently published draft sequence of the wheat genome (The International Wheat Genome

Consortium, 2014) using POPSEQ (Chapman *et al.* 2015). SNPs with minor allele frequency (MAF) of less than 5% were excluded from the dataset. Imputation of the ordered marker data was performed using Beagle v.4 (Browning and Browning 2013). The parameters were set as described by Jordan *et al.*, (2015): window = 5,000 overlap = 500 burnin-its =10 impute-its = 10. SNPs monomorphic for both parents were excluded and finally, a total of 7,427 SNPs were used for GS. The relatedness of lines was assessed with a phylogenetic tree using R package ape with standard settings (Paradis *et al.* 2004).

Genomic Selection

We analyzed the predictive ability of synthetic derived wheat in two sets. The first set was comprised of bi-parental population Synthetic 6 with 242 individuals referred to as Set 1 (old "M6", "SynOpDHs" and "SynOpRILs"). The second set, referred to as Set 2, included 396 lines from all six synthetic populations

Predictive Models

We applied five statistical methods for GS embedded in R package GSwGBS (Gaynor 2015): Ridge Regression of marker effects (RRBLUP) and Reduced Kernel Hilbert Space Regression (RKHS) implemented through r package rrBLUP (Endelman 2011), Partial Least Squares Regression (PLSR) applying r package pls (Mevik and Wehrens 2007), Elastic Net (ELNET) using r package glmnet (Friedman *et al.* 2010), and Random Forest with 1000 trees generated with r package randomForest (RF) (Liaw and Wiener 2002).

RRBLUP is a shrinkage model assuming equal variance of all markers and therefore shrinks all marker effects equally. RKHS is another shrinkage model combining the classical additive

genetic model with kernel functions. RKHS does not assume linearity. Here the Euclidean distance matrix is used for Gaussian kernel predictions and genomic predictions are made by estimating line effects (referred to as G-BLUP in Endelman 2011). PLRS is a dimension reduction method attempting to construct a good model from variables with less well-understood relationships. Latent variables are extracted as linear combination of predictors and used to predict the response (Lorenz *et al.*, 2011). ELNET combines the penalties from ridge regression (shrinkage of marker effects) and the LASSO (least absolute shrinkage and selection operator) cost function. This results in a sparse model that allows for grouping effects of variables. RF is a machine-learning algorithm making use of regression trees grown on bootstrap samples. The prediction of a given observation is obtained by averaging the predictions over trees for which the given observation was not used to build the tree (Lorenz *et al.*, 2011). For a more comprehensive description of the genomic selection prediction models, we refer to Lorenz *et al.*, 2011 and Heslot *et al.*, 2012.

Assessing prediction accuracy

Five-fold cross validation (CV) was performed by randomly assigning 80% of the lines as training population (TP) and the remaining 20% as selection candidates (SC). The whole process was repeated 20 times. The prediction accuracy was measured as Pearson correlation between the BLUE and the GEBV.

Results

Marker data and quality control

Analysis of the genotypic data revealed that 15 of the new SynOpRILs are selfs of the elite cultivar Opata M85, and 18 lines from the old M6 population are genetically unrelated and not progeny of Synthetic W7984 x Opata M85. The two sets of lines split into two distinctive clusters apart from the rest of the population (Figure 3.4). A complete list including line identification of the 33 lines in question is available in Table 3.2. After removing the 33 lines, we retained 396 lines for further analysis.

Heritability and line performance

The phenotypic correlations across two years for DAYSFL, DAYSMT, DTHD, GRYLD and PTHT are shown in Table 3.3. The correlations are high and significant for most traits and environments. Broad sense heritability H2 on an entry mean basis, grand mean, least significant difference (LSD) and coefficient of variation (CV) are shown in Table 3.4. Heritabilities are high for DAYSFL, DAYSMT and DTHD (0.88 – 0.98) across all environments. Heritabilities for GRYLD are lower for the drought (0.42) and irrigated (0.76) trials but remain high for the heat (0.94) trial. The heritability of PTHT was lower for all trials (0.57-0.78). The high heritability for most traits and environments indicate high repeatability of our experiments across years.

Several synthetic derived lines outperformed elite parent Opata M85 in all environments (Figure 3.2 and Figure 3.3). Looking at the entire population, the distribution of the synthetic derived line performance relative to Opata M85 was highest under irrigation and lowest under drought stress (Figure 3.3). Under optimal conditions, 37.1% of synthetic derived lines outperformed elite parent Opata M85. A relative smaller portion of lines outperformed Opata M85 under heat and drought stress. While still 30.6% of synthetic derived lines had a higher

yield than Opata M85 under heat stress, only 19.2% outperformed Opata M86 under drought stress.

Genomic Selection

We used five different statistical models for genomic selection (RRBLUP, RKHS, PLSR, ELNET, RF). The prediction accuracies are reported in Table 3.5 for all models, traits and trials. Generally, the prediction accuracies were highest for the heat trial across both sets. The prediction accuracies were higher in Set 2 (including 6 populations) than Set 1 (only one biparental population) under irrigation and drought. The opposite was observed for the heat trial, all traits had higher prediction accuracies in Set 1 than Set 2.

The GS models predicted traits in the irrigated and drought trials moderately well, while performing slightly better for traits in the heat trial. GS accuracies were highest in all trials for DAYSFL, DAYSMT and DTHT. GRYLD was predicted better than PTHT in the irrigated and heat trials, while the opposite was found for the drought trial. Prediction accuracies varied depending on the applied GS model and data set. ELNET and RF were generally the better performing models (Table 3.5 The performance difference between ELNET, RF and the other three models was largest for DAYSFL, DAYSMT and DTHD. Depending on the GS model used, the prediction accuracies for DAYSFL ranged from 0.41 – 0.54 for irrigation, from 0.39 – 0.49 under drought and from 0.48 – 0.57 under heat stress in Set 2. The prediction accuracies for these traits were generally slightly lower in Set 1 under irrigation and drought. However, the heat trial still had the highest accuracies and the drought trial the lowest. The difference in prediction accuracies between the two sets was more pronounced in the irrigated and drought trials while both sets had similar accuracies in the heat trial. The GS accuracies were in the same range in both sets for DAYSMT and DTHD.

A similar trend was observed for PTHT. Generally, PTHT was poorly predicted across all trials. The highest prediction accuracies for PTHT were obtained with GS models RF, ELNET and RRBLUP. However, the predictive ability of GRYLD was different. While RF still performed best across all trials, ELNET was outperformed by RKHS and RRBLUP in both sets. Generally, RHKS competed very well with RF to predict GRYLD under all conditions. Prediction accuracies for GRYLD using different GS models ranged from 0.22 – 0.35, 0.21 – 0.32 and 0.28 – 37 for the irrigated, drought and heat trial respectively in Set 2. The prediction accuracies of GRYLD in Set 1 for the heat trial were slightly higher than those of Set 2. The opposite was observed for the irrigated and drought trial where higher prediction accuracies where obtained in Set 2.

Discussion

Synthetics contribute yield-promoting alleles

Our study confirms that synthetic derived wheat harbors yield-promoting alleles for wheat grown in different environments (Figure 3.2 and Figure 3.3). The agronomic performance and end-use quality of most synthetic derived lines are rather poor and in most cases, synthetics need to be backcrossed to elite bread wheat. The high yielding material identified here could be utilized for breeding in a broad variety of environments.

The obvious choices are to cross synthetic material with high yield under heat and or drought to elite bread wheat with good end-user quality but poor heat and or drought tolerance. However, the synthetic derived lines outperforming Opata M85 in all three environments are of particular interest. Lines performing well under stressed and optimal conditions could be suitable for several environments. These synthetics could be especially suitable for regions where wheat is grown in rain-fed farming systems with occasional periods of drought or heat.

Elite wheat lines derived from these synthetics would have a high yield potential under well-watered conditions and provide a "safety net" for the farmer in case of a period of heat or drought. However, the production of new synthetic lines and subsequent selection of material suitable for any elite bread-wheat breeding program is tedious and time consuming.

Potential of genomic selection in wheat pre-breeding

Genomic selection has the potential to increase genetic gain in plant breeding by reducing the time per breeding cycle significantly (Meuwissen *et al.* 2001; Heffner *et al.* 2009). GS is an interesting option for pre-breeding because of two reasons, (i) by estimating GEBVs we can reduce the years per cycle and (ii) synthetic derived material has a larger genetic variance then most elite lines and therefore, GS should work well for synthetic derived material. Even though GS prediction worked for our synthetic derived wheat, the prediction accuracies were slightly lower than we expected based on the high heritably for individual traits. The predictability of a GS model depends among other factors on the genetic architecture of the trait we are predicting. The underlying genetic architecture differs between traits and therefore, it is not surprising that different statistical models showed variable performance depending on the trait.

We found that prediction models ELNET and RF best predicted DAYSFL, DAYSMT and DTHD. The same models predicted PTHT best but the predictability of both models was lower than expected. Grain yield, in contrast, is controlled mostly by many loci of small effect. Large effect QTL are often confounded with QTL of grain-yield related traits such as plant height (Maccaferri *et al.* 2008). Here, GRYLD was predicted best by RF, RKHS and RRBLUP. Overall, we found that ELNET and RF are the models with the best predictive ability across all traits and trials in our synthetic derived material.

Another important aspect defining the predictive ability of a GS model is the population structure. Set 1 is comprised of only one biparental population and does *per se* not show any population structure. In Set 2, we included lines from 6 synthetic derived populations, which all share one common parent, Opata M85, and can be regarded as half-sibs. A recent study within a maize diversity panel identified population structure as sole source of prediction accuracy (Windhausen *et al.* 2012). Another study in maize reported expected prediction accuracies within full-sib families, whereas reduced accuracies for predictions within half-sibs. Noticeably, significantly better results were obtained if half-sibs in the TP represented both instead of just one parent (Riedelsheimer *et al.* 2013). According to Meuwissen *et al*. (2009) the prediction of GEBVs should be relatively simple if the TP and SC are from the same biparental population. Similar marker alleles will be found in the TP and SC (Meuwissen 2009). Several studies reported high predictions accuracies for GS in small biparental populations, in one case as small as 35 individuals (Bernardo and Yu 2007; Wong and Bernardo 2008; Heffner *et al.* 2011; Combs and Bernardo 2013). Results reported here are for GS performed across 6 biparental populations with one common parent (Set 2). We ran the analysis separately for population Synthetic 6 comprised of 242 biparental individuals (Set 1). Generally, prediction accuracies in Set 1 were lower than in Set 2 for the irrigated and drought trials, while the prediction accuracies were higher in Set 1 than Set 2 under heat. Furthermore, the smallest differences of prediction accuracies between the two sets were observed in the heat trial for all traits. GS accuracy is a function of the training population size N_{TP}, heritability h^2 on entry-mean basis, number of QTL underlying the trait of interest, the genetic architecture and the number of markers available N_M (Daetwyler *et al.* 2008; Daetwyler *et al.* 2010; Combs and Bernardo 2013).

Combs and Bernardo (2013) studied the effect of N_{TP}, h^2 and N_M in different types of populations and crops. They reported high GS prediction accuracies in a bi-parental barely (*Hordeum vulgare*) population of only 96 individuals with 223 polymorphic markers. Working with bi-parental populations in this study, we could expect higher prediction accuracy for yield based on the relatively heritability. However, it is likely that there are many confounding physiological factors in this population such that there are many combinations of alleles that lead to the same plant architecture and grain yield. Not having a globally optimal maximum in highly diverse populations such as these is likely to lower the accuracy of predicting complex traits.

The heritability for traits measured here, training population size, effective population size and marker number, are sufficient to obtain better prediction accuracies. The synthetic derived lines are very diverse and we observed large differences in plant types in the field. Our results suggest that GS works but might not perform as well as expected with synthetic derived wheat and other exotic materials, due to complex and confounding physiological effects. In an elite wheat breeding program it is likely that we have one global maximum with physiology and plant architecture for high yield. However, it is possible that there are multiple confounding physiologies and different architectures in the synthetic derived wheat. It appears that the GS models do have good predictive ability but are not able to adequately account for this complexity, leading to lower overall predictive ability. While promising, the complexity and extreme diversity may limit the use of GS for rapid introgression of favorable exotic alleles from wild relatives in wheat breeding.

Acknowledgements

The Graduate Research Assistantship of S. Dunckel is supported through the Monsanto Beachell-Borlaug International Scholars Program. This work was funded by the US Agency for International Development (USAID Cooperative Agreement No. AID-OAA-A-13-0005), the United States Department of Agriculture - Agricultural Research Service (Appropriations #5430-21000-006-00D and #3640-21220-021-00) and The Bill & Melinda Gates Foundation through the Durable Rust Resistance in Wheat project to Cornell University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. This work was done under the auspices of the Wheat Genetics Resource Center (WGRC) Industry/University Collaborative Research Center (I/UCRC) supported by NSF grant contract (IIP-1338897) and industry partners.

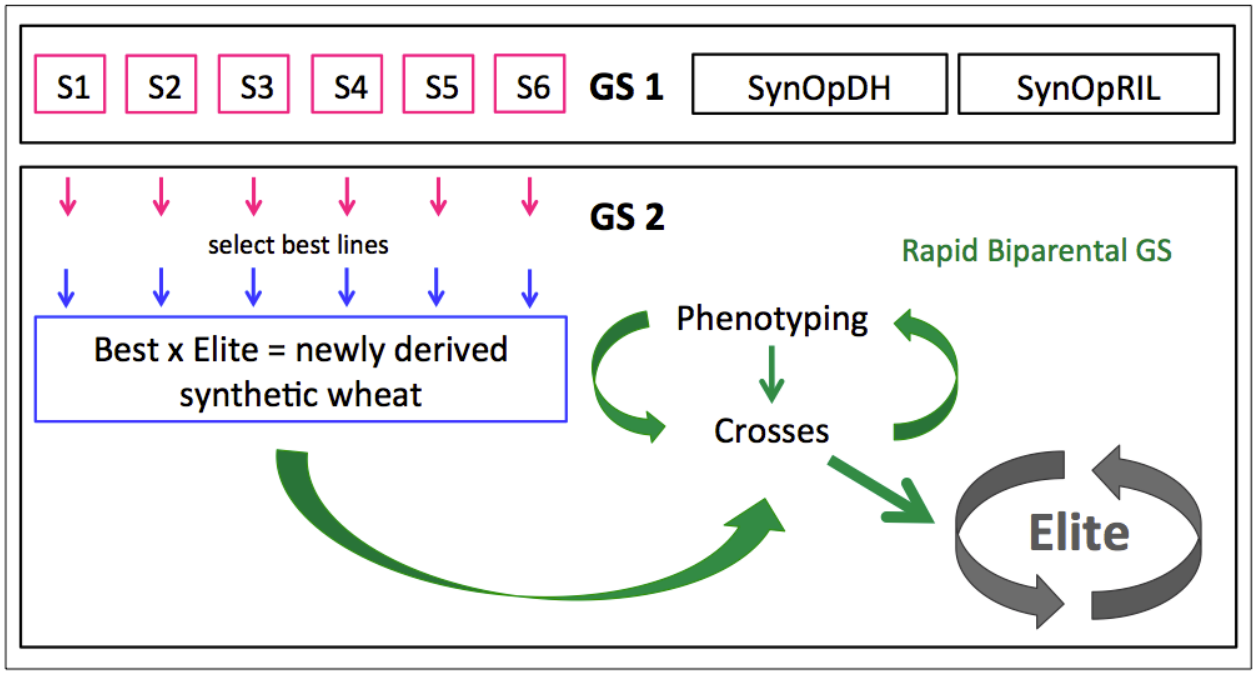

Figure 3.1 Breeding scheme for rapid cycle biparental genomic selection of exotic germplasm

Our six base populations are labeled pink. Population S6 has the same pedigree as SynOpDH and SynOpRIL, which were added to increase the initial population size for genomic selection. GS1 represent the first cycle of our breeding scheme. GS2 focuses on predicting newly derived synthetic material (blue) developed by crossing the top two lines of each population with three elite CIMMYT cultivars.

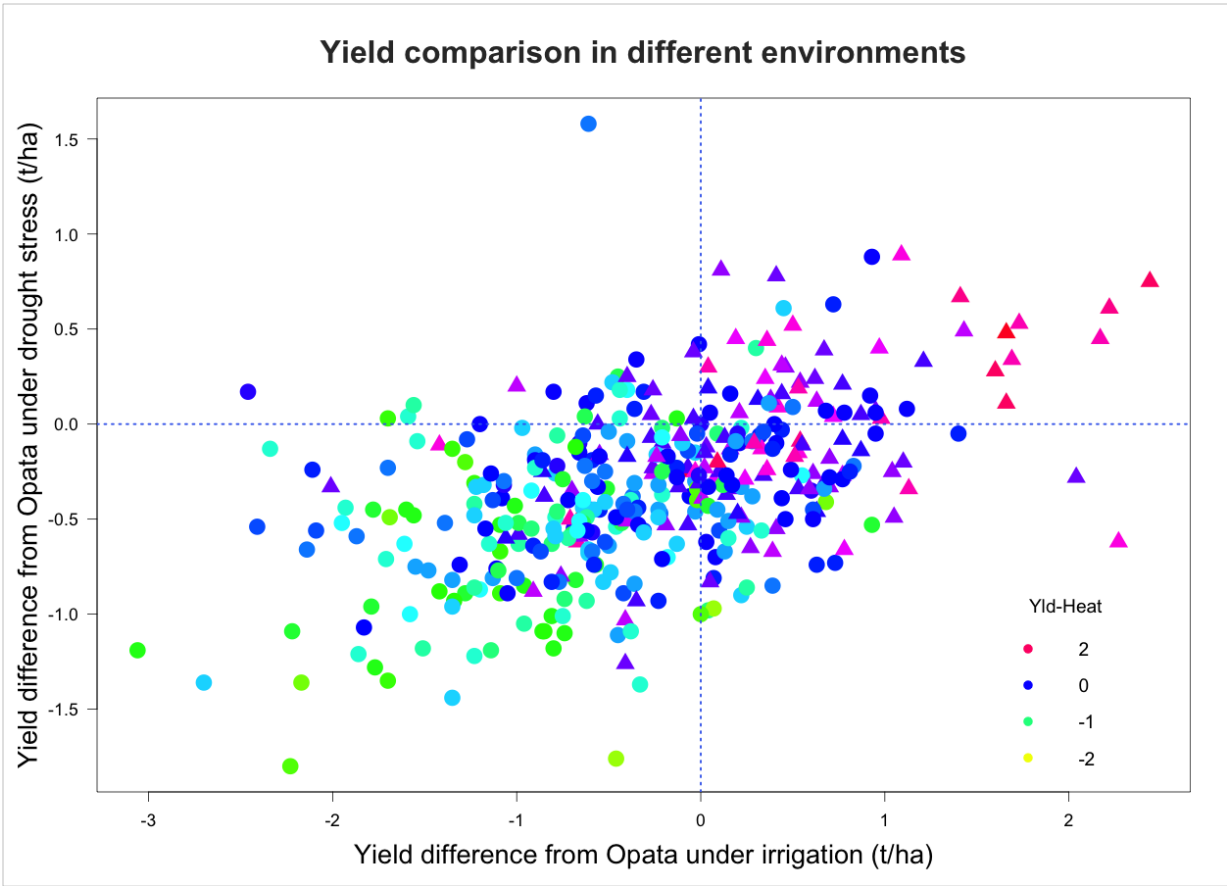

Figure 3.2 Yield comparison of synthetic derived lines with elite parent Opata M85 under heat, drought and irrigation

The x-axis is yield under irrigation, the y-axis yield under drought stress and the color gradient yield under heat stress. The blue lines indicate the yield of Opata M85 (set to 0) and triangles synthetic lines outperforming Opata M85 under heat stress. Yield differences are reported in t/ha.

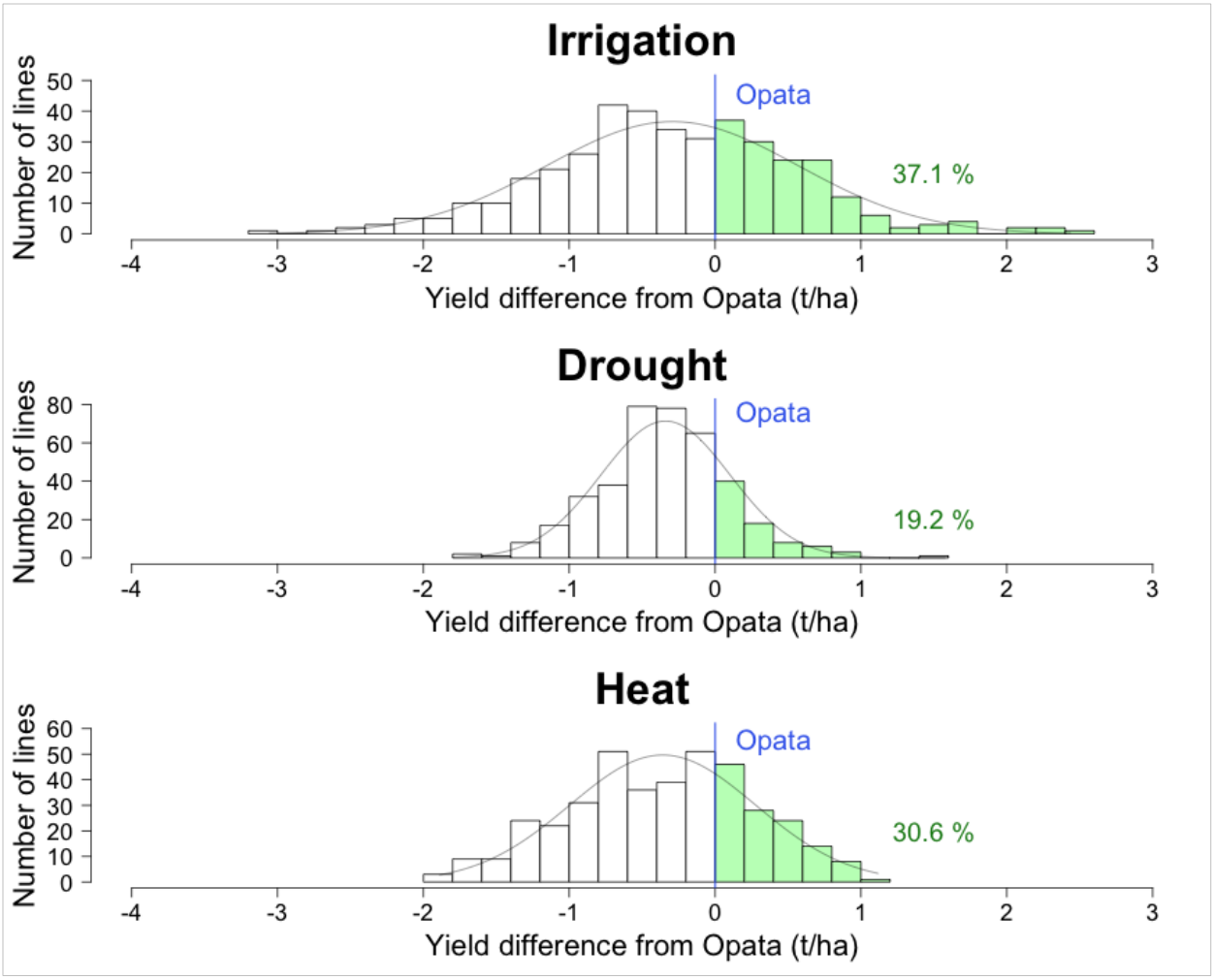

Figure 3.3 Distribution of synthetic derived line performance under heat, drought and irrigation

The blue lines indicate Opata M85 and yield differences are reported in t/ha. The green colored part of each histogram indicates the relative portion of liens outperforming Opata M85.

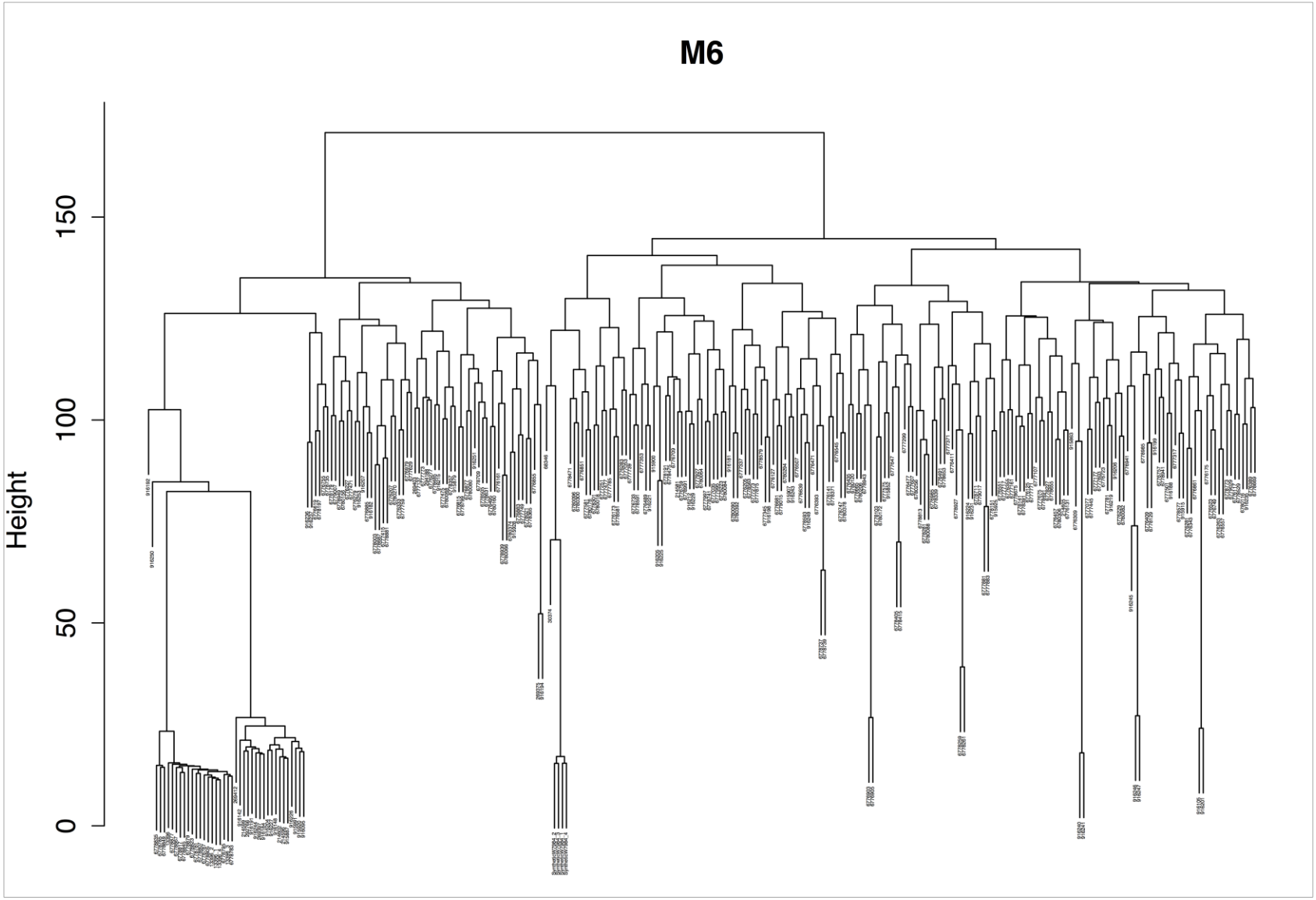

Figure 3.4 Dendrogram of a subset of SynOpRIL, SynOpDH and old M6

The two distinctive clusters on the left show 15 lines that are OpataM85 self and 18 lines that are not part of these populations.

Table 3.1 Plant material included in two years of yield trials

Population	Pedigree	Population size
Synthetic 1	CPI8/GEDIZ/3/GOO//ALB/CRA/4/AE.TAUSCHII (208)/5/OPATA	52
Synthetic 2	YAV_3/SCO//JO69/CRA/3/YAV79/4/AE.TAUSCHII (498)/5/OPATA	38
Synthetic 3	D67.2/P66.270//AE.TAUSCHII (257)/3/OPATA	31
Synthetic 4	GAN/AE.TAUSCHII (897)//OPATA	20
Synthetic 5	DOY1/AE.TAUSCHII(458)//OPATA	13
Synthetic 6	ALTAR 84/AE.TAUSCHII(219)//OPATA	242

Table 3.2 List of 15 possible Opata M85 selfs and 18 unrelated lines.

GID	Pedigree	Line identification	
6778781	ALTAR 84/AE.TAUSCHII(219)//OPATA	MU-1429-SYnOpRIL-1541	OpataM85 self
6778783	ALTAR 84/AE.TAUSCHII(219)//OPATA	MU-1430-SYnOpRIL-1542	OpataM85 self
6778785	ALTAR 84/AE.TAUSCHII(219)//OPATA	MU-1431-SYnOpRIL-1543	OpataM85 self
6778787	ALTAR 84/AE.TAUSCHII(219)//OPATA	MU-1432-SYnOpRIL-1544	OpataM85 self
6778795	ALTAR 84/AE.TAUSCHII(219)//OPATA	MU-1440-SYnOpRIL-1549	OpataM85 self
6778805	ALTAR 84/AE.TAUSCHII(219)//OPATA	MU-1457-SYnOpRIL-1556	OpataM85 self
6778807	ALTAR 84/AE.TAUSCHII(219)//OPATA	MU-1458-SYnOpRIL-1557	OpataM85 self
6778817	ALTAR 84/AE.TAUSCHII(219)//OPATA	MU-1472-SYnOpRIL-1563	OpataM85 self
6778823	ALTAR 84/AE.TAUSCHII(219)//OPATA	MU-1479-SYnOpRIL-1567	OpataM85 self
6778835	ALTAR 84/AE.TAUSCHII(219)//OPATA	MU-1496-SYnOpRIL-1575	OpataM85 self
6778845	ALTAR 84/AE.TAUSCHII(219)//OPATA	MU-1510-SYnOpRIL-1582	OpataM85 self
6778849	ALTAR 84/AE.TAUSCHII(219)//OPATA	MU-1512-SYnOpRIL-1584	OpataM85 self
6778867	ALTAR 84/AE.TAUSCHII(219)//OPATA	MU-1538-SYnOpRIL-1596	OpataM85 self
6778877	ALTAR 84/AE.TAUSCHII(219)//OPATA	MU-1550-SYnOpRIL-1603	OpataM85 self
6778879	ALTAR 84/AE.TAUSCHII(219)//OPATA	MU-1552-SYnOpRIL-1604	OpataM85 self
214596	Unknown	214596	Self
214599	Unknown	214599	Self
214602	Unknown	214602	Self
269397	Unknown	269397	Self
269412	Unknown	269412	Self
292766	Unknown	292766	Self
915887	Unknown	915887	Self
915920	Unknown	915920	Self
915991	Unknown	915991	Self
916095	Unknown	916095	Self
916142	Unknown	916142	Self
916148	Unknown	916148	Self
916184	Unknown	916184	Self
916191	Unknown	916191	Self
916204	Unknown	916204	Self
916226	Unknown	916226	Self
916227	Unknown	916227	Self
916244	Unknown	916244	Self

GBS analysis identified 33 lines that are either Opata M85 selfs or selfs of a more distantly related line to SynOpDHs, SynOpRILs and old M6. The lines were excluded from further analysis.

Table 3.3 Phenotypic correlation across two years per trait and environment

Trial		DAYSFL	DAYSMT	DTHD	GRYLD	PTHT
Drought	Corr	0.65	0.78	0.73	0.38	-0.09
Heat	Corr	0.73	0.71	0.74	0.78	0.55
Irrigation	Corr	0.80	0.70	0.81	0.53	0.61

Table 3.4 Broad sense heritability (H^2) on an entry mean basis, grand mean, least significant difference (LSD) and coefficient of variation (CV) per trait and environment

Trial		DAYSFL	DAYSMT	DTHD	GRYLD	PTHT
Drought	**H^2**	**0.88**	**0.95**	**0.91**	**0.42**	**0.57**
	mean	78.30	106.73	75.56	1.83	69.86
	LSD.05	3.41	2.80	3.17	0.64	8.15
	CV	2.22	1.34	2.14	17.98	5.95
Heat	**H^2**	**0.98**	**0.95**	**0.98**	**0.94**	**0.78**
	mean	59.21	87.35	56.37	1.64	60.44
	LSD.05	1.80	2.60	1.69	0.41	6.04
	CV	1.55	1.52	1.53	12.83	5.09
Irrigation	**H^2**	**0.93**	**0.84**	**0.93**	**0.76**	**0.78**
	mean	85.72	125.73	81.18	5.10	100.13
	LSD.05	4.71	4.79	4.60	1.07	7.26
	CV	2.80	1.94	2.89	10.65	3.70

Table 3.5 Genomic selection prediction accuracies

Environment	GS model	DTHD		DAYSFL		DAYSMT		GRYLD		PTHT	
		Set 1	Set 2	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2
Irrigation	**H^2**	**0.93**	**0.93**	**0.93**	**0.93**	**0.84**	**0.84**	**0.76**	**0.76**	**0.78**	**0.78**
	RRBLUP	0.34	0.41	0.30	0.43	0.33	0.42	0.20	0.28	0.10	0.12
	RHKS	0.34	0.40	0.25	0.41	0.29	0.39	0.13	0.31	0.06	0.10
	PLSR	0.31	0.42	0.28	0.41	0.28	0.42	0.19	0.22	0.11	0.10
	ELNET	0.55	0.54	0.52	0.54	0.41	0.47	0.16	0.27	0.14	0.12
	RF	0.53	0.52	0.47	0.51	0.47	0.52	0.24	0.35	0.18	0.21
Drought	**H^2**	**0.91**	**0.91**	**0.88**	**0.88**	**0.95**	**0.95**	**0.42**	0.42	**0.57**	**0.57**
	RRBLUP	0.30	0.36	0.30	0.39	0.29	0.42	0.15	0.23	0.27	0.21
	RHKS	0.26	0.33	0.27	0.38	0.27	0.39	0.15	0.24	0.27	0.18
	PLSR	0.25	0.31	0.24	0.32	0.27	0.40	0.12	0.21	0.22	0.18
	ELNET	0.43	0.46	0.40	0.48	0.31	0.50	0.13	0.21	0.25	0.29
	RF	0.41	0.48	0.43	0.49	0.38	0.50	0.18	0.32	0.31	0.31
Heat	**H^2**	**0.98**	**0.98**	**0.98**	**0.98**	**0.95**	**0.95**	**0.94**	0.94	**0.78**	**0.78**
	RRBLUP	0.54	0.51	0.54	0.53	0.51	0.52	0.37	0.33	0.27	0.17
	RHKS	0.52	0.50	0.52	0.51	0.47	0.50	0.37	0.35	0.26	0.15
	PLSR	0.51	0.46	0.51	0.48	0.47	0.47	0.34	0.28	0.22	0.14
	ELNET	0.67	0.57	0.66	0.57	0.61	0.59	0.34	0.31	0.24	0.18
	RF	0.55	0.55	0.57	0.55	0.57	0.55	0.38	0.37	0.28	0.25

Accuracies of five different GS models (RRBLUP, RKHS, PLS, ELNET and RF), three environments and five traits (DAYSFL, DAYSMT, DTHD, GRYLD and PTHT) in Set 1 (gray) and Set 2. All correlations are significantly different from 0 with p-values < 0.001 at significance levels $\alpha = 0.05$ and $\alpha = 0.01$

References

Arraiano, L.S., A.J. Worland, C. Ellerbrook, and J.K.M. Brown, 2001 Chromosomal location of a gene for resistance to septoria tritici blotch (Mycosphaerella graminicola)in the hexaploid wheat 'Synthetic 6x'. Theoretical and Applied Genetics 103 (5):758-764.

Bernardo, R., and J. Yu, 2007 Prospects for Genomewide Selection for Quantitative Traits in Maize. Crop Science 47 (3):1082-1090.

Bernardo, R. 2009 Genomewide Selection for Rapid Introgression of Exotic Germplasm in Maize. Crop Science 49: 419-425.

Bernardo, R., 2010 Breeding for Quantitative Traits in Plants. Woodbury, Minnesota: Stemma Press.

Browning, B.L., and S.R. Browning, 2013 Improving the Accuracy and Efficiency of Identity-by-Descent Detection in Population Data. Genetics 194 (2):459-471.

Chapman, J.A., M. Mascher, A. Buluç, K. Barry, E. Georganas, *et al.* 2015 A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome. Genome Biology 16: 26.

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

Crossa, J. and C.O. Gardner, 1987 Introgression of an Exotic Germplasm for Improving an Adapted Maize Population. Crop Science 27: 187-190.

Daetwyler, H.D., B. Villanueva, and J.A. Woolliams, 2008 Accuracy of Predicting the Genetic Risk of Disease Using a Genome-Wide Approach. PLoS ONE 3 (10):e3395.

Daetwyler, H.D., R. Pong-Wong, B. Villanueva, and J.A. Woolliams, 2010 The Impact of Genetic Architecture on Genome-Wide Evaluation Methods. Genetics 185 (3):1021-1031.

Dvorak, J., M.C. Luo, Z.L. Yang, and H.B. Zhang, 1998 The structure of the Aegilops tauschii genepool and the evolution of hexaploid wheat. Theoretical and Applied Genetics 97 (4):657-670.

Ehlers, J.D. and K.W. Foster, 1993 Introgression of agronomic characters from exotic cowpea germplasm into blackeye bean. Field Crops Research 35: 43-50.

Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto *et al*., 2011 A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. PLoS ONE 6 (5):e19379.

Endelman, J.B., 2011 Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. Plant Genome J. 4 (3):250.

Federer, W.T., 1956 Augmented (or hoonuiaku) designs. Hawaiian Planter's Records 55 (191-208).

Feuillet, C., P. Langridge and R. Waugh, 2008 Cereal breeding takes a walk on the wild side. Trends in Genetics 24: 24-32.

Friedman, J., T. Hastie, and R. Tibshirani, 2010 Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of Statistical Software 33 (1):1-22.

Gaynor, R., 2015 Genomic Selection for Kansas Wheat, pp. 57 in Department of Agronomy. Kansas State University Manhattan, Kansas, USA.

Gilmour, A.R., B.J. Gogel, B.R. Cullis, and R. Thompson, 2009 ASReml user guide release 3.0. Hemel Hempstead, UK: VSN Intl. Ltd.

Heffner, E.L., M.E. Sorrells, and J.L. Jannink, 2009 Genomic Selection for Crop Improvement. Crop Science 49 (1):1-12.

Heffner, E.L., A.J. Lorenz, J.L. Jannink, and M.E. Sorrells, 2010 Plant Breeding with Genomic Selection: Gain per Unit Time and Cost. Crop Science 50 (5):1681-1690.

Heffner, E.L., J.-L. Jannink, and M.E. Sorrells, 2011 Genomic Selection Accuracy using Multifamily Prediction Models in a Wheat Breeding Program. Plant Gen. 4 (1):65-75.

Heslot, N., H.-P. Yang, M.E. Sorrells and J.-L. Jannink, 2012 Genomic Selection in Plant Breeding: A Comparison of Models. Crop Science 52: 146-160.

Heslot, N., D. Akdemir, M. Sorrells and J.-L. Jannink, 2014 Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. Theoretical and Applied Genetics 127: 463-480.

Jordan, K., S. Wang, Y. Lun, L.-J. Gardiner, R. MacLachlan *et al.*, 2015 A haplotype map of allohexaploid wheat reveals distinct patterns of selection on homoeologous genomes. Genome Biology 16 (1):1-18.

Kihara, H., 1944 Discovery of the DD-analyser, one of the ancestors of Triticum vulgare. Agric Hort 19:889-890.

Liaw, A., and M. Wiener, 2002 Classification and Regression by randomForest. R News 2 (3):18-22.

Lorenz, A.J., S. Chao, F.G. Asoro, E.L. Heffner, T. Hayashi et al., 2011 Chapter Two - Genomic Selection in Plant Breeding: Knowledge and Prospects, p. 77-123 in Advances in Agronomy, edited by L. S. Donald, Academic Press.

Maccaferri, M., M.C. Sanguineti, S. Corneti, J.L.A. Ortega, M.B. Salem et al., 2008 Quantitative Trait Loci for Grain Yield and Adaptation of Durum Wheat (*Triticum durum* Desf.) Across a Wide Range of Water Availability. Genetics 178: 489-511.

Marcussen, T., S.R. Sandve, L. Heier, M. Spannagl, M. Pfeifer *et al.*, 2014 Ancient hybridizations among the ancestral genomes of bread wheat. Science 345 (6194).

McFadden, E.S., and E.R. Sears, 1946 The origin of Triticum spelta and its free-threshing hexaploid relatives. J Hered 37:81-89.

Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard, 2001 Prediction of total genetic value using genome-wide dense marker maps. Genetics 157 (4):1819-1829.

Meuwissen, T.H.E., 2009 Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. Genetics Selection Evolution 41 (35).

Mevik, B.H., and R. Wehrens, 2007 The pls Package: Principal Component and Partial Least Squares Regression in R. Journal of Statistical Software 18 (2):1-24.

Mujeeb-Kazi, A., R. Delgado, A. Cortes, S. Cano, V. Rosas *et al.*, 2004 Progress in exploiting Aegilops tauschii for wheat improvement. Annual Wheat News Letter 50:79-88.

Narasimhamoorthy, B., B.S. Gill, A.K. Fritz, J.C. Nelson and G.L. Brown-Guedira, 2006 Advanced backcross QTL analysis of a hard winter wheat x synthetic wheat population. Theor Appl Genet 112: 787-796.

Ochanda, N., J. Yu, P.J. Bramel, A. Menkir, M.R. Tuinstra *et al*., 2009 Selection before backcross during exotic germplasm introgression. Field Crops Research 112: 37-42.

Paradis, E., J. Claude, and K. Strimmer, 2004 APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289-290.

Poland, J., J. Endelman, J. Dawson, J. Rutkoski, S. Wu *et al.*, 2012a Genomic Selection in Wheat Breeding using Genotyping-by-Sequencing. Plant Gen. 5 (3):103-113.

Poland, J.A., P.J. Brown, M.E. Sorrells, and J.-L. Jannink, 2012b Development of High-Density Genetic Maps for Barley and Wheat Using a Novel Two-Enzyme Genotyping-by-Sequencing Approach. PLoS ONE 7 (2):e32253.

R Core Team, 2014 R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Reif, J.C., P. Zhang, S. Dreisigacker, M.L. Warburton, M. van Ginkel *et al.*, 2005 Wheat genetic diversity trends during domestication and breeding. Theoretical and Applied Genetics 110 (5):859-864.

Riedelsheimer, C., J.B. Endelman, M. Stange, M.E. Sorrells, J.-L. Jannink *et al*., 2013 Genomic Predictability of Interconnected Biparental Maize Populations. Genetics 194: 493-503

Rife, T., and J. Poland, 2014 Field Book: An open-source application for field data collection on Android. Crop Science 54 (4):1624-1627.

Rutkoski, J.E., J.A. Poland, R.P. Singh, J. Huerta-Espino, S. Bhavani *et al.*, 2014 Genomic Selection for Quantitative Adult Plant Stem Rust Resistance in Wheat. The Plant Genome 7 (3).

Sorrells, M.E., J.P. Gustafson, D. Somers, S. Chao, D. Benscher *et al.*, 2011 Reconstruction of the synthetic W7984 x Opata M85 wheat reference population. Genome 54 (11):875-882.

Spindel, J., H. Begum, D. Akdemir, P. Virk, B. Collard *et al.*, 2015 Genomic Selection and Association Mapping in Rice (*Oryza sativa*): Effect of Trait Genetic Architecture, Training Population Composition, Marker Number and Statistical Model on Accuracy of Rice Genomic Selection in Elite, Tropical Rice Breeding Lines. PLoS Genet 11 (2):e1004982.

Talbert, L.E., L.Y. Smith, and N.K. Blake, 1998 More than one origin of hexaploid wheat is indicated by sequence comparison of low-copy DNA. Genome 41 (3):402-407.

Warburton, M.L., J. Crossa, J. Franco, M. Kazi, R. Trethowan *et al.*, 2006 Bringing wild relatives back into the family: recovering genetic diversity in CIMMYT improved wheat germplasm. Euphytica 149 (3):289-301.

Williams, E.R., J.A. John, and D. Whitaker, 2006 Construction of resolvable spatial row-column designs. Biometrics
62 (1):103-108.

Windhausen, V.S., G.N. Atlin, J.M. Hickey, J. Crossa, J.-L. Jannink *et al.,* 2012 Effectiveness of Genomic Prediction of Maize Hybrid Performance in Different Breeding Populations and Environments. G3: Genes|Genomes|Genetics 2: 1427-1436.

Wong, C.K., and R. Bernardo, 2008 Genomewide selection in oil palm: increasing selection gain per unit time and cost with small populations. Theoretical and Applied Genetics 116 (6):815-824.

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

Zamir, D. 2001. Improving plant breeding with exotic genetic libraries. Nat Rev Genet 2: 983-989.

Zhang, P., S. Dreisigacker, A.E. Melchinger, J.C. Reif, A.M. Kazi *et al.*, 2005 Quantifying novel sequence variation and selective advantage in synthetic hexaploid wheats and their backcross-derived lines using SSR markers. Molecular Breeding 15 (1):1-10.

Zhang, X., P. Perez-Rodriguez, K. Semagn, Y. Beyene, R. Babu *et al.*, 2015 Genomic prediction in biparental tropical maize populations in water-stressed and well-watered environments using low-density and GBS SNPs. Heredity (Edinb) 114 (3):291-299.

Chapter 4 - QTL mapping for improved heat tolerance of bread wheat in Kansas

Sandra Dunckel, Allan Fritz, P.V. Vara Prasad, Jesse Poland

Abstract

Wheat (*Triticum aestivum* L.) is the major cereal crop consumed in many regions of the world. Estimated population growth and climate change present major challenges to agriculture. Heat stress induces pollen sterility and seed abortion resulting in lower seed weight and seed number. To increase heat tolerance in bread wheat in Kansas, several new mapping populations have been developed through single seed decent (SSD). The most promising population with pedigree Overley/Jefimija has been advanced to $F_{5:6}$ recombinant inbred lines (RILs) and used for a growth chamber experiment. Line performance of 203 RILs was assessed under heat stress and optimal conditions and the data used for quantitative trait loci (QTL) analysis. We identified 13 QTL under optimal conditions and 11 QTL under heat stress for biomass, days to heading, grain weight, grain number and thousand-kernel weight. Heat tolerant parent Jefimija contributed all alleles increasing grain related traits under heat stress. Two "hotspots" controlling traits related to heat tolerance were identified on chromosomes 2DS and 5A. The QTL on chromosome 5A might be a functional gene controlling multiple traits. Multiple QTL in very close proximity were mapped on chromosome 2DS. We compared the location of our GBS markers on 2DS with *Ppd-D1* and *Rht8* and concluded that the QTL we identified are *Rht8*, and possibly *Ppd-D1*.

Introduction

Wheat (*Triticum aestivum* L.) is the major cereal crop consumed in many regions of the world providing over 20% of all calories consumed globally (FAO 2015; Shiferaw et al. 2013). The global population is estimated to reach 9.1 billion people by 2050. Together with climate change, this presents a major challenge to agriculture. In the Unites States (U.S.), most wheat is produced in the Great Plains. Estimated long-term climate trends show an increase of temperature across the Great Plains by 2100 of 1.5 – 7 °C depending on emission scenario and climate model (IPCC 2014). The total U.S. wheat production was ~58M tons in 2013 (~8.15% of the global wheat production) (FAO 2015). The production of hard red winter wheat accounts for ~40% of the total U.S. wheat production and is grown primarily from Texas through Montana.

Heat stress during the reproductive development of wheat is a major constraint to wheat production. Reduced photosynthesis and premature senescence are observed in heat susceptible wheat. As a result, yield is reduced by induced pollen sterility and seed abortion leading to lower seed weight, flour yield and quality (Hays et al. 2007). Heat stress can occur either chronically over a long period of time (mean temperature 18 – 25 °C, max. 32 °C during grain filling) or by an abrupt heat-shock (temperatures greater than 32 °C) (Wardlaw and Wrigley 1994). With every 1°C rise above 15 – 20°C yield of heat susceptible wheat decreases 3 – 4%, and over 50% of the total yield can be lost if temperatures reach above 32 – 38 °C (Paulsen 1994; Wardlaw et al. 1989). Heat tolerant wheat is able to maintain photosynthesis and high levels of chlorophyll despite increased temperature. This results in higher number of grains per spike, stem carbohydrate reserves, grain weight and extended period of grain filling (Yang et al. 2002a).

Adaption of wheat to higher temperatures through changes in agronomic practices will not suffice to avoid negative impacts on wheat yield. New cultivars adapted to higher temperatures need to be developed. Combining heat tolerant varieties with varieties harboring

other desirable traits, such as disease resistance and good end-user quality, are a promising strategy of adaptation. Heat tolerance is quantitatively inherited and associated with multiple other traits such as grain weight, grain number, thousand-kernel weight, grain filling duration and chlorophyll content (Hays et al. 2007; Talukder et al. 2014; Yang et al. 2002a, b). Quantitative trait loci (QTL) mapping studies are useful to identify QTL and genes underlying traits for heat tolerance. Identified genetic markers can be used in plant breeding through marker-assisted selection (MAS) to select material harboring genes and QTL improving heat tolerance of wheat.

In this study, we describe the development of new mapping populations and a growth chamber experiment assessing the heat tolerance of the most promising populations with pedigree Overley/Jefimija.

Materials and Methods

Development of mapping populations

Overley is a high yielding, early maturing, semi-dwarf hard red winter wheat cultivar released in 2003 lacking heat tolerance. Several Eastern European wheat cultivars with high heat tolerance were identified by Ristic et al. (2008). Jefimija was the most heat tolerant cultivar. Multiple crosses between Overley and heat tolerant Eastern European varieties Jefimija and Proteinka were made in 2007. Other crosses included heat susceptible hard red winter wheat cultivar Karl 92 and hard white Australian heat tolerant cultivar Ventnor. F_2 seeds from eight different crosses were available and advanced to recombinant inbred lines (RILs) from 2011 – 2015 through single seed decent (SSD) (Table 4.1). All populations were advanced to F_3 in the greenhouse in 2012 and the four most promising with pedigrees Overley/Jefimija (populations U6019 and U6020), Overley/Proteinka, Karl92/(Karl92/Ventnor RIL 73) and Overley/(Karl92/Ventnor RIL

73) were further advanced in head-rows in the field. Furthermore, populations U6019 and U6020 with pedigree Overley/Jefimija were advanced at the same time in the greenhouse to $F_{5:6}$ RILs for a growth chamber experiment. This pedigree was considered the most promising and chosen for a growth chamber experiment to map QTL for heat tolerance.

Plant material for QTL mapping study

Two biparental mapping populations, U6019 and U6020, with pedigree Overley/Jefimija were used for this QTL mapping study. The mapping populations were developed by advancing F_2 plants through SSD in the greenhouse to $F_{5:6}$ RILs. The population size for the heat chamber experiment was reduced by random selection of lines. Population U6019 consisted of 103 $F_{5:6}$ and U6020 of 100 $F_{5:6}$. Both populations were characterized for heat tolerance by comparing line performance under optimal and heat stress conditions.

Experimental Design

Seeds where germinated in 5x5 cm pots containing potting soil (Metro Mix; Hummert Intl, Topeka, KS) in a greenhouse in January 2014. Ten-day old seedlings where vernalized at 4°C for 8 weeks. Subsequent to vernalization, each seedling was transplanted into its own pot of diameter 6.4 cm and depth 25.4 cm. Plants were grown in a greenhouse under optimal conditions and watered regularly to avoid drought stress. Fertilizer, fungicides and systemic insecticides were applied as needed to avoid malnutrition and infestation by fungi or insects. Wheat grown in Kansas is exposed mostly to post-anthesis heat stress. We started our experiment 14 days after heading to assess post-anthesis heat stress tolerance. Plants were tagged at heading of the first spike (defined as main spike, growth stage Feekes 10.3 (Miller 1992)) and moved to growth chambers 14 days after heading. Grouping the plants based on days to heading (DTHD) accounts for genetic variation of DTHD and enables direct comparisons of line performance by reducing

confounding effects of flowering time. The plants were randomly allocated to four growth chambers based on a randomized complete block design (RCBD) with two replicates per temperature treatment. Two temperature treatments were applied: (i) high temperature simulating heat stress, and (ii) optimal temperature as control. Each genotype was included three times per treatment and replicate.

The control group was maintained under optimal growing conditions at day/night temperature of 21/17°C ±1.5°C with 16-hour photoperiod, relative humidity of 70-90%, and light intensity of approximately 1000 $\mu mols^{-2}s^{-1}$. Plants under heat stress were exposed to day/night temperature of 36/30°C ±2.0°C, same settings otherwise. All pots were placed in trays containing 2-3 cm water at all times to avoid dehydration of plants and possible introduction of drought stress. Heat stress was not limited to a set number of days and plants remained in their respective growth chamber until harvest ready (Feekes 11.4).

Assessment of heat tolerance

We collected data on several traits to assess heat tolerance in both mapping populations. Days to heading (DTHD) were collected at heading of the main spike. Tiller number (TINB) and shoot dry biomass (BIOMASS) were counted and measured on a per plant basis, respectively. Data on grain weight (GRWT), grain number (GRNB) and thousand-kernel weight (TKW) was collected separately for the main spike (GRWT_S01, GRNB_S01, TGW_S01) and the rest of the plant (GRWT_S02, GRNB_S02, TGW_S02). 'S01' in the trait name indicates data collected on the main spike and 'S02' data from the rest of the plant. Data was summarized to obtain a whole plant assessment (no addition to trait name). State of the art scale and seed counters were used to weigh (g) and count grain (Kirigwi et al. 2007; Pinto et al. 2010; Wang et al. 2009).

Genotyping

We genotyped all lines with genotyping-by-sequencing (Elshire et al. 2011) using the two-enzyme approach by Poland et al. (2012). Single-nucleotide polymorphisms (SNPs) were called simultaneously for both populations. Briefly, 14,914 unique SNPs with up to 80% missing data were optioned through GBS. SNPs with minor allele frequency (MAF) > 0.2 and less than 50% missing data were retained in the data set. Their tags were aligned to the recently published draft sequence of the wheat genome (The International Wheat Genome Consortium, 2014) using POPSEQ (Chapman et al. 2015). Imputation of the ordered marker data was performed using Beagle v.4 (Browning and Browning 2013). The parameters were set as described by Jordan et al. (2015): window = 5,000 overlap = 500 burnin-its =10 impute-its = 10. A total of 3,500 SNPs for U6019 and 3,280 SNPs for U6020 were used for QTL mapping.

Statistical analysis

The data were analyzed separately for the heat and optimal treatment using JMP Pro 11 Statistical Software (JMP® 1989-2015). The model used to calculate BLUEs was $y_{ij} = \mu + g_i + r_j + e_{ij}$ where y_{ij} is the trait, g_i is the fixed effect for each genotype, r_j is the random effect of the jth replicate, and e_{ij} is the random error with N(0, σ_e^2). Heritability on an entry means basis was calculated as $H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_e^2}{r}}$ where σ_g^2 is the genotypic variance, r is the number of replicates and σ_e^2 is the error variance (Bernardo 2010).

QTL analysis

QTL mapping was performed separately for the heat and control treatment in the R software environment (R Core Team, 2014) applying the R-package R/qtl (Broman K.W, Wu H., et al., 2003). The same methods were applied as described in the stem rust mapping paper by

Dunckel et al., (2015) and the ionomics study described in Chapter 3 of this dissertation. Briefly, QTL were mapped using Single Interval Mapping (SIM) and Composite Interval Mapping (CIM). The most significant markers were identified through stepwise regression. CIM was implemented applying a Haley-Knott regression using forward selection of marker covariates and a window size of 10 cM for all traits. Multiple QTL Mapping (MQM) was used to confirm identified QTL, refine their position, obtain estimated QTL effects and estimated phenotypic variance explained by a QTL (Arends et al., 2010). Furthermore, MQM was applied to identify additional QTL not mapped by SIM and CIM. The allelic state of the markers with the highest LOD score at each QTL was used to represent the allelic state of the QTL. The genome-wide logarithm of the odds value (LOD) for declaring a QTL was determined by 1,000 permutations. The parental alleles for Overley and Jefimija were coded as -1 and 1 respectively (Broman and Saunak, 2009).

Results

To test the repeatability of the experiment heritability on an entry mean basis, mean, standard deviation (Std. Dev.) and coefficient of variation (CV) were calculated for all traits and summarized in Table 4.2. Heritability differs between traits and treatments and was generally higher for data collected on the main spike. The phenotypic distribution of BLUEs shows a normal distribution in both treatments satisfying the condition of continuous traits for QTL mapping (Figure 4.1 and Figure 4.2). The results of the correlation analysis show multiple highly correlated traits. BIOMASS and DTHD were highly correlated in both trials, as well as with most grain related traits (Figure 4.1 - Figure 4.4). GRNB and TGW are, as expected, negatively correlated in both treatments.

We identified 13 QTL for different traits under optimal conditions and 11 QTL under heat stress. More QTL were mapped in U6019 than in U6020 in both treatments, however, several QTL were overlapping across populations and treatments. QTL were mapped applying SIM and CIM and their position refined applying MQM. Several additional QTL were identified through MQM (Table 4.3 and Table 4.4). The LOD profiles for U6019 (Figure 4.5) and U6020 (Figure 4.6) are based on LOD scores obtained through CIM.

We mapped eight QTL for DTHD. Four QTL were identified in U6019 on chromosomes 1BL, 2BL and 5A under optimal conditions explaining 55.51 % of the estimated phenotypic variance. Three QTL on 1DL, 2BL and 5A were mapped under heat stress explaining 36.98% of the estimated phenotypic variance. The eighth QTL for DTHD was mapped on chromosome 2DS in U6020 under heat stress. The estimated allele effect of the allele contributed by Overley was, as expected, negative for most QTL for DTHD.

We identified four QTL for BIOMASS. One QTL was mapped on chromosome 2DS in U6020 explaining 30.29% and 22.08% of the estimated phenotypic variance under optimal and heat conditions, respectively. Three QTL were mapped on chromosomes 2DS, 5A and 6BS in U6019 under heat stress explaining 36.98% of the estimated phenotypic variance.

We mapped two QTL for GRNB in U6019 under optimal conditions on chromosomes 3B and 5AL. Together they explain 31.73% of the estimated phenotypic variance. Based on the allele effect, the QTL on 3B is conferred by Overley and the QTL on 5AL by Jefimija. A QTL for GRNB_S02 was mapped on 3B and is likely the same QTL as reported above for GRNB. One QTL for TGW_S01 was identified on 4BL in U6019 under optimal and heat stress conditions. The QTL mapped under optimal conditions is located at 71.0cM, conferred by Jefimija and explains 19.19% of the estimated phenotypic variance. Under heat stress, the

mapped QTL was located in close proximity at 62.0cM. This QTL explains 14.12% of the phenotypic variance and was also contributed by Jefimija, indicating that they are most likely the same QTL.

Four QTL for grain related traits were mapped in U6020 under optimal conditions. We mapped QTL for GRWT on chromosomes 2DS and 5A. The positive allele of the QTL on 2DS explaining 18.85% of the estimated phenotypic variance was contributed by Overley. The allele increasing GRWT at the QTL on 5AL was conferred by Jefimija and explained 17.30% of the phenotypic variance. Both QTL identified on 5BL for TGW and TGW_S01 mapped to the same GBS marker KSUheat9427 at 43.1cM and are the same. The allele was contributed by Overley explaining 35.11% of the estimated phenotypic variance. Two QTL for GRNB_S01 and GRWT_S01 were mapped on chromosome 5A under heat stress and mapped to GBS marker KSUheat819 at 54.0cM. This QTL was conferred by Jefimija and explained 16.85% and 19.65% of the estimated phenotypic variance of GRNB_S01 and GRWT_S01, respectively.

Discussion

Mapping populations for heat trials available

We developed four RIL mapping populations through SSD in head-rows in the field (Table 4.1). In summer 2015, we harvested one F_5 head per line and will advance all populations one more generation. All lines will be harvested in bulk 2016 and used for yield trials with $F_{6:7}$ materials. Furthermore, we are seed increasing of all parents to include them as replicated checks in yield trials. Data on days to heading (DTHD) and plant height (PTHT) was collected on F_4 plants and will be recollected during season 2015/2016.

QTL for heat tolerance

We identified multiple QTL in both populations under optimal and heat stress conditions for traits related to heat tolerance in wheat. Several QTL for GRNB, GRWT and TGW were mapped on chromosomes 2DS, 3B, 4BL, 5A, 5AL, and 5BL (Table 4.3 and Table 4.4). Eastern European parent Jefimija conferred alleles with positive estimated effect for QTL on 4BL, 5A and 5AL. Overley contributed favorable alleles for QTL on 2DS, 3B and 5BL. However, heat tolerant parent Jefimija contributed all alleles positively associated with GRNB, GRWT and TGW under heat stress (Table 4.4). Other studies have reported QTL for GRNB, GRWT and TGW on chromosomes 3B, 4BL, 5A, 5AL and 5BL (Pinto et al. 2010; Su et al. 2009; Wu et al. 2012; Xu et al. 2014).

We identified two pleiotropic QTL, or "hotspots", on chromosomes 2DS and 5AL. The QTL identified in U6019 on chromosome 5A for BIOMASS under heat stress and DTHD under optimal conditions, was mapped to GBS marker KSUheat13739 at 62.4cM. The QTL for DTHD in U6019 under heat stress was located at 61.8cM and is most likely the same. In U6020, a QTL for GRWT_S01 under optimal conditions explaining 17.30% of the estimated phenotypic variance was also mapped to GBS marker KSUheat13739. Furthermore, the QTL for GRNB_S01 and TGW_S01 identified under heat stress in U6020 is located at 54cM and in close proximity to the QTL described above. Several other studies reported QTL for BIOMASS, GRNB and GRWT on chromosome 5A (Pinto et al. 2010b; Su et al. 2009; Wu et al. 2012; Xu et al. 2014). In all instances, heat tolerant parent Jefimija contributed the alleles related with yield increase. We conclude that this might be a functional gene increasing yield under heat stress.

Another hotspot might be located on chromosome 2DS. We mapped QTL for BIOMASS, GRWT_S01 and DTHD in close proximity. The QTL for BIOMASS on chromosome 2DS mapped in U6020 is located at 23.4cM and at 19.1cM for DTHD. Both QTL are close to the

QTL mapped for BIOMASS in U6019 at 17.5cM. The close proximity of these QTL indicate that they are probably they same QTL. Furthermore, a QTL for GRWT_S01 under optimal conditions was mapped in U6020 at 35.9cM. Other studies reported QTL for BIOMASS on chromosome 2DS (Su et al. 2009). Xu et al. (2014) found that one of their QTL for biomass coincided with *Rht8*. This gene has been mapped on chromosome 2DS in close proximity to *Ppd-D1* (Korzun V et al. 1998). Biomass of wheat is defined by several traits, including plant height and number of tillers. Plant height of winter wheat is controlled largely by the semi-dwarfing genes *Rht-B1* and *Rht-D1* and other QTL of medium and small effect like *Rht8* and photoperiod regulator *Ppd-D1* (Wurschum et al. 2015; Zanke et al. 2014). Worland et al. (1998) located the gene for photoperiod *Ppd-D1* on chromosome 2DS at 20.9cM. Jefimija carries the marker *Xgwm*261-200 for *Rht8* (GRIS 2015). We compared the location of our GBS markers with *Ppd-D1* and *Rht8* and concluded that the QTL we identified are *Rht8*, and possibly *Ppd-D1* (Carollo et al. 2005; Worland et al. 1998; Xu et al. 2014 Langer eta l. 2014). We mapped aditional QTL for DTHD on chromosomes 1BL, 1DL and 2BL. Wang et al. (2009) described QTL for flowering time on chromosomes 1BL and 2BL. The chromosome they describe at 1BL could be the same QTL we mapped here, while the QTL on 2BL is located too distant to be the same.

We developed valuable plant material for testing heat tolerance in the field and identified several QTL underlying heat tolerance in bread wheat. Growth chamber experiments enable identifying QTL without or little genotype-by-environment effect and studying the underlying genetics of traits conferring heat tolerance in wheat. However, we recognize the limitations of growth chamber experiments. The next step will be testing this material in the field to assess its potential for heat tolerance under real field conditions.

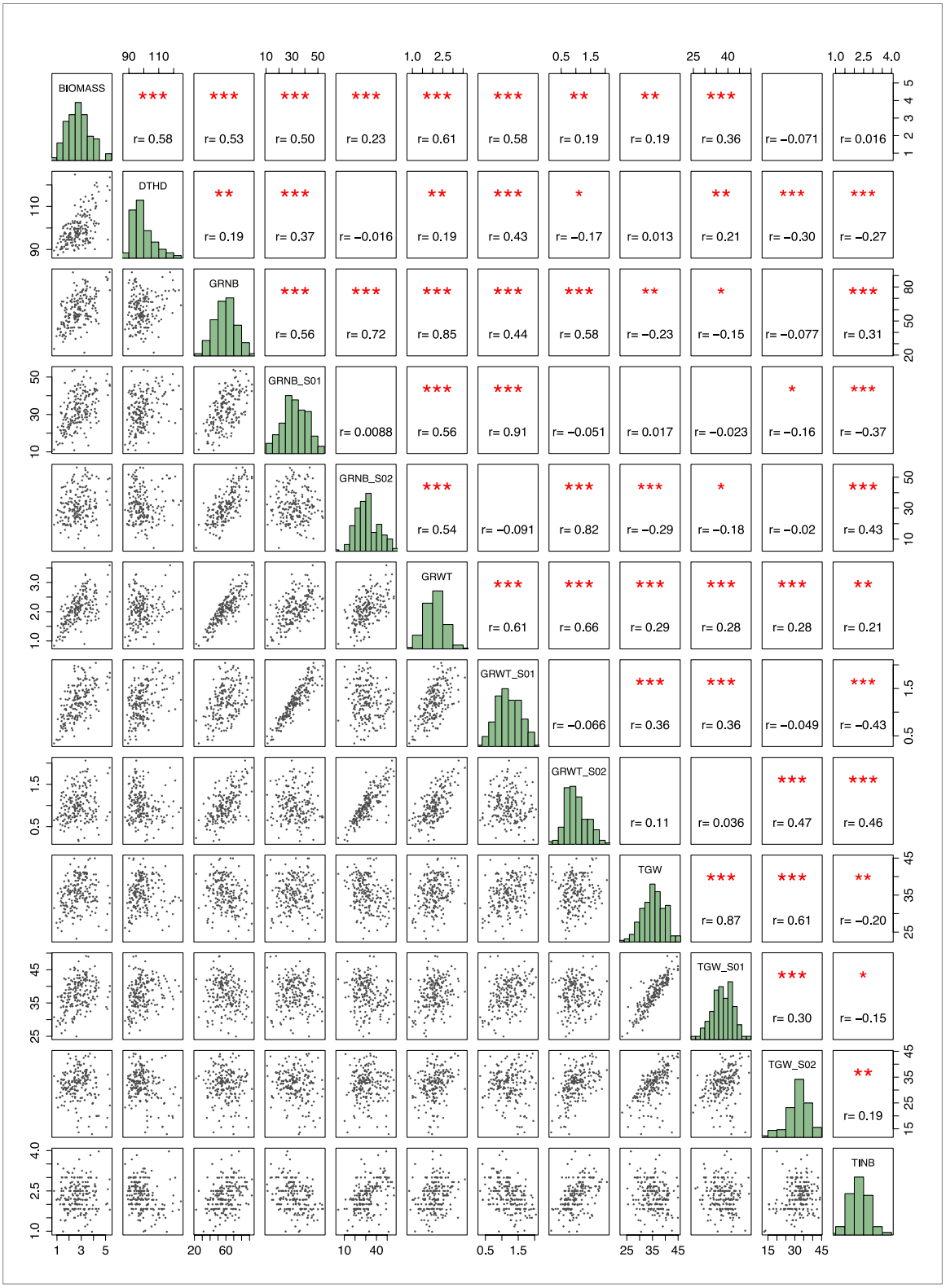

Figure 4.1 Correlation matrix with histograms and significance control treatment

The histogram shows the distribution of BLUEs for all traits. The lower panel of the correlation matrix contains all scatterplots and the upper panel the Pearson correlation coefficient *r* and significance test (*** = p < 0.001, ** = p < 0.01, * = p < 0.05, . = p < 0.1).

Figure 4.2 Correlation matrix with histograms and significance heat treatment

The histogram shows the distribution of BLUEs for all traits. The lower panel of the correlation matrix contains all scatterplots and the upper panel the Pearson correlation coefficient *r* and significance test (*** = p < 0.001, ** = p < 0.01, * = p < 0.05, . = p < 0.1).

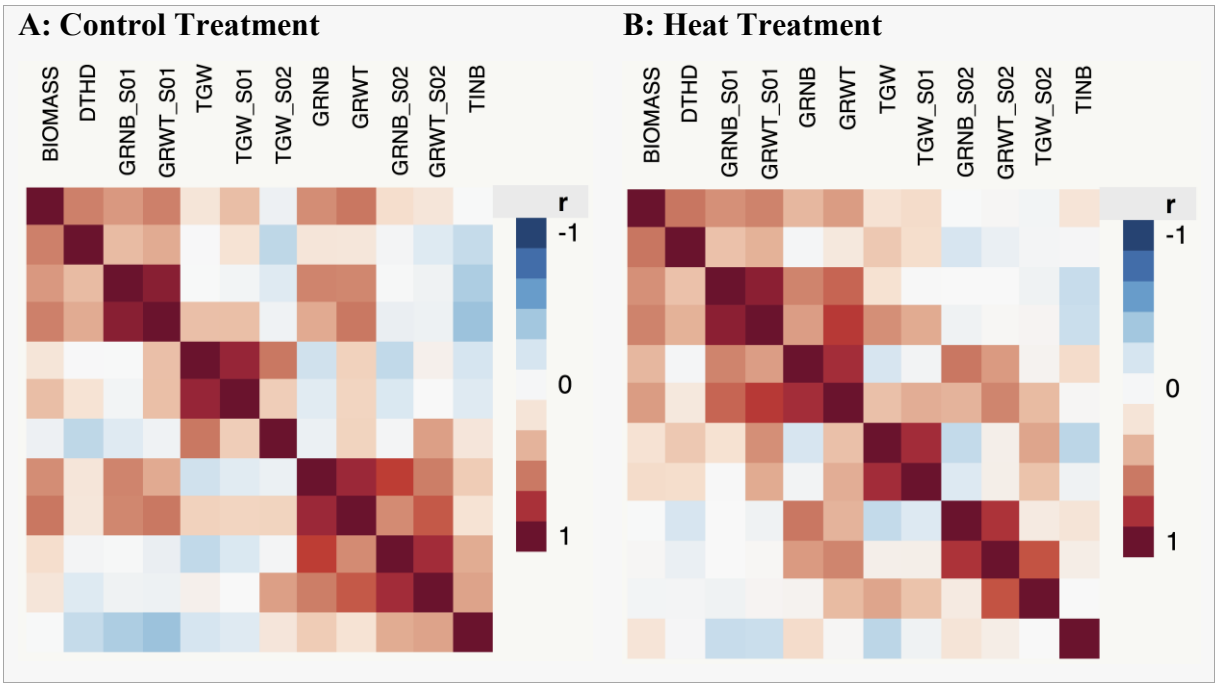

Figure 4.3 Correlation plot control and heat treatment

The correlation analysis is clustered based on correlation. Red indicates strong positive correlation of elements, white no correlation and blue strong negative correlation. BIOMASS and DTHD are correlated in both trials (A: control, B: heat). Other traits such as GRNB, GRWT and TGW are, as expected, correlated as well.

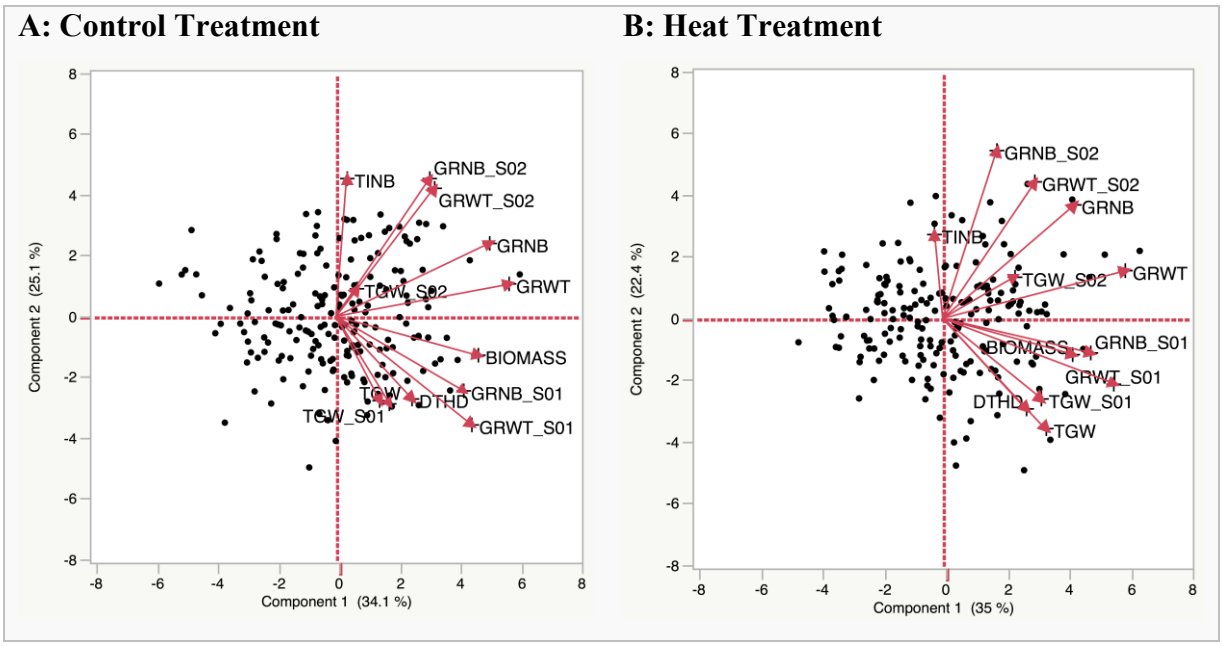

Figure 4.4 PCA biplots based on correlations of all traits in both treatments

The PCA biplot is based on the correlations of all traits. The two major principal components accounted for 59.2 % of the variation in the data set under optimal conditions (A) and 57.4 % under heat stress (B). Both PCI biplots show strong correlations among traits in both environments.

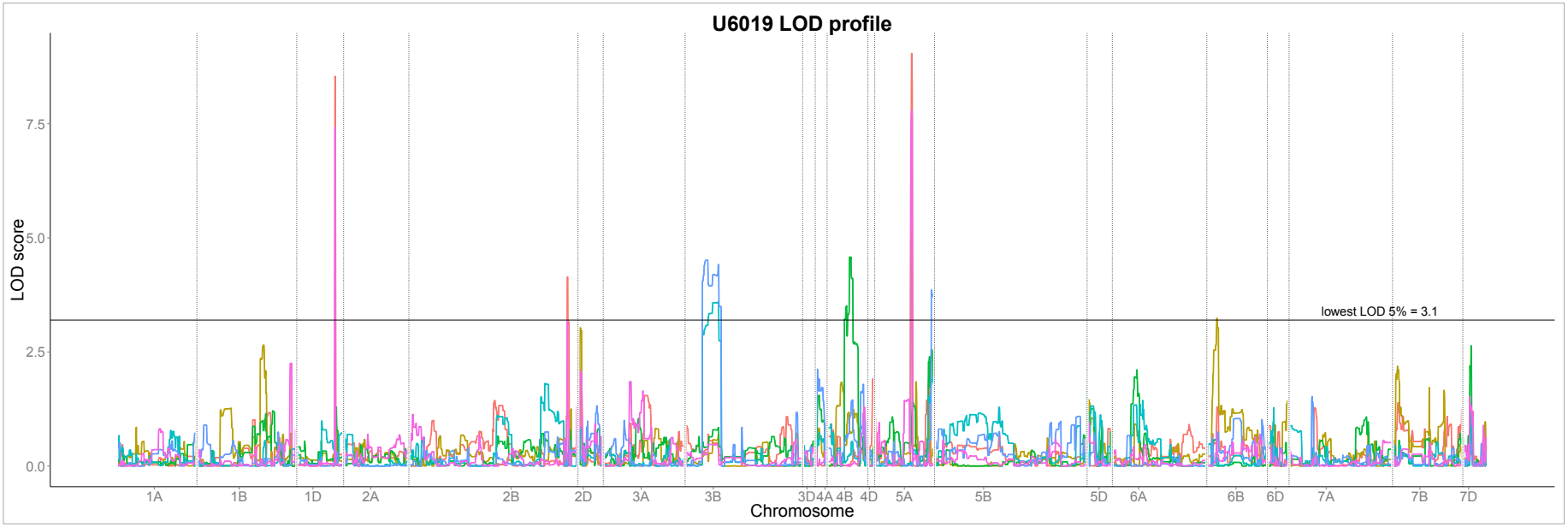

Figure 4.5 LOD profile U6019 heat and control

The LOD profile for population U6019 shows QTL for DTHD on chromosomes 1DL, 2BL and 5A for heat (red) and optimal conditions (pink). Two QTL for GRNB of the whole plant were mapped on chromosomes 3B and 5AL (blue) and one for GRNB_S02 on 3B (turquoise) under optimal conditions. QTL for TGW were mapped on chromosome 4BL under heat and optimal conditions (green). Furthermore, two QTL for biomass were mapped on chromosomes 5A (hidden) and 6BS under heat stress (olive). QTL for DTHD on 1BL were not mapped by CIM but through SIM and confirmed by MQM (not shown). Also, MQM confirmed one more QTL for biomass on 2DS.

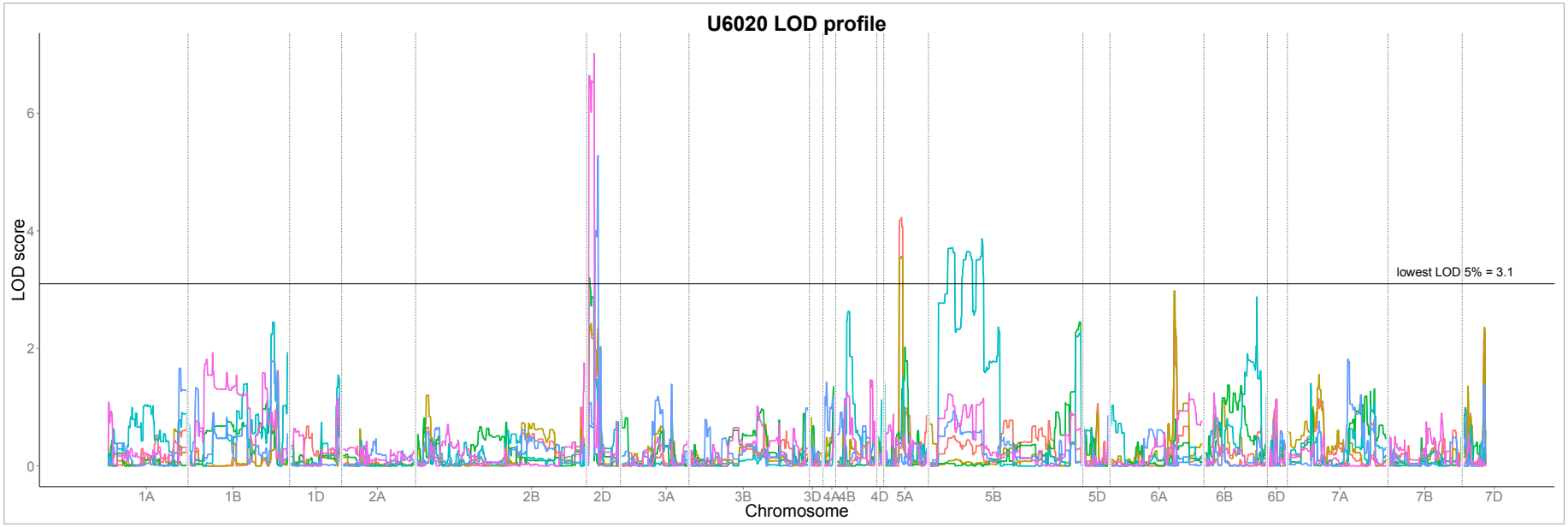

Figure 4.6 LOD profile U6020 heat and control treatment

The LOD profile for population U6020 shows the QTL identified on 2DS for BIOMASS under heat (blue) and optimal conditions (purple) at 24.3cM. This QTL is located in close proximity to a QTL for DTHD at 19.1cM under heat stress (green) and GRWT_S01 under optimal conditions positioned at 35.9cM (red). QTL for GRNB_S01 and GRWT_S01 under heat stress are located at 54.0cM on chromosome 5A (olive), and one QTL for GRWT_S01 under optimal conditions is mapped at 64.2cM (red). Furthermore, the QTL for TGW on 5BL is shown in turquoise.

Table 4.1 Population info and size of four populations available for trials

Population	Pedigree	Status Fall 2015	Rank	Pop. size
U6019	Overley/Jefimija	F_6	1	220
U6020	Overley/Jefimija	F_6		214
U6018	Overley/Proteinka	F_6	2	280
U6021	Overley/Proteinka	F_6		113
U6025	Karl92/(Karl92/Ventnor RIL 73)	F_6	3	423
U6022	Overley/(Karl92/Ventnor RIL 73)	F_6	4	485

Table 4.2 Mean, heritability, standard deviation (Std. Dev.) and coefficient of variation (CV) for all traits in both treatments

Trait	Statistic	Control	Heat
BIOMASS	**H2**	0.57	0.63
	Mean	2.74	2.59
	Std Dev	1.20	1.23
	CV	43.84	47.50
DTHD	**H2**	0.68	0.71
	Mean	100.46	100.67
	Std Dev	10.47	10.58
	CV	10.42	10.51
GRNB_S01	**H2**	0.61	0.57
	Mean	32.24	32.03
	Std Dev	11.94	11.76
	CV	37.03	36.71
GRWT_S01	**H2**	0.60	0.49
	Mean	1.21	1.02
	Std Dev	0.48	0.47
	CV	39.84	45.70
TGW_S01	**H2**	0.49	0.17
	Mean	37.73	31.84
	Std Dev	6.52	8.66
	CV	17.29	27.21
GRNB_S02	**H2**	0.21	0.18
	Mean	32.18	30.23
	Std Dev	17.58	17.35
	CV	54.63	57.37
GRWT_S02	**H2**	0.25	0.27
	Mean	1.03	0.81
	Std Dev	0.61	0.60
	CV	59.40	74.17
TGW_S02	**H2**	0.27	0.23
	Mean	31.98	25.78
	Std Dev	9.17	11.64
	CV	28.68	45.14

Trait	Statistic	Control	Heat
GRNB	**H2**	0.37	0.34
	Mean	59.79	53.67
	Std Dev	21.52	22.06
	CV	36.00	41.10
GRWT	**H2**	0.35	0.35
	Mean	2.10	1.60
	Std Dev	0.78	0.83
	CV	37.03	51.95
TGW	**H2**	0.41	0.18
	Mean	35.42	29.47
	Std Dev	6.42	8.77
	CV	18.12	29.77
TINB	**H2**	0.25	0.23
	Mean	2.31	2.23
	Std Dev	0.91	0.89
	CV	39.14	39.75

Table 4.3 Summary of identified QTL for traits measured in both populations under optimal conditions

Pop	Trait	Chr	GBS Marker	Pos (cM)	LOD	LOD α=0.05	Pheno variance	QTL effect	SE	Effect A	Effect B
U6019	DTHD	1BL	KSUheat2921	63.1	3.48	3.30	7.82	-5.33	1.31	5.33	-5.33
		1BL	KSUheat6741	66.6	7.60	3.30	18.53	8.34	1.32	-8.34	8.34
		2BL	KSUheat9649	135.0	4.87	3.30	11.31	5.35	0.74	-5.35	5.35
		5A	KSUheat13739	62.4	5.45	3.30	12.84	3.03	0.58	-3.03	3.03
	GRNB	3B	KSUheat5790	75.8	4.50	3.29	16.39	-5.49	1.66	5.49	-5.49
		5AL	KSUheat12934	133.5	4.25	3.29	15.43	6.04	1.32	-6.04	6.04
	GRNB_S02	3B	KSUheat2441	81.6	3.60	3.35	15.44	-3.96	0.94	3.96	-3.96
	TGW_S01	4BL	KSUheat6054	71.0	4.60	3.41	19.19	1.76	0.37	-1.76	1.76
U6020	BIOMASS	2DS	KSUheat11830	24.3	6.40	3.43	30.29	-0.43	0.07	0.43	-0.43
	GRWT_S01	2DS	KSUheat10738	35.9	3.99	3.31	15.85	-0.17	0.04	0.17	-0.17
		5A	KSUheat13739	62.4	4.32	3.31	17.30	0.14	0.03	-0.14	0.14
	TGW	5BL	KSUheat9427	43.1	3.92	3.49	18.37	-1.88	0.43	1.88	-1.88
	TGW_S01	5BL	KSUheat9427	43.1	3.54	3.30	16.74	-2.06	0.49	2.06	-2.06

GBS markers, QTL positions Pos (cM), LOD, LOD at the 5% significance threshold score, estimated phenotypic variance explained by QTL, estimated QTL effect and allele effects (effect A and effect B) for all QTL. QTL are sorted based on chromosome and location. The LODs reported here LODs calculated through MQM.

Table 4.4 Summary of identified QTL for traits measured in both populations under heat stress

Pop	Trait	Chr	GBS Marker	Pos (cM)	LOD	LOD α=0.05	Pheno variance	QTL effect	SE	Effect A	Effect B
U6019	BIOMASS	2DS	KSUheat5627	17.5	3.9	3.2	12.63	-0.38	0.09	0.38	-0.38
		5A	KSUheat13739	62.4	3.8	3.2	12.40	0.32	0.07	-0.32	0.32
		6BS	KSUheat340	32.2	3.7	3.2	11.95	0.32	0.08	-0.32	0.32
	DTHD	1DL	KSUheat8204	91.8	8.7	3.4	22.29	3.99	0.58	-3.99	3.99
		2BL	KSUheat9649	135.0	4.1	3.4	9.45	3.45	0.77	-3.45	3.45
		5A	KSUheat2710	61.8	8.2	3.4	20.65	3.87	0.58	-3.87	3.87
	TGW_S01	4BL	KSUheat13198	62.0	3.3	3.3	14.12	1.20	0.30	-1.20	1.20
U6020	BIOMASS	2DS	KSUheat11830	24.3	4.8	3.6	22.08	-0.41	0.08	0.41	-0.41
	DTHD	2DS	KSUheat7222	19.1	3.2	3.1	15.28	-2.19	0.55	2.19	-2.19
	GRNB_S01	5A	KSUheat819	54.0	3.6	3.3	16.85	2.88	0.69	-2.88	2.88
	GRWT_S01	5A	KSUheat819	54.0	3.6	3.3	19.65	0.09	0.02	-0.09	0.09

GBS markers, QTL positions Pos (cM), LOD, LOD at the 5% significance threshold score, estimated phenotypic variance explained by QTL, estimated QTL effect and allele effects (effect A and effect B) for all QTL. QTL are sorted based on chromosome and location. The LODs reported here LODs calculated through MQM.

References

Arends, D., P. Prins, R.C. Jansen and K.W. Broman. 2010. R/qtl: high-throughput multiple QTL mapping. Bioinformatics 26: 2990-2992. doi:10.1093/bioinformatics/btq565.

Bernardo, R. 2010. Breeding for Quantitative Traits in Plants. 2nd ed. Stemma Press, Woodbury, Minnesota.

Broman K.W, Wu H., Sen Ś and C. G.A. 2003. R/ qtl: QTL mapping in experimental crosses. Bioinformatics 19: 889-890.

Broman, K.W. and S. Saunak. 2009. A Guide to QTL Mapping with R/qtlSpringer New York.

Browning, B.L. and S.R. Browning. 2013. Improving the Accuracy and Efficiency of Identity-by-Descent Detection in Population Data. Genetics 194: 459-471. doi:10.1534/genetics.113.150029.

Carollo, V., D.E. Matthews, G.R. Lazo, T.K. Blake, D.D. Hummel, N. Lui, et al. 2005. GrainGenes 2.0. an improved resource for the small-grains community. Plant Physiol 139: 643-651. doi:10.1104/pp.105.064485.

Chapman, J.A., M. Mascher, A. Buluç, K. Barry, E. Georganas, A. Session, et al. 2015. A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome. Genome Biology 16: 26. doi:10.1186/s13059-015-0582-8.

Dunckel, S.M., E.L. Olson, M. Rouse, R. Bowden and J. Poland. 2015. Genetic Mapping of Race-Specific Stem Rust Resistance in the Synthetic Hexaploid W7984 x Opata M85 Mapping Population. Crop Sci. 55: 1-9. doi:doi: 10.2135/cropsci2014.11.0755.

Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, et al. 2011. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. PLoS ONE 6: e19379. doi:10.1371/journal.pone.0019379.

FAO. 2015. FAOSTAT, Statistics Division FAO. Retrieved 08/08/2015 from http://faostat3.fao.org/home/E

GRIS. 2015. Genetic Resources Information System for Wheat and Triticale. Retrieved 08/08/20 from http://wheatpedigree.net/sort/show/34239

Hays, D.B., J.H. Do, R.E. Mason, G. Morgan and S.A. Finlayson. 2007. Heat stress induced ethylene production in developing wheat grains induces kernel abortion and increased

maturation in a susceptible cultivar. Plant Science 172: 1113-1123. doi:http://dx.doi.org/10.1016/j.plantsci.2007.03.004.

IPCC. 2014. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change In: R. K. Pachauri and L. A. Meyer, editors, IPCC, Geneva, Switzerland. p. 151.

JMP®, V. 1989-2015. SAS Institute Inc, Cary, NC.

Jordan, K., S. Wang, Y. Lun, L.-J. Gardiner, R. MacLachlan, P. Hucl, et al. 2015. A haplotype map of allohexaploid wheat reveals distinct patterns of selection on homoeologous genomes. Genome Biology 16: 1-18. doi:10.1186/s13059-015-0606-4.

Kirigwi, F., M. Van Ginkel, G. Brown-Guedira, B. Gill, G. Paulsen and A. Fritz. 2007. Markers associated with a QTL for grain yield in wheat under drought. Molecular Breeding 20: 401-413. doi:10.1007/s11032-007-9100-3.

Korzun V, M.S. Röder, M.W. Ganal, A.J. Worland and C.N. Law. 1998. Genetic analysis of the dwarfing gene (Rht8) in wheat. Part I. Molecular mapping of Rht8 on the short arm of chromosome 2D of bread wheat (Triticum aestivum L.). Theor Appl Genet 96: 1104-1109.

Langer, S.M., C.F.H. Longin and T. Würschum. 2014. Flowering time control in European winter wheat. Frontiers in Plant Science 5: 537. doi:10.3389/fpls.2014.00537.

Paulsen, G.M. 1994. High Temperature Responses of Crop Plants. In: K. J. Boote, J. M. Bennett, T. R. Sinclair and G. M. Paulsen, editors, Physiology and Determination of Crop Yield. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Madison, WI. p. 365-389.

Pinto, R., M. Reynolds, K. Mathews, C. McIntyre, J.-J. Olivares-Villegas and S. Chapman. 2010. Heat and drought adaptive QTL in a wheat population designed to minimize confounding agronomic effects. TAG Theoretical and Applied Genetics 121: 1001-1021. doi:10.1007/s00122-010-1351-4.

Pinto, R.S., M. Reynolds, K. Mathews, C.L. McIntyre, J.-J. Olivares-Villegas and S. Chapman. 2010. Heat and drought adaptive QTL in a wheat population designed to minimize confounding agronomic effects. Theoretical and Applied Genetics 121: 1001-1021. doi:10.1007/s00122-010-1351-4.

Poland, J., P. Brown, M. Sorrells and J.-L. Jannink. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS One 7: e32253.

R Core Team. 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Ristic, Z., U. Bukovnik, I. Momcilovic, J. Fu and V.P.V. Prasad. 2008. Heat-induced accumulation of chloroplast protein synthesis elongation factor, EF-Tu, in winter wheat. Journal of Plant Physiology 165: 192-202. doi:10.1016/j.jplph.2007.03.003.

Shiferaw, B., M. Smale, H.-J. Braun, E. Duveiller, M. Reynolds and G. Muricho. 2013. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. Food Sec. 5: 291-317. doi:10.1007/s12571-013-0263-y.

Su, J.-Y., Q. Zheng, H.-W. Li, B. Li, R.-L. Jing, Y.-P. Tong, et al. 2009. Detection of QTLs for phosphorus use efficiency in relation to agronomic performance of wheat grown under phosphorus sufficient and limited conditions. Plant Science 176: 824-836. doi:http://dx.doi.org/10.1016/j.plantsci.2009.03.006.

Talukder, S.K., M.A. Babar, K. Vijayalakshmi, J. Poland, P.V. Prasad, R. Bowden, et al. 2014. Mapping QTL for the traits associated with heat tolerance in wheat (*Triticum aestivum* L.). Bmc Genet 15: 97. doi:10.1186/s12863-014-0097-4.

Wang, R., L. Hai, X. Zhang, G. You, C. Yan and S. Xiao. 2009. QTL mapping for grain filling rate and yield-related traits in RILs of the Chinese winter wheat population Heshangmai × Yu8679. TAG Theoretical and Applied Genetics 118: 313-325. doi:10.1007/s00122-008-0901-5.

Wardlaw, I. and C. Wrigley. 1994. Heat Tolerance in Temperate Cereals: an Overview. Functional Plant Biology 21: 695-703. doi:http://dx.doi.org/10.1071/PP9940695.

Wardlaw, L.F., I.A. Dawson, P. Manibi and R. Fewster. 1989. The tolerance of wheat to high temperature during reproductive growth. I. Survey procedures and general response patterns. Aust. J. Agric. Res. 40: 1-13.

Worland, A.J., A. Borner, V. Korzun, W.M. Li, S. Petrovic and S.J. Sayers. 1998. The influence of photoperiod genes on the adaptability of European winter wheat. Euphytica 100: 358-394.

Wu, X., X. Chang and R. Jing. 2012. Genetic Insight into Yield-Associated Traits of Wheat Grown in Multiple Rain-Fed Environments. PLoS ONE 7: e31249. doi:10.1371/journal.pone.0031249.

Wurschum, T., S.M. Langer and C.F. Longin. 2015. Genetic control of plant height in European winter wheat cultivars. Theor Appl Genet 128: 865-874. doi:10.1007/s00122-015-2476-2.

Xu, Y., R. Wang, Y. Tong, H. Zhao, Q. Xie, D. Liu, et al. 2014. Mapping QTLs for yield and nitrogen-related traits in wheat: influence of nitrogen and phosphorus fertilization on QTL expression. Theor Appl Genet 127: 59-72. doi:10.1007/s00122-013-2201-y.

Yang, J., R.G. Sears, B.S. Gill and G.M. Paulsen. 2002. Genotypic differences in utilization of assimilate sources during maturation of wheat under chronic heat and heat shock stresses. Euphytica 125: 179-188. doi:10.1023/A:1015882825112.

Yang, J., R.G. Sears, B.S. Gill and G.M. Paulsen. 2002. Growth and senescence characteristics associated with tolerance of wheat-alien amphiploids to high temperature under controlled conditions. Euphytica 126: 185-193. doi:10.1023/a:1016365728633.

Zanke, C.D., J. Ling, J. Plieske, S. Kollers, E. Ebmeyer, V. Korzun, et al. 2014. Whole Genome ssociation Mapping of Plant Height in Winter Wheat (*Triticum aestivum* L.). PLoS ONE 9: e113287. doi:10.1371/journal.pone.0113287.

Appendix A - Copyright Permissions

This appendix contains the copyright permissions required to include in this dissertation material that was previously published research.

ACSESS-Alliance of Crop, Soil, and Environmental Science Societies LICENSE TERMS AND CONDITIONS

Oct 25, 2015

This is a License Agreement between Sandra Dunckel ("You") and ACSESS-Alliance of Crop, Soil, and Environmental Science Societies ("ACSESS-Alliance of Crop, Soil, and Environmental Science Societies") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by ACSESS-Alliance of Crop, Soil, and Environmental Science Societies, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3724830892578
License date	Oct 09, 2015
Licensed Content Publisher	ACSESS-Alliance of Crop, Soil, and Environmental Science Societies
Licensed Content Publication	Crop Science
Licensed Content Title	Genetic Mapping of Race-Specific Stem Rust Resistance in the Synthetic Hexaploid W7984 × Opata M85 Mapping Population
Licensed copyright line	Copyright ©2015 by the Crop Science Society of America, Inc.
Licensed Content Author	Sandra M. Dunckel, Eric L. Olson, Matthew N. Rouse, et al.
Licensed Content Date	Sep 1, 2015
I would like to...	Thesis/Dissertation
Requestor type	Author of requested content
Format	None
Portion	chapter/article
Rights for	Main product
Creation of copies for the disabled	yes
With minor editing privileges	no
For distribution to	Worldwide
In the following language(s)	Original language of publication
With incidental promotional use	no
The lifetime unit quantity of new product	0 to 499
The requesting person/organization is:	Sandra Dunckel, KSU
Order reference number	None
Title of your thesis / dissertation	WHOLE GENOME APPROACHES FOR CHARACTERIZING AND UTILIZING SYNTHETIC

	WHEAT
Expected completion date	Oct 2015
Estimated size (number of pages)	121
Total	**0.00 USD**

Terms and Conditions

Introduction

The Publisher for this copyrighted material is ACSESS. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your CCC account and that are available at any time at <http://myaccount.copyright.com>).

Limited License

Publisher hereby grants to you a non-exclusive license to use this material. Licenses are for one-time use only with a maximum distribution equal to the number that you identified in the licensing process; any form of republication must be completed within 60 days from the date hereof (although copies prepared before then may be distributed thereafter); and any electronic posting is limited to a period of 120 days.

Geographic Rights: Scope

Licenses may be exercised anywhere in the world.

Altering/Modifying Material: Not Permitted

You may not alter or modify the material in any manner, nor may you translate the material into another language.

Reservation of Rights

Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

License Contingent on Payment

While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

Copyright Notice: Disclaimer

You must include the following copyright and permission notice in connection with any reproduction of the licensed material: "Reprinted by Permission, ASA, CSSA, SSSA."

Warranties: None

Publisher makes no representations or warranties with respect to the licensed material.

Indemnity

You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

No Transfer of License

This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

No Amendment Except in Writing

This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

Objection to Contrary Terms

Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated

herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

Jurisdiction: Not Required*

This license transaction shall be governed by and construed in accordance with the laws of Wisconsin. You hereby agree to submit to the jurisdiction of the federal and state courts located in Wisconsin for purposes of resolving any disputes that may arise in connection with this licensing transaction.

Other Terms and Conditions

* If omitted, license will rely on New York law as stated in CCC terms and conditions agreed to by licensee during account creation.

V1.0

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.