DETERMINING THE BREEDING VALUE OF PRIMARY SYNTHETIC BREAD WHEATS FOR INCREASED GRAIN YIELD

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Jafar Jafarzadeh

August 2017

DETERMINING THE BREEDING VALUE OF PRIMARY SYNTHETIC BREAD WHEATS FOR INCREASED GRAIN YIELD

Jafar Jafarzadeh Ph. D.

Cornell University 2017

Long term domestication and breeding of bread wheat increased grain yield, but this increase has slowed down, in part, due to the reduction of genetic variation. To introduce new genetic diversity from wheat progenitors, *Aegilops tauschii* (Coss.) Schmalh and durum wheat (*Triticum. turgidum* L. subsp. *durum*) into the bread wheat gene pool, 20 spring bread wheat parents (BWPs) were crossed to 33 synthetic hexaploid wheat parents (SYNPs) at the International Wheat and Maize Improvement Center. Single spike descent was used to develop 97 synthetic derived populations (SDLs). Yield trials were conducted under irrigated (IRRI), drought (DRO) and heat (HEAT) stress environments from 2011 to 2014 in Ciudad Obregon, Mexico. Genomic estimated breeding values (GEBVs) of genotypes were estimated using a genomic best linear unbiased prediction model with markers. The result of this study addressed that First, SYN lines and SDLs were more diverse than BWPs for A, B and D genomes confirming that the SYN lines are promising genetic resources of novel diversity. Second, grain yield (YLD) increases in SDLs were more frequent under DRO and HEAT stresses and were predominantly in SDLs from first back-cross derived lines. The SYNPs GEBVs for YLD were less negative under DRO and HEAT stresses than those under IRRI indicating SYN lines could increase YLD under stresses. Under DRO and HEAT, the

SYNPs increased plant height (PLH) and days to maturity (DMA). Higher PLH increased YLD but longer DMA decreased YLD. Under IRRI, 29% of SDLs had higher thousand kernel weight (TKW) than BWPs ($P < 0.05$) indicating SYNPs were valuable genetic resources for TKW. Third, a genome-wide association study using SDLs identified associated QTL with TKW, PLH, YLD and DMA and SYNP alleles of these QTL retained in SDLs and increased trait values. Our finding confirmed that SYN lines had positive alleles that can be easily introgressed into cultivated wheat to improve agronomic and phenological traits especially in stress conditions. Therefore, SYN lines should be used in breeding programs to expand the genetic diversity for agronomic traits but selection against undesirable phenology is required to realize the benefit of the novel genetic variation.

BIOGRAPHICAL SKETCH

Jafar Jafarzadeh was born to a small farm family in Viranbagh, a small village in Jolfa, East Azerbaijan, Iran on February 23rd, 1974, to parents Hossein-Ali Jafarzadeh and Roghayyeh Ebrahimi. In 1992, Jafar graduated high school and entered the University of Tabriz. He received a Bachelor of Science degree in Agronomy and Plant Breeding in 1996. In 1999, Jafar evaluated the response of Azerbaijan barley landraces to barley leaf stripe disease (*Pyrenophora graminea*) in the University of Tabriz under the supervision of Dr. Mohammad Moghaddam and received a Master of Science degree in Plant Breeding in 2001. In 2005, Jafar was employed in Dryland Agricultural Research Institute (DARI) as a wheat breeder. In DARI, he gained substantial field experience by conducting research in conventional wheat breeding programs under drought, cold and heat stresses. He gained his first international experiences by participating in wheat improvement and pathology training course at CIMMYT, Mexico in 2010. Jafar gained international collaboration experience by joint international breeding programs with CIMMYT and ICARDA. In 2012, Jafar was awarded Monsanto's Beachell-Borlaug International Scholarship to earn his Ph.D. on the breeding value of primary synthetic wheat genotypes for grain yield under supervision of Dr. Mark E. Sorrells at Cornell University in collaboration with CIMMYT and Dr. David Bonnett. Through this project, he had the opportunity to work with and learn from scientists at CIMMYT. Jafar looks forward to get more experience to integrate conventional and molecular breeding approaches to develop high-yielding cultivars for rainfed conditions and contribute his knowledge to the wheat community.

To my parents and family

ACKNOWLEDGMENTS

I would like to acknowledge all the people who contributed to this project and my development as a plant breeder. I would like to appreciate my committee members; Dr. Mark Sorrells for giving me the opportunity to become a part of his research group and for providing me with guidance, advice, and opportunities that were key to my success; Dr. Jean-Luc Jannink for his valuable guidance and suggestions that improved this research and my knowledge; Dr. David Bonnett for his guidance, sharing information, germplasm, conducting field trails and allowing me to gain practical breeding experience at CIMMYT that have enormously improved my field experience and education. I would like to thank Dr. Susanne Dreisigacker for genotyping my populations. I would like to thank Dr. Mohammad Reza Jalal-Kamali, coordinator of CIMMYT-Iran, for his great assistance in facilitating my Ph.D. application. I want to express my gratitude for Mary Ellen Niederhofer's support during my Ph.D. studies. I would like to thank graduate students and post docs past and present, especially Dr. Deniz Akdemir, Dr. Vahid Edriss, Nicholas Santantonio, Itaraju Junior Baracuhy Brum, Uche Okeke, who have shared their guidance and knowledge through many thoughtful discussions. I am grateful to the Global Wheat Program of the International Maize and Wheat Improvement Center (CIMMYT) for excellent collaboration with Cornell University to do this study. I would like to thank the sources of funding for this project, Monsanto's Beachell-Borlaug International Scholars Program which supports Ph.D. students from developing countries who are working on breeding wheat and rice, "Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA, Mexico)" for supporting the MasAgro Biodiversidad/Seeds of Discovery Project-CIMMYT, USDA-NIFA-AFRI grants numbers 2009-65300-05661, 2011-68002- 30029, and 2005-05130, and Hatch project 149-449. I would like to thank my family, especially my wife Raziyeh, for always supporting me and enriching my life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Importance of wheat

Wheat (*Triticum aestivum* L.) as a third important crop for food security is widely planting around the world. It was estimated that more than 75% of the world's population consumes wheat as part of their diet daily (Lillemo 2005). High nutrient content of wheat grain including proteins, vitamins, lipids, and carbohydrates make wheat as a valuable nutrition source in the population diet. Wheat provides 55% carbohydrate, 19% calories and 21% protein of the global population diet in the world (Gupta et al., 1999; Bagge et al., 2007). The world population rapidly increasing, which is expected to reach 9 billion by 2050 that simultaneously increases global wheat demand, which is expected to reach about 900 million tons by 2050 (Weigand, 2011; Rana et al., 2013). Currently, increasing rate of wheat yield is 0.9% per year, which is less than the 2.4% per year rate required to double global wheat production by 2050 (Ray et al., 2013). Furthermore, other important factors such as the gradual decreasing water sources, arable land area, and climate change are affecting crop production including wheat around the world. All these factors are making a serious concern for food security in the present and future. Therefore, the most important mission is focusing on finding efficient ways to improving yields of all crops to overcome this concern (Yang et al., 2012; Gao et al., 2015).

Domestication of common wheat

One of the earliest cultivated wheat was *Triticum urartu* as a diploid species (2n = 14) with AA genome (einkorn). The limited natural hybridization and chromosome doubling between *Triticum urartu* and *Aegilops speltoides*, a diploid species with BB or SS genome, resulted in developing tetraploid species, *Triticum turgidum* subsp. *dicoccoides*, with AABB genome (2n = 28) as wiled emmer (Figure 1.1) (Mujeeb-Kazi et al., 1996; Van Ginkel and Ogbonnaya, 2007; Dreisigacker et al., 2008; Maxted and Kell, 2009).

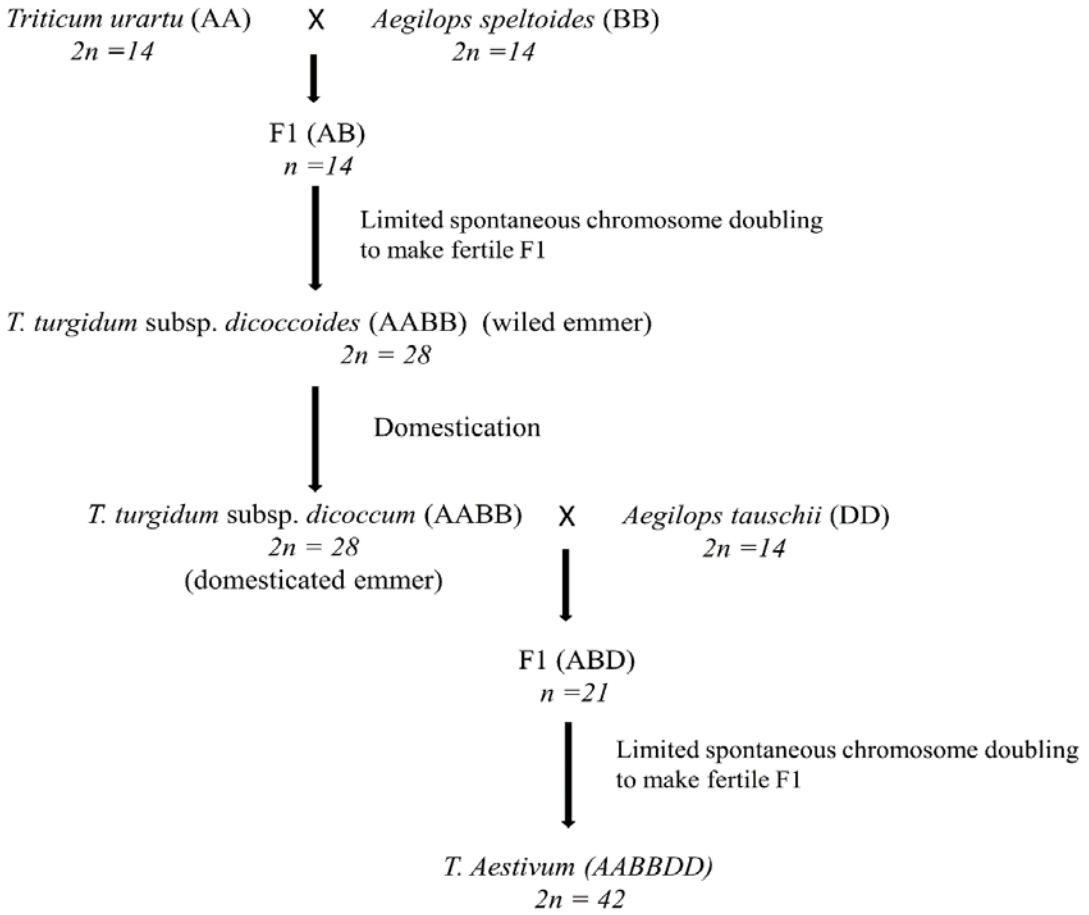


Figure 1.1: Evolutionary of hexaploid bread wheat (Mujeeb-Kazi et al. 1996; Van Ginkel and Ogbonnaya 2007; Dreisigacker et al. 2008; Maxted and Kell 2009; Shewry 2009b)

Both diploid and tetraploid wheat species were originated from the south-eastern part of Turkey (Mujeeb-Kazi et al., 1996; Van Ginkel and Ogbonnaya, 2007; Shewry, 2009a). Wild emmer had been domesticated for some desirable traits such as non-brittle rachis or loss of shattering, larger grains, and free threshing that led to developing domesticated emmer or durum wheat, *T. turgidum* subsp. *dicoccum* (2n = 28, AABB) (Ginkel and Ogbonnaya 2005; Shewry 2009). After then, durum wheat spread out quickly from Fertile Crescent to the northern and northern Africa, Europe, Asia and eventually developed as modern durum wheat *T. turgidum* subsp. *durum* to produce pasta (Ginkel and Ogbonnaya, 2005). The second natural hybridization and chromosome doubling happened between durum wheat and *Aegilops tauschii* (a diploid wild goat grass, also called *Triticum tauschii, Aegilops squarrosa*) to produce hexaploid wheat, *Triticum aestivum* (AABBDD genome and 2n = 42) (Figure 1.1) (Mujeeb-Kazi et al., 1996; Van Ginkel and Ogbonnaya, 2007; Shewry, 2009a). This rare event most likely happened once or a few times about 8000 years ago in the region near the Caspian Sea, Iran (Ginkel and Ogbonnaya, 2005; Curtis and Halford, 2014)(Figure 1.2).

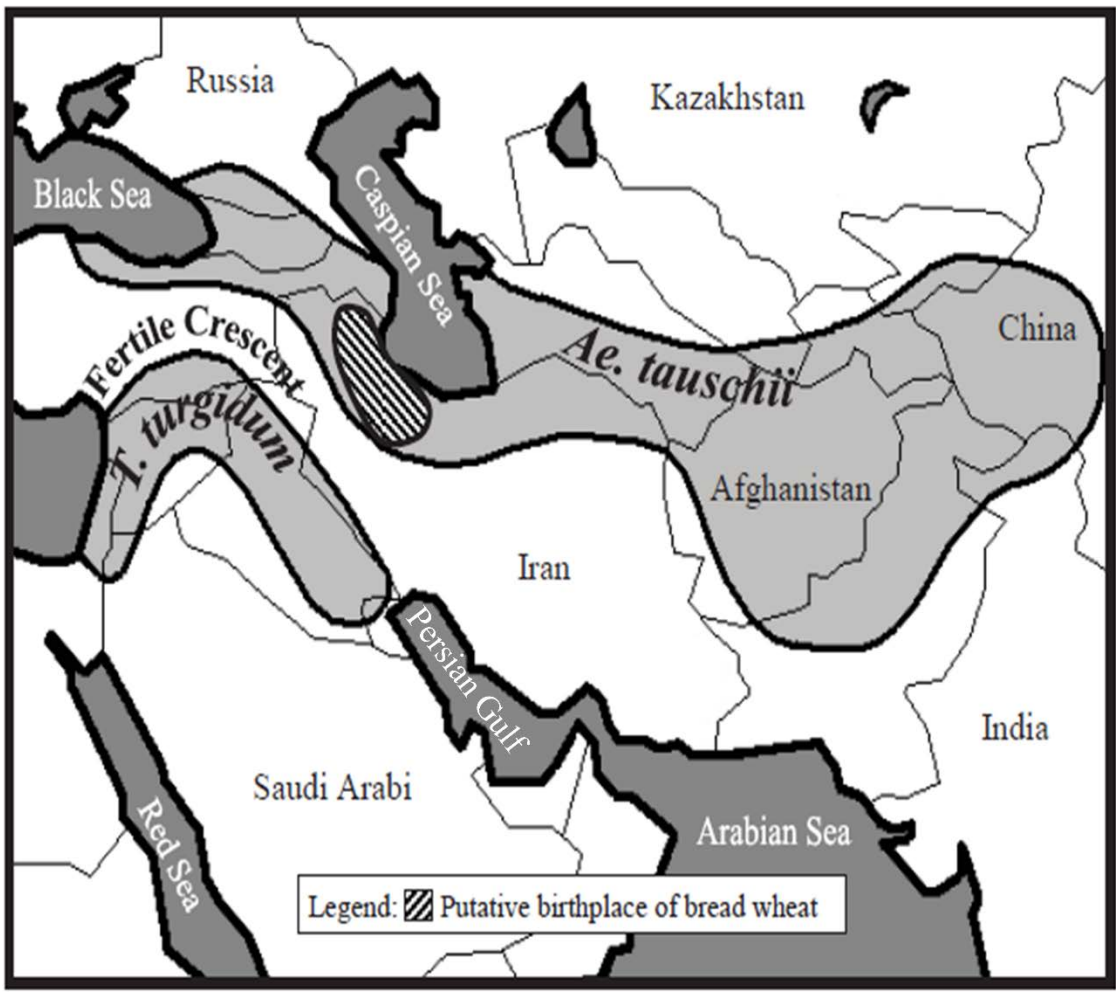


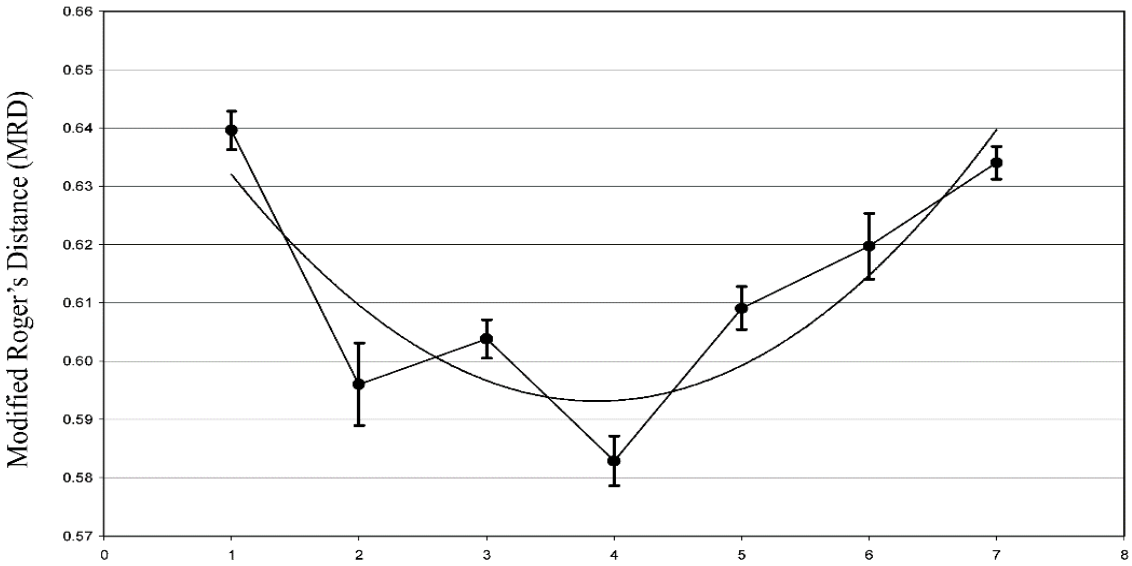
Figure 1.2: The putative birthplace of hexaploid bread wheat relatives to the Fertile Crescent (the domestication region of T. turgidum), and the distribution of Ae. tauschii accessions (Talbot and Hons 2011).

Introgression novel diversity from wheat's wild relatives into common wheat

As above mentioned hexaploid bread wheat has been developed by crossing limited number of three different diploid wild wheats. So potential genetic variation of many other diploid wheats did not contributed to the initial crosses might be contain beneficial genes for different traits (Ginkel and Ogbonnaya, 2005; Curtis and Halford, 2014). Two important bottlenecks happened during wheat evolution, first when wild

emmer domesticated and second when hybridization between domesticated emmer (durum) and one or two *Ae. tauschii* happened (Li et al., 2014). Consequently, modern cultivated bread wheats comprise either one or two alleles of D genome genes, which led to the concept that only one or two *Ae. tauschii* contributed their genome to cultivated hexaploid wheats (Ogbonnaya et al. 2005). The D genome of *A. tauschii* has higher genetic diversity than that of bread wheat for biotic and abiotic stresses (Naghavi and Mardi 2010; Reif et al. 2005; Sohail et al. 2011). Furthermore, long term selection on hexaploid bread wheat either by farmers or by wheat breeders resulted in increasingly narrow genetic diversity of wheat (Zhang et al., 2005; Sohail et al., 2011). For instance, the efforts of International Maize and Wheat Improvement Center (CIMMYT) to breed uniform wheat cultivars and lines was successful in terms of improving yield, yield stability, biotic and abiotic stresses resistance, which was remarkable during green revolution. However, this success made an overall genetic diversity reduction in most of the world's wheat producing regions (Sehgal et al., 2015). Warburton et al. (2006) by studying genetic diversity of wheat landraces, cultivars, advanced breeding lines from 1950 to 2003 reported the average Modified Roger's distances decreased from 0.64 for the landraces to 0.58 for the improved lines in the 1980's (Figure 1.3). This diversity does not satisfy crop production under different biotic and abiotic stresses and breeders have been tried to find new genetic variability (Ogbonnaya et al. 2005). Therefore, to expand the genetic variation of common wheat, CIMYYT have been used landraces and synthetic hexaploid wheats (SHW) (Mujeeb-Kazi et al. 1996; Ogbonnaya et al. 2005; Dreisigacker et al. 2008; Maxted and Kell 2009) in the pedigree of new

advanced lines that lead to increase the overall genetic diversity up to 0.63 that was close to that in the landraces (Figure 1.3) (Warburton et al., 2006).



Year group 1 = Landraces.
Year group 2 = Cultivars released between 1950 & 1966.
Year group 3 = Cultivars released between 1967 & 1974.
Year group 4 = Cultivars released between 1975 & 1982.
Year group 5 = Cultivars released between 1982 & 1989.
Year group 6 = Cultivars released between 1990 & 1997.
Year group 7 = Breeding lines in advanced field trials for the years 2002 - 2003 and performing well; expected to be released within 0 & 3 years as cultivars.

Figure 1.3: Plot of the modified roger's distance over time. This plot shoes trend of genetic diversity from landraces to new advanced lines (Warburton et al., 2006).

The SHWs are amphiploids resulting from interspecific crosses between a diploid *Ae. tauschii*, donor of the D genome and a modern durum or emmer wheat (*Triticum turgidum* L. subsp. *dicoccum*) wheat donor of the A and B genomes. The SHWs have been developing by embryo rescue technic and inducing chromosome doubling using colchicine (Mujeeb-Kazi et al. 1996). The SHWs have been broadly used to introduce new genetic diversity into the cultivated bread wheat gene pool from wheat progenitors at CIMMYT (Mujeeb-Kazi et al. 1996; Ogbonnaya et al. 2005; Dreisigacker et al. 2008).

Molecular analysis of SHWs, synthetic backcross-derived lines (SBLs) (SHW x common wheat) and traditional wheat cultivars at CIMMYT revealed that SHWs and SBLs were more genetically diverse than cultivated wheats. This indicated that developing SHWs and SBLs were successful in expanding genetic diversity of wheat gene pool (specifically D genome) and improving traits of interests simultaneously (Lage et al., 2003; Zhang et al., 2005; Dreisigacker et al., 2008). Also, molecular analysis of Chuanmai-42 (as a cultivated SDL in China) and its parents with a total of 3297 markers was done to identify proportion of SHW parent alleles in Chuanmai-42. Results indicated that only 277 SHW alleles introgressed to Chuanmai-42, which was 15.14%, significantly less than the expected 25% assuming random gene assortment. Also, the distribution of introgressed alleles was not uniform across A, B, and D genomes (B > A > D) (Li et al., 2014). The information about the flux of genetic variation from SHWs and SBLs during several backcross generations with common wheats could help to identify retained favorable chromosomal segments for desirable traits under selection (Zhang et al., 2005).

Success of synthetic hexaploid wheats in breeding program

Synthetic wheats at CIMMYT-Mexico

CIMMYT initially used SHWs in breeding programs to transfer disease and insect resistance genes from *Ae. tauschii* to the common wheat such as Karnal bunt (*Tilletia indica* Mitra)(Villareal et al., 1994). It was, however, soon realized that using SHWs greatly expand overall genetic diversity in the breeding program (Warburton et

al., 2006). CIMMYT, after then, more focused on developing spring habit synthetic wheats to identify and apply genetic diversity of wild wheats in breeding programs through wide crossing program from 1985 and generated 1014 spring habit, and about 200 winter habit SHWs (Mujeeb-Kazi et al., 1996; Ginkel and Ogbonnaya, 2005). Although SHWs had less remarkable phenotypes such as high stature, late maturing, and hard threshing due to hard glume (Dreisigacker et al., 2008; Li et al., 2014), they can easily cross to cultivated wheats and lines and free-threshing can be recovered by one or two back-crossing to common wheat (Dreisigacker et al., 2008). Therefore, SHWs have been crossed to the advanced common wheat to develop thousands of synthetic derivative lines (SDLs). Also to improve grain yield in the low-yielding environments of the world, CIMMYT increasingly incorporated SDLs germplasm in the international yield trials such as Semi-Arid Wheat Yield Trial (SAWYT) to improve adaptation of new advanced lines to drier environments (Ortiz et al., 2007; Lage and Trethowan, 2008; Rattey et al., 2011). It was reported that one-third of new advanced bread wheat lines developed by CIMMYT for irrigated and low rainfed areas were SDLs (Ginkel and Ogbonnaya, 2005; Van Ginkel and Ogbonnaya, 2007).

CIMMYT included SDLs in the 5th Semi-Arid Wheat Yield Trial (SAWYT) for first time, which were trials for worldwide low-rainfall regions, in 1997 for which eight percent of the lines were SDLs. Then the proportion of SDL increased to 46% in the 15th SAWYT (Figure 1.4 A) (Lage and Trethowan, 2008). However, coefficient of parentage of SHWs in SDLs decreased from 75% in the 5th SAWYT to 19% in the 15th SAWYT (Figure 1. 4 A). The average rank of SDLs for yield performance across wide

range environments in SAWYT increased from 30 (out 50 lines) in 5th SAWYT to 23 in 12th SAWYT (Figure 1.4 B). In the 11th and 12th SAWYT, "Vorobey" as a SDL performed equally or slightly better than the best local check from the lowest to highest yielding environments (Lage and Trethowan, 2008). It was concluded that one or two backcrosses are enough to transfer desirable traits from SHWs to common wheat cultivars and lines with minimal linage drag (Lage and Trethowan, 2008).

Lopes and Reynolds (2011) by evaluating a group of elite SDLs under drought stress conditions reported that they outperformed recurrent parents due to increased root mass at depth (60 to 120 cm) to better water extraction, increased water use efficiency, and earliness. Also, they mentioned that "Vorobey" had the highest root mass under drought. They concluded that wild relatives of wheat are valuable genetic source to improve stress-adaptive traits.

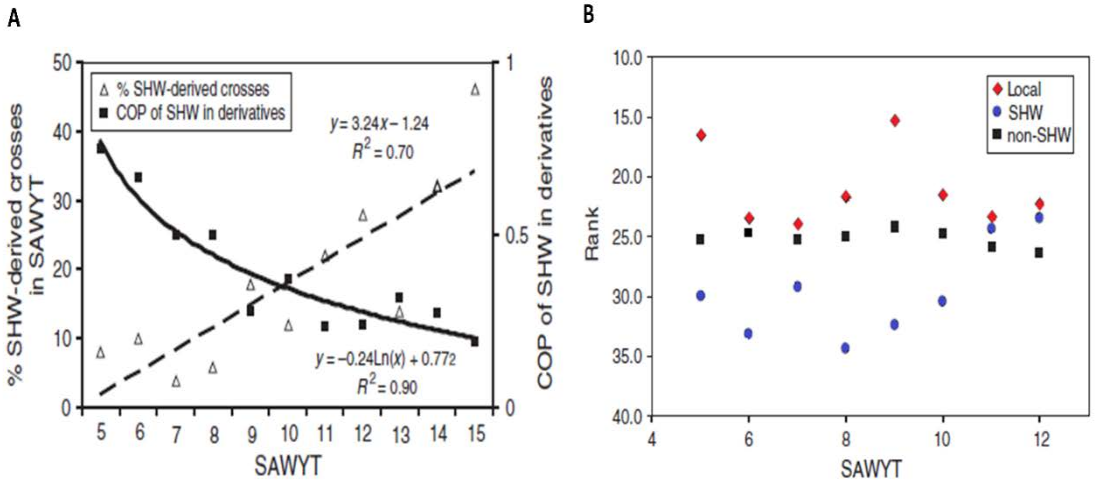


Figure 1.4. Percentage of SHW-derived lines(SDLs) and coefficient of parentage (COP) of SHW in SDLs in Semi-Arid Wheat Yield Trial (SAWYT) (A), and average ranking (1-50) of all SHW, no-SHW, and local checks for SAWYT 5 to 12 (B) (Lage and Trethowan 2008).

Synthetic wheats in Australia

Several studies have been conducted to evaluate performance of SHWs and SDLs from CIMMYT in wheat growing regains in Australia. Gororo et al. (2002) by evaluating a group of SDLs in low-yielding environments in southern Australia, which experienced terminal drought and heat stresses, reported that SDLs out-yielded significantly up to 149% recurrent bread wheat parents due to increased rates of grain filling and larger grain size, which indicated potential of *Ae. tauschii* for improving common wheat.

Trethowan (2004) by studying a set of SDLs in a limited number of environments in north-eastern Australia reported that SDLs performed well under lower yielding and drought environments and mainly out yielded locally adapted wheat varieties.

A research program on SHWs and SDLs called Synthetic-Enriched Resources for Genetic Enhancement (synERGE) program was coordinated through the Department of Primary Industries (DPI) Victoria, Australia (Lillemo 2005). This program designed to improve Australian's wheat varieties and lines for pre-harvest sprouting, drought, and salinity tolerance, cereal cyst nematode (CCN) and yellow leaf spot resistance traits (Lillemo 2005). Through this program it was identified that SDLs had 18 to 30% higher grain yield than commercial varieties in rainfed conditions of Australia.

Rattey et al. (2011a and b) evaluated 273 new conventional hexaploid spring wheats and SDLs from CIMMYT along with 15 locally adapted Australian cultivars

(Oz lines) over four years in a total of 27 environments in north-eastern Australia from 2005 to 2008. Their aim was to compare adaptation of new CIMMYT germplasm and Australian locally adapted spring bread wheat cultivars. They reported that new SDLs were specifically adapted to the lower yielding environments where SDLs' yield was 5 to 13% higher than that of the broadly adapted Oz lines. They mentioned that higher yield of SDLs was due to higher grain weight, canopy temperature depression (CTD), maturity biomass, water soluble carbohydrates, and plant height. They argued that higher CTD of SDLs (cooler canopies) most likely related to increased root in depth that enabled SDLs to extract moisture from deeper soil profiles which was in agreement with the results that reported by Lopes and Reynolds (2010).

Rattey et al.(2011a and b) also reported that across all environments, within low and medium yielding environments, SDLs with two (one back-cross to bread wheat) or fewer doses of bread wheat (bi-parental) had significantly higher grain yield, plant height and seed number m^{-2}, and slightly lower days to anthesis than those with three or more doses. They concluded that may be fewer crosses to bread wheats enable a better balance between integration of beneficial genes from the primary SHWs and preservation of favorable linkage blocks from the high yielding spring bread wheat. Rattey et al. (2011a and b) concluded that their results supported using of SHWs and SDLs to improve adaptation of new wheat lines to drier environments (Ortiz et al. 2007).

Talbot and Hons (2011) evaluated grain yield and yield components of 27 SBLs families in five drought-stressed environments in southern Australia and crossed 14

selected of them to the Australian bread wheat cultivar “Yitpi”. Progenies of these crosses significantly out yielded Yitpi such that in higher-yielding environments, higher grain yield was mostly due to increased grain weight, while in lower-yielding environments, it was mainly because of increased number of grains per square meter.

Synthetic wheats in China

China was the first country that released high-yielding SDLs and commercialized the potential novel genetic diversity of SHWs. China used CIMMYT synthetic wheats (more than 200) in their breeding program in order to improve large kernels, higher spick weight, resistance to new races of stripe rust in the Sichuan province since 1995 (Lillemo, 2005; Yang et al., 2009). Furthermore, about 300 new SHWs including different *T. turgidum* and wild *Ae. tauschii* accessions have been developed in China (Li et al., 2014). Using SHWs in wheat breeding program resulted in releasing four SDLs in China since 2003 and the successful one was Chuanmai 42 that out-yielded commercial variety by 23%. This SDL have been grown more than 100,000 ha since 2006. Using these SDLs as parents in breeding program, 12 new synthetic varieties have been developed (Li et al. 2009; Li et al. 2014). Two of them were Mianmai 367 and Chuanmai 104 that derived from Chuanmai 42 and released in 2011 and 2013 (Tang et al., 2015). They reported that the increased yield of SDLs mainly attributed to seed number m^{-2} and thousand kernel weight, increased biomass and harvest index. Furthermore, stronger vigor in the early growth stage, more above grand dry matter accumulation, and higher spick dry weight at anthesis were physiological components of SDLs that made them to have higher grain yield (Tang et

al., 2015).

A phenotypic and genotypic analysis of SDLs in China showed that primary SHWs and SDLs significantly extended genetic diversity and adaptive evolution of modern hexaploid wheats. Also, they found that introgressed SHW alleles contributed great number of characters to the new varieties including disease resistance, abiotic tolerance, more tillers per plant, more grains spike, larger grains, and higher grain yield. They conclude that using SHWs as a valuable genetic resource was successful in their breeding program (Li et al., 2014).

Other institutions or centers around the world including International Center for Agricultural Research in the Dry Areas (ICARDA) in Syria, Department of Primary Industries (DPI) in Australia, Institute of Plant Genetics and Crop Plant Research (IPK) in Germany, Kyoto University in Japan, and United States Department of Agriculture-Agricultural Research Service (USDA- ARS) in the USA have been developed numerous SHWs, SDLs and included in breeding programs (Ogbonnaya et al., 2013). Results of all these institutions or centers demonstrated that SHWs had outstanding potential to increase world wheat production by extending genetic diversity, introgression of QTL for desirable traits to common wheat, improving grain yield across a diverse range of environments predominantly in moisture-limited or drier environments (Ogbonnaya et al., 2013).

Research objectives

The objectives of this study were i) to determine the capability of synthetic

hexaploid (SYN) lines to increase the genetic diversity of cultivated bread wheat parents ii) to estimate breeding values of SYN lines and bread wheat parents under fully irrigated, heat and drought stress environments, and iii) to evaluate the performance and estimate breeding values of SDLs in fully irrigated, heat and drought stress environments. iv) to identifying key genomic regions or QTL from synthetics hexaploid wheat that should be retained in synthetic derived lines (SDLs) for grain yield and phenological traits

References

Curtis, T., and N.G. Halford. 2014. Food security: The challenge of increasing wheat yield and the importance of not compromising food safety. Ann. Appl. Biol. 164(3): 354–372.

Dreisigacker, S., M. Kishii, J. Lage, and M. Warburton. 2008. Use of synthetic hexaploid wheat to increase diversity for CIMMYT bread wheat improvement. Aust. J. Agric. Res. 59(5): 413–420. http://www.publish.csiro.au/?paper=AR07225.

Gao, F., W. Wen, J. Liu, A. Rasheed, G. Yin, X. Xia, X. Wu, and Z. He. 2015. Genome-Wide Linkage Mapping of QTL for Yield Components, Plant Height and Yield-Related Physiological Traits in the Chinese Wheat Cross Zhou 8425B/Chinese Spring. Front. Plant Sci. 6 (December): 1099. http://www.ncbi.nlm.nih.gov/pubmed/26734019%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4683206.

Van Ginkel, M., and F. Ogbonnaya. 2007. Novel genetic diversity from synthetic wheats in breeding cultivars for changing production conditions. F. Crop. Res. 104(1–3): 86–94. http://www.sciencedirect.com/science/article/pii/S0378429007001153%5Cnhttp://pdn.sciencedirect.com/science?_ob=MiamiImageURL&_cid=271162&_user=1196840&_pii=S0378429007001153&_check=y&_origin=article&_zone=toolbar&_coverDate=31-Dec-2007&view=c&originContentF (verified 28 May 2013).

Ginkel, M. Van, and F. Ogbonnaya. 2005. Invited paper : Using synthetic wheats to breed cultivars better adapted to changing production conditions. : 1–15.

Gororo, N.N.N., H.A.A. Eagles, R.F.F. Eastwood, M.E.E. Nicolas, and R.G.G. Flood. 2002. Use of *Triticum tauschii* to improve yield of wheat in low-yielding environments. Euphytica 123(2): 241–254. http://link.springer.com/article/10.1023/A%3A1014910000128 (verified 27 March 2015).

Lage, J., and R.M. Trethowan. 2008. CIMMYT's use of synthetic hexaploid wheat in breeding for adaptation to rainfed environments globally. Aust. J. Agric. Res. 59(5): 461–469.

Lage, J., M.L. Warburton, J. Crossa, B. Skovmand, and S.B. Andersen. 2003. Assessment of genetic diversity in synthetic hexaploid wheats and their *Triticum dicoccum* and *Aegilops tauschii* parents using AFLPs and agronomic traits. Euphytica 134(3): 305–317.

Li, J., H.S. Wan, and W.Y. Yang. 2014. Synthetic hexaploid wheat enhances variation and adaptive evolution of bread wheat in breeding processes. J. Syst. Evol. 52(6): 735–742. http://doi.wiley.com/10.1111/jse.12110.

LI, J., H.-T. WEI, S.-J. YANG, C.-S. LI, Y.-L. TANG, X.-R. HU, and W.-Y. YANG. 2009. Genetic Effects of 1BS Chromosome Arm on the Main Agronomic Traits in Chuanmai 42. Acta Agron. Sin. 35(12): 2167–2173. http://pub.chinasciencejournal.com/article/getArticleRedirect.action?doiCode=10.3724/SP.J.1006.2009.02167.

Lillemo, M., Van Ginkel, M., Trethowan, R. M., Hernandez, E., & Crossa, J. 2005. Differential adaptation of CIMMYT bread wheat to global high temperature environments. Crop Science, 45(6), 2443-2453.

Lopes, M.S., and M.P. Reynolds. 2010. Partitioning of assimilates to deeper roots is associated with cooler canopies and increased yield under drought in wheat. Funct. Plant Biol. 37(2): 147–156.

Lopes, M.S., and M.P. Reynolds. 2011. Drought adaptive traits and wide adaptation in elite lines derived from resynthesized hexaploid wheat. Crop Sci. 51(4): 1617–1626. https://www.crops.org/publications/cs/abstracts/51/4/1617 (verified 4 June 2013).

Maxted, N., and S. Kell. 2009. Establishment of a Global Network for the In Situ Conservation of Crop Wild Relatives: Status and Needs. FAO Comm. Genet. Resour. Food Agric. Rome, Italy: 266.

Mujeeb-Kazi, A., V. Rosas, and S. Roldan. 1996. Conservation of the genetic variation of *Triticum tauschii* (Coss.) Schmalh. (*Aegilops squarrosa* auct. non L.)

in synthetic hexaploid wheats (*T. turgidum* L. s.lat. x *T. tauschii*; 2n=6x=42, AABBDD) and its potential utilization for wheat improvement. Genet. Resour. Crop Evol. 43(2): 129–134. http://www.springerlink.com/index/10.1007/BF00126756.

Ogbonnaya, F.C., O. Abdalla, A. Mujeeb-Kazi, A.G. Kazi, S.S. Xu, N. Gosman, E.S. Lagudah, D. Bonnett, M.E. Sorrells, and H. Tsujimoto. 2013. Synthetic hexaploids: Harnessing species of the primary gene pool for wheat improvement. Plant Breed. Rev. 37(1): 35–122.

Ogbonnaya F C, Halloran G M., L.E.S. 2005. D genome of wheat - 60 years on from Kihara, Sears and McFadden. p. 205–220. *In* K. Tsunewaki (ed.), In: Frontiers of Wheat Bioscience, Memorial Issue, Wheat Information Service No.100. Kihara Memorial Foundation, Yokohama.

Ortiz, R., R. Trethowan, G.O. Ferrara, M. Iwanaga, J.H. Dodds, J.H. Crouch, J. Crossa, and H.J. Braun. 2007. High yield potential, shuttle breeding, genetic diversity, and a new international wheat improvement strategy. Euphytica 157(3): 365–384.

Rana, R.M., M. Bilal, S.U. Rehman, F. Iqbal, and M.K.N. Shah. 2013. Synthetic wheat; a new hope for the hungry world. Asian J. Agric. Biol. 1(2): 91–94.

Rattey, A.R., R. Shorter, and S.C. Chapman. 2011. Evaluation of CIMMYT conventional and synthetic spring wheat germplasm in rainfed sub-tropical environments. II. Grain yield components and physiological traits. F. Crop. Res. 124(2): 195–204. http://dx.doi.org/10.1016/j.fcr.2011.02.006.

Ray, D.K., N.D. Mueller, P.C. West, and J.A. Foley. 2013. Yield Trends Are Insufficient to Double Global Crop Production by 2050. PLoS One 8(6).

Sehgal, D., P. Vikram, C.P. Sansaloni, C. Ortiz, C. Saint Pierre, T. Payne, M. Ellis, A. Amri, C.D. Petroli, P. Wenzl, and S. Singh. 2015. Exploring and mobilizing the gene bank biodiversity for wheat improvement. PLoS One 10(7): e0132112. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0132112 (verified 13 April 2016).

Shewry, P.R. 2009b. Wheat. J. Exp. Bot. 60(6): 1537–1553. doi:10.1093/jxb/erp058.

Sohail, Q., T. Inoue, H. Tanaka, A.E. Eltayeb, Y. Matsuoka, and H. Tsujimoto. 2011. Applicability of Aegilops tauschii drought tolerance traits to breeding of hexaploid wheat. Breed. Sci. 61: 347–357.

Talbot, S.J., and B. Hons. 2011. Introgression of genetic material from primary

synthetic hexaploids into an Australian bread wheat (*Triticum aestivum* L .) A thesis submitted in fulfilment of the requirements for the degree of Master of Agricultural Science at the University of Adelaide. (March).

Tang, Y., G.M. Rosewarne, C. Li, X. Wu, W. Yang, and C. Wu. 2015. Physiological factors underpinning grain yield improvements of synthetic-derived wheat in Southwestern China. Crop Sci. 55(1): 98–112. https://www.crops.org/publications/cs/abstracts/55/1/98.

Villareal, R.L., A. Mujeeb-Kazi, G. Fuentes-Davila, S. Rajaram, and E. Del Toro. 1994. Resistance to karnal bunt (*Tilletia indica* Mitra) in synthetic hexaploid wheats derived from *Triticum turgidum* x *T. tauschii*. Plant Breed. 112: 63–69.

Warburton, M.L., J. Crossa, J. Franco, M. Kazi, R. Trethowan, S. Rajaram, W. Pfeiffer, P. Zhang, S. Dreisigacker, and M. Van Ginkel. 2006. Bringing wild relatives back into the family: Recovering genetic diversity in CIMMYT improved wheat germplasm. Euphytica 149(3): 289–301.

Weigand, C. 2011. Wheat Import Projections Towards 2050. (January): 1–14. http://www.uswheat.org/USWPublicDocs.nsf/0/80e8bea4307087a58525781e0080f934/$FILE/Wheat Import Projections Towards 2050 - C. Weigand Jan 2011.pdf.

Yang, Z., Z. Bai, X. Li, P. Wang, Q. Wu, L. Yang, L. Li, and X. Li. 2012. SNP identification and allelic-specific PCR markers development for TaGW2, a gene linked to wheat kernel weight. Theor. Appl. Genet. 125(5): 1057–1068.

Yang, W., D. Liu, J. Li, L. Zhang, H. Wei, X. Hu, Y. Zheng, Z. He, and Y. Zou. 2009. Synthetic hexaploid wheat and its utilization for wheat genetic improvement in China. J. Genet. Genomics 36(9): 539–546. http://www.ncbi.nlm.nih.gov/pubmed/19782955.

Zhang, P., S. Dreisigacker, A.E. Melchinger, J.C. Reif, A. Mujeeb-Kazi, M. Van 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. http://repository.cimmyt.org/xmlui/handle/10883/1671 (verified 14 January 2016).

CHAPTER 2

Breeding Value of Primary Synthetic Wheat Genotypes for Grain Yield

Abstract

To introduce new genetic diversity into the bread wheat gene pool from its progenitor, *Aegilops tauschii* (Coss.) Schmalh, 33 primary synthetic hexaploid wheat genotypes (SYN) were crossed to 20 spring bread wheat (BW) cultivars at the International Wheat and Maize Improvement Center. Modified single seed descent was used to develop 97 populations with 50 individuals per population using first back-cross, biparental, and three-way crosses. Individuals from each cross were selected for short stature, early heading, flowering and maturity, minimal lodging, and free threshing. Yield trials were conducted under irrigated, drought, and heat-stress conditions from 2011 to 2014 in Ciudad Obregon, Mexico. Genomic estimated breeding values (GEBVs) of parents and synthetic derived lines (SDLs) were estimated using a genomic best linear unbiased prediction (GBLUP) model with markers in each trial. In each environment, there were SDLs that had higher GEBVs than their recurrent BW parent for yield. The GEBVs of BW parents for yield ranged from -0.32 in heat to 1.40 in irrigated trials. The range of the SYN parent GEBVs for yield was from -2.69 in the irrigated to 0.26 in the heat trials and were mostly negative across environments. The contribution of the SYN parents to improved grain yield of the SDLs was highest under heat stress, with an average GEBV for the top 10% of the SDLs of 0.55 while the weighted average GEBV of their corresponding recurrent BW parents was 0.26. Using the pedigree-based model, the accuracy of genomic prediction for yield was 0.42, 0.43,

and 0.49 in the drought, heat and irrigated trials, respectively, while for the marker-based model these values were 0.43, 0.44, and 0.55. The SYN parents introduced novel diversity into the wheat gene pool. Higher GEBVs of progenies were due to introgression and retention of some positive alleles from SYN parents.

Keywords: genetic diversity, synthetic wheats, genomic estimated breeding value, genomic prediction and grain yield.

Introduction

Domestication and breeding of wheat for many years has increased yield, but recently this increase has slowed down, in part, due to the reduction of genetic variation in the cultivated wheat gene pool (Dreisigacker et al., 2008). Bread wheat (*Triticum aestivum* L.) originated by natural hybridization between durum wheat (*Triticum. turgidum* L. subsp. *durum*) and *Aegilops tauschii* (Coss.) Schmalh, but this probably only happened one or a few times and involved only a few progenitors. Consequently, potential genetic diversity in durum and *Ae. tauschii* was not represented in bread wheat germplasm (Dreisigacker et al., 2008; Li et al., 2014). One approach to introducing new genetic diversity into the cultivated bread wheat gene pool from wheat progenitors is to develop and use synthetic hexaploid wheat (SYN) in breeding (Mujeeb-Kazi et al., 1996). The SYNs are amphiploids resulting from interspecific crosses between a diploid *Ae. tauschii*, donor of the D genome and a modern durum or emmer wheat (*Triticum turgidum* L subsp. *dicoccum)* wheat donor of the A and B genomes. About 1200 winter and spring habit SYN lines have been developed at the International Maize and Wheat Improvement Center (CIMMYT) since the 1980s (Van Ginkel and Ogbonnaya, 2007).

Using SYNs, considerable genetic diversity has been captured from the progenitors of bread wheat (Mujeeb-Kazi et al., 1996; Zhang et al., 2005). The practical value of this diversity can be seen in the resistance to a range of biotic stresses such as Karnal bunt (*Tilletia indica* Mitra) (Villareal et al., 1994), stripe rust (*Puccinia striiformis* f. sp. xtritici) (Kema et al., 1995), Septoria tritici blotch (*Mycosphaerella graminicola* (Fückel.) J. Schröt in Cohn) (Simón et al., 2005), cereal cyst nematode (*Heterodera avenae* Wollenweber) (Mulki et al., 2013) and stem rust (*Puccinia graminis* Pers.: Pers. f. sp. *tritici* Eriks. E. Henn.) (Ogbonnaya et al., 2008). Also, SYNs are a valuable genetic resource for abiotic stress such as drought (Lopes and Reynolds, 2011). Lopes and Reynolds (Lopes and Reynolds, 2011) reported that synthetic derived wheat lines (SDLs) increased drought tolerance which was attributed to traits such as earlier flowering, greater root mass at depth, greater water extraction capacity, and increased water use efficiency at anthesis to produce an average of 26% higher grain yield than the cultivated wheat parents under terminal drought. Hence, crossing SYNs to modern wheat cultivars could result in more productive cultivars for such stress environments. Furthermore, studying yield potential of synthetic backcross-derived lines (SBLs) in the diverse rain-fed environments of Australia showed that SBLs out-yielded the best local checks by 8 to 30% (Ogbonnaya et al., 2007). Cooper et al.(Cooper et al., 2012) backcrossed ten elite primary synthetics to two Texas winter wheat cultivars, TAM111 and TAM112, and evaluated SBLs for yield and yield components. They reported that improved yield in the SBLs was due to an increased number of heads per unit area and grains per head.

In China, SYN lines have been used in breeding programs and four synthetic derived cultivars, Chuanmai 38, Chuanmai 42, Chuanmai 43 and Chuanmai 47 were released and are widely grown by farmers. Of these, Chuanmai 42 had large kernels, resistance to stripe rust, and its grain yield was 16.4 to 22.7% higher than the commercial check, Chuanmai 107 (Yang et al., 2009; Li et al., 2011).

Molecular markers can be used to evaluate the diversity within and among germplasms and to monitor genetic diversity over time (Russell et al., 2000; Christiansen et al., 2002; Heckenberger et al., 2002). Also, molecular markers allow more accurate prediction of breeding values of genotypes through improved estimates of relatedness and estimation of marker effects(Bassi et al., 2015). These values can be used in genomic selection (GS) (Meuwissen et al., 2001) or marker-assisted recurrent selection (MARS) (Eathington et al., 2007). Li et al. (2011) used simple sequence repeat (SSR) markers to transfer a quantitative trait locus (QTL) on chromosome 4D from a synthetic parent, Syn769 to Chuanmai-42. The QTL increased tiller number per plant, number of effective spikes, grains per square meter, harvest index, and grain yield. The authors reported that the average increased grain yield due to this QTL was 8.90%. Additionally, Zhang et al. (2005) studied the genetic variation of SYNs and SBLs using SSR markers and concluded that the novel alleles from SYNs were stably inherited in SBL families and introduced the genetic diversity from *Ae. tauschii* and durum parents to SBLs. They argued that SYNs and SBLs are valuable genetic resources for broadening genetic diversity of wheat breeding germplasm.

The objectives of this study were i) to determine the capability of SYN lines to increase the genetic diversity of cultivated parents ii) to estimate breeding values of SYN lines and bread wheat parents under fully irrigated, heat and drought stress environments, and iii) to evaluate the performance and estimate breeding values of SDLs in fully irrigated, heat and drought stress environments.

Materials and methods

Population development

The populations of SDLs were developed by crossing 20 CIMMYT spring bread wheat (BW) cultivars to 33 primary SYN lines (Table S2.1) using a direct cross (biparental), a first backcross (BC1) and a three-way cross (TC) in 2008. Plants in the segregating populations were selected in a shuttle-breeding program alternating between Yaqui Valley, Ciudad Obregon, north-western Mexico (elevation 38 m, 27°25′ N, 109°54′ W, 320 mm rainfall) and El Batán in the semiarid, subtropical highlands of central Mexico (elevation 2240 m and 19.32°N, 98.51°W, 625 mm rainfall). In the F_1 generation, individuals of some crosses were selected to create biparental families and some of them were crossed to a recurrent BW parent to create BC_1 families as part of routine pre-breeding activities to introgress novel genetic diversity into adapted bread wheat backgrounds. Others were crossed to another BW parent to develop TC families. The breeding scheme thereafter was a modified single seed descent in which 50 individual plants (spikes) per cross were selected in the F_2 generation to plant in F_3 rows (spike to row). In the F_3 generation, a single spike per row was selected for the next generation (50 spikes from 50 rows). In the F_4 and BC_1F_3 generations, rows were bulk

harvested separately for the next year. $F_{4:5}$ and $BC_1F_{3:4}$ bulks were planted in 3m long by 80cm wide raised beds and irrigated to increase seed (bed–channel system) and each family had 50 rows. In the early generations, plants were selected that had semi-dwarf plant height and phenology similar to the adapted parents and in the later generations ($F_{4:5}$ and $BC_1F_{3:4}$), lines were selected for lodging resistance and free threshing. The overall population comprised 97 families with 50 derived $F_{4:5}$ and $BC_1F_{3:4}$ lines. The number of lines per family was reduced in the $F_{4:6}$ and $BC_1F_{3:5}$ due to selection for basic agronomic type and uniformity and family sizes ranged from 1 to 48 and the total number of lines was 2080 in the first year yield trials. In the second and third years the number of families was reduced due to selection for easy threshing, early maturity, plant height, and lodging resulting in 80 families consisting of 13 BW parents and 30 SYN parents. The SYN parents were genotyped but were not planted in the field because of the poor agronomic characteristics and lack of threshability.

Field trials

The selected populations were planted in three parallel trials under the fully irrigated, drought and heat stress conditions at the Norman E. Borlaug Research station (CENEB) in the Yaqui Valley, Ciudad Obregon, northern Mexico (elevation 38 m, 27°25′ N, 109°54′ W) in the year 2011-12. This station is located in an arid region with average precipitation of 320 mm, a mean annual temperature of 24 °C, and its soil was a Hyposodic Vertisol (Calcaric, Chromic) (Verhulst et al., 2011).

The experimental design, for all trials, was a partially replicated design in which 20 percent of genotypes had two replicates and the remainder was unreplicated. The

number of unique genotypes including SDLs, BW parents, and checks in irrigated, drought and heat trials was 2052, 1493, and 1463, respectively, and Vorobey and Quaiu were checks in all trials. The proportions of BC and TC SDLs were 92 and 8%, respectively, for drought and heat trials while for the irrigated trial the BC, biparental, and TC were 68, 27, and 5%, respectively.

The sowing system was bed-channel for the irrigated and heat trials in which each bed (plot) was 3 m long and had two rows 40cm apart with 40cm between beds. Two beds were used for each genotype in the irrigated trial while in the heat trial there was one bed per line. These two trials were fully irrigated. The irrigated trial was planted on December 5th, 2011 while the heat trial was planted on March 23rd, 2012 to coincide with high temperature stress. The drought trial was planted on December 8th, 2011 on a flat plot area without beds and irrigated twice with a drip irrigation system, once at sowing, and again about 45 days later to impose post anthesis drought stress. Plots in the drought trial were wider than the bed system to reduce the relative contribution of plants growing on plot edges and to have a canopy more like in a farmer's field in a drought stressed growing region. Each plot was 1.6 m wide, 3 m long and had 6 rows.

For the second year, 2012-13, the number of lines was decreased based on grain yield in the irrigated, heat, and drought trials, easy threshing, early maturity, plant height, and lodging. Consequently, the number of unique genotypes including SDLs, BW parents, and checks were 1057, 1054, and 1045 in the irrigated, drought, and heat trials, respectively. These were planted in three parallel trials; fully irrigated, drought, and heat stress, respectively. The sizes of beds and plots were the same as in 2011-12

except for the irrigated trial in which one bed was used per line. Planting dates of the irrigated and heat trials were November 25th, 2012 and March 8th, 2013. The irrigated and heat trials were irrigated five and six times through gravity flood-irrigation, respectively. The drought trial was irrigated twice.

In the year 2013-14, the irrigated, drought and heat trials were planted on December 6th, 2013, December 20, 2013, and February 27th, 2014, respectively. The irrigation system and number of irrigations of trials were the same as the second year. Also, the unique number of lines in the irrigated, heat and drought trials was 1056, 1056, and 1054, respectively.

Field experimental design for heat and irrigated trials in the years 2012-13 and 2013-14 was alpha lattice with two replicates while for drought trials it was augmented design. The cultivars Vorobey, Navojoa, Roelfs, Reedling and Quaiu were checks in all trials. The BC and biparental SDLs made up the main part of the population with proportions of 74% and 20%, respectively, followed by 6% TC populations.

Phenotyping

Each year, plant height (PLH), days to heading (DHE), days to flowering (DFL), days to maturity (DMA), and grain yield (YLD t/ha) were measured in all trials according to Pask et al. (Pask AJD, Pietragalla J, Mullan DM, Reynolds MP and Reynolds, 2012). The traits were measured as: DHE; when 50% of the spikes in a plot emerge from the flag leaf sheath, DFL; when 50% of the spikes in a plot reached anthesis, DMA; when 50% of the peduncles in a plot had lost green coloration, and

YLD; measured grain yield (t/ha) from each plot. Thousand kernel weight (TKW) and grain filling duration (GFD) were only measured for the irrigated trial in the year 2011-12 (Pask AJD, Pietragalla J, Mullan DM, Reynolds MP and Reynolds, 2012).

Phenotypic data analysis

The experimental designs were different for each year and trial complicating combined analysis of all trials. To correct for within field heterogeneities spatial analysis was used for each trait/trial combination separately based on row and column orders. The Genstat software (Payne RW, Murray DA, Harding SA, Baird DB, 2009) was used for analysis of the general linear mixed model by the following equation;

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{Z_R u_R} + \boldsymbol{Z_C u_C} + \boldsymbol{\varepsilon}$$

where $\boldsymbol{Y}$ is the response vector, **X** is a design matrix for fixed effects such as overall mean and genotype effects. $\boldsymbol{Z_R}$ is a design matrix for row effects, $\boldsymbol{Z_C}$ is a design matrix for column effects, $\boldsymbol{\beta}$ is a vector for fixed effects, $\boldsymbol{u_R}$ and $\boldsymbol{u_C}$ are vectors for random row and column effects with $\boldsymbol{u_R} \sim N(0, \sigma_R^2\, \boldsymbol{I})$, and $\boldsymbol{u_C} \sim N(0, \sigma_C^2\, \boldsymbol{I})$ correspondingly and $\boldsymbol{\varepsilon}$ is a residual vector with $\boldsymbol{\varepsilon} \sim N(0, \sigma_R^2 R)$, where R is given by $R = Z_\varepsilon[AR1(\rho_R) \otimes AR1(\rho_C)]Z_\varepsilon'$. $AR1(\rho_R)$ is an auto-regressive order one correlation matrix for row effects, $AR1(\rho_C)$ is an auto-regressive order one correlation matrix for column effects and Z_ϵ is a design matrix for row and column combinations. Consequently, row and column effects were removed in each trial and best linear unbiased estimates (BLUEs) of genotypes were generated for subsequent analysis.

Pearson correlation was used to estimate the phenotype correlation coefficients among environments for all traits.

Genotyping

Genomic DNA was extracted from dried leaves collected from a single plant for each line using a modified CTAB (cetyltrimethylammonium bromide) method(Saghai-Maroof et al., 1984) modified as shown in CIMMYT laboratory protocols (Dreisigacker et al., 2013) and quantified using NanoDrop 8000 spectrophotometer V 2.1.0. The genotyping of the samples was accomplished using a genotyping-by-sequencing technique called DArTseq™ developed by DArT Pty. Ltd., Yarralumla, Australia. The detailed protocol is described in Sehgal et al.(Sehgal et al., 2015). A total of 20,468 genotyping–by–sequencing (GBS) markers were used for genotyping of 1991 lines. Marker data were filtered for missing data (NA< 50 %) and minor allele frequency (MAF) (< 1%) for a final number of 10,262 GBS markers selected for subsequent analysis.

Kinship matrices

The genomic relationship matrix, **G** matrix, was generated using 10,262 GBS markers. More specifically, genotyping information was stored in an n x p genotype matrix, X= $\{x_{ik}\}$ where the x_{ik} represents for the i^{th} genotype (i = 1, 2, …, n) and the k^{th} marker (k = 1, 2, ..., p). The biallelic single nucleotide polymorphisms (SNPs) were coded -1, 1, and 0 for A_1A_1, A_2A_2, and A_2A_1, respectively, and NA for missing values. Maximum and average missing values of markers were 30% and 9%, respectively. The

rrBLUP package in R (Endelman and Jannink, 2012) was used to impute the missing data based on expectation maximization (EM) imputation algorithm and generate the **G** matrix as the follows:

$$\boldsymbol{G} = WW'/c$$

where $W = X_{ik} + 1 - 2p_k$, is a centered matrix by mean of allele frequency, $2p_k$, and p_k is the frequency of the 1 allele at marker k, $c = 2\sum_k p_k(1 - p_k)$ is the normalization constant and scales the **G** matrix to be analogous to the numerator relationship matrix **A.**

The numerator relationship matrix, **A** matrix, was created based on pedigree information for populations that included 1986 individuals. More specifically, to generate the **A** matrix, we compared the relatedness of parents and different crosses; biparental, BC and TC for SDLs. For relatedness of SYN lines, f = 0.66 if they had the same durum parents but a different *Ae. squarrosa* parent and f = 0.33 if they had the same *Ae. squarrosa* parent but a different durum parent. For some SYN lines f = 1 if they had the same durum and diploid parents. For BW parents, most of them were unrelated except for two pairs that were identical and f = 1 was used for them.

The heat map of the **G** matrix indicated that there could be some individuals with inconsistencies between the familial relationships given by the **A** matrix and the relationships indicated by the **G** matrix. These individuals were designated as outlier individuals and removed from further study. More specifically, to identify the potential outlier individuals in each family, a distance matrix was created using imputed marker

data. Individuals with a distance larger than Q3+1.5(IQR), where Inter-Quartile Range (IQR) = Q3−Q1, Q1 is the 25th percentile and Q3 is the 75th percentile, within each family were considered outliers. Consequently, 144 individuals belonging to 72 families (from 1 to 7 individuals) were removed from further study. This resulted in the correlation coefficient between off diagonal elements of **A** and **G** matrices increasing from 0.65 to 0.75. Therefore, 1846 genotyped individuals were used for subsequent analyses.

The **H** matrix is a pedigree-marker relationship matrix that modifies the genetic relationship matrix to combine pedigree-based relationship information (VanRaden, 2008; Legarra et al., 2009; Gao et al., 2012). In this study, the **H** matrix was used to combine the pedigree information of 1986 lines with the marker information of 1846 lines. The following covariance matrix was used to create the **H** matrix;

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{A_{11} + A_{12}A_{22}^{-1}(G_w - A_{22})A_{22}^{-1}A_{21}} & \boldsymbol{A_{12}A_{22}^{-1}G_w} \\ \boldsymbol{G_wA_{22}^{-1}A_{21}} & \boldsymbol{G_w} \end{bmatrix},$$

where the pedigree-based relationship matrices $\mathbf{A_{11}}$ and $\mathbf{A_{22}}$ are sub-matrices of **A** matrix for genotyped and non-genotyped individuals, respectively, and $\mathbf{A_{12}}$ or $\mathbf{A_{21}}$ is the covariance matrix between genotyped and non-genotyped individuals. $\mathbf{G_w}$ is the weighted **G** matrix, $\mathbf{G_w = w*G + (1-w)*A_{22}}$, **G** is the genomic relationship matrix and $\boldsymbol{w}$ is the weight for contribution ratio of **A** matrix or portion of genetic variance that was not explained by markers. The ranges of $\boldsymbol{w}$ were from 0 to 1 by 0.1 interval, $\boldsymbol{w}$ =1 represents the **G** matrix and **w =** 0 indicates **A** matrix. In this study different values of $\boldsymbol{w}$ were used to create the **H** matrix and $\boldsymbol{w}$ = 0.1 gave the best overall results in terms of

prediction accuracies in the validation data. Hence, $\boldsymbol{w}$ = 0.1 was used to create the **H** matrix, which included 1986 genotyped and non-genotyped individuals.

Genomic estimated breeding values

The genomic best linear unbiased prediction (GBLUP) model was used to estimate both variance components and genomic estimated breeding values (GEBVs). All analyses were executed with the EMMREML package in R software (Akdemir and Godfrey, 2015). BLUPs were computed using the following univariate mixed model:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{\epsilon},$$

where $\boldsymbol{y}$ is a vector of spatially corrected observations of genotyped individuals for the traits of interest, **X** is a known design matrix for fixed effects which comprised management (Irrigated, heat, and drought environments) and year, **Z** is a known design matrix for random effects (individuals), **β** is a vector for non-genetic fixed effects, **u** is a vector for genetic random effects or breeding values with $\mathbf{u} \sim N(0, \sigma_u^2 \boldsymbol{G})$, **G** is the genomic relationship matrix and $\boldsymbol{\epsilon}$ is a residual vector with $\boldsymbol{\epsilon} \sim N(0, \sigma_e^2 \boldsymbol{I}_n)$ (Piepho et al., 2007). Breeding values were then estimated by solving the mixed model equations. The same model was also fitted by replacing the **G** matrix with **A** and **H** matrices.

Cross Validation and Genomic prediction

The 5-fold cross validation was used to quantify the fidelity of genomic prediction of traits for each trial and all trials together (Mehmani et al., 2015). The accuracy of estimates was based on the correlation between $y - X\beta$ and GEBVs. The

marker, pedigree and pedigree–marker models were used in the training set based on the GBLUP method as described above. Also, mean heritability of traits was estimated using $\frac{\sigma_u^2}{\sigma_u^2+(\frac{\sigma_e^2}{r})}$ in which σ_u^2 and σ_e^2 are genetic and error variances, respectively, and r is the number of replicates for each individual.

Genetic diversity

To measure genetic diversity of BWs, SYNs, SDLs, Nei's genetic diversity, *Hs*, was used (Kosman, 2003). There were 8,612 out of 10,262 SNPs, that had chromosome information, and those were filtered for missing data (NA < 10%) within each group of BWs, SYNs and SDL populations. Nei's genetic diversity was calculated using the formula (Kosman, 2003):

$$\boldsymbol{H_S = \frac{1}{k}\sum_{s=1}^{k} H_{Ss} = \frac{1}{k}\sum_{s=1}^{k}[1 - q_s^2 - (1 - q_s)^2]}$$

Where $\boldsymbol{H_S}$ is genetic diversity index, $\boldsymbol{k}$ is total number of loci, $\boldsymbol{s}$ is diallelic loci, $\boldsymbol{q_s}$, is the allele frequency et the $\boldsymbol{s^{th}}$diallelic locus.

The hierarchical cluster analysis with the Ward method and Euclidean distance (Timm, 2002) was used to classify the BW and SYN parents based on whole genome marker information, 10,262 SNPs.

Results

Phenotypic analysis

The summary information for traits from each trial and year is presented in Table 2.1. Means of the traits in the irrigated trials were similar across the years while means of traits varied widely in the heat and drought trials. For example, DRO.Y13.14 had the lowest mean value, especially for YLD (1.054 t/h), HEAT.Y11.12 had the lowest mean values for PLH and YLD and differed greatly from those in the other two heat trials. This was caused by late planting resulting in very low yield with some genotypes not producing any grain. For this year, YLD ranged from 0 to 2.40 t/h and PLH ranged from 20 to 70 cm. Thus, it was considered to be an outlier environment and the data were not used in subsequent analyses (Table 2.1).

Table 2.1: Mean and range of traits in different trials in years 2011-14 in Ciudad Obregon, CIMMYT, Mexico.

Trial\Trait	DHE	DFL	DMA	PLH (cm)	YLD (t/h)
IRRI.Y11.12	-	81[a] (61-95)[b]	128 (119-36)	114 (87-150)	6.34 (2.90-8.50)
IRRI.Y12.13	73 (58-93)	78 (63-97)	126 (117-36)	102(82-121)	5.95 (2.78-8.94)
IRRI.Y13.14	75 (65-88)	79 (69-92)	121 (107-33)	102 (86-121)	5.55 (3.18-7.59)
DRO.Y11.12	-	81 (72-99)	117 (104-30)	84 (58-120)	2.42 (1.09-3.56)
DRO.Y12.13	75 (65-87)	78 (66-92)	-	-	2.30 (1.55-2.95)
DRO.Y13.14	67 (58-79)	69 (60-80)	100 (91-109)	70 (50-96)	1.05 (0.49-1.40)
HEAT.Y11.12	-	-	-	42 (20-70)	0.57 (0.00-2.40)
HEAT.Y12.13	50 (45-59)	-	81 (78-89)	61 (45-75)	1.96 (0.29-3.18)
HEAT.Y13.14	56 (50-66)	59 (54-69)	87 (82-96)	59 (41-89)	2.07 (0.33-3.26)

DHE: Days to heading, DFL: Days to flowering, DMA: Days to maturity, PLH: Plant height, and YLD: Grain Yield t/h.
IRRI: Irrigated, DRO: Drought, HEAT: Heat trials, Y11.12: Year 2011-12, Y12.13: Year 2012-13, and Y13.14: Year 2013-14 (e.g. IRRI.Y11.12: irrigated trial in the year 2011-12).
[a]; Mean of the trait, [b]; Range of the trait.

All phenotypic correlation coefficients among environments for PLH and YLD were significant (Table 2.2). For YLD, correlations within treatments (irrigated, heat or

drought) across the three years ranged from 0.54 to 0.60 for irrigated trials, 0.42 to 0.61 for heat trials, and 0.42 to 0.49 for drought trials while, correlations between different treatments ranged from 0.13 to 0.59. Over all the trials, correlation coefficients for YLD ranged from 0.13 to 0.61 for HEAT.Y11.12 with IRRI.Y12.13 and HEAT.Y11.12 with HEAT.Y13.14, respectively (Table 2.2 below diagonal). For PLH, correlations within treatments across the three years ranged from 0.68 to 0.78 for irrigated trials, 0.38 to 0.50 for heat trials, and 0.52 for drought trials while, correlations between different treatments ranged from 0.33 to 0.65. Among treatments, correlations for PLH ranged from 0.33 to 0.65 for HEAT.Y11.12 with DRO.Y13-14 and IRRI.Y11.12 with DRO.Y12.13, respectively (Table 2.2 above diagonal).

Phenotypic correlations for DFL (Table 2.3 below diagonal), DMA (Table 2.3 above diagonal), and DHE (Table 2.4) were significant and ranged from 0.26 to 0.84. For these traits, correlations between and within trials for the three years were medium to high except for some low correlations observed for DMA between HEAT.Y12.13 with IRRI.Y11.12 and HEAT.Y12.13 with DRO.Y11.12 (Table 2.3 above diagonal).

Table 2.2: Phenotypic correlations for PLH (above diagonal) and YLD (below diagonal) within and among environments.

Trial/Trait	PLH								
IRRI.Y11.12	1	0.78*	0.69	0.65	-	0.47	0.43	0.44	0.58
IRRI.Y12.13	0.54	1	0.68	0.55	-	0.41	0.41	0.44	0.56
IRRI.Y13.14	0.60	0.54	1	0.58	-	0.43	0.36	0.39	0.56
DRO.Y11.12	0.36	0.14	0.22	1	-	0.52	0.35	0.39	0.51
DRO.Y12.13	0.39	0.34	0.42	0.48	1	-	-	-	-
DRO.Y13.14	0.27	0.17	0.26	0.42	0.49	1	0.33	0.36	0.51
HEAT.Y11.12	0.35	0.13	0.20	0.33	0.33	0.26	1	0.38	0.45
HEAT.Y12.13	0.45	0.41	0.45	0.38	0.59	0.40	0.42	1	0.50
HEAT.Y13.14	0.38	0.28	0.35	0.31	0.52	0.44	0.61	0.59	1
YLD	IRRI. Y11.12	IRRI. Y12.13	IRRI. Y13.14	DRO. Y11.12	DRO. Y12.13	DRO. Y13.14	HEAT. Y11.12	HEAT. Y12.13	HEAT Y13.14

IRRI: Irrigated, DRO: Drought, HEAT: Heat trials, Y11.12: Year 2011-12, Y12.13: Year 2012-13, and Y13.14: Year 2013-14.
*: All correlation coefficients were significant.

Table 2.3: Phenotypic correlation for DMA (above diagonal) and DFL (below diagonal) within and among environments.

Trial/Trait	DMA							
IRRI.Y11.12	1	0.56*	0.56	0.54	-	0.58	0.26	0.40
IRRI.Y12.13	0.70	1	0.58	0.50	-	0.71	0.36	0.59
IRRI.Y13.14	0.73	0.82	1	0.48	-	0.62	0.35	0.41
DRO.Y11.12	0.61	0.51	0.56	1	-	0.62	0.28	0.38
DRO.Y12.13	0.51	0.50	0.52	0.44	1	-	-	-
DRO.Y13.14	0.75	0.84	0.81	0.57	0.55	1	0.40	0.58
HEAT.Y12.13	-	-	-	-	-	-	1	0.46
HEAT.Y13.14	0.54	0.72	0.62	0.39	0.36	0.68	-	1
DFL	IRRI. Y11.12	IRRI. Y12.13	IRRI. Y13.14	DRO. Y11.12	DRO. Y12.13	DRO. Y13.14	HEAT. Y12.13	HEAT. Y13.14

*: All correlation coefficients were significant.

Table 2.4: Phenotypic correlation for DHE within and among environments.

Trial/Trait	DHE					
IRRI.Y12.13	1					
IRRI.Y13.14	0.82*	1				
DRO.Y12.13	0.54	0.59	1			
DRO.Y13.14	0.84	0.82	0.62	1		
HEAT.Y12.13	0.44	0.42	0.31	0.48	1	
HEAT.Y13.14	0.73	0.63	0.41	0.68	0.54	1
	IRRI. Y12.13	IRRI. Y13.14	DRO. Y12.13	DRO. Y13.14	HEAT. Y12.13	HEAT. Y13.14

*: All correlation coefficients were significant.

The range for TKW for the IRRI.Y11.12 trial was from 40 to 65 gr for SDL populations while for 13 BW parents the range was from 41 to 54 gr and for the top 10% of the populations (the top 10% was based on YLD) it was 41 to 58 gr (Table S2.2). Sixty seven percent of SDLs had higher TKW than their corresponding recurrent BW parents. Furthermore, among 26 biparental families, the TKW mean decreased by -2 to -3.92% for four populations, while it increased from 0.67 to 24.39% for 22 populations compared to the TKW mean of the BW parents. The same comparison for 38 BC populations showed that TKW of six populations decreased by – 0.44 to -5.40% while TKW for 32 of them increased from 3.3 to 16.1%. Among the four TC populations, one had the highest reduction for TKW (-17.9%) but TKW for the other three populations increased from 6.83 to 12.68% (Table S2.2).

The range of GFD was from 48 to 62 days over all genotypes in the IRRI.Y11.12 trial. For the 13 BW parents it ranged from 49 to 60 days and for the top 10% of the SDL populations it ranged from 48 to 61 days (Table S2.2).

Relationships between TKW and GFD were significantly positive over the all populations ($y = 0.21x + 44;\ P < 0.001,\ R^2 = 0.17$) and for the top 10% of the SDL populations ($y = 0.15x + 46;\ P < 0.001,\ R^2 = 0.05$) in the IRRI.Y11.12 trial.

Relationships between YLD and GFD were significantly negative over all populations ($y = -0.032x + 8.40;\ P < 0.001, R^2 = 0.02$) while it was not significant for the top 10% of the SDL populations ($y = -0.022x + 8.40,\ R^2 = 0.009$). Also, significant a negative relationship was observed between YLD and TKW overall and for the top 10% of the SDL populations ($y = -0.017x + 7.50;\ P < 0.001, R^2 = 0.02$) and ($y = -0.017x + 8;\ P < 0.05, R^2 = 0.08$), respectively.

Clustering of bread wheat and synthetic parents

As expected, the dendrogram of the hierarchal cluster analysis revealed that SYN lines were more genetically diverse than BW parents (Figure 2.1). For instance, using an arbitrarily cut off, BW parents made one group, cluster 1, while SYN lines grouped into five different clusters.

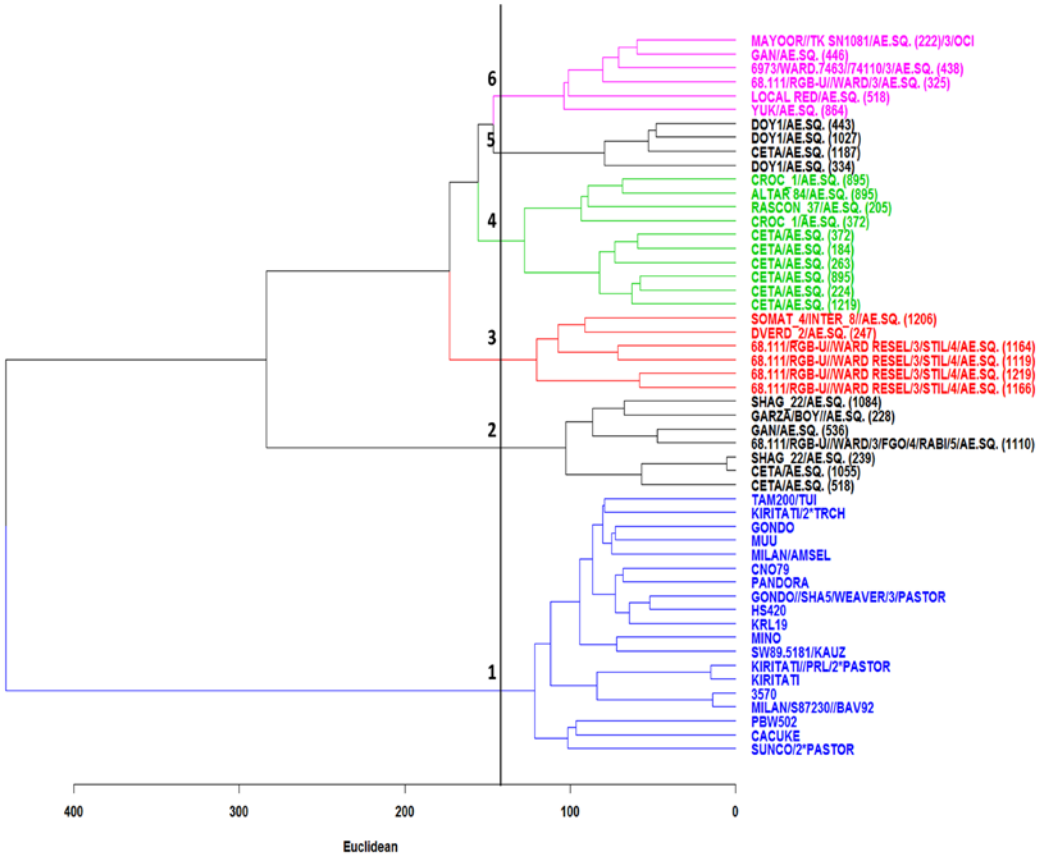


Figure 2.1: Dendrogram of the classification of BW parents (Blue color) and SYN lines using the Ward method based on polymorphic SNP markers.

Marker information for A+B and D genomes were used to investigate the genetic diversity of durum (Figure 2.2 A) and *Ae. squarrosa* parents (Figure 2.2 B) that were used to develop SYN parents. Seventeen durum parents were grouped into four clusters. Cluster 1 comprised five unrelated durum parents. Cluster 2 had only two durum parents CETA and SHAG_22 crossed to AE.SQUARROSA 239, however the durum parent named SHAG_22 was likely to be CETA. Cluster 3 had two durum parents DOY1 and CETA crossed to AE.SQUARROSA 1187, however the durum parent named CETA was likely to be DOY1. Cluster 4 comprised 11 unrelated durum parents (Figure 2.2 A).

Based on D genome markers, 28 AE.SQUARROSA parents were grouped into three clusters. Cluster 1 included seven AE.SQUARROSA that were closely related (Figure 2.2 B). Cluster 2 comprised four unrelated AE.SQUARROSA parents. Cluster 3 included 22 AE.SQUARROSA parents in which some of them were highly related or identical.

Based on whole genome marker information, most of the BW parents of this study were not closely related except for two pairs of lines (Figure 2.2 C). For KIRITATI and KIRITATI//PRL/2*PASTOR BW parents, this could have resulted from being sister lines or from selfed progenies of KIRITATI. For MILAN/S87230//BAV92 with BW line 3570, an error in labeling or seed packaging is more likely. Errors in pedigrees will affect predictions when using the pedigree based relationship **A** matrix or **H** matrix. However, we corrected these errors when generating the **A** matrix.

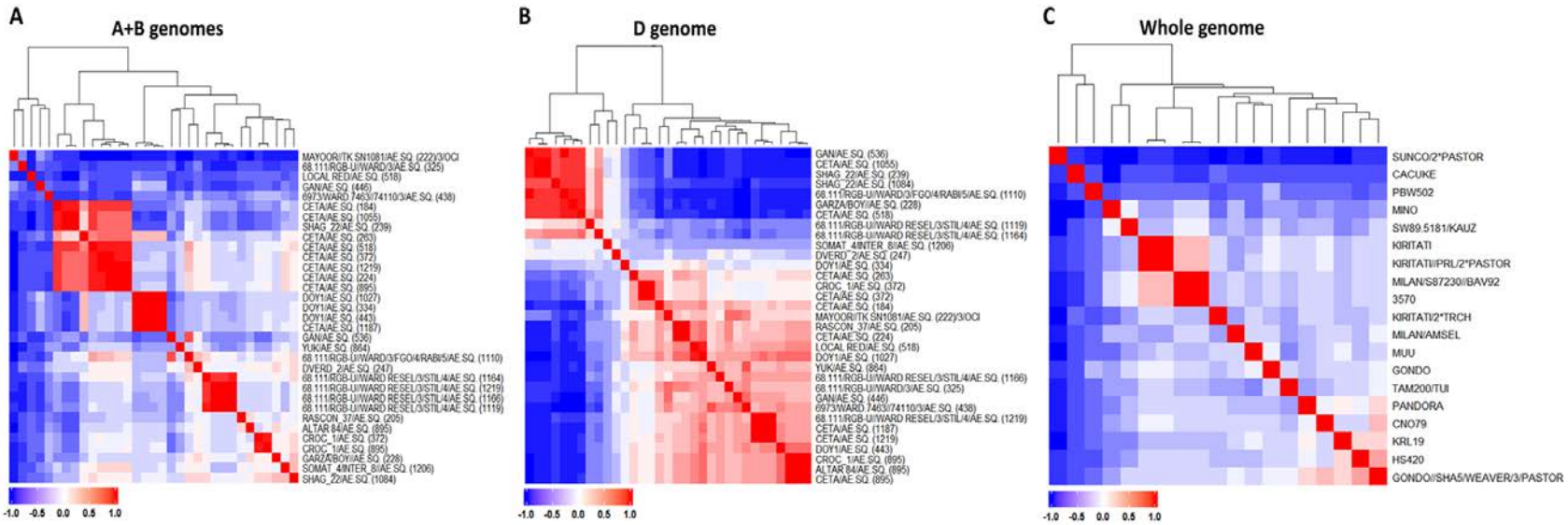


Figure 1.2: Heat map for SYN and BW parents based on genome-specific marker information. (A) Clustering of SYN parents using A+B genomes and (B) D genome, (C) Clustering of BW parents based on whole genome.

Genome distribution of the markers and Nei's genetic diversity (*Hs*) for each genome for BW, SYN parents, and SDLs are shown in Table 2.5. SNP markers were not evenly distributed in the three genomes. The D genome with 3691 had the most markers and the A genome with 2333 had the lowest. For SYNs, *Hs* for A, B, and D genomes were 0.35, 0.38, and 0.40, respectively, and they were greater than those for the BW parents, which were 0.27, 0.26, 0.06 (Table 2.5). For SDLs, *Hs* was 0.36 for A and B genomes and 0.19 for the D genome, all greater than those for BW parents. The mean genetic diversity was 0.19 for BWs, 0.38 for SYNs and 0.28 for SDLs (Table 2.5).

Table 2.5. Distribution of markers and diversity index (H_S) in each genome for BWs, SYNs and SDLs.

		No. marker after filtering for NA < 10%			H_S		
Genome	No. marker	BWs	SYNs	SDLs	BWs	SYNs	SDLs
A	2333	1443	404	1595	0.27	0.35	0.36
B	2587	1584	468	1747	0.28	0.38	0.36
D	3691	2073	929	2630	0.06	0.40	0.19
Total/Mean	8612	5100	1801	5972	0.19	0.38	0.28

Estimating genomic breeding value of parents

The GEBV values of cultivated wheat parents

Most of the BW parents had positive GEBVs for grain yield across all environments and their values ranged from -0.16 to 1.40 under irrigated, -0.15 to 0.43 under drought, and -0.33 to 0.65 under heat environments (Figure 2.3 A and Table S2.3). Among BW parents, MILAN/S87230//BAV92 and BW line 3570 were the best parents and had the highest GEBVs across three environments while MUU, SUNCO/2*PASTOR and MILAN/AMSEL were the poorest parents with very small

positive values in one environment and negative values in the other environments. Parents reflected genotype by environment interaction (GEI) and they usually had the highest GEBVs in the irrigated trials except for KIRITATI/2*TRCH, SUNCO/2*PASTOR, and MUU that had negative values. Generally, GEBVs of parents decreased in stress conditions except for PBW502 and GONDO that had almost the same positive value in irrigated and heat environments. However, SUNCO/2*PASTOR and KIRITATI/2*TRCH had negative yield GEBVs in the irrigated trials and positive values in the heat and drought stress trials.

For DHE, almost all the BW parents had positive GEBVs across environments except KIRITATI/2*TRCH, CACUKE, KRL19. The GEBVs ranged from -7.88 to 9.88 for irrigated, from -3.89 to 5.61 for drought, and from -2.33 to 5.05 for heat environments. SW89.5181/KAUZ had the highest positive GEBVs across all environments while KIRITATI/2*TRCH had the highest negative GEBVs under irrigated and drought conditions. For this trait, GEI was observed and HS420 had very low GEI across environments (Figure 2.3 B and Table S2.3).

For DMA, the trend for GEBVs of BW parents was similar to those for DHE but the values decreased for all parents except for HS420 which increased in drought and irrigated conditions (Figure 2.3 C and Table S2.3). Also, MILAN/S87230//BAV92 and BW line 3570 showed less GEI for DMA than for DHE.

The PLH GEBVs were nearly all negative for BW parents except for CACUKE that had positive values in all environments and four other parents that had at least one positive value in one environment (Figure 2.3 D and Table S2.3).

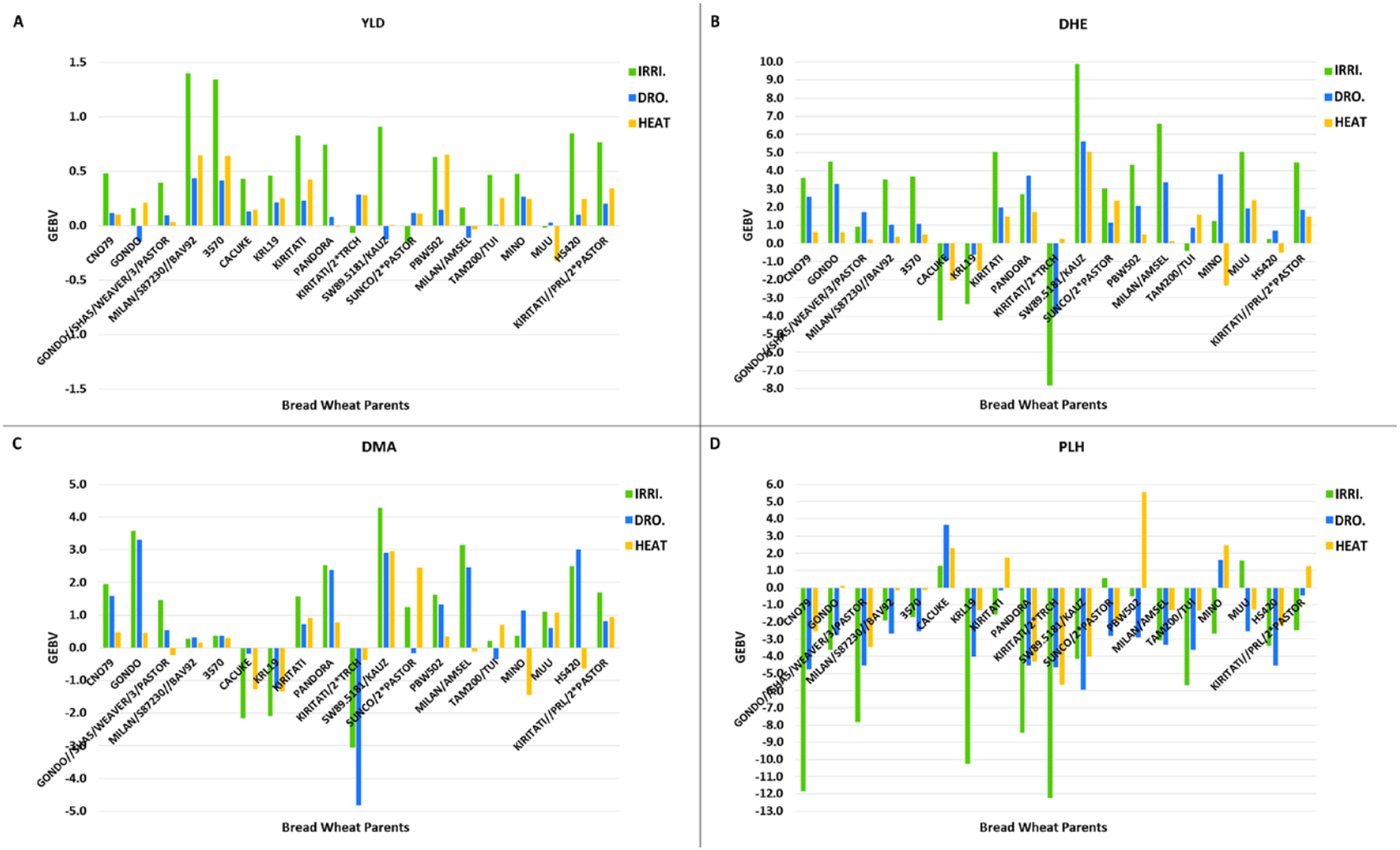


Figure 2.4: GEBVs of BW parents for traits in three contrasting environments. Irrigated (IRRI.), Drought (DRO.), and Heat (HEAT): (A) grain yield (YLD) GEBVs, (B) days to heading (DHE) GEBVs, (C) days to maturity (DMA) GEBVs and (D) plant height (PLH) GEBVs across three environments.

The GEBV values of synthetic lines

All the SYN lines had negative GEBVs for grain yield across all environments except SYNP12, SYNP26, SYNP27, and SYNP36 that had small positive values under the heat stress. Predominantly, they had the lowest GEBVs in irrigated condition (-0.25 to -2.69) while their value ranged from -0.10 to -1.02 for drought and from 0.26 to -1.74 for heat stress (Figure 2.4 A and Table S2.4). However, these results were expected, because SYN lines have very low grain yield.

For DHE, GEBVs of all SYN lines were negative and decreased DHE except for six SYN parents that had positive values in all or at least one environment (Figure 2.4 B and Table S2.4). GEBVs for DHE ranged from -8.04 to 1.55 under irrigated, -6.61 to 1.48 under drought and -4.31 to 1.21 under heat conditions (Table S2.4). Most of the SYN lines had less strongly negative GEBVs under heat stress indicating that they strongly influenced them to head earlier. For DFL, most of the SYN lines showed similar trends across all environments (Table S2.4).

For DMA, breeding values of SYN parents were more variable than those for DHE and many parents had positive GEBVs in one or more environments (Figure 2.4 C and Table S2.4). Also, SYN parents had overall lower negative GEBVs for DMA than DHE and increased DMA. Under irrigated environments, the range of GEBVs was -3.50 to 3.04, -3.46 to 1.87 for drought and -2.5 to 1.33 for heat stress trials. GEI for DMA was greater than that for DHE.

All SYN lines contributed to increased PLH in all environments except SYNP44, which had negative GEBVs. Their GEBVs were higher in irrigated trials and ranged from -6.53 to 20.51 while they had lower values in heat stress trials ranging from -0.55 to 5.84 (Figure 2.4 D and Table S2.4).

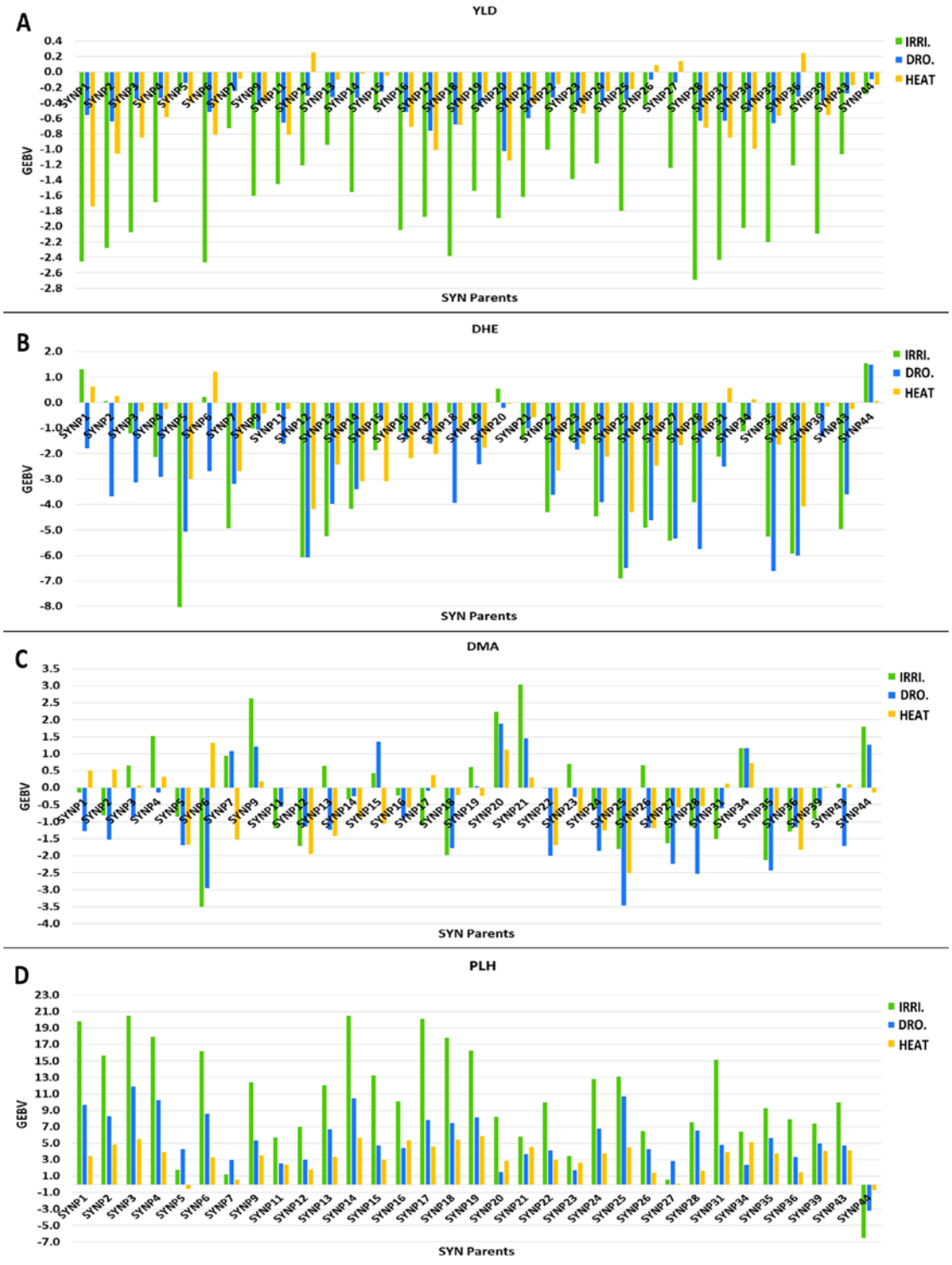


Figure 2.5: GEBVs of SYN parents for traits in three contrasting environments. Irrigated (IRRI.), Drought (DRO.), and Heat: (A) grain yield (YLD) GEBVs, (B) days to heading (DHE) GEBVs, (C) days to maturity (DMA) GEBVs and (D) plant height (PLH) GEBVs across three environments.

Correlation of parent GEBV values across environments

All GEBV correlation coefficients among environments for BW parents were significant (Table 2.6 above diagonal). The correlations between GEBVs for drought stress and those for irrigated environments were lower than those between irrigated and heat, and drought and heat environments. For SYN lines, correlations between different environments were significant (Table 2.6, below diagonal) and they showed lower GEI than BW parents.

Table 2.6: Pearson correlation coefficients of parent GEBVs across environments for yield.

	BW parents		
Environments	IRRI.	DRO.	HEAT
IRRI.	1	0.47	0.62
DRO.	0.73	1	0.68
HEAT	0.70	0.72	1
SYN parents			

IRRI.: Irrigated, DRO.: Drought and HEAT: Heat.

Performance of synthetic-derived lines in different environments

Crossing SYN lines to BW parents extended their genetic diversity for measured traits. The variation for grain yield GEBVs was greatest under irrigation and ranged from –2.02 to 1.69 for SDLs, while it ranged from -0.16 to 1.34 for BWs (Figure 2.5 A1). Variation in yield GEBVs was least under drought stress ranging from -0.91 to 0.54 for SDLs and from -0.15 to 0.43 for BWs (Figure 2.5 C1). Under heat stress, GEBV variation ranged from -1.28 to 0.88 for SDLs and -0.326 to 0.649 for BWs (Figure 2.5 B1).

To determine how many SYN parents were able to improve the YLD of BW parents in different environments, the top 10% of SDLs was selected and the average GEBV values for each cross or family was compared to their corresponding recurrent BW parent's GEBV values (Figure 2.5). This top 10% included progenies of 13 BW and 23 SYN parents in which MILAN/S87230//BAV92, SUNCO/2*PASTOR, PANDORA, SYNP4, SYNP5, SYNP17, SYNP20, SYNP21, SYNP23, SYNP27, SYNP39, and SYNP43 had major contributions across all environments (Tables S2.5-S2.7).

Heat stress

The top 10% SDLs in heat stress comprised 175 SDLs and the average GEBVs of SDLs in each cross was higher than those of their corresponding recurrent BW parents except for SDLs in crosses with MILAN/S87230//BAV92, PBW502, and BW line 3570 (Table S2.5). The increased GEBVs for SDLs compared to their BW parents ranged from 2 to 427% and included mainly BC progenies. However, there was also one TC and six biparental crosses in which the progenies had higher GEBVs than the BW parent (Table S2.5). Under heat stress the average yield GEBVs of the top 10% of SDLs was 0.55 while the weighted average GEBV of their recurrent BW parents was 0.26 (Figure 2.5 B1).

Irrigated environment

In the irrigated trials, the average yield GEBVs of the top 10% of SDLs ranged from 0.69 to 1.09 while these values ranged from -0.16 to 1.40 for BW parents (Table

S2.6). The average GEBVs of SDLs of crosses with BW line 3570, HS420, MILAN/S87230//BAV92, KIRITATI, PANDORA (in 3 crosses), and SW89.5181/KAUZ (in 1 cross) decreased by -2 to -49% while GEBVs of SDLs of crosses with other BW parents increased by 8 to 111%. In the irrigated trials, BC progenies had generally higher GEBVs but there was one TC and three biparental crosses whose progenies had higher GEBVs (Table S2.6). The average GEBV of the top 10% of SDLs was 0.94 while the weighted average GEBV was 0.90 for their recurrent BW parents (Figure 2.5 A1).

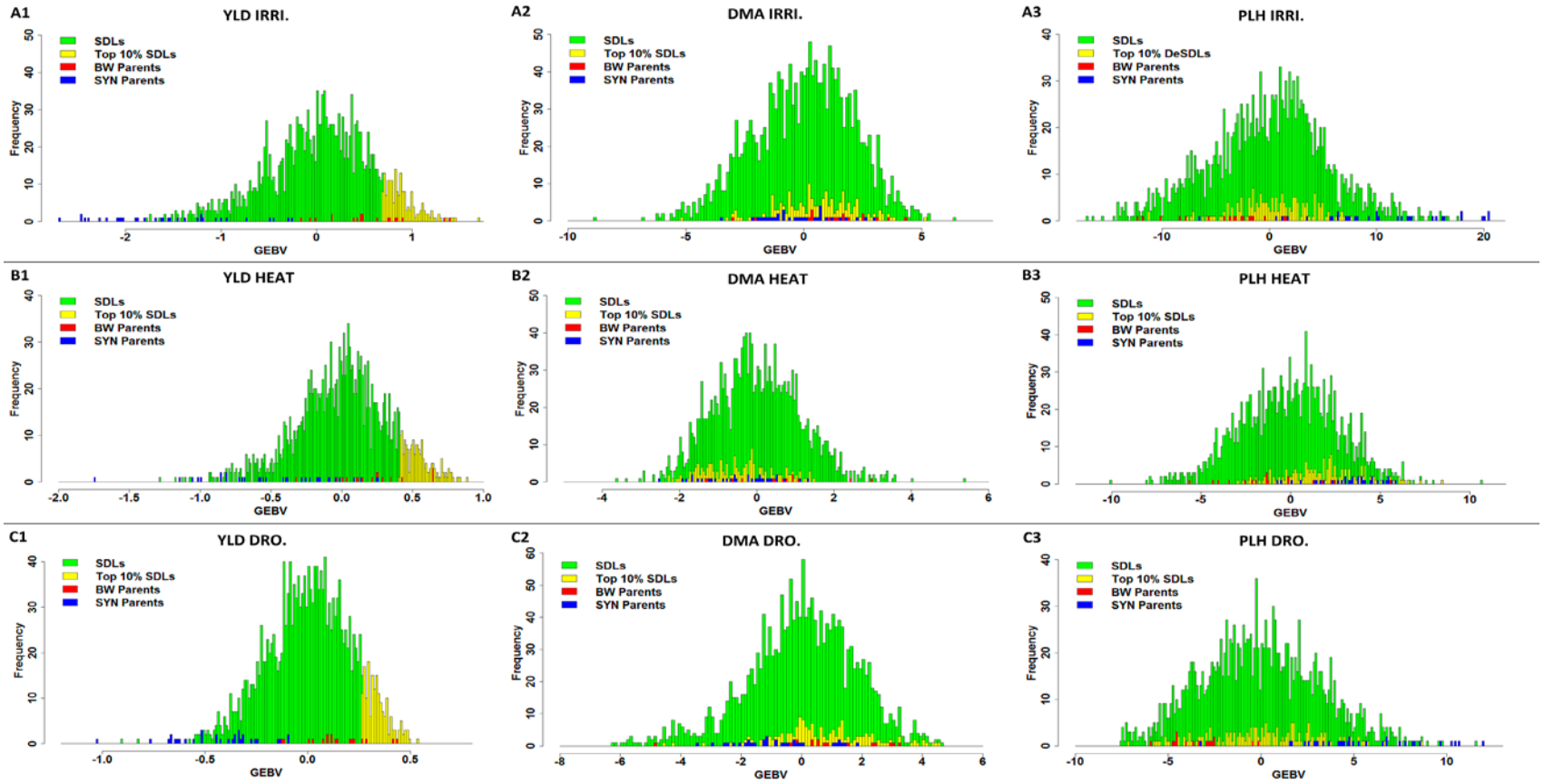


Figure 2.6: Distribution of GEBVs for the SDLs, SYN and BW parents in different trials. This figure compares the top 10% of SDLs (yellow tail) to BW and SYN parents that are constant in each trial for three traits (YLD, DMA and PLH): (A1) distribution of YLD GEBVs in irrigated trials, (A2) distribution of DMA GEBVs in irrigated trials in which, GEBVs of the top 10% SDLs are in the same range of the parents, (A3) distribution of PLH's GEBVs in irrigated trials in which PLH of the top 10% SDLs were skewed toward the BW parents, (B1) distribution of YLD GEBVs in heat trials, (B2) DMA GEBVs in heat trials where GEBVs of the top 10% SDLs were placed within the range of GEBVs of the parents. (B3) distribution of PLH GEBVs in heat trials. GEBVs of the top 10% SDLs were skewed toward the SYN parents, (C1), (C2) and (C3) are for drought trials.

Drought stress

The top 10% of the SDLs of populations grown under drought stress involved 179 SDLs for which the average GEBVs of crosses ranged from 0.26 to 0.44 while the range for BW parents was -0.11 to 0.44. The increased GEBVs for SDLs compared to the corresponding recurrent BW parents ranged from 12 to 422% (Table S2.7). However, the cross of SYN parents to MILAN/S87230//BAV92 did not improve the GEBVs of its SDLs. Also, the average yield GEBV of the top 10% of SDLs was 0.34 while the weighted average GEBV was 0.30 for their corresponding recurrent BW parents (Figure 2.5 C1).

Across all environments, it was observed that the SYN lines most significantly increased grain yield of low yielding BW parents in both stress and normal conditions (Tables 2.S5- 2.S7). For example, SUNCO/2*PASTOR was a low-yielding BW parent across all environments. In crosses with SYNs it contributed 59 progenies in the top 10% of SDLs and all of them outperformed the BW parent. Their yield GEBVs ranged from 0.27 to 1.20 under drought stress and irrigated conditions, respectively, while the range of yield GEBVs for SUNCO/2*PASTOR was from -0.16 to 0.12 under irrigated and drought environments, respectively. The high-yielding BW parents, MILAN/S87230//BAV92, had 173 progenies among the top 10% of the SDLs, but only 25 of them had higher GEBVs than the BW parent. Their GEBVs for yield ranged from 0.44 to 1.69 under drought and irrigated conditions, respectively. This pattern is similar for the other low- and high-yielding BW parents.

To determine if the high yield of SDLs is related to the phenological traits, the correlation coefficients between GEBVs of YLD and those for other traits was calculated. Squared correlations of YLD with DMA were 0.08, 0.06, and 0.04 in drought, heat and irrigated, respectively, and these values for PLH were 0.01, 0.05, and 0.06 in drought, heat and irrigated, respectively which indicated that DMA and PLH did not affect the yield.

Genomic Prediction

The univariate, random five-fold cross validation was used for genomic prediction of traits for each trial and for all trials together. As was previously mentioned, the heat trial in year 2011-12 experienced extreme temperatures and when using the phenotypic observations from this trial in genomic prediction models, both heritability and prediction accuracy of the traits decreased across environments. Consequently, this trial was excluded from cross validation.

Broad-sense heritabilities of traits in different environments based on pedigree, marker and pedigree-marker models are shown in Table 2.7. Estimated heritabilities for all traits using the corrected pedigree model were slightly higher in each environment except for DFL under the heat stress environments (Table 2.7). The differences in heritabilities could be due to 1) the artificially high genetic variance assigned to unrelated parents that are actually related 2) the differences in the amount of estimated genetic variances using **A** or **G** matrices in the model (Loberg et al., 2015). We observed that estimated genetic variances using the **G** matrix (gVarG) were smaller than those using the **A** matrix (gVarA) for all traits under drought stress. Under heat stress, gVarG

for all traits were smaller than gVarA except for DFL, and under the irrigated environment, the trend was similar except for DFL and PLH. The genetic variances estimated using the **G** matrix explained 66 to 96% (gVarG/ gVarA) of those estimated using the **A** matrix under drought stress. This ratio ranged from 81 to 131% and from 76 to 118% for the heat and irrigated environments, respectively 3) Sampling error due to finite markers can affect the estimation of the **G** matrix as reported by Haile- Mariam et al. (2003) and Powell et al (2010). 4) All the diagonal elements of the **A** matrix were 2 while the average of the diagonal elements of the marker based relationship matrix (**G** matrix) was 1.86 (0.25 to 9.99). However, scaling the **G** matrix did not change the results (data not shown).

Table 2.7: Mean heritability for traits in each trial.

Trials	DRO.			IRRI.			HEAT		
Model/ Trait	Ped.	Mar.	Ped.-Mar.	Ped.	Mar.	Ped-Mar.	Ped.	Mar.	Ped.-Mar.
DHE	0.68	0.63	0.64	0.71	0.67	0.67	0.64	0.58	0.59
DFL	0.68	0.64	0.64	0.77	0.75	0.75	0.24	0.34	0.37
DMA	0.70	0.62	0.64	0.69	0.66	0.66	0.54	0.49	0.50
PLH	0.60	0.51	0.53	0.80	0.79	0.79	0.61	0.53	0.56
YLD	0.57	0.52	0.52	0.70	0.64	0.65	0.68	0.63	0.64

Ped.: pedigree, Mar.: marker, Ped-Mar.: pedigree-marker.

The trait heritabilities were consistently higher in irrigated than drought and heat stress environments using the three models. Heritabilities of DHE, DMA and DFL were higher under irrigated and drought environments but lower under heat stress especially for DFL (Table 2.7). This could be related to the lower number of observations for these traits. DHE and DMA had two years of data but DFL had only one year of data. PLH had the highest heritability under irrigated environments (0.79 – 0.80) and decreased under drought and heat stress (0.51 to 0.61) (Table 2.7). Also, the highest heritability

for YLD was observed under irrigated (0.64 – 0.70) followed by heat (0.63 – 0.68) and drought stresses (0.52 - 0.57) (Table 2.7).

Predictability was assessed as the correlation between GEBVs and observed phenotypes and were corrected for fixed effects by cross-validation. Our results showed that the marker model gives higher genetic prediction accuracy than the pedigree model for all traits either in the single environments (e.g. Irrigated, heat, and drought) (Table 2.8) or combined environments (Table 2.9). Mean accuracy of the three models ranged from 0.30 to 0.64 across all environments. The highest prediction accuracy was obtained in irrigated environments while lower accuracies were mostly observed in heat stress environments. Increased prediction accuracy using the marker model ranged from 2% for YLD to 5% for DHE under drought stress. This range was 5% for PLH to 9% for DFL under irrigation and 5% for YLD to 12% for PLH and DHE in heat stress. Using the marker-pedigree model did not improve the prediction accuracy (Table 2.8).

Table 2.8: Mean genomic prediction accuracy of traits for each trial in cross validation.

Trials	DRO.			IRRI.			HEAT		
Model/Trait	Ped.	Mar.	Ped.-Mar.	Ped.	Mar.	Ped.-Mar.	Ped.	Mar.	Ped.-Mar.
DHE	0.49	0.54	0.55	0.54	0.61	0.64	0.33	0.45	0.45
DFL	0.45	0.48	0.50	0.49	0.58	0.60	0.43	0.52	0.55
DMA	0.44	0.46	0.48	0.36	0.43	0.43	0.30	0.39	0.37
PLH	0.41	0.44	0.45	0.53	0.58	0.57	0.34	0.46	0.44
YLD	0.42	0.44	0.46	0.49	0.55	0.55	0.43	0.48	0.49

Ped.: pedigree, Mar.: marker, Ped-Mar.: pedigree-marker.

Table 2.9: Mean heritability and accuracy of genomic prediction of traits across environments in cross validation.

	Heritability			Accuracy		
Model/ Trait	Ped.	Mar.	Ped.-Mar.	Ped.	Mar.	Pedi.-Mar.
DHE	0.50	0.48	0.48	0.36	0.42	0.42
DFL	0.70	0.70	0.69	0.40	0.47	0.47
DMA	0.27	0.30	0.29	0.26	0.31	0.31
PLH	0.87	0.89	0.88	0.40	0.46	0.47
YLD	0.64	0.57	0.58	0.36	0.42	0.42

Ped.: pedigree, Mar.: marker, Ped-Mar.: pedigree-marker.

Combining environments, the mean prediction accuracy of all traits was decreased in all models except for PLH for which accuracy was almost equal or higher than that in drought and heat stresses. The greatest reduction in accuracy occurred in the irrigated environment, which on average was 0.13% (0.09 to 0.22%) while the lowest reduction was observed under the heat stress by on average 0.03% (0 to 0.08%) (Table 2.9).

Combining environments also decreased the heritability of DHE and DMA compared to single environments in all models, while it increased the heritability of PLH. Furthermore, heritability of DFL was increased compared to heat and drought stresses but it decreased compared to irrigated environments. For YLD, heritability was lower for drought stress compared to irrigated and heat environments (Table 2.9).

Discussion

Results of this study revealed that SYN parents are more diverse than cultivated BW wheat cultivars used in this study as shown in Figure 2.1. Also, based on Nei's genetic diversity, SYN parents had higher genetic diversity than BW parents across all three genomes, specifically for D genome (H_s = 0.40) (Table 2.5). This was because 28

different *A. tauschii* accessions and 17 durums were used to develop the SYNs. The Nei's genetic diversity indicated that SDL populations were more diverse than BW parents for A, B and D genomes in which D genome had the highest increased diversity (H_s = 0.19) (Table 2.5). Therefore, SYN lines are promising genetic resources to introduce novel genetic variation into the cultivated wheat gene pool. Similarly, Huang et al. (2006) and Hoisington et al. (1999) reported that SYN lines were used to improve quality, disease resistance, grain yield, and grain yield components of elite lines. One of the successful synthetic derived cultivars was Chuanmai-42 which increased grain yield by 0.45 to 0.75 t ha^{-1} in southwestern China compared to contemporary cultivars (Yang et al., 2009; Li et al., 2014). The SHW and SDLs are now widely used to develop modern wheat cultivars in China (Li et al., 2014).

Equally important is the question of whether SYN lines can contribute to increased grain yield. The current study shows that the yield increases were predominantly in SDLs from BC1 derived lines (Tables S2.5-S2.7). However, there were a few SDLs from biparental and TC crosses whose yield was higher than their corresponding BW parents. The potential of SDLs from BC1 derived lines to improve yield in both stress and normal conditions was reported in previous studies (Ogbonnaya et al., 2007; Del Blanco et al., 2001; Dreccer et al., 2007 and Van Ginkel and Ogbonnaya, 2007). However, those studies did not have genotypes of the parents and derived lines.

Our results show that while SYN parents mostly have negative GEBVs for grain yield, they have less negative values under stress conditions and can increase grain yield

of recurrent BW parents especially under drought and heat stress conditions (Figure 2.4 A and Table 2.S4). Yield increases were more frequent under heat stress and the average yield GEBVs of the top 10% SDLs was 0.55 while for their recurrent BW parents it was 0.26 (Figure 2.5 C1). Consequently, these results indicate that SYN lines are useful genetic resources for increasing grain yield in stress environments. Similar results were observed by Gororo et al.(2002) who evaluated SDLs in drought and irrigated conditions and reported that SDLs exhibit higher yield potential over the recurrent parents in drought stress. Also, Reddy et al (1996) evaluated common wheat lines and *T. tauschii* under drought stress and found that some *T. tauschii* lines represented were more tolerant than drought tolerant wheat lines. Furthermore, Ogbonnaya et al. (2007) investigated the yield potential of SDLs (derived from BC1) in rainfed environments of Australia and reported that many of them out-yielded both recurrent parents and commercial varieties from 8 to 30% in different environments. They concluded that SDLs could improve yield in more diverse and stressed environments. For heat tolerance, Sharma et al.(2014) evaluated 24 SYN lines under heat stress and identified three highly tolerant SYN lines. Using polymorphic inter-simple sequence repeat (ISSR) markers, they found that the genetic basis of heat tolerance in SYN lines is different and these new sources of genetic diversity could be used to improve heat tolerance of cultivated wheats. Furthermore, Cossani and Reynolds (2015) by comparing six advanced synthetic derivative (ASD) lines with their BW and synthetic derivative (Syn-Der) parents under normal, heat-stress and extreme heat-stress environments reported that the ASD lines outperformed their best parent (Syn-Der) by

on average 5, 15 and 13% for yield under normal, heat and extreme heat stress, respectively.

The higher yield of SDLs could be attributed to introgression of some positive alleles from the SYN lines that increase grain yield. For instance, Li et al. (2011b) used 705 polymorphic SSR markers and found four QTLs (*Barc1183, Barc241, Xcfe25, and Xcfd223*) from the SYN parent in Chuanmai-42 that had significant positive effects on grain yield. *Barc1183*, which is located on the long arm of chromosome 4D, increased grain yield by 7.00 to 11.30%. Similarly, Gororo et al. (2002) investigated yield performance of SDLs derived from direct hybridization of wheat with *T. tauschii* and concluded that the increased yield in SDLs was caused by genes introduced from *T. tauschii*. Also, Liu et al., (2006) using introgression lines (ILs), crossed a SYN line, Am3, to common wheat, Laizhou953. Using 205 SSR markers they detected two QTLs (*Xgwm113* and *Xgwm159*) of Am3 on chromosomes 4B and 5B of the ILs that increased spikes per plant (0.65 to 1.18) and thousand kernel weight (6.10 to 6.30 gr), respectively. These findings support the introgression and retention of some positive yield QTLs from SYN lines in SDLs.

This study showed that the SYN lines contributed significantly more to increased grain yield of lower yielding BW parents in both stress and irrigated conditions. For example, SUNCO/2*PASTOR is one of the lower-yielding BW parents across all environments but all of its progenies that contributed to the top 10% of SDLs had higher GEBVs than the BW parents. Also of the high-yielding BW parents, MILAN/S87230//BAV92, produced 173 progenies among the top 10% SDLs, and 14%

of them had higher GEBVs than the BW parent indicating that the SYN parents contributed positive alleles in crosses to all of the BWs.

In this study, SYN parents extended genetic diversity of the populations for three related traits, DHE, DFL and DMA, in the same direction across environments. As shown in Figures 2.5 A2, 2.5 B2, and 2.5 C2, GEBVs of the top high-yielding SDLs for DMA are similar to the range of BW parent GEBVs. While the GEBVs of the SDLs are more diverse than those of the BW parents, the difference is small. This is because during segregating generations, populations were under selection for maturity approximating that of the BW parents. Since late maturing progenies were not included in the populations, these results did not represent the true diversity of the populations for these three traits. However, these results are likely to be more relevant to a wheat breeding program.

In this study, there was a low correlation between GEBVs for yield and DMA, DFL and DHE, suggesting that the higher GEBVs of SDLs compared to their corresponding recurrent BW parents were not due to their phenology such as late or early maturity. This result differs from other studies. For example, Cooper et al., (2012) reported that almost all high-yielding SDLs were earlier than their recurrent BW parents. In contrast to this study, they concluded that SYN lines contributed to yield because of their earlier maturity.

For PLH, diversity of populations was increased across environments (Figures 2.5 A3, B3, and C3), but because of selection, diversity introduced from SYN lines was

reduced. The GEBVs of top high-yielding SDLs for PLH were similar to those for BW parents in the heat stress (Figure 2.5 B3), whereas most of them were taller than BW parents in irrigated and drought environments (Figures 2.5 A3, and C3). Correlation coefficients of GEBVs of PLH and YLD were low across three environments (r = 0.04 to 0.29), suggesting that higher GEBVs of SDLs were not the result of increased plant height.

Our analyses showed that GFD values for SDLs were within the range of those for BW parents. However, this was due to selection of SDLs for maturity approximating that of the BW parents. So, these values did not show the true diversity of SYN lines for this trait. Also, the negative relationship between YLD and GFD indicated that there was no advantage of selecting genotypes for longer GFD. Increased YLD of SDLs was not associated with variation in GFD.

Results of this study indicated that SYN lines contributed to increased TKW of SDLs and increased the family mean from 0.67 to 24.39%. However, this contribution was not consistent for all SYN parents used in this study, such that family mean TKW of 11 SDLs were lower than the corresponding recurrent BW parents. Moreover, some SYN parents decreased TKW of SDLs in biparental populations while they increased TKW in the same BC populations. Our analyses for these specific populations indicated that, although 67% of SDLs had higher TKW than recurrent parents, the negative relationship and very low R^2 values between TKW and YLD, indicated that phenotypic variation of YLD was not generally associated with TKW. Therefore, increased yield of SDLs was not a result of increased seed weight. In contrast to our finding, Cooper et

al.(2012) backcrossed ten elite primary synthetics to two Texas winter wheat cultivars, TAM111 and TAM112, and reported that all SYN lines contributed to high yielding SDLs through an increase in seed weight. Also, Röder et al. (2008), using ILs from crossing a synthetic line, W-7984, to a German winter wheat, 'Prinz', reported a QTL for grain weight, *QTgw.ipk-7D*, which was associated with microsatellite marker, *Xgwm1002-7D*. They reported that the ILs had 10% increased TKW compared to 'Prinz' and checks and 84.70% of the phenotypic variance could be explained by the segregation of *Xgwm1002-7D.*

GEBV values of SYN lines and cultivated wheat

High-throughput genotyping technologies provide an opportunity to estimate breeding value of genotypes more accurately using a genomic relationship matrix (Goddard and Hayes, 2007). These tools can improve the accuracy of parental selection in the breeding program. In this study, BW parents showed positive GEBVs for yield across all environments. Nevertheless, they reflected higher GEI in drought vs. irrigated, heat vs. irrigated, and drought vs. heat (Table 2.6). Some of the BW parents such as MILAN/S87230//BAV92 and BW line 3570 had high GEBV values in all environments (Figure 2.3 A and Table2.S3) and are good candidates to be used in breeding for diverse environments. On the other hand, almost all of the SYN lines had negative GEBVs across all environments for yield (Figure 2.4 A and Table 2.S4). This was expected because SYN lines are exotic lines that have a durum variety and a wild diploid accession as parents and they have not been directly bred for yield. Only by evaluating

populations of segregants from SYN crosses with BWs can we identify their positive and novel yield alleles for improving the yield of BW parents.

For PLH, most BW parents had negative GEBVs (Figure 2.3 D and Table 2. S3) that can be attributed to dwarfing or semi-dwarfing genes in their genetic background. Generally, in irrigated environments plants with short to average height are favored to avoid lodging. Thus, parents with lower GEBVs for PLH are best suited for irrigated environments. Under stress conditions, taller plants are more tolerant as observed in this study (Figures 2.5 B3 and C3). They can store more assimilates in their stems for remobilization during the grain filling stage. Thus, parents with high positive GEBVs would be better for production in stress environments. Although populations were under selection for PLH, all SYN lines had highly positive GEBV values for PLH (Figure 2.4 D and Table 2. S4). This was because SYN lines are very tall genotypes and have many genes for PLH and selection did not remove all of them.

Our findings indicate that the majority of BW parents have positive GEBVs for DHE and DMA (Figures 2.3 B and C), while nearly all SYN parents have negative GEBVs for DHE and decreased this trait (Figure 2.4 B). For DMA, there are more SYN lines that have positive GEBVs in one or more environments (Figure 2.4 C). We expected their positive GEBV values for these traits because SYN lines tend to be late maturing genotypes.

Genomic prediction

In traditional genetic evaluation, linear mixed models with the pedigree relationship matrix have been used for genomic prediction and selection in breeding programs (Piepho et al., 2007; Gao et al., 2012). New genotyping technologies provided dense genome-wide molecular markers that have been used to derive more accurate genomic relationships to increase selection accuracy in breeding programs (Nejati-Javaremi, 1997; Meuwissen et al., 2001; Goddard and Hayes, 2007; Crossa et al., 2010). Our results indicated that using marker data improved genomic prediction accuracy over the pedigree method. Improvement rates varied based on the different traits and environments and ranged from 2 to 12% (Table 2.8). The greatest improvement in prediction accuracy was mainly observed in heat stress (5 to 12%) and the lowest rate was observed in drought environments (2 to 5%) indicating that environments affect the relative prediction accuracy of pedigree - vs. marker – based prediction (Table 2.8). The higher prediction accuracy using the genomic relationship matrix is attributed to: 1) exploiting Mendelian sampling variation during gamete formation and 2) including relationship information from genotypes that the pedigree classified as unrelated genotypes 3) the **G** matrix provides better coverage of the genetic rearrangements that occur during SYN and SDL development that are not covered by the pedigree. (Nejati-Javaremi, 1997; Goddard and Hayes, 2007; Zapata-Valenzuela et al., 2013). A simulation study confirmed that using genomic relationship instead of pedigree relationship to estimate GEBVs increased selection accuracy (Villanueva et al., 2005). Similar results were reported by Nejati-Javaremi (1997) and VanRaden VanRaden, (2008). However, despite the potential mistakes in the pedigrees, genomic prediction

accuracies from the pedigree model were reasonable and close to those of the marker model, in part because of the relatively small family sizes that limit the Mendelian segregation. This was because marker information was additionally used to identify incorrect pedigrees (removing outlier genotypes). In this study our results showed that, using the pedigree-marker method called the single–step blending approach by Gao et al.(2012), that uses information from both genotyped and non-genotyped lines simultaneously to do genomic prediction did not improve genomic prediction accuracies (Tables 2.8 and2.9).

Cross validation using combined environments decreased prediction accuracies of traits in all models (Table 2.9). However, the decreasing trend was not similar for all traits. The highest average decrease was observed for DHE and DMA (0.11%) while the lowest average reduction was for PLH (0.03%). These results were due to GEI interaction as shown in Tables 2.2, 2.3, and 2.4 such that phenotypic correlations within treatments (irrigated, heat or drought) were overall greater than those among treatments. However, this was not consistent for all traits and for some of them among treatment correlations were greater than those for within treatments (e.g. PLH). These results confirmed that GEI affects the genomic prediction accuracy and traits with high GEI had lower prediction accuracy. Similarly, Zapata-Valenzuela et al. (2013) argued that the accuracy of GEBVs using either **A** or **G** matrices would be lower in cases where there is strong GEI. This could lead to prediction models developed in one environment that lose their prediction power in other environments (Resende et al., 2012 and Zapata-Valenzuela et al., 2013).

In this study estimated heritabilities using the pedigree model were consistently slightly higher than those using marker models (Table 2.7). This differences could be due to the differences in the amount of estimated genetic variances using **A** or **G** matrices in the model (Loberg et al., 2015) as we observed in this study. Similarly, Loberg et al., (2015) reported that the genetic variances estimated using the **A** matrix were greater than those estimated by the **G** matrix. Hence, estimated heritabilities using the **A** matrix were greater. They reported that gVarG, explained 10-60% of gVarA. Also, Powell et al., (2010) mentioned that incomplete linkage disequilibrium between the markers and the causal variants can reduce the genetic variance using the marker model. They concluded that the difference between the estimated gVarA and genetic variance explained by SNPs estimated using the **G** matrix was the missing heritability.

Conclusion

These findings confirm that synthetic hexaploid wheat germplasm is a valuable genetic resource for improving grain yield and other traits. Synthetic hexaploid wheat lines have positive, novel alleles that can be easily introgressed into cultivated wheat to improve yield, especially in stress conditions. Therefore, SYN lines should be used in breeding programs to expand the genetic diversity for agronomic traits but selection against undesirable phenology is required to realize the benefit of the novel genetic variation.

Acknowledgments

We are grateful to the Monsanto's Beachell-Borlaug International Scholars Program which supports Ph.D. students from developing countries who are working on breeding of wheat and rice. We are grateful to the Global Wheat Program of the International Maize and Wheat Improvement Center (CIMMYT) for excellent collaboration with Cornell University and to "Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA, Mexico)" for supporting the MasAgro Biodiversidad/Seeds of Discovery Project-CIMMYT to conduct this study.

References

Akdemir, D., O.U. Godfrey .2015. EMMREML: Fitting Mixed Models with Known Covariance Structures. R package version 3.1. https://CRAN.R-project.org/package=EMMREML.

Bassi, F.M., A.R. Bentley, G. Charmet, R. Ortiz, and J. Crossa. 2015. Breeding schemes for the implementation of genomic selection in wheat (*Triticum* spp.). Plant Sci. 242: 23–36. http://dx.doi.org/10.1016/j.plantsci.2015.08.021.

Del Blanco, I.A.A., S. Rajaram, and W.E.E. Kronstad. 2001. Agronomic potential of synthetic hexaploid wheat-derived populations. Crop Sci. 41(3): 670–676. http://www.mendeley.com/research/agronomic-potential-synthetic-hexaploid-wheatderived-populations/ (verified 7 January 2015).

Christiansen, M.J., S.B. Andersen, and R. Ortiz. 2002. Diversity changes in an intensively bred wheat germplasm during the 20th century. Mol. Breed. 9(1): 1–11. http://www.springerlink.com/content/l874u25h11h27783.

Cooper, J.K., A.H. Ibrahim, J. Rudd, S. Malla, D.B. Hays, and J. Baker. 2012. Increasing hard winter wheat yield potential via synthetic wheat: I. path-coefficient analysis of yield and its components. Crop Sci. 52(5): 2014–2022.

Cossani, C.M., and M.P. Reynolds. 2015. Heat stress adaptation in elite lines derived from synthetic hexaploid wheat. Crop Sci. 55(6): 2719–2735.

Crossa, J., G. De Los Campos, P. Pérez, D. Gianola, J. Burgueño, J.L. Araus, D.

Makumbi, R.P. Singh, S. Dreisigacker, J. Yan, V. Arief, M. Banziger, and H.J. Braun. 2010. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. Genetics 186(2): 713–724.

Dreccer, M.F., M.G. Borgognone, F.C. Ogbonnaya, R.M. Trethowan, and B. Winter. 2007. CIMMYT-selected derived synthetic bread wheats for rainfed environments: Yield evaluation in Mexico and Australia. F. Crop. Res. 100(2–3): 218–228. http://linkinghub.elsevier.com/retrieve/pii/S037842900600150X (verified 22 September 2014).

Dreisigacker, S., M. Kishii, J. Lage, and M. Warburton. 2008. Use of synthetic hexaploid wheat to increase diversity for CIMMYT bread wheat improvement. Aust. J. Agric. Res. 59(5): 413–420. http://www.publish.csiro.au/?paper=AR07225.

Dreisigacker, S., R. Tiwari, and S. Sheoran. 2013. ICAR-CIMMYT molecular breeding course in wheat Directorate of Wheat Research. ICAR/BMZ: p 36. http://repository.cimmyt.org/xmlui/handle/10883/3221#.

Eathington, S.R., T.M. Crosbie, M.D. Edwards, R.S. Reiter, and J.K. Bull. 2007. Molecular markers in a commercial breeding program. Crop Sci. 47(SUPPL. DEC.).

Endelman, J.B., and J.-L. Jannink. 2012. Shrinkage estimation of the realized relationship matrix. G3 (Bethesda). 2(11): 1405–13. http://www.g3journal.org/content/2/11/1405.full.

Gao, H., O.F. Christensen, P. Madsen, U.S. Nielsen, Y. Zhang, M.S. Lund, and G. Su. 2012. Comparison on genomic predictions using three GBLUP methods and two single-step blending methods in the Nordic Holstein population. Genet. Sel. Evol. 44(1): 8. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3400441&tool=pmcentrez&rendertype=abstract (verified 30 November 2015).

Van Ginkel, M., and F. Ogbonnaya. 2007. Novel genetic diversity from synthetic wheats in breeding cultivars for changing production conditions. F. Crop. Res. 104(1–3): 86–94. http://www.sciencedirect.com/science/article/pii/S0378429007001153%5Cnhttp://pdn.sciencedirect.com/science?_ob=MiamiImageURL&_cid=271162&_user=1196840&_pii=S0378429007001153&_check=y&_origin=article&_zone=toolbar&_coverDate=31-Dec-2007&view=c&originContentF (verified 28 May 2013).

Goddard, M.E., and B.J. Hayes. 2007. Genomic selection. J. Anim. Breed. Genet. 124(6): 323–30. http://www.ncbi.nlm.nih.gov/pubmed/18076469.

Gororo, N.N.N., H.A.A. Eagles, R.F.F. Eastwood, M.E.E. Nicolas, and R.G.G. Flood. 2002. Use of *Triticum tauschii* to improve yield of wheat in low-yielding environments. Euphytica 123(2): 241–254. http://link.springer.com/article/10.1023/A%3A1014910000128 (verified 27 March 2015).

Haile-Mariam, M., J.M. Morton, and M.E. Goddard. 2003. Estimates of genetic parameters for fertility traits of Australian Holstein-Friesian cattle. Anim. Sci. 76: 35–42.

Heckenberger, M., M. Bohn, J.S. Ziegle, L.K. Joe, J.D. Hauser, M. Hutton, and A.E. Melchinger. 2002. Variation of DNA fingerprints among accessions within maize inbred lines and implications for identification of essentially derived varieties . I . Genetic and technical sources of variation in SSR data. Mol. Breed. 10: 181–192.

Hoisington, D., M. Khairallah, T. Reeves, J.M. Ribaut, B. Skovmand, S. Taba, and M. Warburton. 1999. Plant genetic resources: what can they contribute toward increased crop productivity? Proc. Natl. Acad. Sci. U. S. A. 96(11): 5937–5943.

Huang, X.Q., S. Cloutier, L. Lycar, N. Radovanovic, D.G. Humphreys, J.S. Noll, D.J. Somers, and P.D. Brown. 2006. Molecular detection of QTLs for agronomic and quality traits in a doubled haploid population derived from two Canadian wheats (*Triticum aestivum* L.). Theor. Appl. Genet. 113(4): 753–766.

Kema, G.H.J., W. Lange, and C.H. Vansilfhout. 1995. Differential Suppression of Stripe Rust Resistance in Synthetic Wheat Hexaploids Derived from *Triticum turgidum* subsp *dicoccoides* and *Aegilops squarrosa*. Phytopathology 85(Iss 4): 425–429.

Kosman, E. 2003. Nei's gene diversity and the index of average differences are identical measures of diversity within populations. Plant Pathol. 52(5): 533–535.

Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. J. Dairy Sci. 92(9): 4656–4663. http://dx.doi.org/10.3168/jds.2009-2061.

Li, J., H.S. Wan, and W.Y. Yang. 2014. Synthetic hexaploid wheat enhances variation and adaptive evolution of bread wheat in breeding processes. J. Syst. Evol. 52(6): 735–742. http://doi.wiley.com/10.1111/jse.12110.

Li, J., H.T. Wei, X.R. Hu, C.S. Li, Y.L. Tang, D.C. Liu, and W.Y. Yang. 2011. Identification of a High-Yield Introgression Locus in Chuanmai 42 Inherited from Synthetic Hexaploid Wheat. Acta Agron. Sin. 37(2): 255–262. http://dx.doi.org/10.1016/S1875-2780(11)60007-2.

Liu, S., R. Zhou, Y. Dong, P. Li, and J. Jia. 2006. Development, utilization of introgression lines using a synthetic wheat as donor. Theor. Appl. Genet. 112(7): 1360–73. http://www.ncbi.nlm.nih.gov/pubmed/16550399 (verified 17 January 2016).

Loberg, A., J.W. Dürr, W.F. Fikse, H. Jorjani, and L. Crooks. 2015. Estimates of genetic variance and variance of predicted genetic merits using pedigree or genomic relationship matrices in six Brown Swiss cattle populations for different traits. J. Anim. Breed. Genet. 132(5): 376–385. http://doi.wiley.com/10.1111/jbg.12142.

Lopes, M.S., and M.P. Reynolds. 2011. Drought adaptive traits and wide adaptation in elite lines derived from resynthesized hexaploid wheat. Crop Sci. 51(4): 1617–1626. https://www.crops.org/publications/cs/abstracts/51/4/1617 (verified 4 June 2013).

Mayer, K.F.X., J. Rogers, J. Dole el, C. Pozniak, K. Eversole, C. Feuillet, B. Gill, et al. (2014). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science. 345(6194): 1251788–1251788. http://www.sciencemag.org/cgi/doi/10.1126/science.1251788 (verified 18 July 2014).

Mehmani, A., S. Chowdhury, and A. Messac. 2015. Predictive quantification of surrogate model fidelity based on modal variations with sample density. Struct. Multidiscip. Optim. 52(2): 353–373. http://link.springer.com/10.1007/s00158-015-1234-z.

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. http://www.genetics.org/content/157/4/1819.short (verified 7 January 2015).

Mujeeb-Kazi, A., V. Rosas, and S. Roldan. 1996. Conservation of the genetic variation of *Triticum tauschii* (Coss.) Schmalh. (*Aegilops squarrosa* auct. non L.) in synthetic hexaploid wheats (*T. turgidum* L. s.lat. x *T. tauschii*; 2n=6x=42, AABBDD) and its potential utilization for wheat improvement. Genet. Resour. Crop Evol. 43(2): 129–134. http://www.springerlink.com/index/10.1007/BF00126756.

Mulki, M.A., A. Jighly, G. Ye, L.C. Emebiri, D. Moody, O. Ansari, and F.C. Ogbonnaya. 2013. Association mapping for soilborne pathogen resistance in synthetic hexaploid wheat. Mol. Breed. 31(2): 299–311.

Nejati-Javaremi, A. 1997. Effect of total allelic relationship on accuracy of evaluation and response to selection. J. Anim. Sci. 75(7): 1738–1745. http://www.journalofanimalscience.org/content/75/7/1738.short.

Ogbonnaya, F.C., M. Imtiaz, H.S. Bariana, M. McLean, M.M. Shankar, G.J. Hollaway, R.M. Trethowan, E.S. Lagudah, and M. Van Ginkel. 2008. Mining synthetic hexaploids for multiple disease resistance to improve bread wheat. Aust. J. Agric. Res. 59(5): 421–431.

Ogbonnaya, F.C., G. Ye, R. Trethowan, F. Dreccer, D. Lush, J. Shepperd, and M. van Ginkel. 2007. Yield of synthetic backcross-derived lines in rainfed environments of Australia. Euphytica 157(3): 321–336. http://link.springer.com/10.1007/s10681-007-9381-y (verified 22 September 2014).

Pask A.J.D., J. Pietragalla, D.M. Mullan, M.P. Reynolds and M. Reynolds (Ed.). 2012. Physiological breeding II: A field guide to wheat phenotyping. Mexic, D.F.:CIMMYT. http://libcatalog.cimmyt.org/download/cim/96144.pdf.

Payne R.W., D.A. Murray, S.A. Harding, D.B.S.D. Baird .2009. GenStat for Windows (12th Edition) Introduction.

Piepho, H.P., J. Möhring, A.E. Melchinger, and A. Büchse. 2007. BLUP for phenotypic selection in plant breeding and variety testing. Euphytica 161(1–2): 209–228. http://link.springer.com/10.1007/s10681-007-9449-8 (verified 10 July 2014).

Powell, J.E., P.M. Visscher, and M.E. Goddard. 2010. Reconciling the analysis of IBD and IBS in complex trait studies. Nat. Rev. Genet. 11(11): 800–805. http://www.nature.com/doifinder/10.1038/nrg2865.

Reddy N, Halloran G M, Nicolas M E. 1996. Agronomic assessment of lines derived from a direct cross of wheat with *T. tauschii* L. p. 24–26. *In* Proc 8th Assembly of Wheat Breed. Soc. of Australia.

Resende, M.F.R.D. V, P. Muñoz, J.J. Acosta, G.F. Peter, J.M. Davis, D. Grattapaglia, et al. 2012. Accelerating the domestication of trees using genomic selection: Accuracy of prediction models across ages and environments. New Phytol. 193(3): 617–624. http://dx.doi.org/10.3168/jds.2007-0403.

Roder, M.S., X.Q. Huang, and A. Borner. 2008. Fine mapping of the region on wheat chromosome 7D controlling grain weight. Funct. Integr. Genomics 8(1): 79–86.

Russell, J.R., R.P. Ellis, W.T.B. Thomas, R. Waugh, J. Provan, A. Booth, J. Fuller, P. Lawrence, G. Young, and W. Powell. 2000. A retrospective analysis of spring barley germplasm development from “foundation genotypes” to currently successful cultivars. Mol. Breed. 6: 553–568.

Saghai-Maroof, M.A., K.M. Soliman, R.A. Jorgensen, and R.W. Allard. 1984.

Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. Proc. Natl. Acad. Sci. 81(24): 8014–8018. http://www.pnas.org/content/81/24/8014.abstract (verified 16 April 2016).

Sehgal, D., P. Vikram, C.P. Sansaloni, C. Ortiz, C. Saint Pierre, T. Payne, M. Ellis, A. Amri, C.D. Petroli, P. Wenzl, and S. Singh. 2015. Exploring and mobilizing the gene bank biodiversity for wheat improvement. PLoS One 10(7): e0132112. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0132112 (verified 13 April 2016).

Sharma, P., S. Sareen, M. Saini, A. Verma, B.S. Tyagi, and I. Sharma. 2014. Assessing genetic variation for heat tolerance in synthetic wheat lines using phenotypic data and molecular markers. 8(4): 515–522. http://www.cropj.com/sharma_8_4_2014_515_522.pdf.

Simón, M.R., A.J. Worland, and P.C. Struik. 2005. Chromosomal location of genes encoding for resistance to septoria tritici blotch (Mycosphaerella graminicola) in substitution lines of wheat. NJAS - Wageningen J. Life Sci. 53(2): 113–129. http://linkinghub.elsevier.com/retrieve/pii/S1573521405800016.

Timm, N.H. 2002. Applied Multivariate Analysis. 2nd ed. Springer-Verlag, New York.

VanRaden, P.M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91(11): 4414–23. http://www.ncbi.nlm.nih.gov/pubmed/18946147 (verified 10 December 2014).

Verhulst, N., K.D. Sayre, M. Vargas, J. Crossa, J. Deckers, D. Raes, and B. Govaerts. 2011. Wheat yield and tillage-straw management system by year interaction explained by climatic co-variables for an irrigated bed planting system in northwestern Mexico. F. Crop. Res. 124(3): 347–356.

Villanueva, B., R. Pong-Wong, J. Fernández, and M.A. Toro. 2005. Benefits from marker-assisted selection under an additive polygenic genetic model. J. Anim. Sci. 83(8): 1747–1752.

Villareal, R.L., A. Mujeeb-Kazi, G. Fuentes-Davila, S. Rajaram, and E. Del Toro. 1994. Resistance to karnal bunt (*Tilletia indica* Mitra) in synthetic hexaploid wheats derived from *Triticum turgidum* x *T. tauschii*. Plant Breed. 112: 63–69.

Yang, W., D. Liu, J. Li, L. Zhang, H. Wei, X. Hu, Y. Zheng, Z. He, and Y. Zou. 2009. Synthetic hexaploid wheat and its utilization for wheat genetic improvement in China. J. Genet. Genomics 36(9): 539–546. http://www.ncbi.nlm.nih.gov/pubmed/19782955.

Zapata-Valenzuela, J., R.W. Whetten, D. Neale, S. McKeand, and F. Isik. 2013. Genomic estimated breeding values using genomic relationship matrices in a cloned population of loblolly pine. G3 (Bethesda). 3(5): 909–16. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3656736&tool=pmcentrez&rendertype=abstract.

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

Supplemental Information

Table S2.1: List of cultivated wheat and synthetic hexaploid lines used to develop the SDL populations.

GID	Bread Wheat parents
5653842	MILAN/S87230//BAV92
6763520	3570
6763547	CACUKE
6763623	KRL19
6763636	KIRITATI
6763790	PANDORA
6763911	KIRITATI/2*TRCH
6763917	SW89.5181/KAUZ
6764000	SUNCO/2*PASTOR
6764117	PBW502
6764191	MILAN/AMSEL
6764276	TAM200/TUI
6764815	MINO
6764954	MUU
6765077	HS420
6765130	KIRITATI//PRL/2*PASTOR
4747309	GONDO
4885783	GONDO//SHA5/WEAVER/3/PASTOR
4248	CNO79
	OCI

Table S2.1 (continue). List of cultivated wheat and synthetic hexaploid lines used to develop the SDL populations.

GID	SYNP NO.	Synthetic parents
5989409	SYNP1	68.111/RGB-U//WARD RESEL/3/STIL/4/AE.SQUARROSA (1119)
5989410	SYNP2	68.111/RGB-U//WARD RESEL/3/STIL/4/AE.SQUARROSA (1164)
5989411	SYNP3	68.111/RGB-U//WARD RESEL/3/STIL/4/AE.SQUARROSA (1166)
5989414	SYNP4	68.111/RGB-U//WARD RESEL/3/STIL/4/AE.SQUARROSA (1219)
180150	SYNP5	68.111/RGB-U//WARD/3/AE.SQUARROSA (325)
5989418	SYNP6	68.111/RGB-U//WARD/3/FGO/4/RABI/5/AE.SQUARROSA (1110)
227752	SYNP7	6973/WARD.7463//74110/3/AE.SQUARROSA (438)
5989403	SYNP9	ALTAR 84/AE.SQUARROSA (895)
192513	SYNP11	CETA/AE.SQUAROOSA (263)
227787	SYNP12	CETA/AE.SQUARROSA (1055)

GID	SYNP NO.	Synthetic parents
5989458	SYNP13	CETA/AE.SQUARROSA (1187)
5989460	SYNP14	CETA/AE.SQUARROSA (1219)
192496	SYNP15	CETA/AE.SQUARROSA (184)
5989431	SYNP16	CETA/AE.SQUARROSA (224)
5989440	SYNP17	CETA/AE.SQUARROSA (372)
5989446	SYNP18	CETA/AE.SQUARROSA (518)
5989450	SYNP19	CETA/AE.SQUARROSA (895)
5989390	SYNP20	CROC_1/AE.SQUARROSA (372)
5989393	SYNP21	CROC_1/AE.SQUARROSA (895)
174363	SYNP22	DOY1/AE.SQUARROSA (1027)
174317	SYNP23	DOY1/AE.SQUARROSA (334)
2447504	SYNP24	DOY1/AE.SQUARROSA (443)
180022	SYNP25	DVERD_2/AE.SQUARROSA (247)
172285	SYNP26	GAN/AE.SQUARROSA (446)
2447515	SYNP27	GAN/AE.SQUARROSA (536)
3562369	SYNP28	GARZA/BOY//AE.SQUARROSA (228)
5989495	SYNP31	LOCAL RED/AE.SQUARROSA (518)
5989487	SYNP34	RASCON_37/AE.SQUARROSA (205)
4061131	SYNP35	SHAG_22/AE.SQUARROSA (1084)
227726	SYNP36	SHAG_22/AE.SQUARROSA (239)
5989505	SYNP39	SOMAT_4/INTER_8//AE.SQUARROSA (1206)
172922	SYNP43	YUK/AE.SQUARROSA (864)
4254403	SYNP44	MAYOOR//TK SN1081/AE.SQUARROSA (222)/3/OCI

Table S2.2: TKW and GFD of BW parents and different BP, BC and TC SDL populations in IRRI.Y11.12.

Populations	Cross [1]	BW[2] parents	SYN[3] parents	Mean GFD[4] (days)	Range GFD (days)	BWP GFD (days)	Inc./dec.[5] GFD (%)	Mean TKW[6] (gr)	Range TKW (gr)	BWP TKW (gr)	Inc./dec. TKW (%)
68	BP	CACUKE	SYNP16	56	54-59	57	-1.8	52	41-61	54	-3.70
60	BP	CACUKE	SYNP5	58.5	57-60	57	2.6	57.8	51-63	54	7.04
78	BP	KIRITATI	SYNP24	55	55-55	52	5.8	49	49-49	50	-2.00
71	BP	KRL19	SYNP18	54.5	54-55	56	-2.7	44	44-44	41	7.32
86	BP	KRL19	SYNP36	58	58-58	56	3.6	51	51-51	41	24.39
70	BP	MILAN/S87230//BAV92	SYNP17	53.5	53-54	52	2.9	46.5	46-47	43	8.14
75	BP	MILAN/S87230//BAV92	SYNP20	52.9	49-56	52	1.7	52	44-59	43	20.93
77	BP	MILAN/S87230//BAV92	SYNP21	54.8	52-57	52	5.4	51.5	46-57	43	19.77
83	BP	MILAN/S87230//BAV92	SYNP27	51	49-54	52	-1.9	48	44-51	43	11.63
88	BP	MILAN/S87230//BAV92	SYNP39	53.3	50-57	52	2.5	50.8	43-60	43	18.14
58	BP	MILAN/S87230//BAV92	SYNP4	56	54-57	52	7.7	51	49-54	43	18.60
84	BP	MUU	SYNP34	55.5	50-60	51	8.8	53	42-62	47	12.77
67	BP	PANDORA	SYNP14	56.2	55-60	55	2.2	50.6	45-60	45	12.44
72	BP	PANDORA	SYNP18	55.8	52-59	55	1.5	46.6	46-48	45	3.56
73	BP	PANDORA	SYNP19	55.6	53-58	55	1.1	50.8	48-55	45	12.89
53	BP	PANDORA	SYNP1	53.3	50-56	55	-3.1	45.3	44-46	45	0.67
76	BP	PANDORA	SYNP21	57.3	56-61	55	4.2	55.5	50-64	45	23.33
54	BP	PANDORA	SYNP2	57	54-60	55	3.6	43.5	43-44	45	-3.33
89	BP	PANDORA	SYNP39	54.1	50-56	55	-1.6	48	45-51	45	6.67
56	BP	PANDORA	SYNP3	58	58-58	55	5.5	51	51-51	45	13.33
80	BP	PBW502	SYNP25	61	61-61	52	17.3	58	58-58	51	13.73
90	BP	PBW502	SYNP43	58	58-58	52	11.5	52	52-52	51	1.96
61	BP	PBW502	SYNP5	58	56-60	52	11.5	49	49-49	51	-3.92
82	BP	SUNCO/2*PASTOR	SYNP27	53.4	51-56	53	0.8	46	42-50	41	12.20
62	BP	SUNCO/2*PASTOR	SYNP5	56.7	56-57	53	7.0	47	43-49	41	14.63
63	BP	SW89.5181/KAUZ	SYNP6	50.3	49-52	49	2.7	46	43-48	43	6.98

[1]: Cross; BP: Bi-parental, BC: Back-cross, TC: Three-way cross, [2]: BW: Bread wheat, [3]: SYN; Synthetic, [4]: GFD; Grain filling duration, [6]: Increase/decrease [6]: TKW; Thousand kernel weight.

Table S2.3: GEBVs of cultivated wheat for measured traits in three contrasting environments.

Trait	YLD			PLH			DHE			DFL			DMA		
Environments BWP	IRRI.	DRO.	HEAT	IRRI.	DRO.	HEAT	IRRI.	DRO.	HEAT	IRRI.	DRO.	HEAT	IRRI.	DRO.	HEAT
CNO79	0.482	0.116	0.100	-11.843	-4.774	-2.542	3.611	2.553	0.603	3.035	2.364	1.170	1.953	1.579	0.464
GONDO	0.162	-0.147	0.210	-3.638	-2.789	0.089	4.521	3.279	0.623	6.201	2.921	2.782	3.572	3.296	0.447
GONDO//SHA5/WEAVER/3/PASTOR	0.394	0.097	0.030	-7.832	-4.536	-3.451	0.908	1.733	0.197	1.989	1.563	0.130	1.462	0.543	-0.220
MILAN/S87230//BAV92	1.400	0.435	0.647	-1.926	-2.664	-0.163	3.529	1.008	0.363	2.735	0.851	1.622	0.284	0.316	0.161
BW Line 3570	1.344	0.416	0.643	-1.720	-2.542	-0.139	3.663	1.084	0.502	2.849	0.828	1.725	0.370	0.369	0.296
CACUKE	0.427	0.132	0.147	1.276	3.656	2.295	-4.237	-0.941	-2.050	-2.984	-0.261	-2.903	-2.160	-0.182	-1.263
KRL19	0.458	0.212	0.251	-10.248	-4.009	-1.304	-3.341	-0.624	-1.504	-2.901	-0.480	-1.980	-2.091	-1.258	-1.335
KIRITATI	0.824	0.226	0.422	-1.556	-0.152	1.740	5.033	1.967	1.469	3.302	0.842	2.324	1.570	0.718	0.913
PANDORA	0.741	0.078	-0.011	-8.458	-4.549	-4.306	2.716	3.731	1.720	3.024	3.171	1.585	2.523	2.376	0.786
KIRITATI/2*TRCH	-0.067	0.283	0.282	-12.240	-4.633	-5.656	-7.833	-3.893	0.234	-5.538	-2.449	-1.693	-3.046	-4.821	-0.377
SW89.5181/KAUZ	0.906	-0.128	0.003	-4.146	-5.945	-4.030	9.883	5.608	5.047	7.713	4.233	5.046	4.279	2.906	2.951
SUNCO/2*PASTOR	-0.160	0.117	0.110	0.548	-2.816	-1.917	3.014	1.122	2.357	1.712	1.175	1.101	1.242	-0.166	2.444
PBW502	0.633	0.146	0.650	-0.536	-2.916	5.560	4.328	2.078	0.476	3.775	2.154	2.656	1.624	1.320	0.356
MILAN/AMSEL	0.165	-0.111	-0.033	-3.063	-3.320	-1.316	6.567	3.358	0.109	5.312	3.162	0.790	3.155	2.453	-0.111
TAM200/TUI	0.466	0.007	0.253	-5.665	-3.619	-1.351	-0.400	0.858	1.567	0.453	0.996	1.901	0.213	-0.361	0.684
MINO	0.475	0.263	0.242	-2.680	1.615	2.446	1.243	3.810	-2.333	0.948	2.427	-1.272	0.366	1.131	-1.438
MUU	-0.021	0.026	-0.326	1.560	-2.548	-1.257	5.019	1.901	2.385	2.780	1.119	2.974	1.108	0.598	1.077
HS420	0.848	0.098	0.244	-3.405	-4.523	-2.199	0.246	0.702	-0.506	2.680	1.772	-0.802	2.496	3.019	-0.625
KIRITATI//PRL/2*PASTOR	0.764	0.205	0.342	-2.483	-0.469	1.247	4.457	1.861	1.475	2.905	0.793	2.357	1.681	0.809	0.928

Table S2.4: GEBVs of SYN lines for measured traits in three contrasting environments.

Trait	YLD			PLH			DHE			DFL			DMA		
Env. SYNP NO.	IRRI.	DRO.	HEAT	IRRI.	DRO.	HEAT	IRRI.	DRO.	HEAT	IRRI.	DRO.	HEAT	IRRI.	DRO.	HEAT
SYNP1	-2.46	-0.55	-1.74	19.84	9.69	3.42	1.30	-1.80	0.62	-0.41	-2.23	-1.98	-0.13	-1.28	0.51
SYNP2	-2.28	-0.64	-1.06	15.65	8.28	4.83	0.07	-3.68	0.25	-1.86	-3.39	-1.30	-0.81	-1.52	0.54
SYNP3	-2.07	-0.35	-0.85	20.51	11.91	5.49	-1.20	-3.14	-0.34	-0.39	-3.32	-2.41	0.65	-0.88	0.07
SYNP4	-1.69	-0.34	-0.58	17.92	10.20	3.95	-2.13	-2.91	-0.26	-1.03	-2.96	-2.51	1.51	-0.13	0.32
SYNP5	-0.25	-0.15	-0.39	1.76	4.31	-0.55	-8.04	-5.06	-2.99	-5.49	-4.07	-3.83	-0.85	-1.68	-1.67
SYNP6	-2.46	-0.52	-0.81	16.20	8.58	3.30	0.22	-2.68	1.21	-1.85	-2.42	0.50	-3.50	-2.95	1.33
SYNP7	-0.73	-0.24	-0.09	1.20	2.94	0.54	-4.92	-3.19	-2.68	-1.46	-1.27	-1.71	0.95	1.08	-1.52
SYNP9	-1.60	-0.42	-0.39	12.40	5.30	3.49	-1.01	-1.07	-0.43	0.57	-0.81	-2.41	2.62	1.21	0.18
SYNP11	-1.45	-0.66	-0.81	5.66	2.54	2.40	-0.31	-1.62	-0.24	-1.79	-1.51	-1.32	-1.18	-0.57	-0.01
SYNP12	-1.21	-0.31	0.26	7.04	3.06	1.84	-6.08	-6.08	-4.19	-5.00	-2.67	-1.28	-1.71	-1.15	-1.94
SYNP13	-0.94	-0.32	-0.10	12.05	6.73	3.32	-5.23	-3.97	-2.44	-2.30	-3.35	-3.41	0.64	-1.23	-1.42
SYNP14	-1.55	-0.34	-0.02	20.46	10.44	5.63	-4.16	-3.41	-3.07	-1.76	-3.37	-2.86	-0.34	-0.25	-0.88
SYNP15	-0.49	-0.26	-0.05	13.25	4.69	3.01	-1.86	-0.70	-3.07	-0.24	-0.06	-1.35	0.43	1.35	-1.06
SYNP16	-2.05	-0.52	-0.71	10.09	4.41	5.31	-1.16	-1.44	-2.18	-1.59	-2.24	-2.36	-0.22	-0.90	-0.58
SYNP17	-1.88	-0.76	-1.01	20.10	7.80	4.60	-0.67	-1.61	-2.03	-2.19	-2.19	-1.48	-1.12	-0.09	0.37
SYNP18	-2.38	-0.67	-0.69	17.85	7.45	5.42	-0.42	-3.93	-1.10	-0.92	-3.16	-1.53	-1.98	-1.77	-0.21
SYNP19	-1.54	-0.44	-0.28	16.28	8.18	5.84	-0.88	-2.41	-1.77	-0.82	-2.36	-2.51	0.61	0.04	-0.24
SYNP20	-1.89	-1.02	-1.15	8.20	1.52	2.84	0.54	-0.22	-0.04	2.37	0.14	-0.66	2.24	1.87	1.12
SYNP21	-1.62	-0.60	-0.41	5.82	3.69	4.53	-1.44	-1.02	-0.57	2.11	-0.56	-2.21	3.04	1.45	0.30
SYNP22	-1.00	-0.33	-0.16	9.98	4.11	3.00	-4.30	-3.62	-2.67	-3.13	-2.99	-2.56	0.00	-2.01	-1.68
SYNP23	-1.39	-0.44	-0.54	3.43	1.68	2.61	-1.54	-1.84	-1.61	-0.51	-1.48	-2.04	0.70	-0.27	-0.76
SYNP24	-1.19	-0.38	-0.25	12.80	6.77	3.74	-4.46	-3.92	-2.13	-2.93	-3.41	-2.85	-0.61	-1.86	-1.25
SYNP25	-1.80	-0.36	-0.22	13.12	10.70	4.48	-6.90	-6.48	-4.31	-2.97	-4.64	-3.57	-1.81	-3.46	-2.50
SYNP26	-0.44	-0.10	0.09	6.44	4.30	1.37	-4.92	-4.62	-2.47	-4.24	-2.75	-2.82	0.67	-1.17	-1.19
SYNP27	-1.24	-0.14	0.14	0.58	2.82	0.02	-5.42	-5.34	-1.67	-3.21	-2.21	0.28	-1.62	-2.24	-0.57
SYNP28	-2.69	-0.63	-0.72	7.54	6.55	1.67	-3.91	-5.75	-1.14	-1.69	-3.41	-1.96	-1.15	-2.54	-0.55
SYNP31	-2.43	-0.63	-0.85	15.18	4.76	3.91	-2.13	-2.51	0.56	-4.07	-3.49	-1.39	-1.52	-0.61	0.12
SYNP34	-2.02	-0.51	-0.99	6.43	2.37	5.11	-1.14	-0.67	0.12	-0.29	-0.68	-1.20	1.16	1.17	0.72
SYNP35	-2.20	-0.66	-0.57	9.26	5.64	3.76	-5.26	-6.61	-1.65	-2.12	-3.56	-2.60	-2.12	-2.43	-0.68
SYNP36	-1.21	-0.32	0.24	7.94	3.35	1.47	-5.92	-5.99	-4.07	-4.60	-2.66	-1.30	-1.29	-1.16	-1.82
SYNP39	-2.09	-0.37	-0.56	7.36	4.97	4.04	-0.46	-1.34	-0.16	0.35	-0.62	-0.92	-0.90	-0.47	0.03
SYNP43	-1.06	-0.28	-0.17	9.96	4.72	4.12	-4.97	-3.60	-0.25	-5.26	-3.54	-1.12	0.12	-1.72	0.09
SYNP44	-0.29	-0.10	-0.16	-6.53	-3.21	-0.70	1.55	1.48	0.06	1.99	1.72	0.14	1.79	1.26	-0.14

Env.: environment., SYNP.: synthetic parent

Table S2.5: YLD GEBVs of BW parents and the top 10% of the SDLs within the population under heat stress (HEAT).

BW Parents	SYN Parents	Cross	Ave. YLD GEBV	% Increase/Decrease GEBV
BW 3570			**0.64**	
3570	SYNP14	BC	0.55	-14
CACUKE			**0.15**	
CACUKE	SYNP5	BC	0.46	215
CACUKE	SYNP16	BC	0.50	243
CACUKE	SYNP43	BC	0.62	322
GONDO// SHA5/WEAVER/3/PASTOR			**0.03**	
GONDO//SHA5/WEAVER/3/PASTOR	SYNP7	BP	0.45	42
HS420			**0.24**	
HS420	SYNP13	BC	0.59	142
KIRITATI			**0.42**	
KIRITATI	SYNP5	BC	0.43	2
KRL19			**0.25**	
KRL19	SYNP18	BC	0.49	95
KRL19	SYNP36	BC	0.42	67
MILAN/S87230//BAV92			**0.65**	
MILAN/S87230//BAV92	SYNP4	BC	0.59	-9
MILAN/S87230//BAV92	SYNP4	BP	0.45	-30
MILAN/S87230//BAV92	SYNP17	BC	0.49	-24
MILAN/S87230//BAV92	SYNP17	BP	0.57	-12
MILAN/S87230//BAV92	SYNP20	BC	0.53	-18
MILAN/S87230//BAV92	SYNP20	BP	0.41	-37
MILAN/S87230//BAV92	SYNP21	BP	0.61	-6
MILAN/S87230//BAV92	SYNP21	BC	0.62	-4
MILAN/S87230//BAV92	SYNP23	BC	0.58	-10
MILAN/S87230//BAV92	SYNP27	BC	0.58	-10

BW Parents	SYN Parents	Cross	Ave. YLD GEBV	% Increase/Decrease GEBV
MILAN/S87230//BAV92	SYNP39	BC	0.51	-21
MILAN/S87230//BAV92	SYNP39	BP	0.49	-24
MINO			**0.24**	
MINO	SYNP36/4/GONDO//SHA5/WEAVER/3/PASTOR	TC	0.50	108
MUU			**-0.33**	
MUU	SYNP34	BP	0.55	69
PANDORA			**-0.01**	
PANDORA	SYNP1	BC	0.64	63
PANDORA	SYNP14	BP	0.59	58
PANDORA	SYNP19	BP	0.43	42
PANDORA	SYNP19	BC	0.46	45
PANDORA	SYNP21	BP	0.53	52
PANDORA	SYNP26	BC	0.49	48
PBW502			**0.65**	
PBW502	SYNP22//KIRITATI	TC	0.48	-26
PBW502	SYNP5	BC	0.42	-35
PBW502	SYNP25	BC	0.52	-20
SUNCO/2*PASTOR			**0.11**	
SUNCO/2*PASTOR	SYNP27	BC	0.58	427
SUNCO/2*PASTOR	SYNP5	BC	0.47	327
SUNCO/2*PASTOR	SYNP43	BC	0.49	345
TAM200/TUI			**0.25**	
TAM200/TUI	SYNP2	BC	0.48	92
TAM200/TUI	SYNP3	BP	0.41	64
TAM200/TUI	SYNP4	BC	0.42	68

Table S2.5 compares GEBVs of BW parents (Gray row) with average GEBVs of its corresponding top 10% SDLs (White row) for grain yield (YLD) under heat stress.

Table S2.6: GEBVs of BW parents and the top 10% of the SDLs within the population under irrigated conditions.

BW Parents	SYN Parents	Cross	Ave. YLD GEBVs	% increase/decrease GEBV
3570			**1.34**	
3570	SYNP14	BC	0.98	-27
CACUKE			**0.43**	
CACUKE	SYNP5	BC	0.73	72
HS420			**0.85**	
HS420	SYNP13	BC	0.78	-8
KIRITATI			**0.82**	
KIRITATI	SYNP5	BC	0.78	-6
MILAN/S87230//BAV92			**1.40**	
MILAN/S87230//BAV92	SYNP4	BC	0.92	-34
MILAN/S87230//BAV92	SYNP4	BP	0.78	-44
MILAN/S87230//BAV92	SYNP20	BC	0.94	-33
MILAN/S87230//BAV92	SYNP20	BP	0.89	-37
MILAN/S87230//BAV92	SYNP21	BC	0.90	-36
MILAN/S87230//BAV92	SYNP21	BP	0.83	-41
MILAN/S87230//BAV92	SYNP21	BP	0.86	-38
MILAN/S87230//BAV92	SYNP23	BC	1.02	-27
MILAN/S87230//BAV92	SYNP27	BC	1.05	-25
MILAN/S87230//BAV92	SYNP27	BP	0.78	-45
MILAN/S87230//BAV92	SYNP27	BP	0.72	-49
MILAN/S87230//BAV92	SYNP39	BC	0.86	-38
MINO			**0.48**	
MINO	SYNP36/4/GONDO//SHA5/WEAVER/3/ PASTOR	TC	0.84	77
MUU			**-0.02**	
MUU	SYNP34	BP	0.76	78
PANDORA			**0.74**	
PANDORA	SYNP3	BC	1.00	35
PANDORA	SYNP11	BC	0.72	-2

BW Parents	SYN Parents	Cross	Ave. YLD GEBVs	% increase/decrease GEBV
PANDORA	SYNP14	BC	0.80	8
PANDORA	SYNP18	BP	0.83	12
PANDORA	SYNP19	BC	0.96	29
PANDORA	SYNP23	BC	0.71	-4
PANDORA	SYNP26	BC	0.82	11
PANDORA	SYNP39	BC	0.69	-7
SUNCO/2*PASTOR			**-0.16**	
SUNCO/2*PASTOR	SYNP5	BC	0.95	111
SUNCO/2*PASTOR	SYNP5	BP	0.71	87
SUNCO/2*PASTOR	SYNP27	BC	0.80	96
SUNCO/2*PASTOR	SYNP43	BC	0.90	106
SW89.5181/KAUZ			**0.91**	
SW89.5181/KAUZ	SYNP6	BC	1.09	20
SW89.5181/KAUZ	SYNP35	BC	0.76	-17
TAM200/TUI			**0.47**	
TAM200/TUI	SYNP2	BC	0.78	67
TAM200/TUI	SYNP3	BC	0.80	71

Table S2.6 compares GEBVs of BW parents (Gray row) with average GEBVs of its corresponding top 10% SDLs (White row) for grain yield (YLD) under irrigated condition.

Table S2.7: GEBVs of BW parents and the top 10% of the SDLs within the population under drought stress.

BW Parents	SYN Parents	Cross	Ave. YLD GEBVs	% Increase/Decrease GEBVs
CACUKE			**0.13**	
CACUKE	SYNP43	BC	0.27	104
CNO79			**0.12**	
CNO79	SYNP44	TC	0.32	177
KIRITATI			**0.23**	
KIRITATI	SYNP5	BC	0.31	36
KIRITATI/2*TRCH			**0.28**	
KIRITATI/2*TRCH	SYNP1	BC	0.32	12
KRL19			**0.21**	
KRL19	SYNP18	BC	0.33	53
MILAN/AMSEL			**-0.11**	
MILAN/AMSEL	SYNP7/4/GONDO// SHA5/WEAVER/3/PASTOR	TC	0.34	411
MILAN/S87230//BAV92			**0.44**	
MILAN/S87230//BAV92	SYNP4	BC	0.35	-20
MILAN/S87230//BAV92	SYNP4	BP	0.32	-27
MILAN/S87230//BAV92	SYNP17	BP	0.41	-5
MILAN/S87230//BAV92	SYNP21	BC	0.34	-23
MILAN/S87230//BAV92	SYNP21	BP	0.29	-33
MILAN/S87230//BAV92	SYNP23	BC	0.34	-21
MILAN/S87230//BAV92	SYNP27	BC	0.38	-12
MILAN/S87230//BAV92	SYNP27	BP	0.34	-22
MILAN/S87230//BAV92	SYNP39	BP	0.44	0
MILAN/S87230//BAV92	SYNP39	BC	0.34	-21
MILAN/S87230//BAV92	SYNP39	BP	0.39	-10
MINO			**0.26**	
MINO	SYNP36/4/GONDO//SHA5/WEAVER/3/ PASTOR	TC	0.32	22
PANDORA			**0.08**	

BW Parents	SYN Parents	Cross	Ave. YLD GEBVs	% Increase/Decrease GEBVs
PANDORA	SYNP1	BC	0.30	282
PANDORA	SYNP3	BC	0.27	252
PANDORA	SYNP14	BP	0.41	422
PANDORA	SYNP18	BP	0.28	260
PANDORA	SYNP19	BC	0.27	252
PANDORA	SYNP31	BC	0.26	239
PANDORA	SYNP39	BP	0.28	259
PBW502			**0.15**	
PBW502	SYNP5	BC	0.34	134
PBW502	SYNP25	BC	0.27	82
PBW502	SYNP43	BC	0.31	110
SUNCO/2*PASTOR			**0.12**	
SUNCO/2*PASTOR	SYNP5	BC	0.35	196
SUNCO/2*PASTOR	SYNP27	BC	0.37	218
SUNCO/2*PASTOR	SYNP43	BC	0.33	182
SW89.5181/KAUZ			**-0.13**	
SW89.5181/KAUZ	SYNP35	BC	0.34	363

Table S2.7 compares GEBVs of BW parents (Gray row) with average GEBVs of its corresponding top 10% SDLs (White row) for grain yield (YLD) under drought stress.

CHAPTER 3

Genome-wide Association Study for Grain Yield and Phenological Traits using Synthetic-Derived Wheat Populations

Abstract

Genome-wide association study (GWAS) is a widespread method to identify quantitative trait loci (QTL) in crops. To identify QTL alleles from synthetic hexaploid wheat parents (SYNPs) retained in synthetic derived lines (SDLs), a GWAS was performed at the International Wheat and Maize Improvement Center using 97 SDL populations consisting of 1 to 48 lines each. Yield trials were conducted under irrigated, drought, and heat-stress environments from 2011 to 2014 in Ciudad Obregon, Mexico and grain yield (YLD), days to heading (DHE), days to flowering (DFL), days to maturity (DMA), plant height (PLH), and thousand kernel weight (TKW) and seed number per square meter (SN/m^2) were measured for bread wheat parents (BWPs) and SDLs. Over three years, 1.15% of the SDLs out-yielded the BWP under irrigated conditions, 2.38% under drought, and 6.18% under heat stress ($P < 0.05$). Under irrigation 29% of SDLs had higher TKW than BWPs ($P < 0.05$). GWAS was performed for measured traits using genotyping-by-sequencing markers. A total of 13 QTL were identified for traits on chromosomes 4A, 5A, 7A, 1B, 3B, 4B,5B, 6B, 2D, 3D, 4D, and 6D. The phenotypic variance of traits explained by QTL ranged from 1.51 to 3.7% for YLD, 0.32 to 4.91% for DHE, DFL, and DMA, 3.37 to 4.70% for TKW, 1 to 3.52% for SN/m^2 and 1.65 to 4.7% for PLH. SYNPs contributed positive alleles for increased

YLD, TKW, PLH and maturity traits indicating that synthetic hexaploid wheats can be a valuable source of variation for agronomic and phenological traits and for extending genetic diversity in breeding programs.

Keywords: Synthetic hexaploid wheat, Bread wheat, GWAS, and QTL

Introduction

Increasing genetic diversity is one of the factors that can be used to improve grain yield and other agronomic traits. Genetic diversity has been reduced by long term domestication and selection of crops such as wheat. Hexaploid wheat (*Triticum aestivum* L.) evolved by rare a hybridization between tetraploid wheat and *Aegilops tauschii* L. about 10,000 years ago (Curtis and Halford, 2014). This happened only once or a few times and created an evolutionary bottleneck in hexaploid wheat. Therefore, there is limited genetic diversity for desirable traits in hexaploid wheat and wild relatives could be useful in breeding programs (Mujeeb-Kazi et al., 1996; Dreisigacker et al., 2008). Hybridization of synthetic hexaploid wheat (SYN) with cultivated hexaploid wheat to develop synthetic derived lines (SDL) is an efficient approach to using this diversity. The SYNs are interspecific crosses between *Ae. tauschii* (Coss) Schmalh, donor of the D genome and a modern durum (*Triticum turgidum* L. subsp. *durum*) wheat donor of the A and B genomes (Mujeeb-Kazi et al. 1996). The use of SYNs in breeding programs has been shown to improve disease resistance (Villareal et al., 1994; Kema et al., 1995; Simón et al., 2005; Mulki et al., 2013), drought and heat tolerance (Lopes and Reynolds, 2011; Sharma et al., 2014; Cossani and Reynolds, 2015) yield and its components and other agronomic traits (Ogbonnaya et al., 2007; Yang et

al., 2009; Li et al., 2011b; Cooper et al., 2012; Jafarzadeh et al., 2016). Synthetics have extensive genetic variation for seedling vigor, straw strength, plant height, phenological cycle (Villareal et al., 1994b, 1994c), grain characteristics (Rasheed et al., 2014a; Wu et al., 2015), and thousand kernel weight (Calderini and Reynolds 2000).

Progress in genomic technologies and genome analysis methods have enabled genome wide association studies (GWAS) in animal and plant genetics to identify significant marker–trait associations (MTA) (Neumann et al., 2011; Jighly et al., 2016) and estimates of the effects of quantitative trait loci (QTL) However, GWAS in wheat is challenging because it is a polyploid crop with a large genome and an incomplete genome sequence, which make it difficult to assign the markers to the homoeologous chromosomes (Sukumaran and Yu, 2014).

In a GWAS for grain yield (YLD) and yield related traits using 123 Pakistani historical wheat cultivars under rainfed field conditions, 44 MTAs were located on nine chromosomes (Ain et al., 2015). Yu et al. (2014) using SDL populations developed by crossing SHW-L1 to Chuanmai32, identified six QTL for days to heading (DHE) on chromosomes 2D, 5A, 6B, 7A (2 QTL) and 7D. For these QTL, three of the alleles from the SYN parent increased DHE. They, also, identified seven QTL for thousand kernel weight (TKW) on chromosomes 1B (2 QTL), 2D, 5A (2 QTL), and 7D (2 QTL) for which, four of the alleles from the SYN parent increased TKW. QTLs on chromosomes 2D, 5A, and 7D had pleiotropic effects on both DHE and TKW. Börner et al. (2002), using 114 RILS of the International Triticeae Mapping Initiative (ITMI) population,

identified four major QTLs for plant height (PLH) on chromosomes arms 1AS, 2DS, 4AL, and 6AS, one major QTL for days to flowering (DFL) on chromosome arm 2DS and one major QTL on arm 5AS and one minor QTL on chromosome 5B for grain filling duration (GFD), three major QTLs on chromosome arms 5AL, 3AS, and 6BS for TKW. Huang et al. (2003) used a BC_2F_2 population derived from W-7984 (SYNP) x Prinz mapped with SSR markers in advance backcross QTL analysis and identified 11 QTL for YLD, 16 for yield components, five for PLH, and eight for DHE. For 60% of these QTL, alleles from SYNP were favorable for agronomic traits. Börner et al.(2002) reported one major QTL and one minor QTL for GFD on chromosomes 5AL and 5B, respectively and mentioned that SYN parent alleles extended GFD.

The objective of this study was to identify desirable QTL alleles for grain yield and phenological traits in SDL populations that could be useful in wheat breeding programs using marker-assisted selection (MAS).

Materials and methods

Plant materials and field trials

In this study, synthetic derived line (SDL) populations were used to conduct yield trials under the fully irrigated (IRRI), drought (DRO), and heat (HEAT) stress environments at the Norman E. Borlaug Research station (CENEB) in the Yaqui Valley, Cuidad Obregon, northern Mexico (elevation 38 m, 27°25′ N, 109°54′ W) from 2011 to 2014. The number of genotypes including SDLs, bread wheat parents (BWPs) and checks, planting and harvesting dates for each trial across three years is shown in Table

3.1. The SDL populations were developed by crossing synthetic hexaploid wheat parents (SYNPs) (33) to BWPs (20) at the International Wheat and maize Improvement Center (CIMMYT), and inbred by single spike descent as described in Jafarzadeh et al. (2016). The entire population comprised 97 families in the year 2011-12 yield trials in which family sizes ranged from 1 to 48. In the second and third years, the number of families was reduced due to selection for easy threshing, early maturity, reduced plant height, and lodging resulting in 80 families consisting of 13 BWPs and 30 SYNPs. The SYNPs were genotyped but were not planted in the field because of the poor agronomic characteristics and lack of threshability (Jafarzadeh et al., 2016).

The experimental design followed a partially replicated design (20% of individuals had two replicates and the rest only had one observation) in the year 2011-12 while in the years 2012-13 and 2013-14 it was an alpha lattice with two replicates for heat and irrigated trials and an augmented design for drought trials. The sowing system was bed-channel for the irrigated and heat trials and flat plot area without beds for drought trials across all years. The irrigated and heat trials were irrigated five and six times through gravity flood-irrigation, respectively. The drought trial was irrigated twice, the first at planting time and the second one about 45 days later using a drip irrigation system.

Table 3.1: Information for yield trials in years 2011-14 in Ciudad Obregon, CIMMYT, Mexico.

Years	Trial	Planting data	Harvesting date	No. unique lines
2011-12	DRO.	8-Dec. -2011	14-May-2012	1493
	IRRI.	5-Dec. -2011	30-May-2012	2052
	HEAT	23–March-2012	4-July-2012	1463
2012-13	DRO.	20-Dec. -2012	13-May-2013	1054
	IRRI.	25-Nov.-2012	22-April-2013	1057
	HEAT	8- March -2012	17-June-2013	1045
2013-14	DRO.	20-Dec.-2013	12-May-2014	1056
	IRRI.	6-Dec.-2013	19-21-May-2014	1056
	HEAT	27-Feb.-2014	16-18-June-2014	1054

DRO.: drought trial, IRRI.: irrigated trial, and HEAT trial.

Phenotypic data

Traits measured in yield trials were PLH, DHE, DFL, days to maturity (DMA), GFD (DMA – DHE), YLD gr/m^2 (converted to t/ha) for each year. Thousand kernel weight (TKW), and seed number per square meter (SN/m^2) were measured for the irrigated trial in the year 2011-12 according to (Pask, 2012).

Genotypic data

Genomic DNA was extracted from dried leaves collected from a single plant per line using a modified CTAB (cetyltrimethylammonium bromide) method (Saghai-Maroof et al., 1984) modified as described in CIMMYT laboratory protocols (Dreisigacker et al., 2013) and quantified using a NanoDrop 8000 spectrophotometer V2.1.0. The genotyping of the samples was accomplished using a genotyping-by-sequencing technique called DArTseq™ developed by DArT Pty. Ltd., Yarralumla, Australia. The detailed protocol is described in Sehgal et al. (2015). A total of 20,468 genotyping–by–sequencing (GBS) markers were used for genotyping of 1991 lines.

Marker data were filtered for missing data (NA < 50 %) and minor allele frequency (MAF < 5%). Finally, 9,496 GBS markers were imputed for missing data using the rrBLUP package in R software (Endelman and Jannink, 2013) based on expectation maximization (EM) imputation algorithm to generate the genomic relationship matrix, **G** matrix.

To obtain anchor information for markers, local blasting of SNP's query using the Triticeae Toolbox (T3) (https://triticeaetoolbox.org/wheat/) and the 64K consensus map provided by DArT Pvt. Ltd., Australia (Sehgal et al., 2015) were used. However, local blasting and the consensus map did not locate all markers and the number of anchored markers for the A, B and D genomes was 2277 (24%), 2528 (27%) and 3300 (35%), respectively.

Statistical analysis

Phenotypic analysis of field trial

Since the experimental designs were different for each year and trial, that complicated a combined analysis of all trials. Therefore, to correct for within field heterogeneities spatial analysis was used for each trait/trial combination separately based on row and column orders. The Genstat software (Payne, 2009) was used for analysis of the general linear mixed model by the following equation (Jafarzadeh et al., 2016);

$$\boldsymbol{Y = X\beta + Z_R u_R + Z_C u_C + \varepsilon}$$

where $\boldsymbol{Y}$ is the response vector, **X** is a design matrix for fixed effects such as over mean and genotype effects. $\boldsymbol{Z_R}$ is a design matrix for row effects, $\boldsymbol{Z_C}$ is a design matrix for column effects, is a vector for fixed effects, $\boldsymbol{u_R}$ and $\boldsymbol{u_C}$ are vectors for random row and column effects with $\boldsymbol{u_R} \sim N(0, \sigma_R^2\, \boldsymbol{I})$, and $\boldsymbol{u_C} \sim N(0, \sigma_C^2\, \boldsymbol{I})$ correspondingly and $\boldsymbol{\varepsilon}$ is a residual vector with $\boldsymbol{\varepsilon} \sim N(0, \sigma_{\boldsymbol{R}}^2 R)$, where R is given by $R = Z_\varepsilon[AR1(\rho_R) \otimes AR1(\rho_C)]Z_\varepsilon'$. $AR1(\rho_R)$ is an auto-regressive order one correlation matrix for row effects, $AR1(\rho_C)$ is an auto-regressive order one correlation matrix for column effects and Z_ϵ is a design matrix for row and column combinations. Consequently, row and column effects were removed in each trial and best linear unbiased estimates (BLUEs) of genotypes were generated for subsequent analysis.

Comparison of SDLs with BWPs for traits

A linear model was fitted using the *lm* function in R software (Douglas et al., 2015) for unbalanced data for each environment (DRO, HEAT, and IRRI) and over years. In this model genotypes and years were regarded as fixed effects. Residual standard error from the model was adjusted for the number of observations for each genotype to use in a t-test. Average YLD, PLH, and DMA of SDLs over years were compared with average values of these traits for their respective BWP. For multiple test correction, the Benjamini-Hochberg False Discovery Rate (FDR) method (Benjamini and Yekutieli, 2001) was performed to t-test to avoid type-I errors deriving from the large number of tests. This method was applied calling the p.adjust function incorporated in the R program (R Core Team, 2016). To avoid

false positives, the adjusted p-values were used to determine significance.

For unreplicated data for TKW and SN/m^2 in IRRI.Y11.12, a mixed model was used to estimate variance components using EMMREML package in R software (Akdemir and Godfrey, 2015). Variance component were computed using the following univariate mixed model:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{\epsilon}$$

where y is a vector of raw phenotypic data (not BLUEs) for the genotyped individuals for TKW and SN/m2 (Because these two traits were measured on a subset of the individuals, spatial analysis did not apply to them for estimating BLUEs), **X** is a known design matrix for fixed effects which comprised families, **Z** is a known design matrix for random effects (individuals), **β** is a vector for non-genetic fixed effects, **u** is a vector for genetic random effects or breeding values with $\mathbf{u} \sim N(0, \sigma_u^2 \boldsymbol{K})$, $\boldsymbol{K}$ is the genomic relationship matrix and $\boldsymbol{\epsilon}$ is a residual vector with $\boldsymbol{\epsilon} \sim N(0, \sigma_e^2 \boldsymbol{I}_n)$(Piepho et al., 2007). Residual sstandard error from the mixed model was used for the t-test.

Generating joint genetic linkage map and marker coverage

In this project, the number of individuals for each biparental, BC1 and TC SDL families ranged from 1 to 48 and were too small to construct a genetic map. Therefore, a subset of populations in which progenies of one common BWP crossed to several SYNPs was used to generate a genetic map similar to the nested association mapping procedure (NAM) described by Li et al. (2011). As an example, the PANDORA NAM

population had 12 BC families sharing a common BWP PANDORA. In this NAM population, the number of individuals for each family ranged from 11 to 31, resulting in a total of 245 lines. Markers in each family of the NAM population were filtered for < 40% missing data and only markers that were polymorphic between PANDORA and all the other 12 SYNPs were selected. Also, a chi-squared test was used to identify segregation distortion in each family for the expected 3:1 ratio of BWPs and SYNPs alleles in BC and 1:1 in biparental families. Although, the BWPs and SYNPs alleles in the SDL populations were selected for traits resembling that of the BWPs, the ratio of most parental alleles approximated 3:1. Heterozygous loci were converted to missing data. Markers that were not polymorphic in some specific families were replaced with NA. Map construction was performed using QTL IciMapping v4.0.6.0 software with Kosambi mapping function (Meng et al., 2015) (freely available from www.isbreeding.net). Grouping of markers used anchored marker information and a minimum LOD value of 6.0 for unanchored markers. The nnTwoOpt algorithm was used for optimal ordering of markers within each linkage group where it followed a path through the genome that returned the shortest genetic distance map. The sum of adjacent recombination frequencies criterion was used for rippling of markers (Meng et al., 2015). It was assumed that all families within the PANDORA NAM population had the same recombination frequency. The genetic map was used an approximate map for GWAS using all lines and markers.

Linkage disequilibrium

The pairwise r^2 for markers on each chromosome and unlinked markers (between chromosomes) were calculated for linkage disequilibrium (LD) between mapped SNPs (Singh and Singh 2015) and plotted against the genetic distance (in cM) between pairs of the markers within each chromosome. Then a locally weighted polynomial regression curve (LOESS) of r^2 values was fitted on the genetic distance (Singh and Singh, 2015) using the statistical program R (http://www.r-project.org). A critical value for r^2 was estimated using 95^{th} percentile of the distribution of unlinked r^2 values (pairwise r^2 values for markers on different chromosomes) according to Breseghello and Sorrells (2006). The extent of LD in the chromosome was the intersection of the loess curve with the baseline.

GWAS for identification of QTL

GWAS was performed with the mixed linear model (MLM) including both fixed and random effects using TASSEL Standalone v.5.2.30 which implemented the Efficient Mixed-Model Association (EMMA) and the population parameters previously determined (P3D) algorithm to reduce computing time. In the MLM model the kinship matrix (K) was used to control population structure (familial relationship) (Bradbury et al., 2007);

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{e}$$

where $\boldsymbol{y}$ is a vector of spatially corrected observations of genotyped individuals for traits, **X** is a known design matrix for fixed effects including markers, **Z** is a known

design matrix for random effects comprising individuals, **β** is a vector for fixed effects, **u** is a random vector for additive genetic effects for individuals with $\mathbf{u} \sim N(0, \sigma_a^2 \boldsymbol{K})$, σ_a^2 is unknown genetic variance and $\boldsymbol{K}$ is the genomic relationship matrix, and $\boldsymbol{e}$ is a residual vector with $\mathbf{e} \sim N(0, \sigma_e^2 \boldsymbol{I}_n)$ (Bradbury et al., 2007). Bonferroni test at 5% significance level was used as a threshold to identify the significant marker-trait associations.

For the GWAS analysis over years, for each management, a weighted linear model was used for unbalanced data and weights were the number of observations for each individual (Bates et al., 2015). In this model $\boldsymbol{y}$ was a vector of the BLUEs for genotyped individuals for traits for each single year and management combinations (e.g. Y11.12.DRO, Y12.13.DRO, …). Genotypes and years were regarded as fixed effects. Not all individuals had phenotypic and genotypic data in all environments across three years. The final population size for each of the trait-environment combinations used in GWAS is shown in Table 3.2.

Table 3.2: Population size for six traits across environments and years.

Manag.	DRO.				HEAT			IRRI.			
Trait/Year	Y11.12	Y12.13	Y13.14	Over years	Y12.13	Y13.14	Over years	Y11.12	Y12.13	Y13.14	over years
DHE	-	912	913	912	913	913	913	-	913	913	913
DFL	1283	912	913	1473	-	913	913	1787	913	913	1811
DMA	1283	-	913	1473	913	913	913	1787	913	913	1811
PLH	1283	-	913	1473	913	913	913	1787	913	913	1811
YLD	1283	912	913	1473	913	913	913	1787	913	913	1811
TKW	-	-	-	-	-	-	-	808	-	-	-

Manag.: management, DHE: days to heading, DFL: days to flowering, DMA: days to maturity, YLD: grain yield, PLH: plant height, TKW: thousand kernel weight, DRO: drought stress, IRRI: irrigated condition and HEAT for heat stress, Y11.12: year 2011-2012, Y12.13: year 2012-2013, Y13.14: year 2013-2014.

QTL by environment interaction

Significant MTAs from GWAS (single year and over years) were used to identify QTL by environment interaction (QEI) using the same model (MLM used in GWAS). Here QEI was split in two terms; marker by management interaction (DRO, HEAT and IRRI) (MMI) and marker by year interaction (MYI). In MLM markers, managements, years, MMI, and MYI were included in X matrix as fixed effects. The EMMREML package (Akdemir and Godfrey, 2015) was used to estimate and test interaction effects. This package uses Wald test statistics for testing whether the fixed effect coefficients are equal to zero and obtains probability (*P*) values from large sample theory for the fixed effects. The kinship relationship matrix included in the model to control population structure.

Population structure

To identify subpopulation structure, the hierarchical cluster analysis with the Ward method and Euclidean distance (Timm, 2002) was used to classify the BWPs and SYNPs and SDL populations based on the whole genome marker information of 9,496 SNPs. To account for the population structure principal components analysis (PCA) was performed using 9,496 SNPs and the first four principal components were used to identify subpopulations (Timm, 2002).

Results

Phenotypic results

The summary information for traits from each trial and year is presented in Table 3.3 and shows that parents and SDLs had a wide range of values for measured traits in each trial/year. Genotypes indicated higher coefficient of variation (CV) for YLD than the other traits and HEAT trials showed the highest variation (CV = 25.62 to 65.87%) for YLD followed by DRO and IRRI trials. The highest CV for the HEAT.Y11.12 was caused by late planting resulting in very low yield with some genotypes not producing any grain. The range of average YLD was from 5.55 to 6.34 t/ha for IRRI, 1.05 to 2.42 t/ha for DRO and 0.57 to 2.07 t/ha HEAT, respectively. The number of SDLs that out-yielded their respective BWPs were 96 (6.18%) under HEAT, 37 (2.38%) under DRO and 23 (1.15%) under IRRI (FDR $P < 0.05$) (Table S3.1). Under HEAT, the highest number of SDL that out-yielded their BWPs was for populations whose BWPs were PANDORA followed by SUNCO/2*PASTOR, MUU, and TAM200/TUI. Under DRO, the highest number of SDL that out-yielded their BWPs was for populations whose BWPs were PANDORA, SW89.5181/KAUZ, TAM200/TUI. Under IRRI, the highest number of SDL that out-yielded their BWPs was for populations whose BWPs were SUNCO/2*PASTOR, KIRITATI/2*TRCH, MUU, and PANDORA (Table S3.1).

For PLH also HEAT trials had higher CV followed by DRO and IRRI (Table 3.3). The number of SDLs that had significantly greater PLH than their respective BWPs

were 32 (1.60%) under IRRI, 491 (31.64%) under DRO and 271 (27.37%) under HEAT (FDR $P < 0.05$) (Table S3.2).

The genotypes had almost the same CV for DMA except for IRRI.Y13.14 that had the highest CV (4.48%) (Table 3.3). In contrast, CVs for DHE and DFL were higher than those for DMA and varied across trials and years. For DMA, the number of SDLs with significant higher DMA than their recurrent BWPs were 199 (12.56%) under DRO, 29 (2.93%) under HEAT and 16 (0.80%) under IRRI (FDR $P < 0.05$) (Table S3.3).

Table 3.3: Mean and range of traits in different trials for years 2011-14 in Ciudad Obregon, CIMMYT, Mexico.

Trial/Trait	DHE	CV%	DFL	CV%	DMA	CV%	PLH (cm)	CV%	YLD (t/ha)	CV%
IRRI.Y11.12	-	-	81[a] (61-95)[b]	4.55	128 (119-136)	2.42	114 (87-150)	7.64	6.34 (2.90-8.50)	12.66
IRRI.Y12.13	73 (58-93)	8.95	78 (63-97)	8.41	126 (117-136)	2.20	102(82-121)	5.70	5.95 (2.78-8.94)	15.58
IRRI.Y13.14	75 (65-88)	4.52	79 (69-92)	4.46	121 (107-133)	4.48	102 (86-121)	4.29	5.55 (3.18-7.59)	13.95
DRO.Y11.12	-	-	81 (72-99)	2.94	117 (104-130)	2.66	84 (58-120)	7.37	2.42 (1.09-3.56)	20.82
DRO.Y12.13	75 (65-87)	4.21	78 (66-92)	4.27	-	-	-	-	2.30 (1.55-2.95)	20.84
DRO.Y13.14	67 (58-79)	5.35	69 (60-80)	5.42	100 (91-109)	2.69	70 (50-96)	6.99	1.05 (0.49-1.40)	18.04
HEAT.Y11.12	-	-	-	-	-	-	42 (20-70)	8.89	0.57 (0.00-2.40)	65.87
HEAT.Y12.13	50 (45-59)	6.08	-	-	81 (78-89)	2.32	61 (45-75)	8.42	1.96 (0.29-3.18)	25.62
HEAT.Y13.14	56 (50-66)	6.47	59 (54-69)	6.09	87 (82-96)	2.98	59 (41-89)	7.89	2.07 (0.33-3.26)	26.91

DHE: Days to heading, DFL: Days to flowering, DMA: Days to maturity, PLH: Plant height, and YLD: Grain Yield t/h, CV: coefficient of variation in percent.

IRRI: Irrigated, DRO: Drought, HEAT: Heat trials, Y11.12: Year 2011-12, Y12.13: Year 2012-13, and Y13.14: Year 2013-14 (e.g. IRRI.Y11.12: irrigated trial in the year 2011-12).

[a]; Mean of the trait, [b]; Range of the trait.

The range for TKW for the IRRI.Y11.12 trial was from 41 to 54 gr for BWPs and on average from 44 to 56 gr for SDLs (Table S3.4). Under IRRI condition 29% of SDLs (237 out of 808) had significantly higher TKW than their respective BWPs (FDR $P < 0.05$) and the range of significant TKW differences was from 5 to 19 gr. The largest differences in TKW means (19 gr) occurred in SDL populations that had PANDORA as the recurrent BWP (Table S3.4).

The range for SN/m^2 for the IRR.Y11.12 trial was from 12098 to 18842 for BWPs while this range was from 10388 to 19311 (average 11593) for 68 SDL populations. Only for three populations in which PANDORA and KIRITATI//PRL/2*PASTOR were recurrent BWPs, two and six SDLs had higher SN/m^2, respectively (FDR $P < 0.05$) (Tables S3.5 and S3.6). The correlation between SN/m^2 and YLD was significantly positive over the all populations ($y = 0.00024x + 3.4; P < 0.001. R^2 = 0.57$) for the IRRI.Y11.12 trial.

Pearson's correlation coefficients were calculated for traits based on the data averaged across years for each environment. Under DRO stress, YLD had negative correlations with DHE, DMA, DFL, and GFD (not meaningful), while it had a positive correlation with PLH. Under HEAT stress, the trend was the same except that YLD was positively correlated with GFD (Table 3.4). In contrast, under IRRI conditions, YLD had positive significant correlations with DHE, DMA, and DFL and a negative correlation with GFD (Table 3.5). Moreover, YLD had a significant negative correlation with TKW, while it was positively correlated with SN/m^2 (Table 3.5). GFD and PLH

had positive correlations with TKW and negative correlations with SN/m^2 under IRRI. TKW had negative correlations with DHE, DFL, and DMA but had a positive correlation with GFD (Table 3.5).

Table 3.4: Correlation coefficients among traits based on average data across years under DRO (above diagonal), and HEAT (below diagonal) stresses.

Trial	DRO.					
Trait	DHE	DFL	DMA	GFD	PLH	YLD
DHE		**0.93**	**0.71**	**-0.21**	**-0.38**	**-0.36**
DFL	**0.88**		**0.74**	**-0.08**	**-0.38**	**-0.38**
DMA	**0.91**	**0.82**		**0.54**	**-0.32**	**-0.38**
GFD	**-0.82**	**-0.67**	**-0.50**		0.00	**-0.09**
PLH	**-0.44**	**-0.34**	**-0.38**	**0.40**		**0.30**
YLD	**-0.30**	**-0.19**	**-0.29**	**0.23**	**0.33**	
Trial	HEAT					

Values in bold were significant ($P < 0.05$).

Table 3.5: Correlation coefficients among traits based on average data across years under IRRI conditions.

Trial	IRRI							
Trait	DHE	DFL	DMA	GFD	PLH	TKW*	SN/m^{2*}	YLD
DHE	1							
DFL	**0.97**	1						
DMA	**0.73**	**0.75**	1					
GFD	**-0.74**	**-0.69**	**-0.09**	1				
PLH	**-0.09**	**-0.1**	**-0.16**	**-0.04**	1			
TKW	**-0.45**	**-0.44**	**-0.32**	**0.34**	**0.23**	1		
SN/m^2	**0.41**	**0.41**	**0.28**	**-0.32**	**-0.18**	**-0.73**	1	
YLD	**0.18**	**0.19**	**0.13**	**-0.13**	-0.03	**-0.13**	**0.77**	1

Values in bold were significant ($P< 0.01$).
*: TKW and SN/m^2 were based one-year data (IRRI.Y11.12)

Marker coverage and polymorphism

Using 245 SDLs and 2595 polymorphic SNPs, the linkage map had a total length of 2610 cM, with an average density of one SNP per 0.99 cM. The minimum and maximum length of chromosomes was for chromosome 6B (35 cM) and chromosome 3A (188 cM). The maps for the A, B, and D genomes were 957, 823, and 830 cM, with

a density of one SNP per 0.85, 0.79, and 1.96 cM, respectively. The numbers of SNPs for the A, B, and D were 1126 (43.39%), 1045 (40.27%), and 424 (16.34%), respectively. The numbers of SNPs per chromosome ranged from 35 (chromosome 4D) to 217 (chromosome 2B) (Table S3.7, Figures S3.1 and S3.2).

Population stratification

The hierarchical cluster analysis indicated, by an arbitrary cutoff at the Euclidean distance of 600, SDLs, BWPs and SYNPs grouped in six different clusters. Therefore, this number was used to select subpopulations using PC values (Figure 3.1). Although Figure 3.1 shows a distinctive subpopulation structure, low values for PC1 = 6.6% and PC2 = 5.3% revealed that there was a low population structure. The first four PCs, which explained 20.2% variation in populations, generated six subpopulations (Figure 3.2 A and B). These subpopulations grouped by BWPs in which CACUKE, PANDORA, MILAN/S87230//BAV92, and TAM200/TUI had a major contribution in the population stratification and their names were used for labeling of the clusters. However, two clusters were labeled A and B because these two clusters comprised different parents and their progenies. Cluster A was the biggest cluster with 569 SDLs, 12 BWPs, and two SYNPs. In this cluster, collectively, progenies of BWPs KIRITATI, KRL19, PBW502, and SUNCO/2*PASTOR comprised 62% of SDLs (Fig3.2. A and B). Cluster B was not a distinctive cluster and most of its individuals were distributed in other clusters. This cluster included 30 SYNPs, two BWPs and, in total, 72% of its SDLs were progenies of BWPs MILAN/S87230//BAV92, PANDORA, and SW89.5181/KAUZ (Fig3.2. A and B). The other four clusters were labeled

TAM200/TU, CACUKE, MILAN/S87230//BAV92, and PANDORA and included 95, 98, 95, and 98% of SDLs of these parents, respectively (Fig3.2. A and B). However, some individuals in cluster B (red cluster) were placed very close to the cluster PANDORA (purple cluster) and were biparental SDLs of this BWP.

The Quantile-Quantile (Q-Q) plots for all traits in each year/trial (Figure S3.3) and over years across three trials (DRO, HEAT, and IRRI) (Figure S3.4) indicated that subpopulation structure was effectively controlled by including the kinship matrix (K) in the MLM model.

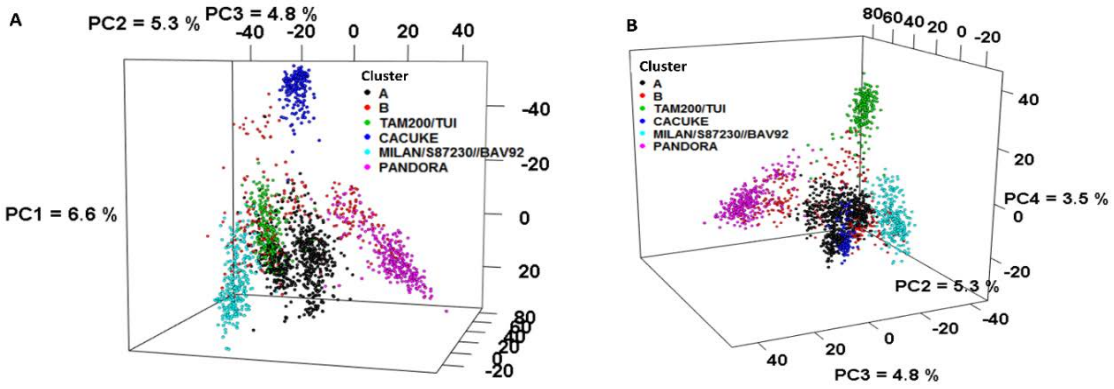


Figure 3.1: 3D plots of PCs for population structure of SDLs, BW and SYN parents. Clustering of SDLs and parents based on PC1, PC2, and PC3 (A) and PC2, PC3, PC4 (B) resulted in six clusters. In clusters A (black) and B (red), none of the BWPs or SYNPs had a major contribution while four other clusters were labeled with the primary BWP; TAM200/TUI (green), CACUKE (blue), MILAN/S87230//BAV92 (light blue) and PANDORA (purple).

Linkage Disequilibrium

The pairwise r^2 values averaged 0.124, 0.128, 0.120, and 0.126 for genomes A, B, D, and the whole genome, respectively. The critical value of r^2, at the 95 percentile of the distribution of unlinked pairs, was estimated to be 0.0609 and $r^2 > 0.0609$ values were considered to be due to genetic linkage. The percentage of pairwise estimated r^2 values above the baseline was 41%, 45%, 43%, and 43% for genome A, B, D (Figure S3.5) and whole genome (Figure 3.2). The extent of LD was 32 cM, for D genome 30 cM for A genome, 28 cM for B genome (Figure S3.5), and 30 cM for the whole genome (Figure 3.2).

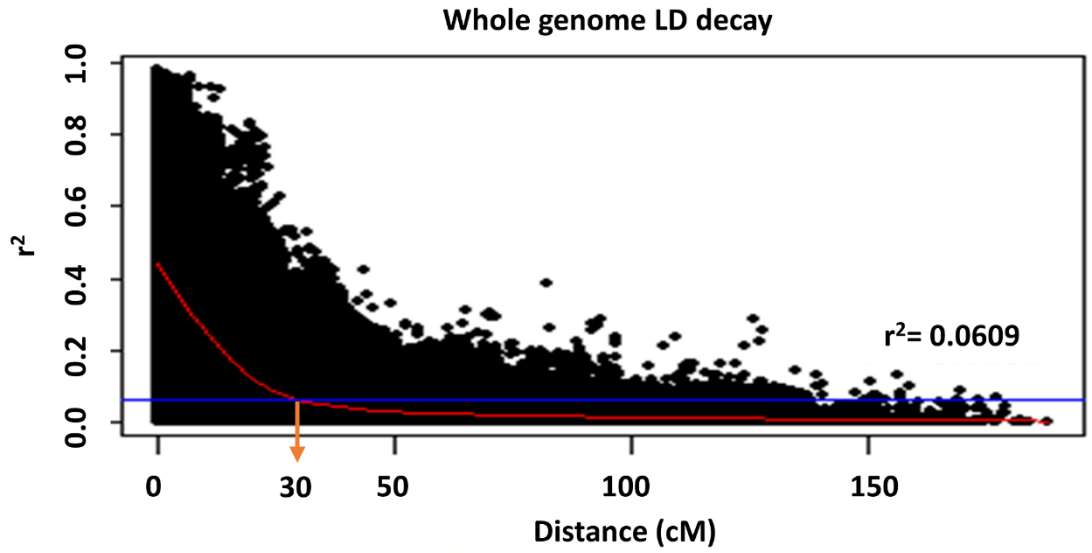


Figure 3.2: LD decay plot for whole genome (all chromosomes). Estimated pairwise r^2 were plotted against the genetic distance (cM). Red curve is LOESS smooth line and horizontal blue line is the 95 percentile of the distribution of unlinked pairwise r^2. The intersect of smooth line and baseline as the extent LD was 30 cM.

Identified QTL by GWAS

Using MLM and a 0.05 significance threshold with Bonferroni correction, a total of 13 QTL were identified including three for DHE, DFL and DMA, two for PLH, six for TKW, and four for YLD across all environments (DRO, HEAT, and IRRI and years) (Tables 3.6, 3.7 and 3.8). However, one of PLH QTL was coincident with YLD and TKW on chromosomes 4B, one QTL for DFL was coincident with YLD on chromosome 2D and one QTL overlapped with a QTL for TKW on chromosome 3D. The D and B genomes each had five QTL and the A genome had three QTL. The phenotypic variance of traits explained by QTL (R^2) ranged from 0.32 to 4.90%. To assign a QTL for a trait, correlations among SNPs associated with the trait and genetic position of SNPs in the linkage map were used. To explain features of each QTL, the SNP with the highest *P* value was used as the best predictor of the QTL. In this study, the origin of the SNP associated with the trait was important and we had to determine which parent allele had a positive effect on the trait. Therefore, in each family, for each SNP, the BWP allele was initially assigned -1 and the SYNP allele +1. It is important to note that the SYNP or the BWP did not always have the negative (reference allele) or positive allele (alternate allele). Therefore, phasing differed for different families. In each family, for a given SNP, if the BWP and SYNP alleles had -1 and +1, respectively, it was kept. If the signs for the parent's allele of the SNP did not follow the above-mentioned order, the signs were reversed such that the BWP allele was changed from +1 to -1 and SYNP allele was changed from -1 to +1. This change was applied to all SDLs within that family. To identify which allele had a positive or negative effect on

trait, the additive effect of allele substitution was used (e.g. in T/C SNP, substitution of BWP allele, T, by the C (the SYNP allele)). If the additive effect was negative, the BWP allele increase the trait value and if it was positive, the SYNP allele increased the trait value. For example, in Table 3.6 the SYNP allele for the T/C SNP for DFL is C and that allele increased DFL by 1.1 days.

Segregation ratios of the parental alleles for SNPs associated with traits in the SDL populations were compared to the theoretical ratio of 3:1 (BWP allele: SYNP allele) in BC1 and 1:1 in biparental families using the chi-squared test (Table S3.8). Nine markers including SNP 1093512 for YLD (on chromosome 2D), SNPs 2244579, 2242893 for YLD and DFL (on chromosome 2D), SNP 2251719 for TKW (on chromosome 2D), SNP 985496 for PLH (on chromosome 3B), SNPs 1216917, 983836, 1088389 for PLH, TKW and YLD (on chromosome 4B), SNP 1102535 for YLD (on chromosome 4D) deviated from 3:1 and 1:1 ratio in the BC1 and biparental SDL populations ($P < 0.05$) except SNP 985496 for PLH (on chromosome 3B) that did not deviate from 1:1 ratio (Table S3.8). These SNPs had an excess of BWP allele scores. This distortion most likely occurred due to selection of SDLs for traits approximating that of the BWPs.

Heat stress

Under heat stress, five QTL, one associated QTL with DFL, DHE, and DMA, one with PLH, and three with YLD, were detected (Table 3.6). The threshold was 5.679E-06 at a significance level of 5% after Bonferroni multiple test correction.

For maturity traits, a QTL detected was on chromosome 5A coincident with DHE, DFL, and DMA was designated as qDHE-5A-1 (Figure S3.6 A), qDFL-5A-1, and qDMA-5A-1 (Table 3.6). The SYNP allele of the qDFL-5A-1 explained 4.51% of the phenotypic variance of DFL (R^2) and its additive effect was 1.1 day on DFL (DFL was not recorded for Y12.13) (Table 3.6). The SYNP allele of these QTL with positive additive effect on maturity traits delayed SDLs in Y13.14. For DHE, the SYNP allele of the qDHE-5A-1 had an additive effect of 1.1 day and R^2 of 4.6%. For DMA, the additive effect of SYNP allele of the qDMA-5A-1 was 0.9 day with R^2 of 4.81% (Table 3.6).

For YLD, three QTL were identified on chromosomes 2D, 4D, and 6D and designated as qYLD-2D-1, qYLD-4D-1, and qYLD-6D-1 (Table 3.6, Figure S3.7 A and B). The additive effect of BWP allele for qYLD-2D-1 was -0.171 t/ha in Y12.13 with R^2 of 2.52% which is indicating that the BWP allele increased YLD in Y13.14 (Table 3.6). The BWP allele of qYLD-4D-1 had additive effects of -0.127, -0.159 and -0.142 t/ha and R^2 of 2.82, 2.87 and 1.63% in Y12.13, Y13.14 and over years, respectively. For qYLD-6D-1, additive effects of SYNP allele was 0.146 t/ha in Y13.14 with R^2 of 3.70% indicating that the SYNP allele increased YLD.

For PLH, a QTL was detected on chromosome 3B and designated as qPLH-3B-1. The BWP allele of this QTL had an additive effect of -1.6 cm with R^2 of 1.73% indicating that the BWP allele increased PLH (Table 3.6).

Table 3.6: Summary of GWAS results for different traits under heat stress (HEAT) for each year and over years.

Trait	SNP NO.	SNP	Chr.	Chr. pos. (cM)	QTL	Y12.13			Y13.14			Y12.14 (over years)			MAF
						P value	Add. Effect*	R^2 %	P value	Add. Effect	R^2 %	P value	Add. Effect	R^2 %	
DFL	1141498	T/C	5A	59	qDFL-5A-1				2.49E-06	1.1	4.51				0.27
DHE	1141498	T/C	5A	59	qDHE-5A-1				1.97E-06	1.1	4.60				0.27
DMA	1141498	T/C	5A	59	qDMA-5A-1				4.52E-07	0.9	4.81				0.27
PLH	985496	G/T	3B	NA	qPLH-3B-1							1.37E-06	-1.6	1.73	0.33
YLD	2242893	C/T	2D	112	qYLD-2D-1	5.38E-06	-0.171	2.52							0.11
YLD	1102535	G/A	4D	89	qYLD-4D-1	4.41E-06	-0.127	2.83	9.19E-07	-0.159	2.87	1.85E-07	-0.142	1.63	0.16
YLD	1067078	C/T	6D	111	qYLD-6D-1				1.10E-08	0.146	3.70				0.39

SNP: single nucleotide polymorphism, Chr.pos: chromosome position, Add. Effect: additive effect, MAF: minor allele frequency, R^2: phenotypic variance of trait explained by QTL, DHE: days to heading, DFL: days to flowering, DMA: days to maturity, PLH: plant height, YLD: grain yield.
*: "-":Add. Effects indicates increasing trait value by the BWP allele and "+" indicates increasing trait value by the SYNP allele.
Additive effect for a SNP (e.g. T/C SNP) for a given trait = (mean of homozygous SDLs for TT – mean of homozygous SDLs for CC)/2
Units for Add. Effect for DFL, DHE, and DMA is day, for PLH is cm and for YLD is t/ha.

Drought stress

Under drought stress, six QTL, two QTL associated with DFL, DHE, one with DMA, one with PLH and two YLD, were identified (Table 3.7). The threshold was 5.56E-06 at a significance level of 5% for Bonferroni multiple test correction.

For DHE, qDHE-5A-1 and qDHE-4A-1 were detected on chromosome 5A and 4A in Y13.14 (Figure S3.6 B) and over years. The additive effect of the SYNP allele of the qDHE-5A-1 was 1.1 and 0.9 day on DHE with R^2 of 3.91and 0.33% in Y13.14 and over years, respectively (Table 3.7). The BWP allele of the qDHE-4A-1 had additive effect of -0.9 and -0.7 day on DHE and R^2 of 3.33 and 0.32% in Y13.14 and over years, respectively. For DMA, a QTL was detected on chromosome 5B in Y13.14 and designated as qDMA-5B-1. The SYNP allele of this QTL had additive effect of 0.7 day on DMA and R^2 of 4.10% (Table 3.7).

For PLH, the qPLH-4B-1was detected on chromosome 4B in Y11.12 and over years (Figure S3.8 A). The SYNP allele of this QTL had an additive effect of 2.4 and 2.0 cm on PLH and R^2 of 2.06 and 1.87% in Y11.12 and over years, respectively (Table 3.7).

For YLD, two QTL, qYLD-2D-1 and qYLD-4D-1, were detected on chromosome 2D and 4D in over years and its BWP allele had an additive effect of -0.136 and -0.095 t/ha on YLD with R^2 of 2.11 and 1.96%, respectively (Table 3.7 and Figure S3.9).

Table 3.7: Summary of GWAS results for different traits under drought stress (DRO) for each year and over years.

Trait	SNP NO.	SNP	Chr.	Chr. Pos. (cM)	QTL	Y11.12			Y13.14			Y11.14 (over years)			MAF
						P value	Add. Effect*	R^2 %	P value	Add. Effect	R^2 %	P value	Add. Effect	R^2 %	
DFL	1141498	T/C	5A	59	qDFL-5A-1				4.37E-06	1.0	3.27				0.24
DFL	3027878	G/A	4A	35	qDFL-4A-1				4.19E-06	-0.9	3.19				0.33
DHE	1141498	T/C	5A	59	qDHE-5A-1				6.38E-07	1.1	3.91	1.30E-06	0.9	0.33	0.24
DHE	3027878	G/A	4A	35	qDHE-4A-1				2.16E-06	-0.9	3.33	4.10E-06	-0.7	0.32	0.33
DMA	1092197	C/T	5B	NA	qDMA-5B-1				4.72E-06	0.7	4.10				0.26
PLH	983836	T/C	4B	33	qPLH-4B-1	6.86E-07	2.4	2.06				1.87E-06	1.9	1.65	0.14
YLD	1093512	T/G	2D	112	qYLD-2D-1							4.64E-08	-0.136	2.11	0.12
YLD	1102535	G/A	4D	89	qYLD-4D-1							1.59E-07	-0.093	1.94	0.18

SNP: single nucleotide polymorphism, Chr.pos: chromosome position, Add. Effect: additive effect, MAF: minor allele frequency, R^2: phenotypic variance of trait explained by QTL, DHE: days to heading, DFL: days to flowering, DMA: days to maturity, PLH: plant height, YLD: grain yield.
*: "-":Add. Effects indicates increasing trait value by the BWP allele and "+" indicates increasing trait value by the SYNP allele.
Units for Add. Effect for DFL, DHE, and DMA is day, for PLH is cm and for YLD is t/ha.

Irrigated condition

Under irrigated conditions, 12 QTL, three for DFL, one for DHE, one for PLH were coincident with YLD, six for TKW and one for YLD, were identified (Table 3.8). The threshold was 5.68E-06 at a significance level of 5% for Bonferroni multiple test correction.

For DFL, qDFL-4A-1, qDFL-5A-1 and qDFL-2D-1 were detected on chromosomes 4A, 5A and 2D in Y11.12, Y12.13 and over years, respectively. The additive effect of SYNP allele of the qDFL-4A-1 was 0.7 day on DFL with R^2 of 1.24%. For the qDFL-5A-1, SYNP allele had an additive effect of 2.07 days on DFL and R^2 of 4.91%. The SYNP allele of qDFL-2D-1 (coincident with QTL for YLD) had an additive effect of 1.01 days on DFL with R^2 of 1.32% (Table 3.8, and Figure S3.10). For DHE, qDHE-5A-1 (coincident with QTL for DFL) was identified on chromosome 5A in Y12.13 and over years (Figure S3.6 C). The SYNP allele of this QTL had an additive effect of 2.05 and 1.4 days on DHE with R^2 of 4.87 and 1.80% in Y12.13 and over years, respectively (Table 3.8).

For PLH, qPLH-4B-1 (coincident with QTL for YLD) was detected on chromosome 4B in Y11.12, Y13.14, and over years (Figure S3.8 B). Its SYNP allele had an additive effect of 4.2, 2.3, and 3.8 cm on PLH with R^2 of 3.34, 4.70, and 3.52% in in Y11.12, Y13.14, and over years, respectively (Table 3.8).

For TKW, six QTL, qTKW-1B-1, qTKW-4B-1, qTKW-6B-1, qTKW-2D-1, qTKW-3D-1 and qTKW-7A-1 were identified on chromosomes 1B, 2B, 6B, 2D, 3D

and 7A, respectively. However, our genetic map did not cover their genetic positions. The SYNP alleles of these QTL had positive additive effect on TKW and increased the TKW value (Table 3.8). For qTKW-1B-1, qTKW-4B-1, qTKW-6B-1, additive effects of SYNP alleles were 2.20, 1.79 and 1.76 gr on TKW with R^2 of 3.98, 4.54 and 3.96%, respectively. The qTKW-4B-1 most likely was coincident with QTL for YLD and PLH (qYLD-4B-1 and qPLH-4B-1) and placed in the same chromosome position because squared correlations among SNP 1208575 T/C (linked to the qTKW-4B-1) and SNPs 1216917 G/T, 983836 T/C and 1088389 A/G (linked SNPs to the qYLD-4B-1 and qPLH-4B-1) were 0.25, 0.27 and 0.24. For qTKW-2D-1 and qTKW-3D-1, SYNP alleles had additive effects of 1.42, 1.70, 2.24 gr and R^2 of 3.56, 3.84 and 4.90%, respectively. For qTKW-7A-1, the additive effect of SYNP alleles was 1.79 gr with R^2 of 3.37% (Table 3.8). The qTKW-2D-1 was different QTL than the qYLD-2D-1 and qDFL-2D-1 for YLD and DFL because squared correlations among SNP 2251719 A/G (linked to the qTKW-2D-1) and SNPs 1093512 T/G, 2242893 C/T and 2244579 A/G (linked SNPs to the qYLD-2D-1 and qDFL-2D-1) were 0.01, 0.02 and 0.01 which were smaller than 0.0609 (the 95 percentile of the distribution of unlinked pairwise r^2).

For SN/m^2, qSN-3D-1 (coincident with QTL for TKW) was identified on chromosome 3D and its BWP allele had an additive effect of 686 SN/m^2 with $R^2$of 3.52% (Table 3.8).

For YLD, two QTL, qYLD-4B-1 (coincident with QTL for PLH and TKW) and qYLD-2D-1 (coincident with QTL for DFL) were identified on chromosomes 4B and

2D (Table 3.8 and Figure S3.10). For the qYLD-4B-1, additive effect of BWP allele was -0.262, -0.401 and -0.269 t/ha on YLD with R^2 of 1.62, 3.44 and 1.73% in Y11.12, Y13.14 and over years, respectively (Table 3.8). The SYNP allele of this QTL increased PLH. For the qYLD-2D-1, BWP allele had additive effect of -0.240 and -0.204 t/ha on YLD with R^2 of 2.10 and 1.51% in Y11.12 and over years. The SYNP allele the qYLD-2D-1 increased DHE (Table 3.8).

Table 3.8: Summary of GWAS results for different traits under irrigated condition (IRRI) for each year and over years.

Trait	SNP NO.	SNP	Chr.	Chr. pos. (cM)	QTL	Y11.12			Y12.13			Y13.14			Y11.14 (over years)			MAF
						P value	Add.* Effect	R^2 %	P value	Add. Effect	R^2 %	P value	Add. Effect	R^2 %	P value	Add. Effect	R^2 %	
DFL	2256857	C/T	4A	32	qDFL-4A-1	5.21E-06	-0.7	1.24										0.31
DFL	1141498	T/C	5A	59	qDFL-5A-1				7.94E-08	2.07	4.91							0.24
DFL	2244579	A/G	2D	112	qDFL-2D-1										2.37E-06	1.01	1.32	0.12
DHE	1141498	T/C	5A	59	qDHE-5A-1				1.07E-07	2.05	4.87				6.27E-07	1.4	1.80	0.24
PLH	1088389	A/G	4B	33	qPLH-4B-1	5.15E-14	4.2	3.34				2.63E-10	2.3	4.70	1.59E-14	3.8	3.52	0.21
TKW	3064689	T/C	1B	NA	qTKW-1B-1	4.72E-08	2.20	3.98										0.23
TKW	1208575	T/C	4B	33	qTKW-4B-1	4.34E-08	1.79	4.54										0.28
TKW	1002527	G/A	6B	NA	qTKW-6B-1	2.22E-06	1.76	3.96										0.32
TKW	2251719	A/G	2D	NA	qTKW-2D-1	5.41E-07	1.42	3.56										0.21
TKW	2259110	C/A	3D	NA	qTKW-3D-1	9.51E-09	2.24	4.90										0.35
TKW	3064692	C/G	7A	NA	qTKW-7A-1	3.11E-07	1.79	3.37										0.17
SN/m^2	2259110	C/A	3D	NA	qSN-3D-1	7.84E-07	-686	3.52										0.35
YLD	1216917	G/T	4B	33	qYLD-4B-1	2.93E-07	-0.262	1.62				4.35E-08	-0.401	3.44	1.49E-07	-0.269	1.73	0.21
YLD	2242893	C/T	2D	112	qYLD-2D-1	2.97E-09	-0.240	2.10							4.09E-07	-0.204	1.51	0.16

SNP: single nucleotide polymorphism, Chr.pos: chromosome position, Add. Effect: additive effect, MAF: minor allele frequency, R^2: phenotypic variance of trait explained by QTL, DHE: days to heading, DFL: days to flowering, DMA: days to maturity, PLH: plant height, TKW: thousand kernel weight, SN/m^2: seed number per square meter, YLD: grain yield.
*: "-":Add. Effects indicates increasing trait value by the BWP allele and "+" indicates increasing trait value by the SYNP allele.
Units for Add. Effect for DFL, DHE, and DMA is day, for PLH is cm for TKW is gr and for YLD is t/ha.

QTL by environment interaction

GWAS results indicated that there were quantitative trait loci (QTL)-by-environment interactions (QEI) for PLH and YLD QTL. Here QEI was split in two terms; marker by management interaction (DRO, HEAT and IRRI) (MMI) and marker by year interaction (MYI). For YLD, the qYLD-4B-1 had significant positive MMI effect under DRO and HEAT and its SYNP allele increased YLD 0.130 and 0.094 t/ha , respectively, and under IRRI its BWP allele increased YLD 0.225 t/ha (Table 3.9). However, the main effect of the qYLD-4B-1 was not significant under DRO and HEAT stresses. The qYLD-6D-1 showed significant MMI effect only under HEAT stress and its SYNP allele increased YLD 0.055 t/ha (Table 3.9). For qYLD-2D-1, significant MYI was observed for YLD in Y11.12 and Y13.14 and the BWP allele of this QTL increased YLD 0.084 t/ha in Y11.12 while its SYNP allele increased YLD by 0.112 t/ha in Y13.14 (Table 3.10). The qYLD-6D-1 had significant MYI for YLD in Y12.13 and its BWP allele increased YLD 0.043 t/ha in this year (Table 3.10).

Table 3.9: Marker by management interaction (MMI) effect under DRO, HEAT, and IRRI environments for YLD.

		Marker x Management Interaction (MMI)					
SNP	QTL	DRO x SNP	*P* value	HEAT x SNP	*P* value	IRRI x SNP	*P* value
1216917 G/T	qYLD-4B-1	0.130	**6.86E-09**	0.094	**4.53E-03**	-0.225	**2.31E-25**
1067078 C/T	qYLD-6D-1	-0.027	8.88E-02	0.055	**1.18E-04**	-0.028	7.62E-02

Bonferroni alpha = 0.0083333 (Boldfaced are significant). DRO: drought stress, HEAT: heat stress, and IRRI: irrigated environment.

For PLH, the qPLH-4B-1 showed significant MYI and its SYNP allele increased PLH 1.4 cm in Y11.12 while its BWP allele increased PLH 0.8 and 0.6 cm in Y12.13 and Y13.14. respectively (Table 3.10).

Table 3.10: Marker by year interaction effect from 2011 to 2014 for YLD and PLH.

		Marker x Year Interaction (MYI)					
SNP	QTL	Y11.12 x SNP	*P* value	Y12.13 x SNP	*P* value	Y13.14 x SNP	*P* value
2242893 C/T	qYLD-2D-1	-0.084	**1.49E-03**	-0.028	1	0.112	**5.25E-07**
1067078 C/T	qYLD-6D-1	0.012	1	-0.043	**4.78E-04**	0.031	2.95E-02
983836 T/C	qPLH-4B-1	1.4	**1.14E-08**	-0.8	**9.11E-03**	-0.6	**1.56E-02**

Bonferroni alpha = 0.0166667 (Boldfaced are significant). Y11.12: year 2011-12, Y12.13: year 2012-13, and Y13.14: year 2013-14.

Discussion

The objectives of this study were to characterize yield and phenological trait responses to irrigated and stress environments and the contributions of SYNP alleles to these responses. In this study, SDL populations and BWPs displayed a wide range of phenotypic diversity for traits measured under DRO, HEAT, and IRRI conditions (Table 3.3). The SYNP alleles were more frequently increased BWPs YLD under HEAT (96 superior SDLs) and DRO (37 superior SDLs) stresses (Table S3.1) and superior SDLs had 29 to 99% higher YLD than their recurrent BWP under DRO and 29 to 144% under HEAT stress. Becker (2014) evaluated 90 SDLs selected from crossing three synthetic hexaploid wheats (SHW) to "Hatcher" as a winter wheat in three IRRI and DRO environments in Colorado in year 2012-13 and reported that four SDLs outranked Hatcher for YLD in all three environments. The highest yielding SDL reached 115% of the recurrent BWP. In this study, under IRRI, also 23 SDLs significantly out yielded their respective BWPs (Table S3.1) and superior SDLs had 13 to 43% higher YLD than BWP. Del Blanco et al. (2001) reported that eight lines out of 282 BC_2F_2 SDLs (2.84%) had significantly higher grain yield than their respective BWP and superior SDLs had up to 11% higher YLD than their recurrent BWP under irrigated conditions. Jafarzadeh

et al. (2016) reported earlier that SYNPs contributed to increased YLD, TKW and phenological traits specifically under drought and heat stresses. Also, SYNPs increased genetic diversity of SDLs in A, B and D genomes relative to BWPs.

The correlation of YLD with different traits differed across environments. In IRRI, correlations of YLD were positive with DHE, DFL, and DMA and negative with GFD (Table 3.5), while, in DRO and HEAT, these correlations were all negative except for GFD in HEAT (Table 3.4). The positive correlation between YLD and DHE was in disagreement with Pranger (2012) who reported a negative correlation (r = - 0.20) under IRRI. The correlation of YLD with PLH was also variable in different environments. The correlation of PLH was positive with YLD in DRO and HEAT, while in IRRI there was no correlation with YLD (Tables 3.4, and 3.5). The positive correlation of YLD with PLH (r = 0.25) was reported by Pranger (2012) under moderate moisture stress. It was expected that early maturity and tall plants would have advantages under stress conditions (DRO and HEAT) due to avoiding or minimizing exposure to stress under drought and high temperatures by storing more assimilates in stems for remobilization during grain filling. A similar relationship was suggested by (Richards, 1992) who stated that grain yield depends on an optimum PLH and by Butler et al. (2005) who found that shorter plants had significantly lower YLD in the irrigated and drought stress conditions while tall plants had the highest YLD under drought stress.

Under IRRI, a negative correlation between YLD and TKW suggested that phenotypic variation of YLD was not explained by TKW. In contrast, a positive

correlation between SN/m^2 and YLD indicated that phenotypic variation of YLD was associated with SN/m^2 (Table 3.5). Similar results were reported by Becker (2014), Pranger (2012), Huang et al. (2003) and Villareal et al. (1994a). In this environment, PLH had a positive correlation with TKW and a negative correlation with SN/m^2 in agreement with Börner et al. (2002), Huang et al. (2003) Zhang et al. (2013) and Gao et al. (2015) and in disagreement with Pranger (2012) who reported negative correlation between PLH and TKW (r= -0.28) in IRRI. Also, the correlation between TKW and SN/m^2 was negative as expected because yield components tend to compensate (Table 3.5). In this study, SDL populations showed that YLD was associated with higher SN/m^2, reduced PLH (Xiao et al. 2012), longer growing season, shorter GFD, and lower TKW under IRRI conditions. TKW and SN/m^2 were not recorded under DRO and HEAT stresses.

For TKW, Calderini and Reynolds (2000) reported that synthetic wheats were a valuable source of genetic variability for TKW compared to cultivated bread wheat and their TKWs were as high as 67 gr. In this study, TKW values ranged from 40 to 65 gr for SDL populations and 41 to 54 gr for recurrent BWPs. Our findings indicated that 29% of SDLs had significantly (FDR $P < 0.05$) higher TKW than their recurrent BWPs under IRRI conditions supporting the important contribution of SYNPs to TKW. Also, Del Blanco et al. (2001) reported that 83% of SDLs were significantly superior to their recurrent BWPs for TKW. In this study, the SYNP allele of six QTL for TKW on chromosomes 1B, 4B, 6B, 2D, 3D and 7A had positive additive effects on TKW and increased TKW values from 1.42 to 2.24 gr with range of R^2 from 3.37 to 4.90%.

However, TKW on chromosome 3D overlapped with a QTL for SN/m^2, qSN-3D-1, and its BWP allele increased the SN/m^2 value (Table 3.8). SDL populations also had a range of 0.13 to 0.25 frequencies for the SYNP alleles of this QTL that confirming SYNPs were able to increase SDL's TKW (Table 3.8). In regards to the importance of SYNPs for TKW, Yu et al. (2014) using an SDL population of SHW-L1 x Chuanmai32, identified seven QTL for TKW for which SYNP alleles for QTL on chromosomes 7D, 5A, and 1B increased TKW (2.67 to 3.15 gr, with R^2 = 6.84-11.19%) while SYNP alleles for QTL on 1B, 2D, 5A decreased this trait (-5.16 to -3.10 gr, R^2 = 9.18-26.35%). Also, Rasheed et al. (2014b) identified a QTL for TKW on chromosome 3D (160 cM) using 231 synthetic hexaploid wheat and DArT markers with R^2 = 6%. Cumulatively, these six TKW QTL explained 11.20% of phenotypic variation in IRRI conditions. across Our QTL for TKW needs to be verified under DRO and HEAT stresses.

PLH is trait that affects plant lodging, harvest index, and grain yield. In wheat, Rht-B1, and Rht-D1, located on chromosomes 4BS and 4DS, respectively, are the major dwarfing genes that reduce PLH and affect grain number and yield (Cadalen et al., 1998; Zanke et al., 2014). Although, SDL populations were selected for PLH resembling BWPs, our result indicated that more SDLs were taller than their recurrent BWPs under DRO (31.64%) and HEAT (27.37%) than IRRI (1.60%) (Table S3.2) confirming the contribution of SYNPs for PLH under stress conditions. In this study a QTL for PLH, qPLH-3B-1, identified only in HEAT stress and over years and its BWP allele increased PLH (Table 3.6). The second QTL, qPLH-4B-1, mapped on chromosome 4B and was detected in three out of eight environments and it is near to *Rht-B1* according to a genetic

composite map 2004 (http://www.gramene.org). The SYNP allele of this QTL had positive additive effect and increased PLH. However, its R^2 and additive effects on PLH differed in different environments (Tables 3.7 and 3.8) and it showed significant MYI with the highest SYNP allele additive effect in Y11.12 followed by the BWP allele effect in Y12.13 and Y13.14 (Table 3.10). This QTL is very likely to be the same minor QTL for PLH (QHt.ipk-4B) that was reported by Börner et al. (2002), which was in a comparable position on chromosome arm 4BL in the ITMI population. This QTL also overlapped with a QTL for YLD under IRRI conditions and the SNYP allele decreased YLD. The qPLH-4B-1 also overlapped with a QTL for TKW, qTKW-4B-1, under IRRI and SYNP allele increased TKW. PLH had a positive correlation with TKW confirming the importance of tall plants for TKW. Furthermore, Börner et al. (2002), using 114 RILS of the ITMI population, mapped four major QTL on chromosomes arms 1AS, 2DS, 4AL, and 6AS for PLH. The positive correlation of PLH with YLD in DRO, HEAT and with TKW in IRRI, and overlapping PLH QTL, qPLH-4B-1, with a QTL for YLD and TKW, suggesting the usefulness of this QTL for stress conditions.

The complexity of grain yield as a quantitative trait and its high GEI effect complicates the analysis. In this study four QTL were identified on chromosomes 2D, 4D, 6D, and 4B for yield across all environments (managements and years). Three of these QTL showed MMI and MYI. The qYLD-4B-1, which was detected in Y11.12, Y13.14 and over years under IRRI (Table 3.8) and Y11.12 and over years under DRO, only indicated the MMI effect and it had a positive MMI effect. The SYNP allele increased YLD under HEAT and DRO stresses while its BWP allele had negative MMI

effect and increased YLD under IRRI condition (Table 3.9). However, it should be mentioned that the main additive effect of the qYLD-4B-1 was not significant under HEAT and DRO stresses. The YLD-4B-1 QTL interval was coincident with QTL for PLH, TKW, and YLD and genetic correlations between these traits were present in IRRI, but these correlations disappeared in the rest of the environments. The genetic correlations among traits could be due to linked or pleiotropic QTLs and could be consistence or inconsistent across environments (Malosetti et al. 2008). Also, the direction of the additive effect of the qYLD-4B-1 was different for these three traits, such that the SYNPs allele increased PLH and TKW but decreased YLD. Since there was a positive correlation between PLH and YLD in DRO and HEAT, it can be concluded that positive interaction of the SYNP allele of this QTL with DRO and HEAT for YLD most likely was because of increased PLH. The qYLD-6D-1 showed both MMI and MYI and its SYNP allele increased YLD under HEAT stress while its BWP allele increased YLD in Y12.13 (Tables 3.9 and 3.10). The qYLD-2D-1 identified across three managements (Tables 3.6,3. 3.7, and 3.8) showed only MYI and its BWP allele increased YLD in Y11.12 and Y12.13 while its SYNP allele increased YLD in Y13.14 (Table 3.10). The qYLD-2D-1 overlapped with a QTL for DFL over years under IRRI and its SYNP allele increased DFL (Table 3.8). This QTL most likely was different than the QTL that was reported by Huang et al. (2003) on chromosome 2D in the BC2F2 population derived from W-7984 (SYNP) x Prinz (BWP) for which the SYNP allele increased YLD by 3%. Also, Huang et al. (2003) identified other QTL for YLD on chromosomes 1B, 2A and 5B where the W-7984 allele increased YLD by 5%, 15%, and 14.5%, respectively. The qYLD-4D-1 was identified over years in DRO, Y12.13,

Y13.14 and over years in HEAT and its BWP allele increased YLD. Pranger (2012) evaluated 188 BC_2F_2-derived lines of Ankor x Sokoll (winter wheat x synthetic-derived spring wheat) under IRRI and moderate moisture stress and detected two YLD QTL, one on chromosome 7A for which the SYNP allele (Sokoll) had 293.1 and 233.9 kg/ha additive effect with R^2 = 13.7 and 12.9% in DRO stress in two years and another on chromosome 3A for which the SYNP allele had 231.9 kg/ha additive effect and R^2 of 9.8%. In summary, QTL identified for YLD were minor and their R^2 ranged from 1.51 to 3.70%. The highest R^2 was for qYLD-6D-1 in HEAT and originated from SYNPs. Cumulatively, these four YLD QTL explained 10.34% of phenotypic variation across all environments (years and managements). They also showed QEI as was expected for trials involving contrasting environmental conditions.

Maturity traits (DHE, DFL, and DMA) are important phenological traits in wheat affecting environment adaptability, grain yield and quality (Zhou et al., 2016) and are mainly controlled by three major groups of genes: vernalization response genes (*vrn-A1, vrn-B1 and vrn-D1*) on homoeologous group 5 (Trevaskis et al., 2003) and photoperiod response genes (*Ppd-A1, Ppd-B1 and Ppd-D1*) on homoeologous group 2 (Beales et al., 2007; Distelfeld et al., 2009), and *"earliness per se" (Eps)* genes on homoeologous group 2 and 4, 7B, 6B and 3A (Bullrich et al., 2002; Zhou et al., 2016). In this study, QTL for maturity traits were minor and mapped on chromosomes 4A, 5A, 4B, 5B and 2D. The qDHE-5A-1 (which was associated with DFL and DMA) with a higher additive effect and R^2 across three managements (HEAT, DRO and IRRI) was more import than the other detected QTL and was located near *Vrn-A1* according to

composite genetic map 2004 (http://www.gramene.org). The SYNP allele of this QTL increased DHE, DMA, and DFL (Tables 3.6, 3.7 and 3.8). However, expression of this QTL and its effect size varied across years and managements. This QTL was likely to be the same QTL reported by Yu et al. (2014) which was in a comparable position on chromosome 5A in a SDL population with an additive effect of 4.09 and R^2 = 8.68% for DHE. The second QTL was qDFL-4A-1 on chromosome 4A (associated with DHE as well) in IRRI and DRO and its BWP allele increased DHE and DFL. This QTL was different from flowering time QTL on chromosome arm 4AL (Araki et al. 1999; http://www.gramene.org) and a minor QTL for flowering time on chromosome arm 4AL (Börner et al. ,2002). The third QTL for DFL, qDFL-2D-1 was detected only in IRRI over years and its SYNP allele increased DFL (Table 3.8). This QTL was not located near *Ppd-D1* (Liu et al. 2014) or near QFLt.ipk-2D (QTL for flowering time) (Börner et al. ,2002). its SYNP allele increased DFL (Table 3.8). The SYNP allele of qDMA-5B on long arm chromosome 5B increased DMA in DRO.Y13.14 and probably related to *Vrn-B1* on chromosome 5BL (Leonova et al. 2003). However, the genetic position of qDMA-5B was not clear. These results indicate SYNP alleles increased maturity traits across all environments specifically under DRO and HEAT (Table S3.3). In the current study, we did not detect major QTL for maturity traits, possibly because the SDL populations were selected for a narrow range of phenology to facilitate measurement of other traits affected by maturity. In this regard, Zhou et al. (2016) stated that some QTL associated with DFL cannot easily be identified, even when two parents with large differences in DFL are used to develop the mapping population and could be

due to genetic interactions between genes on different chromosomes, which are common in wheat.

Conclusion

In this study SYNPs increased grain yield and plant height under stress conditions and thousand kernel weight under irrigated environments. They had QTL for agronomic and phenological traits, which were minor with coincident QTL and QEI. However, these specific SDL populations were selected for maturity traits, and PLH approximating that of the BWPs. Also, the majority of these populations were BC1 derived lines. Therefore, it is possible that many of the SYNP alleles were discarded or retained in a lower frequency which reduced the power to detect their effects or association with traits of interest. This study identified SDLs with high performance for yield under stress conditions and high TKW and can be used in breeding programs to improve common wheat cultivars and elite lines Therefore, synthetic hexaploid wheats are a valuable source for agronomic and phenological traits and genetic diversity that should be used in breeding programs while simultaneously selecting against undesirable traits.

Acknowledgments

This project was funded by the Monsanto's Beachell-Borlaug International Scholars Program which supports Ph.D. students from developing countries who are working on breeding of wheat and rice. This project was conducted in the Global Wheat Program (GWP) of the International Maize and Wheat Improvement Center (CIMMYT), Mexico that developed the

experimental populations and supported two years of field trials. Funding to conduct the first year of phenotypic experiments and to generate the sequencing data used in this project was provided by “Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA Mexico)” through the MasAgro Biodiversidad/Seeds of Discovery Project-CIMMYT. This research was supported in part by USDA-NIFA-AFRI grants, award numbers 2009-65300-05661, 2011-68002-30029, and 2005-05130, and by Hatch project 149-449.

References

References

Ain, Q.-U., A. Rasheed, A. Anwar, T. Mahmood, M. Imtiaz, X. Xia, Z. He, and U.M. Quraishi. 2015. Genome-wide association for grain yield under rainfed conditions in historical wheat cultivars from Pakistan. Front. Plant Sci. 6: 743. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4585131&tool=pmcentrez&rendertype=abstract (verified 15 July 2016).

Akdemir, D., O.U. Godfrey .2015. EMMREML: Fitting Mixed Models with Known Covariance Structures. R package version 3.1. https://CRAN.R-project.org/package=EMMREML.

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Softw. 67(1). http://www.jstatsoft.org/v67/i01/.

Beales, J., A. Turner, S. Griffiths, J.W. Snape, and D.A. Laurie. 2007. A Pseudo-Response Regulator is misexpressed in the photoperiod insensitive Ppd-D1a mutant of wheat (*Triticum aestivum* L.). Theor. Appl. Genet. 115(5): 721–733. http://www.ncbi.nlm.nih.gov/pubmed/17634915.

Becker, S.R. 2014. Exploiting Drought Tolerance Traits and Genetic Diversity of Synthetic Hexaploid Wheat in Winter Wheat Breeding (Doctoral dissertation, Colorado State University).

Bickel, P.J., Y. Ritov, and A.B. Tsybakov. 2008. Simultaneous analysis of Lasso and Dantzig selector. BMC Bioinformatics 9(4): 114. http://dx.doi.org/10.1214/08-AOS620.

Del Blanco, I.A.A., S. Rajaram, and W.E.E. Kronstad. 2001. Agronomic potential of

synthetic hexaploid wheat-derived populations. Crop Sci. 41(3): 670–676. http://www.mendeley.com/research/agronomic-potential-synthetic-hexaploid-wheatderived-populations/ (verified 7 January 2015).

Börner, A., E. Schumann, A. Fürste, H. Cöster, B. Leithold, M.S. Röder, and W.E. Weber. 2002. Mapping of quantitative trait loci determining agronomic important characters in hexaploid wheat (*Triticum aestivum* L.). Theor. Appl. Genet. 105(6–7): 921–936.

Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, and E.S. Buckler. 2007. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23(19): 2633–5. http://bioinformatics.oxfordjournals.org/content/23/19/2633.full.pdf+html (verified 22 May 2016).

Breseghello, F., and M.E. Sorrells. 2006. Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. Genetics 172(2): 1165–77. http://www.genetics.org/content/172/2/1165 (verified 7 September 2016).

Butler, J.D., P.F. Byrne, V. Mohammadi, P.L. Chapman, and S.D. Haley. 2005. Agronomic Performance of Alleles in a Spring Wheat Population across a Range of Moisture Levels. Crop Sci. 45(3): 939. https://www.crops.org/publications/cs/abstracts/45/3/0939.

Cadalen, T., P. Sourdille, G. Charmet, M.H. Tixier, G. Gay, C. Boeuf, S. Bernard, P. Leroy, and M. Bernard. 1998. Molecular markers linked to genes affecting plant height in wheat using a doubled-haploid population. TAG Theor. Appl. Genet. 96(6–7): 933–940. http://link.springer.com/10.1007/s001220050823.

Calderini, D.F., and M.P. Reynolds. 2000. Changes in grain weight as a consequence of de-graining treatments at pre- and post-anthesis in synthetic hexaploid lines of wheat (*Triticum durum* x *T. tauschii*). Funct. Plant Biol. 27(3): 183-191. https://doi.org/10.1071/PP99066.

Cooper, J.K., A.H. Ibrahim, J. Rudd, S. Malla, D.B. Hays, and J. Baker. 2012. Increasing hard winter wheat yield potential via synthetic wheat: I. path-coefficient analysis of yield and its components. Crop Sci. 52(5): 2014–2022. https://www.crops.org/publications/cs/abstracts/52/5/2014

Cossani, C.M., and M.P. Reynolds. 2015. Heat Stress Adaptation in Elite Lines Derived from Synthetic Hexaploid Wheat. Crop Sci. 55(6): 2719. https://dl.sciencesocieties.org/publications/cs/abstracts/55/6/2719.

Curtis, T., and N.G. Halford. 2014. Food security: The challenge of increasing wheat

yield and the importance of not compromising food safety. Ann. Appl. Biol. 164(3): 354–372. doi: 10.2135/cropsci2015.02.0092.

Distelfeld, A., C. Li, and J. Dubcovsky. 2009. Regulation of flowering in temperate cereals. Curr. Opin. Plant Biol. 12(2): 178–184. http://linkinghub.elsevier.com/retrieve/pii/S1369526608002227.

Douglas B., M. Maechler, B. Bolker, and S.W. 2015. Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Softw. 67(1): 1–48. http://www.jstatsoft.org/v67/i01/.

Dreisigacker, S., M. Kishii, J. Lage, and M. Warburton. 2008. Use of synthetic hexaploid wheat to increase diversity for CIMMYT bread wheat improvement. Aust. J. Agric. Res. 59(5): 413–420. http://www.publish.csiro.au/?paper=AR07225.

Dreisigacker, S., R. Tiwari, and S. Sheoran. 2013. ICAR-CIMMYT molecular breeding course in wheat Directorate of Wheat Research. ICAR/BMZ: p 36. http://repository.cimmyt.org/xmlui/handle/10883/3221#.

Endelman, J.B., and J.-L. Jannink. 2013. Shrinkage Estimation of the Realized Relationship Matrix. G3: Genes, Genomes, Genetics 2(11): 1405–1413. http://g3journal.org/cgi/doi/10.1534/g3.112.004259.

Gao, F., W. Wen, J. Liu, A. Rasheed, G. Yin, X. Xia, X. Wu, and Z. He. 2015. Genome-Wide Linkage Mapping of QTL for Yield Components, Plant Height and Yield-Related Physiological Traits in the Chinese Wheat Cross Zhou 8425B/Chinese Spring. Front. Plant Sci. 6(December): 1099. http://www.ncbi.nlm.nih.gov/pubmed/26734019%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4683206.

Huang, X.Q., H. Coster, M.W. Ganal, and M.S. Roder. 2003. Advanced backcross QTL analysis for the identification of quantitative trait loci alleles from wild relatives of wheat (*Triticum aestivum* L.). Theor Appl Genet 106(8): 1379–1389. http://www.ncbi.nlm.nih.gov/pubmed/12750781.

Jafarzadeh, J., D. Bonnett, J. Jannink, D. Akdemir, S. Dreisigacker, and M.E. Sorrells. 2016. Breeding Value of Primary Synthetic Wheat Genotypes for Grain Yield (SK Parida, Ed.). PLoS One 11(9): e0162860. http://dx.plos.org/10.1371/journal.pone.0162860.

Jighly, A., M. Alagu, F. Makdis, M. Singh, S. Singh, L.C. Emebiri, and F.C. Ogbonnaya. 2016. Genomic regions conferring resistance to multiple fungal pathogens in synthetic hexaploid wheat. Mol. Breed. 36(9): 127.

http://link.springer.com/10.1007/s11032-016-0541-4 (verified 6 September 2016).

Kema, G.H.J., W. Lange, and C.H. Vansilfhout. 1995. Differential Suppression of Stripe Rust Resistance in Synthetic Wheat Hexaploids Derived from *Triticum turgidum* subsp *dicoccoides* and *Aegilops squarrosa*. Phytopathology 85(Iss 4): 425–429 oi: 10.1094/Phyto-85-425.

L., B., A. M., T. G., L. S., and D. J. 2002. Mapping of a thermo-sensitive earliness per se gene on *Triticum monococcum* chromosome 1A m. TAG Theor. Appl. Genet. 105(4): 585–593. http://link.springer.com/10.1007/s00122-002-0982-5.

Li, H., P. Bradbury, E. Ersoz, E.S. Buckler, and J. Wang. 2011a. Joint QTL Linkage Mapping for Multiple-Cross Mating Design Sharing One Common Parent (D Kliebenstein, Ed.). PLoS One 6(3): e17573. http://dx.plos.org/10.1371/journal.pone.0017573.

Li, J., H.T. Wei, X.R. Hu, C.S. Li, Y.L. Tang, D.C. Liu, and W.Y. Yang. 2011b. Identification of a High-Yield Introgression Locus in Chuanmai 42 Inherited from Synthetic Hexaploid Wheat. Acta Agron. Sin. 37(2): 255–262. http://dx.doi.org/10.1016/S1875-2780(11)60007-2.

Lopes, M.S., and M.P. Reynolds. 2011. Drought adaptive traits and wide adaptation in elite lines derived from resynthesized hexaploid wheat. Crop Sci. 51(4): 1617–1626. https://www.crops.org/publications/cs/abstracts/51/4/1617 (verified 4 June 2013).

Malosetti, M., J.M. Ribaut, M. Vargas, J. Crossa, and F.A. van Eeuwijk. 2008. A multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (*Zea mays* L.). Euphytica 161(1–2): 241–257. http://link.springer.com/10.1007/s10681-007-9594-0.

Meng, L., H. Li, L. Zhang, and J. Wang. 2015. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. Crop J. 3(3): 269–283. http://www.sciencedirect.com/science/article/pii/S2214514115000161 (verified 4 September 2016).

Mujeeb-Kazi, A., V. Rosas, and S. Roldan. 1996. Conservation of the genetic variation of *Triticum tauschii* (Coss.) Schmalh. (*Aegilops squarrosa auct*. non L.) in synthetic hexaploid wheats (*T. turgidum* L. s.lat. x *T. tauschii*; 2n=6x=42, AABBDD) and its potential utilization for wheat improvement. Genet. Resour. Crop Evol. 43(2): 129–134. http://www.springerlink.com/index/10.1007/BF00126756.

Mulki, M.A., A. Jighly, G. Ye, L.C. Emebiri, D. Moody, O. Ansari, and F.C. Ogbonnaya. 2013. Association mapping for soilborne pathogen resistance in synthetic hexaploid wheat. Mol. Breed. 31(2): 299–311. doi:10.1007/s11032-012-9790-z.

Neumann, K., B. Kobiljski, S. Denčić, R.K. Varshney, and A. Börner. 2011. Genome-wide association mapping: A case study in bread wheat (*Triticum aestivum* L.). Mol. Breed. 27(1): 37–58. doi:10.1007/s11032-010-9411-7.

Ogbonnaya, F.C., G. Ye, R. Trethowan, F. Dreccer, D. Lush, J. Shepperd, and M. van Ginkel. 2007. Yield of synthetic backcross-derived lines in rainfed environments of Australia. Euphytica 157(3): 321–336. http://link.springer.com/10.1007/s10681-007-9381-y (verified 22 September 2014).

Pask A.J.D., J. Pietragalla, D.M. Mullan, M.P. Reynolds and M. Reynolds (Ed.). 2012. Physiological breeding II: A field guide to wheat phenotyping. Mexic, D.F.:CIMMYT. http://libcatalog.cimmyt.org/download/cim/96144.pdf.

Payne R.W., D.A. Murray, S.A. Harding, D.B.S.D. Baird .2009. GenStat for Windows (12th Edition) Introduction.

Piepho, H.P., J. Möhring, A.E. Melchinger, and A. Büchse. 2007. BLUP for phenotypic selection in plant breeding and variety testing. Euphytica 161(1–2): 209–228. http://link.springer.com/10.1007/s10681-007-9449-8 (verified 10 July 2014).

Pranger, A.L. 2012. Synthetic Hexaploid Wheat Contributes Favorable Alleles for Yield and Yield Components in an Advanced Backcross Winter Wheat Population. http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&res_dat=xri:pqdiss&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&rft_dat=xri:pqdiss:1511047.

R Core Team. 2016. R: A language and environment for statistical computing. https://www.r-project.org/.

Rasheed, A., X. Xia, F. Ogbonnaya, T. Mahmood, Z. Zhang, A. Mujeeb-Kazi, and Z. He. 2014a. Genome-wide association for grain morphology in synthetic hexaploid wheats using digital imaging analysis. BMC Plant Biol. 14(1): 128. http://bmcplantbiol.biomedcentral.com/articles/10.1186/1471-2229-14-128.

Rasheed, A., X. Xia, F. Ogbonnaya, T. Mahmood, Z. Zhang, A. Mujeeb-Kazi, and Z. He. 2014b. Genome-wide association for grain morphology in synthetic hexaploid wheats using digital imaging analysis. BMC Plant Biol. 14(1): 128.

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4057600&tool=pmcentrez&rendertype=abstract (verified 4 September 2014).

Richards, R.A. 1992. The effect of dwarfing genes in spring wheat in dry environments. I. Agronomic characteristics. Aust. J. Agric. Res. 43(3): 517–527 http://www.publish.csiro.au/cp/AR9920517.

Saghai-Maroof, M.A., K.M. Soliman, R.A. Jorgensen, and R.W. Allard. 1984. Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. Proc. Natl. Acad. Sci. 81(24): 8014–8018. http://www.pnas.org/content/81/24/8014.abstract (verified 16 April 2016).

Sehgal, D., P. Vikram, C.P. Sansaloni, C. Ortiz, C. Saint Pierre, T. Payne, M. Ellis, A. Amri, C.D. Petroli, P. Wenzl, and S. Singh. 2015. Exploring and mobilizing the gene bank biodiversity for wheat improvement. PLoS One 10(7): e0132112. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0132112 (verified 13 April 2016).

Sharma, P., S. Sareen, M. Saini, A. Verma, B.S. Tyagi, and I. Sharma. 2014. Assessing genetic variation for heat tolerance in synthetic wheat lines using phenotypic data and molecular markers. 8(4): 515–522. http://www.cropj.com/sharma_8_4_2014_515_522.pdf.

Simón, M.R., A.J. Worland, and P.C. Struik. 2005. Chromosomal location of genes encoding for resistance to *septoria tritici* blotch (*Mycosphaerella graminicola*) in substitution lines of wheat. NJAS - Wageningen J. Life Sci. 53(2): 113–129. http://linkinghub.elsevier.com/retrieve/pii/S1573521405800016.

Singh, D., and K. Singh. 2015. Marker-Assisted Plant Breeding : Principles and Practices. New Delhi, India: Springer.

Timm, N.H. 2002. Applied Multivariate Analysis. 2nd ed. Springer-Verlag, New York.

Trevaskis, B., D.J. Bagnall, M.H. Ellis, W.J. Peacock, and E.S. Dennis. 2003. MADS box genes control vernalization-induced flowering in cereals. Pnas 100(22): 13099–104. http://www.pnas.org/content/100/22/13099.short.

Villareal, R.L., A. Mujeeb-Kazi, G. Fuentes-Davila, S. Rajaram, and E. Del Toro. 1994. Resistance to karnal bunt (*Tilletia indica* Mitra) in synthetic hexaploid wheats derived from *Triticum turgidum* x *T. tauschii*. Plant Breed. 112: 63–69.

Wu, Q.H., Y.X. Chen, S.H. Zhou, L. Fu, J.J. Chen, et al. 2015. High-density genetic

linkage map construction and QTL mapping of grain shape and size in the wheat population Yanda1817 x Beinong6. PLoS One 10(2): e0118144. https://doi.org/10.1371/journal.pone.0118144.

Xiao, Y.G., Z.G. Qian, K. Wu, J.J. Liu, X.C. Xia, W.Q. Ji, and Z.H. He. 2012. Genetic gains in grain yield and physiological traits of winter wheat in Shandong province, China, from 1969 to 2006. Crop Sci. 52(1): 44–56. doi:10.2135/cropsci2011.05.0246.

Yang, W., D. Liu, J. Li, L. Zhang, H. Wei, X. Hu, Y. Zheng, Z. He, and Y. Zou. 2009. Synthetic hexaploid wheat and its utilization for wheat genetic improvement in China. J. Genet. Genomics 36(9): 539–546. http://www.ncbi.nlm.nih.gov/pubmed/19782955.

Yu, M., G. yue Chen, L. quan Zhang, Y. xi Liu, D. cai Liu, J. rui Wang, Z. en Pu, L. Zhang, X. jin Lan, Y. ming Wei, C. ji Liu, and Y. liang Zheng. 2014. QTL Mapping for Important Agronomic Traits in Synthetic Hexaploid Wheat Derived from *Aegiliops tauschii* ssp. *tauschii*. J. Integr. Agric. 13(9): 1835–1844. https://doi.org/10.1016/S2095-3119(13)60655-3.

Zanke, C.D., J. Ling, J. Plieske, S. Kollers, E. Ebmeyer, V. Korzun, O. Argillier, G. Stiewe, M. Hinze, K. Neumann, M.W. Ganal, and M.S. Röder. 2014. Whole genome association mapping of plant height in winter wheat (*Triticum aestivum* L). PLoS One 9(11): e113287. https://doi.org/10.1371/journal.pone.0113287.

Zhang, J., B. Dell, B. Biddulph, F. Drake-Brockman, E. Walker, N. Khan, D. Wong, M. Hayden, and R. Appels. 2013. Wild-type alleles of *Rht-B1* and *Rht-D1* as independent determinants of thousand-grain weight and kernel number per spike in wheat. Mol. Breed. 32(4): 771–783. doi:10.1007/s11032-013-9905-1.

Zhou, W., S. Wu, M. Ding, J. Li, Z. Shi, W. Wei, J. Guo, H. Zhang, Y. Jiang, and J. Rong. 2016. Mapping of *Ppd-B1*, a major candidate gene for late heading on wild emmer chromosome arm 2BS and assessment of its interactions with early heading QTLs on 3AL. PLoS One 11(2): e0147377. https://doi.org/10.1371/journal.pone.0147377.

Supplemental Information

Table S3.1: Number of SDLs over their respective BWP for average YLD from 2011 to 2014 ($P < 0.05$).

Trial	IRRI			DRO			HEAT		
BWP	Ave. YLD (t/ha)	No. SDL	SDL > BWP No. (%)	Ave. YLD (t/ha)	No. SDL	SDL > BWP No. (%)	Ave. YLD (t/ha)	No. SDL	SDL > BWP No. (%)
BWP3570	8.293	17	0 (0.00)	2.635	8	0 (0.00)	0.967	8	0 (0.00)
CACUKE	6.246	217	1 (0.46)	2.041	190	2 (1.05)	1.829	192	0 (0.00)
HS420	7.532	25	0 (0.00)	2.427	24	0 (0.00)			0 (0.00)
KIRITATI	7.262	130	0 (0.00)	2.025	109	0 (0.00)	1.954	110	0 (0.00)
KIRITATI//PRL/2*PASTOR	6.544	29	0 (0.00)	2.867	27	0 (0.00)	1.504	58	0 (0.00)
KIRITATI/2*TRCH	5.486	38	3 (7.89)	2.048	18	1 (5.56)	1.862	16	0 (0.00)
KRL19	6.391	99	0 (0.00)	2.262	75	0 (0.00)	1.830	74	0 (0.00)
MILAN/AMSEL	5.749	30	0 (0.00)	1.846	26	1 (3.85)	1.364	26	2 (7.69)
MILAN/S87230//BAV92	7.368	341	0 (0.00)	2.444	266	0 (0.00)	1.957	265	1 (0.38)
MINO	6.269	35	0 (0.00)	2.351	34	0 (0.00)	1.578	34	0 (0.00)
MUU	6.041	45	3 (6.67)	1.887	37	2 (5.40)	1.229	37	4 (10.81)
PANDORA	6.262	428	2 (0.47)	1.691	306	23 (7.52)	1.184	308	70 (22.73)
PBW502	6.513	89	0 (0.00)	2.074	76	0 (0.00)	1.985	74	0 (0.00)
SUNCO/2*PASTOR	6.142	121	14 (11.57)	2.595	104	0 (0.00)	1.575	103	15 (14.56)
SW89.5181/KAUZ	7.426	103	0 (0.00)	1.636	80	5 (6.25)	1.520	78	0 (0.00)
TAM200/TUI	6.482	254	0 (0.00)	1.651	172	3 (1.74)	1.521	170	4 (2.35)
Total		**2001**	**23 (1.15%)**		**1552**	**37 (2.38%)**		**1553**	**96 (6.18%)**

IRRI: irrigated trial, DRO: drought trial, HEAT: heat trial, SDL: synthetic derived line, BWP: bread wheat parent, Ave. YLD: average grain yield.

Table S3.2: Number of SDLs over their respective BWP for average PLH from 2011 to 2014 ($P < 0.05$).

Trial	IRRI			DRO			HEAT		
BWP	Ave. PLH (cm)	No. SDL	SDL > BWP No. (%)	Ave. PLH (cm)	No. SDL	SDL > BWP No. (%)	Ave. PLH (cm)	No. SDL	SDL > BWP No. (%)
BWP3570	116	17	0 (0.00)	78	8	5 (62.50)	-	-	-
CACUKE	107	217	5 (2.30)	83	190	12 (6.30)	64	130	1 (0.77)
HS420	111	25	0 (0.00)	82	24	0 (0.00)	-	-	-
KIRITATI	105	130	0 (0.00)	75	109	5 (4.60)	69	59	0 (0.00)
KIRITATI//PRL/2*PASTOR	100	29	0 (0.00)	78	27	0 (0.00)	60	30	0 (0.00)
KIRITATI/2*TRCH	89	38	3(7.89)	70	18	12 (66.70)	51	1	0 (0.00)
KRL19	93	99	0 (0.00)	71	75	38 (50.70)	55	41	17 (41.46)
MILAN/AMSEL	101	30	1 (3.33)	72	26	16 (61.50)	60	11	1 (39.09)
MILAN/S87230//BAV92	103	341	0 (0.00)	73	266	91 (34.20)	58	224	69 (30.80)
MINO	103	35	0 (0.00)	82	34	0 (0.00)	65	18	0 (0.00)
MUU	100	45	2 (4.44)	72	37	2 (5.40)	53	36	27 (75)
PANDORA	94	428	11(2.57)	69	306	193 (53.90)	53	187	111 (22.73)
PBW502	105	89	0 (0.00)	72	76	28 (36.80)	69	39	29 (59.35)
SUNCO/2*PASTOR	106	121	10 (8.26)	75	104	37 (35.60)	57	74	15 (20.27)
SW89.5181/KAUZ	100	103	0 (0.00)	72	80	17 (21.30)	54	44	0 (0.00)
TAM200/TUI	99	254	0 (0.00)	72	172	63 (36.6)	59	96	5 (5.21)
Total		**2001**	**32 (1.60%)**		**1552**	**491 (31.64%)**		**990**	**271 (27.37%)**

IRRI: irrigated trial, DRO: drought trial, HEAT: heat trial, SDL: synthetic derived line, BWP: bread wheat parent, Ave. PLH: average plant height.

Table S3.3: Number of SDLs over their respective BWP for average DMA from 2011 to 2014 (P < 0.05).

Trial	IRRI			DRO			HEAT		
BWP	Ave. DMA (day)	No. SDL	SDL > BWP No. (%)	Ave. DMA (day)	No. SDL	SDL > BWP No. (%)	Ave. DMA (day)	No. SDL	SDL > BWP No. (%)
BWP3570	131	17	0 (0.00)	118	8	0 (0.00)	-	-	-
CACUKE	123	217	7 (3.22)	108	190	21 (11.05)	83	130	4 (3.08)
HS420	130	25	0 (0.00)	122	24	0 (0.00)	-	-	-
KIRITATI	127	130	0 (0.00)	110	109	3 (2.75)	86	59	0 (0.00)
KIRITATI//PRL/2*PASTOR	128	29	0 (0.00)	118	27	0 (0.00)	86	30	0 (0.00)
KIRITATI/2*TRCH	120	38	3(7.89)	102	18	14 (77.78)	85	1	0 (0.00)
KRL19	122	99	0 (0.00)	108	75	9 (12.00)	83	41	0 (0.00)
MILAN/AMSEL	130	30	0 (0.00)	110	26	26 (15.38)	84	11	(0.00)
MILAN/S87230//BAV92	125	341	0 (0.00)	108	266	47 (17.67)	84	224	15 (6.70)
MINO	126	35	0 (0.00)	110	34	5 (14.71)	82	18	3 (16.67)
MUU	128	45	2 (4.44)	110	37	0 (0.00)	85	36	5 (13.89)
PANDORA	128	428	3(0.70)	110	306	14 (4.58)	86	187	1 (0.53)
PBW502	128	89	0 (0.00)	110	76	0 (0.00)	85	39	0 (0.00)
SUNCO/2*PASTOR	129	121	1 (0.83)	117	104	0 (0.00)	89	74	0 (0.00)
SW89.5181/KAUZ	130	103	0 (0.00)	111	80	5 (6.25)	87	44	1 (2.27)
TAM200/TUI	126	254	0 (0.00)	107	172	73 (42.44)	87	96	0 (0.00)
Total		**2001**	**16 (0.80%)**		**1552**	**195 (12.56%)**		**990**	**29 (2.93%)**

IRRI: irrigated trial, DRO: drought trial, HEAT: heat trial, SDL: synthetic derived line, BWP: bread wheat parent, Ave. DMA: average days to maturity.

Table S3.4: TKW for SDLs and BWPs and number of SDLs that had higher TKW than their respective BWPs under IRRI in year 2011-12 ($P < 0.05$).

BWP	# SYNPs crossed to BWP	Total SDLs	BWP TKW gr)	SDLs TKW (gr) Ave. (range)	# SDL > BWP for TKW (gr)	% SDL > BWP for TKW (gr)	CV%
PANDORA	12	162	45	49 (40 – 64)	46	28	9.26
MILAN/S87230//BAV92	7	208	43	48 (40 – 60)	86	41	8.60
CACUKE	4	125	54	56 (41 – 65)	16	13	7.26
KRL19	3	36	41	46 (42 – 54)	11	31	7.15
KIRITATI	4	51	50	49 (41 – 57)	0	0	7.51
SW89.5181/KAUZ	2	39	43	46 (40 – 56)	5	13	8.43
SUNCO/2*PASTOR	3	67	41	47 (41 – 57)	37	55	6.64
PBW502	4	37	51	53 (45 – 65)	7	19	8.30
MILAN/AMSEL	1	10	44	49 (43 – 55)	4	40	7.36
MINO	1	17	41	46 (41 – 57)	6	35	7.84
MUU	2	35	47	52 (41 – 62)	19	54	9.91
KIRITATI//PRL/2*PASTOR	1	21	54	44 (40 – 53)	0	0	8.77
Total		808			237	29%	

BWP: bread wheat parent, SYNP: synthetic hexaploid wheat parent, SDL: synthetic derived line, TKW: thousand kernel weight. CV: coefficient of variation in percent.

The second column shows how many SYNPs crossed to each BWP. The third column shows the total number of SDLs for each BWP crossed to different SYNPs. Column six indicates the number of SDLs that had significantly higher TKW than their respective BWP ($P < 0.05$).

Table S3.5: Summary information for SN/m^2 for BWPs and SLD populations in the IRRI.Y11.12.

BWP	No. SDLs	BWP SN/m^2	Mean (range) SN/m^2 SDLs	SDL > BWP for SN/m^2 No. (%)*
PANDORA	162	14922	14053 (9969-19202)	2 (1.2)
MILAN/S87230//BAV92	208	18842	14774 (10340-19311)	0 (0)
CACUKE	125	12948	11435 (7835-14943)	0 (0)
KRL19	36	17307	14197 (11840-16724)	0 (0)
KIRITATI	51	14526	13749 (10725-16422)	0 (0)
SW89.5181/KAUZ	39	17298	14826 (8355-18435)	0 (0)
SUNCO/2*PASTOR	67	17602	15025 (9339-18898)	0 (0)
PBW502	37	13937	12081 (9155-15049)	0 (0)
MILAN/AMSEL	10	14452	13208 (12124-15821)	0 (0)
MINO	17	15749	14762 (11144-16910)	0 (0)
MUU	35	15579	11642 (5275-16249)	0 (0)
KIRITATI//PRL/2*PASTOR	21	12098	13666 (11245-15900)	6 (28.6)
Total	**808**			**8 (0.99%)**

BWP: Bread wheat parent, No. SDL: Number of Synthetic derived lines for each BWP, SN/m^2: Seed number per square meter.
*: Number of SDLs hat had significantly higher SN/m^2 than their respective BWP (adjusted FDR $P < 0.05$).

Table S3.6: SDLs that had higher SN/m^2 than their respective BWP in the IRRI.Y11.12 ($P < 0.05$).

SDL pop.	GID (SDL)	BWP	SYNP	Cross	SN/m^2 BWP	SN/m^2 SDL
38	6763703	PANDORA	GAN/AE.SQUARROSA (446)	BC	14922	19202
27	6765281	PANDORA	CETA/AE.SQUARROSA (895)	BC	14922	18040
94	6670257	KIRITATI//PRL/2*PASTOR	MAYOOR//TK SN1081/AE.SQUARROSA (222)/3/OCI	TC	12098	15343
94	6670263	KIRITATI//PRL/2*PASTOR	MAYOOR//TK SN1081/AE.SQUARROSA (222)/3/OCI	TC	12098	15176
94	6670295	KIRITATI//PRL/2*PASTOR	MAYOOR//TK SN1081/AE.SQUARROSA (222)/3/OCI	TC	12098	14970
94	6670298	KIRITATI//PRL/2*PASTOR	MAYOOR//TK SN1081/AE.SQUARROSA (222)/3/OCI	TC	12098	15895
94	6670299	KIRITATI//PRL/2*PASTOR	MAYOOR//TK SN1081/AE.SQUARROSA (222)/3/OCI	TC	12098	15109
94	6670300	KIRITATI//PRL/2*PASTOR	MAYOOR//TK SN1081/AE.SQUARROSA (222)/3/OCI	TC	12098	15900

SDL pop.: Synthetic derived line population, GID: Genotype identity, BWP: Bread wheat parent, SYNP: Synthetic parent, SN/m^2: Seed number per square meter.

Table S3.7: Summary of linkage map for genomes A, B, D, and each chromosome.

Chromosome Name	# Markers	Length (cM)
1A	140	90
2A	161	122
3A	198	188
4A	104	122
5A	209	182
6A	98	123
7A	216	130
1B	175	142
2B	217	146
3B	149	123
4B	37	78
5B	170	162
6B	98	35
7B	199	138
1D	36	85
2D	65	112
3D	71	112
4D	35	106
5D	85	178
6D	52	113
7D	80	124
Whole Genome	**2595**	**2610**

Table S3.8: Chi-squared test for segregation ratios of the parental alleles for SNPs associated with traits in the SDL populations compared to the theoretical ratio of 3:1 in BC1 and 1:1 in biparental.

				BC families	**Biparental families**	**Trait/Environment**			
SNP NO.	**SNP**	**Chr.**	**pos.**	***P* value**	***P* value**	**DRO**	**HEAT**	**IRRI**	**Linked SNPs**
3064689	T/C	1B	NA	0.25112^{ns}	5.19E-03^{ns}			TKW	
1093512	T/G	2D	112	1.04E-11***	1.78E-30***	YLD			2
2244579	A/G	2D	112	1.39E-20***	3.25E-34***			DFL-YLD	2
2242893	C/T	2D	112	5.82E-12***	1.71E-19***		YLD	YLD-DFL	2
2251719	A/G	2D	NA	7.91E-04***	5.47E-05***			TKW	
985496	G/T	3B	NA	1.92E-03*	1.06E-02^{ns}		PLH		
2259110	C/A	3D	NA	2.33E-01^{ns}	1.85E-01^{ns}			TKW	
2256857	C/T	4A	32	5.28E-01^{ns}	1.75E-02^{ns}			DFL	3
3027878	G/A	4A	35	6.21E-02^{ns}	5.08E-01^{ns}	DFL-DHE^{+}			3
1216917	G/T	4B	33	5.09E-08***	7.43E-06***			YLD	1
983836	T/C	4B	33	1.77E-15***	1.08E-06***	PLH		PLH	1
1088389	A/G	4B	33	3.12E-11***	3.64E-06***			PLH	1
1208575	T/C	4B	33	1.32E-01^{ns}	1.14E-01^{ns}			TKW	1
1102535	G/A	4D	89	4.50E-13***	1.27E-07***	YLD	YLD		
1141498	T/C	5A	59	1.58E-02^{ns}	1.31E-01^{ns}	DFL-DHE	DFL-DHE-DMA	DFL-DHE	
1092197	C/T	4A	35	6.11E-01^{ns}	7.41E-02^{ns}	DMA			
1002527	G/A	6B	NA	4.21E-01^{ns}	2.95E-02^{ns}			TKW	
1067078	C/T	6D	111	1.64E-02^{ns}	6.66E-01^{ns}		YLD		
3064692	C/G	7A	NA	4.68E-03^{ns}	1.94E-01^{ns}			TKW	

Chr.: chromosome, pos.: chromosome position in cM, P value: probability value, DRO: drought, HEAT: heat, IRRI: irrigated, DHE: days to heading, DFL: days to flowering, DMA: days to maturity, PLH: plant height, YLD: grain yield, TKW: thousand kernel weight.

^{ns}: non-significant, *, ***: significant at P < 0.05 and 0.001. (The threshold was 0.0025 at a significance level of 5% for Bonferroni multiple test correction).

$^{+}$: some SNPs associated with multiple traits. In the "Linked SNPs" column, SNPs with common numbers were in the same position or were linked.

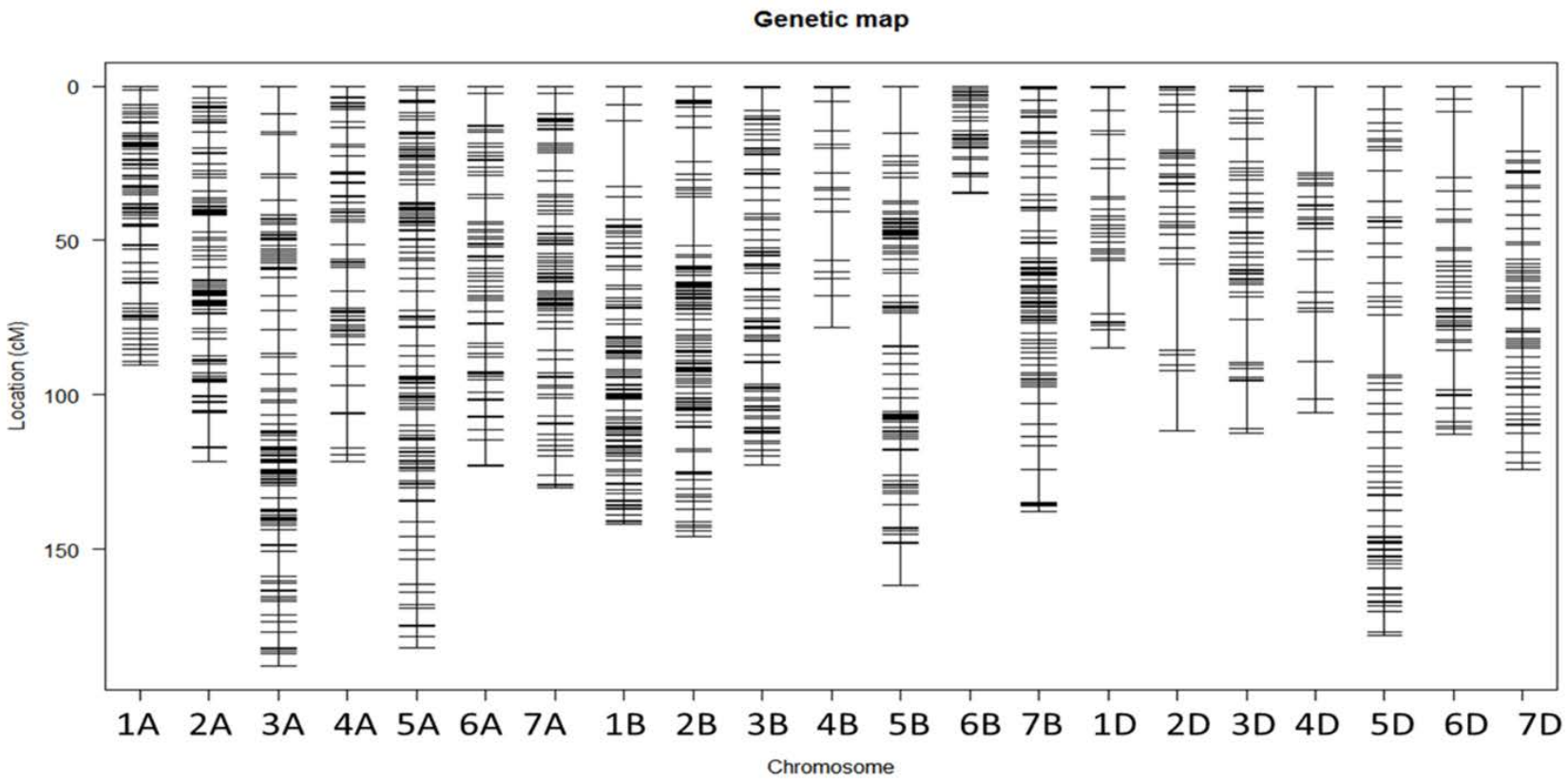


Figure S3.1: Genetic linkage map and distribution of SNP markers on chromosomes in genomes A (1A to 7A), B (1B to 7B), and D (1D to 7D)

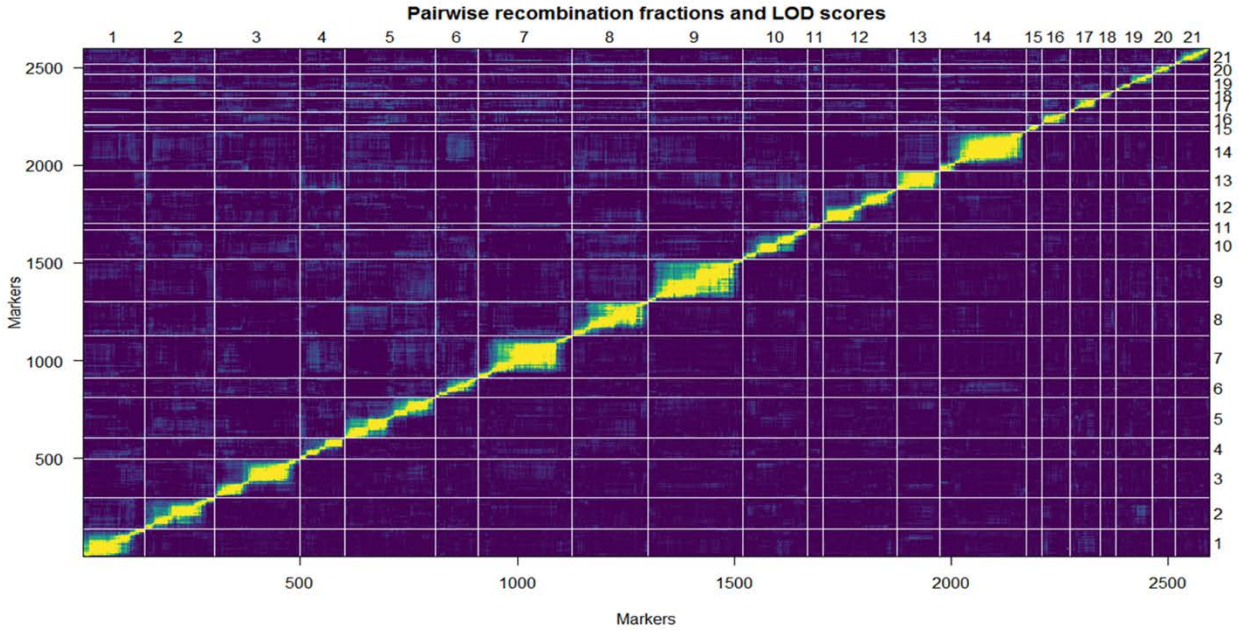


Figure S3.2: Pairwise recombination fractions (above diagonal) and logarithm of the odds (LOD) scores (below diagonal) for SNP markers on each chromosome in genomes A (chr. 1 to 7), B (chr. 8 to 14) and D (chr. 15 to 21).

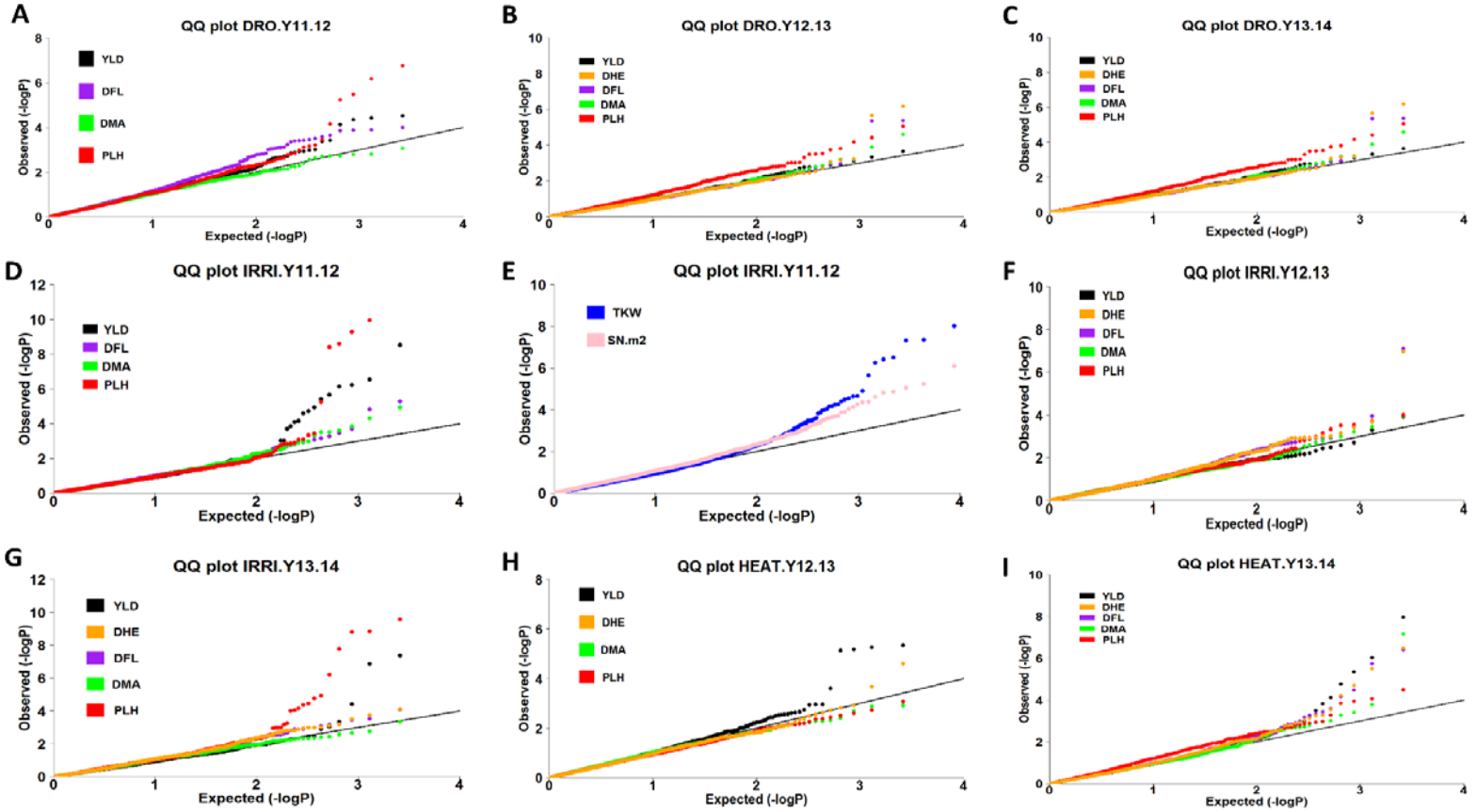


Figure S3.3: Q-Q plot for all traits across eight environments (trial/year). (A to C) Q-Q plot for drought (DRO), (D to G) Q-Q plot for irrigated (IRRI), and (H and I) Q-Q plot for heat (HEAT). DHE: days to heading, DFL: days to flowering, DMA: days to maturity, YLD: grain yield, PLH: plant height, TKW: thousand kernel weight, and SN.m2: seed number per square meter.
Y11.12: year 2011-12, Y12.13: year 2012-13, and Y13.14: year 2013-14.

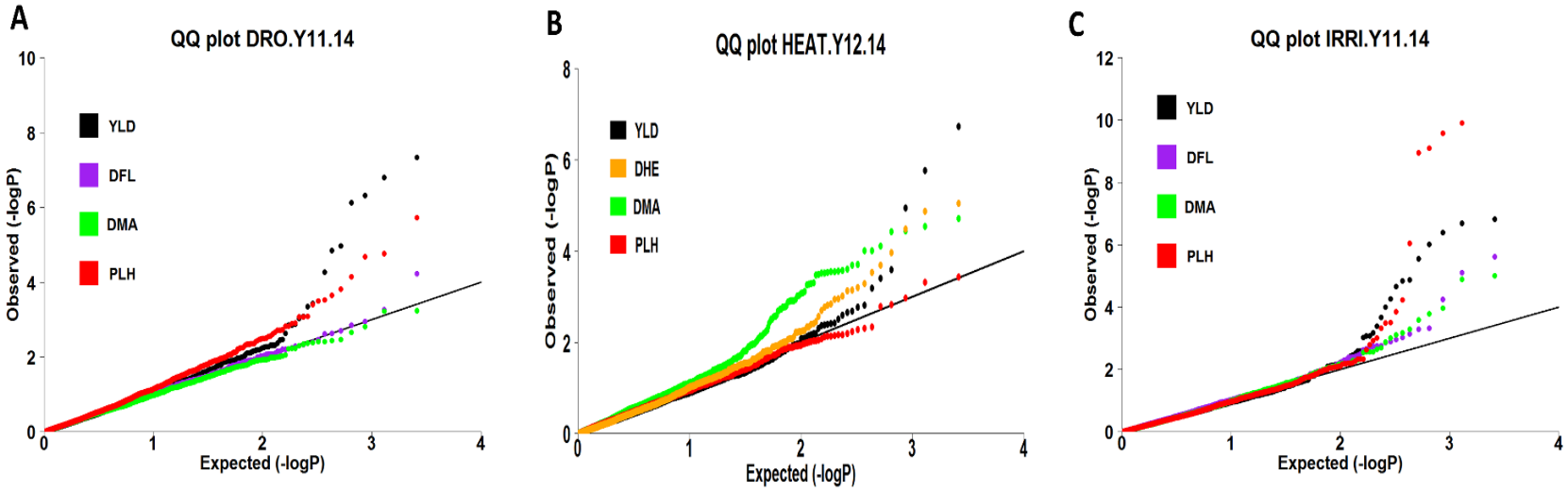


Figure S3.4: Q-Q plot for all traits across three environments. (A) Q-Q plot for drought (DRO), (B) Q-Q plot for heat, and (C) Q-Q plot for irrigated (IRRI). DHE: days to heading, DFL: days to flowering, DMA: days to maturity, YLD: grain yield, and PLH: plant height.

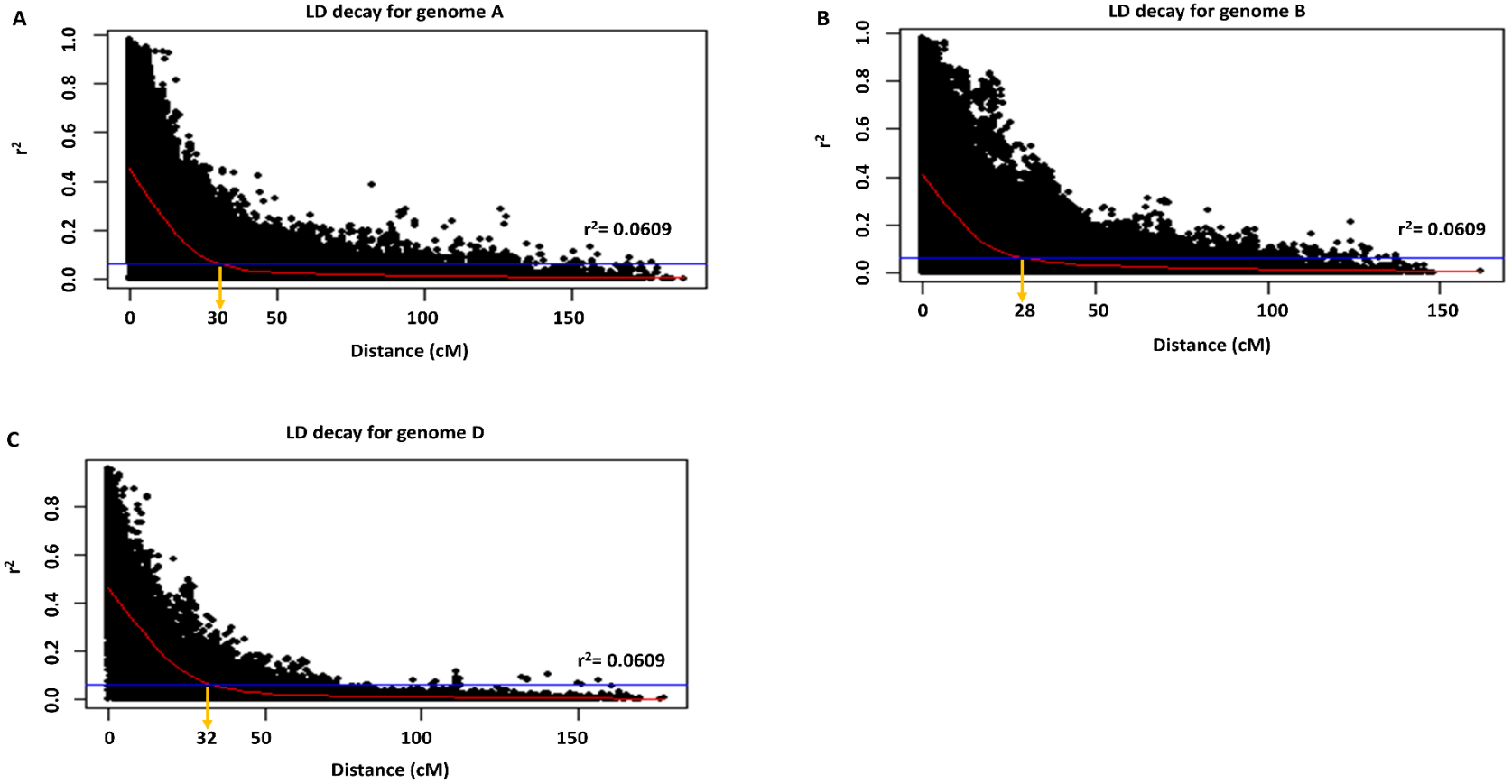


Figure S3.5: LD decay plot for genome A (A), B (B), and D (C). Estimated pairwise r^2 were plotted against the genetic distance (cM). Red curve is LOESS smooth line and horizontal blue line is the 95 percentile of the distribution of unlinked pairwise r^2. The intersect of smooth line and baseline as the extent LD was 30 cM for genome A (A), 28 cM for genome B (B), and 32 cM for genome D (C).

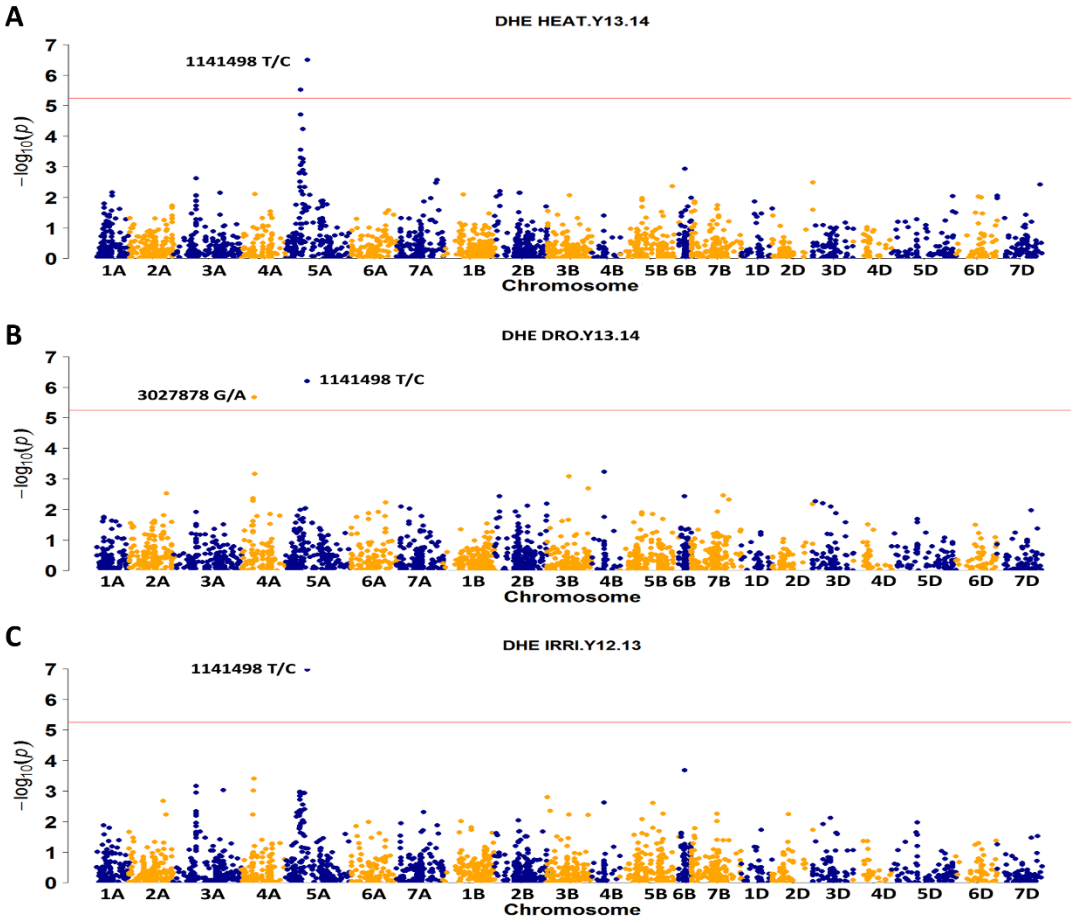


Figure S3.6: Manhattan plot for QTL associated with days to heading (DHE), qDHE-5A-1 (1141498 T/C SNP), on chromosome 5A in the HEAT.Y13.14 (A) in the DRO.Y13.14 (B) and in the IRRI.Y12.13 (C). The qDHE-5A-1 was also associated with DFL and DMA under HEAT and with DFL under DRO. The SYNP allele of this QTL increased trait's values under HEAT, DRO and IRRI. The qDHE-4A-1 (3027878 G/A) on chromosome 4A associated with DHE (and DFL as well) in DRO.Y13.14 (B) and BWP allele this QTL increased DHE and DFL.

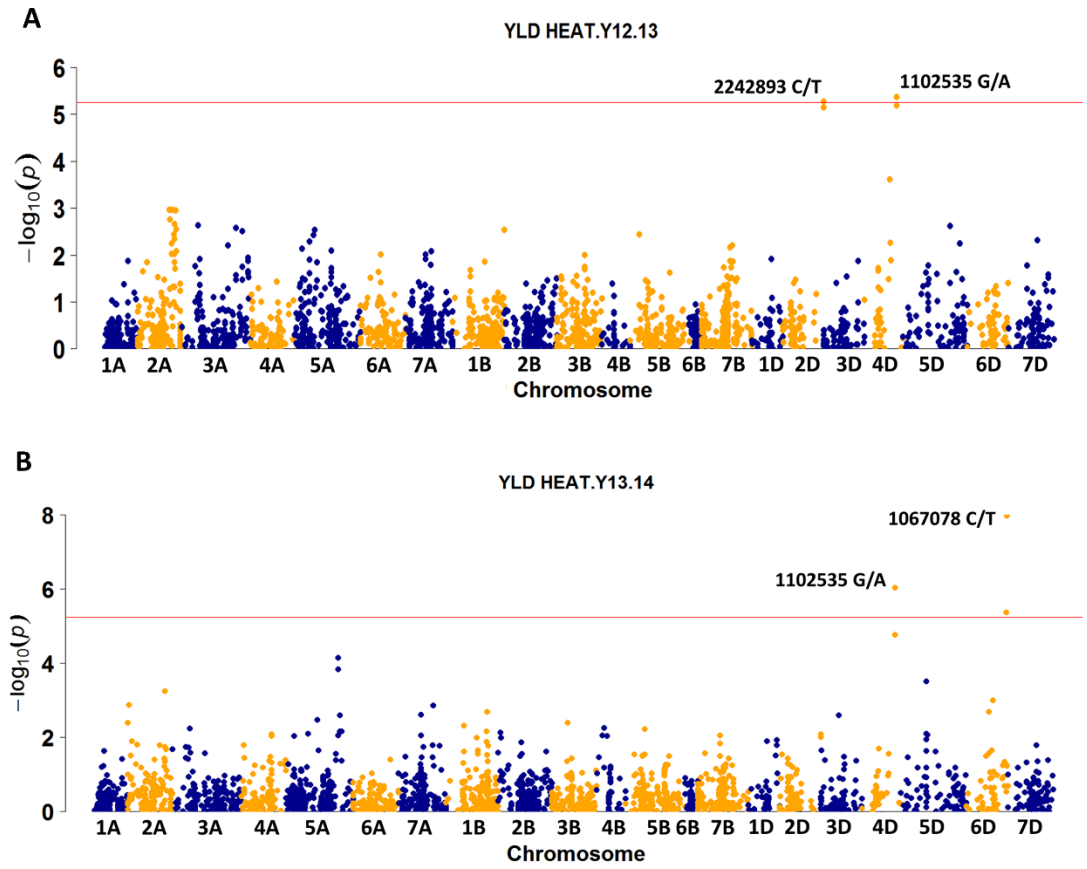


Figure S3.7: Manhattan plots for grain yield (YLD) QTL on chromosomes 2D, 4D, and 6D. A) The qYLD-2D-1 (2242893 C/T SNP) and qYLD-4D-1 (1102535 G/A SNP) on chromosomes 2D and 4D, respectively, in the HEAT.Y12.13 (HEAT in year 2012-13), B) The qYLD-4D-1 (1102535 G/A SNP) and qYLD-6D-1 (1067078 C/T SNP) on chromosomes 4D and 6D, respectively, in the HEAT.Y13.14. The BWP alleles of qYLD-2D-1 and qYLD-4D-1 increased YLD while the SYNP allele of qYLD-6D-1 increased YLD.

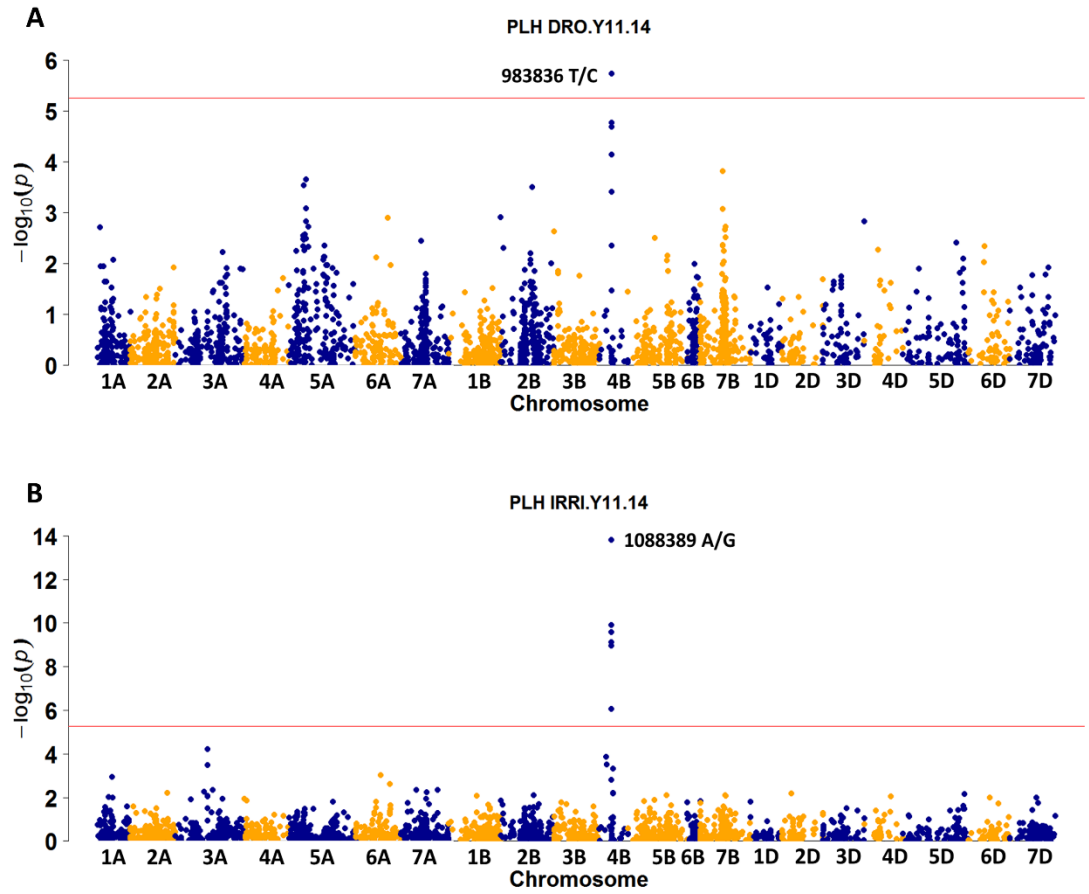


Figure S3.8: Manhattan plot for plant height (PLH) QTL on chromosome 4B, qPLH-4B-1 (983836 T/C and 1088389 A/G SNPs were linked), A) in the DRO.Y11.14 (over years), and B) in the IRRI.Y11.14 (over years). The SYNP allele of this QTL increased PLH.

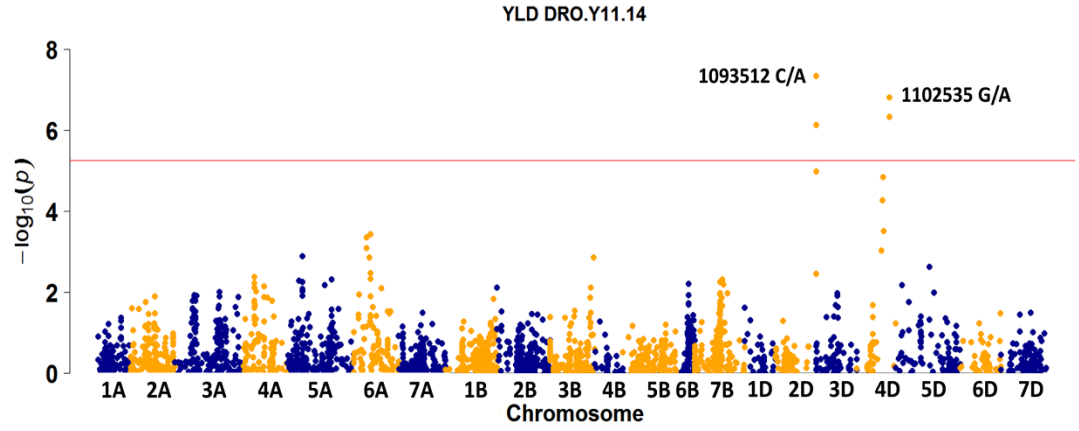


Figure S3.9. Manhattan plot for YLD QTL on chromosomes 2D, qYLD-2D-1 (1093512 C/A SNP) and 4D, qYLD-4D-1 (1102535 G/A SNP), in the DRO.Y11.14 (over years). The qYLD-2D-1 showed a MYI (marker by year interaction) and in the Y11.12 BWP allele increased YLD while in the Y13.14 SYNP allele increased YLD. For the qYLD-4D-1, BWP allele increased YLD.

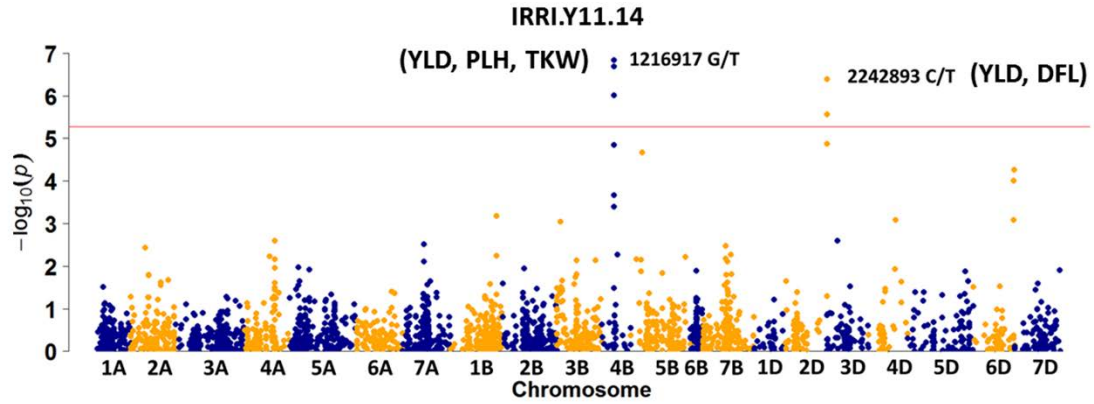


Figure S3.10: Manhattan plot for QTL associated with YLD, PLH, and TKW on chromosome 4B (1216917 G/T SNP) and QTL associated with YLD and DFL on chromosome 2D (2242893 C/T SNP) in the IRRI.Y11.14 (over years). The QTL on chromosome 4B (QTL-4B) showed a MMI (marker by management interaction) for YLD and the SYNP allele of this QTL increased YLD under DRO and HEAT stresses while its BWP allele increased YLD under IRRI. The SYNP allele of QTL-4B increased PLH and TKW. The SYNP allele of the QTL on chromosome 2D increased DFL while its BWP allele increased YLD.

CHAPTER 4

CONCLUSION

The overall objective of this study was using synthetic heaxploid wheat to introduce novel genetic diversity from wild relatives of hexaploid wheat to wheat gene pool and increase genetic variation of grain yield and phenological traits. Study's finding showed that SYN lines were more diverse than BW cultivars for A, B and specifically D genomes. Equally important is the question of whether SYN lines can contribute to increased grain yield and phenological traits. The SDL populations displayed a wide range of phenotypic diversity for traits measured under DRO, HEAT and IRRI conditions. The yield increases were predominantly in SDLs from BC1 derived lines under DRO and HEAT stress indicating that SYN lines could increase grain yield of elite wheat varieties under stress conditions. This was confirmed with estimating GEBVs of SYNPs under three environments for grain yield for which SYNPs had less negative GEBVs under DRO stress and with some positive GEBVs under HEAT stress. The SYNPs also increased TKW under IRRI confirming that SYN line were valuable genetic source for this trait. Under HEAT and DRO, SYNPs increased PLH that indirectly increased YLD.

The SYNPs had QTL for agronomic and phenological traits, which were minor with coincident QTL and QEI. The SYNP allele were more frequently increased YLD, PLH and DMA under HEAT and DRO stresses and TKW under IRRI conditions. However, these specific SDL populations were selected for maturity traits, and PLH

approximating that of the BWPs. Also, the majority of these populations were BC1 derived lines. Therefore, it is possible that many of the SYNP alleles were discarded or retained in a lower frequency which reduced the power to detect their effects or association with traits of interest.

This project results confirmed that synthetic hexaploid wheat germplasms are valuable genetic resource for improving agronomic and phenological traits and SYN lines should be used in breeding programs to expand the genetic diversity for agronomic traits but selection against undesirable phenology is required to realize the benefit of the novel genetic variation.