Genetic basis of eco-geographic adaptation in wild relatives for wheat improvement

by

Elina Adhikari

B.S., Tribhuvan University, 2011
M.S., North Dakota State University, 2015

AN ABSTRACT OF A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Interdepartmental Genetics
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2020

Abstract

Wheat production has been improved significantly through plant breeding and agronomic practices. Nonetheless, the yield drive came with the cost of reduced genetic diversity in elite varieties. Bottleneck due to domestication, complexity in gene sharing due to polyploidization, loss of genetic diversity during hybridization, and years of selective breeding during commercialization narrowed the genetic base of the hexaploid wheat as compared to its wild counterparts.

Availability of ancestral genomes in the wild relatives provides a fallback option for finding new allelic diversity required in wheat breeding. The wheat wild relatives are adapted to a wide range of climatic conditions. However, only a small fraction of the existing genetic diversity has been used in improving the adaptive potential of the hexaploid wheat. In my dissertation, I study how wheat evolved to adapt to its local climate using two immediate progenitors of wheat, Ae. tauschii and wild emmer wheat.

In my first study, I used wild emmer, a tetraploid wild relative of the hexaploid wheat to study the genetics of clinal adaptation in wild emmer. The major objective of this study is to understand the role of geography and climate in shaping the allele frequency distributions indicating clinal adaptation. This study identifies the genomic signatures of adaptation and uses them for predicting adaptive potential. I use 444 geo-referenced wild emmer accessions collected throughout the wide climatic range of Israel. Genotyping is carried out using the 90K iSelect SNP array (90K) and the sequence-based genotyping (GBS) and is mapped to a wild emmer reference genome. The GBS and 90 K respectively identify 341,228 SNPs and 26,548 SNPs, among which, 26,697 SNPs and 9,175 SNPs are retained for respective genotyping platforms. The analysis of population stratification revealed four genetically distinct groups of wild emmer

accessions largely reflecting their geographic distribution. Pearson correlation among 103 historical bioclimatic variables identifies twenty-five unique climatic variables that are noncollinear (Person correlation 0.85). Partitioning of genetic variance shows that geography and climate together explain 44% of genetic variations among emmer accessions with climatic factors accounting for 10% of the SNP variance. The eco-geographic adaptation alleles identified through environmental association scans of historical on-site climatic data improved the prediction accuracy of the heading date by 9%. This research finds that geography and climate play a vital role in shaping the genetic diversity of the wild emmer wheat.

In my second study, I examine the environmental drivers guiding the clinal adaptation in Ae. tauschii, one of the diploid wild relative of wheat. A genome-wide environmental scan is used to identify the climate associated alleles (CAA) and select a representative set of 21 Ae. tauschii accessions enriched for CAAs. Using the top six Kansas adapted wheat varieties as recurrent parents and the selected Ae. tauschii accessions as donor parents, inter-specific population carrying introgression of the Ae. tauschii alleles in the D genome. Three hundred and fifty-one inter-specific BC1F3:5 lines developed through direct hybridization are used in studying the genomic pattern of alien introgression. Using 136 georeferenced Ae. tauschii accessions collected throughout the wide range of climatic conditions, I find that the Ae. tauschii broad lineages segregate with the eco-geographic climatic gradients. Out of 103 historical bioclimatic variables, twelve unique non-collinear climatic variables were identified. Climate, geography, and climate together with geography, respectively, explain 28%, 55%, and 65% of the SNPs variation.

A genome-wide environmental scan identified a total of 508 CAAs, out of which 322 were successfully transferred into the Ae. tauschii-wheat inter-specific population. A separate analysis

using the same data finds that hybrid sterility, reduced introgression in the pericentromeric regions, and reduced retention of the introgressed alleles in regions carrying domestication gene contribute to low retention of alien introgression in the inter-specific population. The field-based evaluation of 351 introgression lines in drought vs rainfed trials showed that 25 % of these lines outperform the best checks, implying successful introgression of beneficial variants positively affecting wheat performance under stress conditions. However, more in-depth analysis is required to understand the effect of these alien introgressions on wheat performance. Overall, my research demonstrates the tremendous potential of wild relatives' genetic diversity for breeding climate-resilient wheat.

Genetic basis of eco-geographic adaptation in wild relatives for wheat improvement

by

Elina Adhikari

B.S., Tribhuvan University, 2011
M.S., North Dakota State University, 2015

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Interdepartmental Genetics
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2020

Approved by:

Major Professor
Eduard Akhunov

Copyright

Abstract

Wheat production has been improved significantly through plant breeding and agronomic practices. Nonetheless, the yield drive came with the cost of reduced genetic diversity in elite varieties. Bottleneck due to domestication, complexity in gene sharing due to polyploidization, loss of genetic diversity during hybridization, and years of selective breeding during commercialization narrowed the genetic base of the hexaploid wheat as compared to its wild counterparts.

Availability of ancestral genomes in the wild relatives provides a fallback option for finding new allelic diversity required in wheat breeding. The wheat wild relatives are adapted to a wide range of climatic conditions. However, only a small fraction of the existing genetic diversity has been used in improving the adaptive potential of the hexaploid wheat. In my dissertation, I study how wheat evolved to adapt to its local climate using two immediate progenitors of wheat, Ae. tauschii and wild emmer wheat.

In my first study, I used wild emmer, a tetraploid wild relative of the hexaploid wheat to study the genetics of clinal adaptation in wild emmer. The major objective of this study is to understand the role of geography and climate in shaping the allele frequency distributions indicating clinal adaptation. This study identifies the genomic signatures of adaptation and uses them for predicting adaptive potential. I use 444 geo-referenced wild emmer accessions collected throughout the wide climatic range of Israel. Genotyping is carried out using the 90K iSelect SNP array (90K) and the sequence-based genotyping (GBS) and is mapped to a wild emmer reference genome. The GBS and 90 K respectively identify 341,228 SNPs and 26,548 SNPs, among which, 26,697 SNPs and 9,175 SNPs are retained for respective genotyping platforms. The analysis of population stratification revealed four genetically distinct groups of wild emmer

accessions largely reflecting their geographic distribution. Pearson correlation among 103 historical bioclimatic variables identifies twenty-five unique climatic variables that are noncollinear (Person correlation 0.85). Partitioning of genetic variance shows that geography and climate together explain 44% of genetic variations among emmer accessions with climatic factors accounting for 10% of the SNP variance. The eco-geographic adaptation alleles identified through environmental association scans of historical on-site climatic data improved the prediction accuracy of the heading date by 9%. This research finds that geography and climate play a vital role in shaping the genetic diversity of the wild emmer wheat.

In my second study, I examine the environmental drivers guiding the clinal adaptation in Ae. tauschii, one of the diploid wild relative of wheat. A genome-wide environmental scan is used to identify the climate associated alleles (CAA) and select a representative set of 21 Ae. tauschii accessions enriched for CAAs. Using the top six Kansas adapted wheat varieties as recurrent parents and the selected Ae. tauschii accessions as donor parents, inter-specific population carrying introgression of the Ae. tauschii alleles in the D genome. Three hundred and fifty-one inter-specific BC1F3:5 lines developed through direct hybridization are used in studying the genomic pattern of alien introgression. Using 136 georeferenced Ae. tauschii accessions collected throughout the wide range of climatic conditions, I find that the Ae. tauschii broad lineages segregate with the eco-geographic climatic gradients. Out of 103 historical bioclimatic variables, twelve unique non-collinear climatic variables were identified. Climate, geography, and climate together with geography, respectively, explain 28%, 55%, and 65% of the SNPs variation.

A genome-wide environmental scan identified a total of 508 CAAs, out of which 322 were successfully transferred into the Ae. tauschii-wheat inter-specific population. A separate

analysis using the same data finds that hybrid sterility, reduced introgression in the pericentromeric regions, and reduced retention of the introgressed alleles in regions carrying domestication gene contribute to low retention of alien introgression in the inter-specific population. The field-based evaluation of 351 introgression lines in drought vs rainfed trials showed that 25 % of these lines outperform the best checks, implying successful introgression of beneficial variants positively affecting wheat performance under stress conditions. However, more in-depth analysis is required to understand the effect of these alien introgressions on wheat performance. Overall, my research demonstrates the tremendous potential of wild relatives' genetic diversity for breeding climate-resilient wheat.

Table of Contents

List of Figures

List of Tables

Acknowledgements

There is a very popular saying “it takes a village to raise a child”. A village helped me to get a Ph.D. I would like to thank everyone who cultivated a curiosity to learn and most importantly to those who believed in my curiosity and supported along the way.

I would like to thank my major adviser Dr. Eduard Akhunov for all the support and guidance. Eduard, thank you for taking the time to teach me some of the important concepts. In future, I will invest time in training the next generation of scientists like you did with me. I am very thankful to Dr. Allan K. Fritz for his constant guidance. Allan, thank you for teaching me to think like a breeder. Thank you to my other committee members; Dr. Christopher Toomajian and Dr. Robert Aiken, for being instrumental in my overall development.

Big thank you to everyone in Eduard’s and Allan’s labs. Each of you made this journey more meaningful and I am very grateful for your support and help. To my friends in Plant Pathology, Agronomy, and Nepalese student community, and Plant Pathology office staff thank you all for the beautiful memories we created together and best wishes for your future endeavors.

To my family, you all make me a very fortunate person. I am very grateful for the love and support from my father Binod Adhikari, my sister Pratikshya Adhiakri, my brother in-law Suman Pokhrel, my sweetheart nice Prisha Pokhrel, my brother Ashish Adhikari and my sister in-law Prakrita Bhandari, my mother in-law Rina Regmi, and my whole Regmi family.

If there is one person who deserves more credits for this achievement even more than me then it is my mother Gyanu Adhikari. I don’t have enough words to thank my mother; you are my hero and my backbone in this journey. Last but not least, to the love of my life; my husband Madhav Regmi, and my soul; my son Elon Regmi, thank you for your unconditional love and support.

Dedication

This dissertation is dedicated to my family and every little girl chasing their big dreams.

Chapter 1 - Role of wild relatives in wheat improvement

Breeding germplasm that better adapts to future environments can help alleviate the threats of climate change. Improvements in adaptive potential in wheat depend on the integration of diverse wild progenitors and landraces into the breeding programs. Previous studies show that wild progenitors and landraces from stressful environments are a valuable reservoir of new allelic diversity for improving adaptive potential in wheat (Reynolds et al., 2007; Peleg et al., 2005). Introgression of adaptive variants to develop climate-resilient wheat can help increase the genetic gains to meet exponentially rising food demands.

The selection of superior plants in terms of desirable characteristics dates back to early human civilization. This kind of preferential selection resulted in the domestication of the plant species. Cultivation is believed to have started about 10,000 years ago in the Neolithic Revolution (Shewry, 2009). Selection imposed during domestication altered genes involved in seed shattering, non-threshing or tough glume, indeterminate growth habit, seed dormancy, and small seed size in the wild (Harlan et al., 1973). Origin and domestication of most of the cereal crops are believed to have occurred in Fertile Crescent, a horseshoe shaped area spreading from Israel, Turkey to Iran (Harlan and Zohary, 1966). The first cultivated wheat was Einkorn; a diploid (AA genome), and emmer wheat, a tetraploid (AABB genome), identified in that same region (Dubcovsky and Dvorak, 2007). A rare natural hybridization between tetraploid wheat (*Triticum turgidum* (2n = 4x = 28; AABB genome) and a diploid goatgrass *Aegilops tauschii* (2n = 2x = 14; DD genome), formed the allohexaploid bread wheat; *Triticum aestivum* (2n = 6x = 42; AABBDD genome), about 9000 years ago (McFadden and Sears, 1946; Jordan et al., 2015). Since then, hexaploid wheat superseded diploid and tetraploid wheat in genome plasticity which

allowed it to adapt to a wide range of environments throughout the world with improved stress tolerance and superior grain quality (Dubcovsky and Dvorak, 2007).

In self-pollinated crops like wheat, wild relatives play a major role in conserving genetic diversity. Genetic bottleneck due to domestication, polyploidization, interspecific hybridization, and modern plant breeding resulted in narrow genetic variability in bread wheat genomes as compared to its wild counterparts which increases its vulnerability to biotic (insect-pests and diseases) and abiotic (environmental) stresses (Akhunov et al., 2010; Ladizinsky,1985; Cox, 1997). Genetic diversity in wild relatives of wheat has been used in improving adaptative potential (He et al., 2019; Jia et al., 2013), disease resistance (Zhang, et al., 2018; Friesen et al., 2008), yield (Reynolds et al., 2017; Fritz et al., 1995), and quality traits in the bread wheat (Li et al., 2012; Hsam et al., 2001). For example, *Aegilops tauschii* and wild emmer (*Triticum turgidum* subsp. *dicoccoides*) has long been used in breeding programs to improve the genetic diversity of the wheat (Dreisigacker et al., 2008; Merchuk-Ovnat et al., 2016). Interspecific hybridization not only allows introgression of favorable allelic variants from wild species into the hexaploid gene pool (Mujeeb-Kazi et al., 1996; Singh et al., 2019) but also helps reduce the bottleneck effect (Dubcovsky and Dvorak, 2007).

Ae. tauschii is reported to spread around a wide geographical range from Georgia, Turkey, and Syria in the west up to China in the east (Wang et al., 2013); however, the genetic diversity in the D genome of bread wheat is highly restricted compared to its wild counterparts. Only a handful of *Ae. tauschii* accessions from a small area near the Caspian Sea is believed to have hybridized with wheat leading to polyploidy bottleneck (Dubcovsky and Dvorak, 2007). In addition, the limited hybridization of *Ae. tauschii* with hexaploid wheat due to change in ploidy, domestication, and selection imposed by targeted breeding hindered the gene flow resulted in

narrow genetic base of wheat D genome (Akhunov *et al.*, 2010; Wang et al., 2013). The diversity level in the hexaploid wheat D genome is traced back to a single *Ae. tauschii subsp strangulata* gene pool coming from Armenia and SW Caspian sea (Dvorak et al., 1998). This makes the wild D genome counterpart a valuable reservoir of untapped beneficial genetic diversity (Singh et al., 2019). Further, it encourages breeding programs to incorporate *Ae. tauschii* in their breeding scheme to capture the novel genetic diversity present in the wild D-genome.

The ease of using wild relatives in the breeding scheme depends on the availability of a homologous genome for pairing during meiosis (Martínez-Pérez et al., 1999). *Aegilops tauschii* shares one homologous genome (D genome) and the wild emmer shares two (A and B) homologous genomes with wheat. They can either be crossed directly with wheat or through the development of synthetic wheat (Cox, 1997; Cox et al., 2017). However, introgression from the distant wild relative such as the ones in the tertiary gene pool requires biotechnological intervention. For example, the *Ph1* gene located at chromosome 5B of bread wheat prevents homoeologous chromosomes pairing and recombination (Riley and Chapman, 1959). When individuals in tertiary gene pool are crossed with the individuals in primary gene pool the hybrids produced are sterile. Transfer of novel genetic variants from landraces and wild progenitors strongly depends on the frequency of recombination along the chromosome (Wijnker and de Jong. 2008; Jordan et al., 2018; Akhunov et al., 2003). The reshuffling of genes during meiotic recombination breaks the linkage drag allowing plant breeders to pyramid desired alleles in one germplasm. However, the lower recombination with possible deleterious effect alleles in the peri-centromeric regions (Jordan et al., 2018), wide cross hybrid sterility (Akhunov et al., 2010), reduced introgression around the domestication genes (He et al., 2019) infertility of the segregating generations (Nyine et al., 2020), cross incompatibility between the wild species and

the cultivated crop, and linkage drag (Zamir, 2001) highlights some of the limitations in using wild relatives for crop improvements. Despite these limitations, wild relatives are critical resources for improving modern-day varieties.

Traditionally, *Ae. tauschii* introgression is done following two distinct classical breeding approaches; introgression through the development of synthetic wheat, and direct hybridization with the hexaploid wheat (Cox et al., 2017). Usually, parental lines were selected based on pre-screening, subjecting lines through appropriate stress (Reynold et al., 2007). During synthetic wheat development, segregating amphiploids, also known as synthetic wheat, are produced through hybridization between *T. turgidum* (tetraploid) and *Ae. tauschii* (diploid). The resulted F1 hybrid's chromosome number is doubled using colchicine (mitotic agent) or through the union of unreduced gametes producing hexaploid seeds which are known as synthetics. All three genomes of the newly developed synthetic lines segregate for favorable and deleterious allele's introgressed from each donor parents. Besides increasing genetic diversity, the interaction between two uniquely adapted genomes causes masking and epistasis of some useful genes, hybrid necrosis, loss of free threshability (Cox et al., 2017), and oftentimes down-regulated the expression of *Ae. tauschii* genes (He et al., 2003).

In the direct hybridization approach, *Ae. tauschii* is directly crossed with hexaploid wheat. The resulting F1 hybrid is embryo rescued and backcrossed with hexaploid wheat-producing amphiploids that are segregating for chromosome numbers (Gill and Raupp, 1987; McFadden and Sears, 1946; Cox, 1997). This process allows complete conservation of A and B genomes, and rapid introgression of allelic variants from *Ae. tauschii* into hexaploid D genome gene pool only (Cox et al., 2017). While crossing wheat with synthetic wheat lines is relatively easy as compared to crossing wheat with *Ae. tauschii,* genetic gain form synthetic methods might

be considerably low due to its long breeding cycle. Inter-specific population developed through the direct cross of *Ae. tauschii* and *T. aestivum,* improved the grain yield, kernel weight, and grain quality traits (Fritz et al.,1995). As the "AA" and "BB" genomes of *Ae. tauschii*-wheat inter-specific hybrids are the same as that of the recurrent parent and only "DD" genome is mosaic of the two parents it is arguably less wild alleles to deal with. Which eventually will result in fewer breeding cycles than the synthetics.

My first research identified SNPs that are associated with the major climatic factors, and quantified the usefulness of those alleles in predicting adaptation traits using the wild emmer accessions collected across a wide geographic range of Israel. Results suggest that climate and geography play a major role in shaping the genetic architecture of wild emmer populations contributing to higher fitness in their local environments. I discovered that SNPs favored during adaptation increase the heading date prediction accuracy by 9% as compared to a random set of SNPs. This result validates climate adaptive alleles detected using Redundancy analysis (RDA), a multivariate model, and BAYENV, and suggest that SNPs identified using environmental association scans can be the useful resource for breeding crop varieties resilient to variation in climatic factors.

In my second research study, I examined the genetic basis of adaptation in *Ae. tauschii*, the diploid wild relatives of the wheat. The main objective of this research is to identify the individual accessions of *Ae. tauschii* that are genetically diverse and enriched for heat and drought associated SNPs using the genome-wide environmental scan. These accessions are then used for developing a inter-specific population by directly crossing *Ae. tauschii* with Kansas adapted wheat to understand the usefulness of climate associated SNPs for improving wheat. Using the historical climatic data, I find that the genetic variation within the broad lineages

of *Ae. tauschii* are significantly influenced by the environment and geography. Multivariate linear regression in RDA identified that the *Ae. tauschii* accessions coming from drought-prone areas are enriched for adaptive variation. Twenty-one of these *Ae. tauschii* accessions were used in direct hybridization with the six elite Kansas adapted wheat lines to produce 351 $BC1F_{3:5}$ inter-specific lines. This introgression population is being used to characterize the utility of *Ae. tauschii* introgression for breeding hexaploid wheat better adapted to drought prone environments. Results based on one year of data indicate that about 25% of the newly developed lines out-performed the best check yield in bussels per acerage (bu/ac) suggestive of the positive effect of *Ae. tauschii* introgression on wheat performance. However, a detailed quantitative analysis with more data is needed to assess the contribution of the introgressed adaptation alleles to increasing yield under stress conditions.

References

Akhunov, Eduard D., Alina R. Akhunova, Olin D. Anderson, James A. Anderson, Nancy Blake, Michael T. Clegg, Devin Coleman-Derr et al. "Nucleotide diversity maps reveal variation in diversity among wheat genomes and chromosomes." *BMC Genomics* 11, no. 1 (2010): 702.

Akhunov, Eduard D., Andrew W. Goodyear, Shu Geng, Li-Li Qi, Benjamin Echalier, Bikram S. Gill, J. Perry Gustafson et al. "The organization and rate of evolution of wheat genomes are correlated with recombination rates along chromosome arms." *Genome Research* 13, no. 5 (2003): 753-763.

Cox, T. S. "Deepening the wheat gene pool." *Journal of Crop Production* 1, no. 1 (1997): 1-25.

Cox, Thomas S., Jizhong Wu, Shuwen Wang, Jin Cai, Qiaofeng Zhong, and Bisheng Fu. "Comparing two approaches for introgression of germplasm from *Aegilops tauschii* into common wheat." *The Crop Journal* 5, no. 5 (2017): 355-362.

Dreisigacker, Susanne, Masahiro Kishii, Jacob Lage, and Marilyn Warburton. "Use of synthetic hexaploid wheat to increase diversity for CIMMYT bread wheat improvement." *Australian Journal of Agricultural Research* 59, no. 5 (2008): 413-420.

Dubcovsky, Jorge, and Jan Dvorak. "Genome plasticity a key factor in the success of polyploid wheat under domestication." *Science* 316, no. 5833 (2007): 1862-1866.

Dvorak, J., M-C. Luo, Z-L. Yang, and H-B. Zhang. "The structure of the Aegilops tauschii genepool and the evolution of hexaploid wheat." *Theoretical and Applied Genetics* 97, no. 4 (1998): 657-670.

Friesen, T. L., S. S. Xu, and M. O. Harris. "Stem rust, tan spot, Stagonospora nodorum blotch, and Hessian fly resistance in Langdon durum–Aegilops tauschii synthetic hexaploid wheat lines." *Crop science* 48, no. 3 (2008): 1062-1070.

Fritz, A. K., T. S. Cox, B. S. Gill, and R. G. Sears. "Marker-Based Analysis of Quantitative Traits in Winter Wheat× Triticum tauschii Populations." *Crop Science* 35, no. 6 (1995): 1695-1699.

Gill, Bikram S., and W. J. Raupp. "Direct Genetic Transfers from Aegilops squarrosa L. to Hexaploid Wheat 1." *Crop Science* 27, no. 3 (1987): 445-450.

Harlan, Jack R., and Daniel Zohary. "Distribution of wild wheats and barley." *Science* 153, no. 3740 (1966): 1074-1080.

Harlan, Jack R., J. M. J. De Wet, and E. Glen Price. "Comparative evolution of cereals." *Evolution* 27, no. 2 (1973): 311-325.

He, Fei, Raj Pasam, Fan Shi, Surya Kant, Gabriel Keeble-Gagnere, Pippa Kay, Kerrie Forrest et al. "Exome sequencing highlights the role of wild-relative introgression in shaping the adaptive landscape of the wheat genome." *Nature Genetics* 51, no. 5 (2019): 896-904.

He, Ping, Bernd R. Friebe, Bikram S. Gill, and Jian-Min Zhou. "Allopolyploidy alters gene expression in the highly stable hexaploid wheat." *Plant Molecular Biology* 52, no. 2 (2003): 401-414.

Hsam, S. L. K., R. Kieffer, and F. J. Zeller. "Significance of Aegilops tauschii glutenin genes on breadmaking properties of wheat." *Cereal Chemistry* 78, no. 5 (2001): 521-525.

Jia, Jizeng, Shancen Zhao, Xiuying Kong, Yingrui Li, Guangyao Zhao, Weiming He, Rudi Appels et al. "Aegilops tauschii draft genome sequence reveals a gene repertoire for wheat adaptation." *Nature* 496, no. 7443 (2013): 91-95.

Jordan, Katherine W., Shichen Wang, Yanni Lun, Laura-Jayne Gardiner, Ron MacLachlan, Pierre Hucl, Krysta Wiebe et al. "A haplotype map of allohexaploid wheat reveals distinct patterns of selection on homoeologous genomes." *Genome Biology* 16, no. 1 (2015): 1-18.

Ladizinsky, Gideon. "Founder effect in crop-plant evolution." *Economic Botany* 39, no. 2 (1985): 191-199.

Li, Yulian, Ronghua Zhou, Jin Wang, Xiangzheng Liao, Gerard Branlard, and Jizeng Jia. "Novel and favorable QTL allele clusters for end-use quality revealed by introgression lines derived from synthetic wheat." *Molecular Breeding* 29, no. 3 (2012): 627-643.

Martínez-Pérez, Enrique, Peter Shaw, Steve Reader, Luis Aragón-Alcaide, Terry Miller, and Graham Moore. "Homologous chromosome pairing in wheat." *Journal of Cell Science* 112, no. 11 (1999): 1761-1769.

McFadden, Edgar S., and Ernest Robert Sears. "The origin of Triticum spelta and its free-threshing hexaploid relatives." *Journal of Heredity* 37, no. 3 (1946): 81-89.

Merchuk-Ovnat, Lianne, Vered Barak, Tzion Fahima, Frank Ordon, Gabriel A. Lidzbarsky, Tamar Krugman, and Yehoshua Saranga. "Ancestral QTL alleles from wild emmer wheat improve drought resistance and productivity in modern wheat cultivars." *Frontiers in Plant Science* 7 (2016): 452.

Mujeeb-Kazi, A., V. Rosas, and S. Roldan. "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." *Genetic Resources and Crop Evolution* 43, no. 2 (1996): 129-134.

Nyine, Moses, Elina Adhikari, Marshall Clinesmith, Katherine W. Jordan, Allan K. Fritz, and Eduard Akhunov. "Genomic patterns of introgression in interspecific populations created

by crossing wheat with its wild relative." *G3: Genes, Genomes, Genetics* 10, no. 10 (2020): 3651-3661.

Peleg, Zvi, T. Fahima, S. Abbo, T. Krugman, E. Nevo, D. Yakir, and Y. Saranga. "Genetic diversity for drought resistance in wild emmer wheat and its ecogeographical associations." *Plant, Cell & Environment* 28, no. 2 (2005): 176-191.

Reynolds, Matthew, Fernanda Dreccer, and Richard Trethowan. "Drought-adaptive traits derived from wheat wild relatives and landraces." *Journal of Experimental Botany* 58, no. 2 (2007): 177-186.

Riley, Ralph, Victor Chapman, and Gordon Kimber. "Genetic control of chromosome pairing in intergeneric hybrids with wheat." *Nature* 183, no. 4670 (1959): 1244-1246.

Shewry, Peter R. "Wheat." *Journal of experimental botany* 60, no. 6 (2009): 1537-1553.

Singh, Narinder, Shuangye Wu, Vijay Tiwari, Sunish Sehgal, John Raupp, Duane Wilson, Mehraj Abbasov, Bikram Gill, and Jesse Poland. "Genomic analysis confirms population structure and identifies inter-lineage hybrids in Aegilops tauschii." *Frontiers in Plant Science* 10 (2019): 9.

Wang, Jirui, Ming-Cheng Luo, Zhongxu Chen, Frank M. You, Yuming Wei, Youliang Zheng, and Jan Dvorak. "Aegilops tauschii single nucleotide polymorphisms shed light on the origins of wheat D-genome genetic diversity and pinpoint the geographic origin of hexaploid wheat." *New Phytologist* 198, no. 3 (2013): 925-937.

Wijnker, Erik, and Hans de Jong. "Managing meiotic recombination in plant breeding." *Trends in plant science* 13, no. 12 (2008): 640-646.

Zamir, Dani. "Improving plant breeding with exotic genetic libraries." *Nature Reviews Genetics* 2, no. 12 (2001): 983-989.

Zhang, Dale, Yun Zhou, Xinpeng Zhao, Linlin Lv, Cancan Zhang, Junhua Li, Guiling Sun, Suoping Li, and Chunpeng Song. "Development and utilization of introgression lines using synthetic octaploid wheat (aegilops tauschii× hexaploid wheat) as Donor." *Frontiers in Plant Science* 9 (2018): 1113.

Chapter 2 - Eco-geographic Adaptation Signals Improve Heading Date Prediction Accuracy in Wild Emmer Wheat

Abstract

In the face of global climate change, breeding for varieties that adapt better to the future climatic condition has become one of the primary components of germplasm improvement. As a rich source of genomic variation, the wild relatives of wheat have been broadly used in breeding to develop varieties adapted to diverse environments. I identified accessions of wild emmer wheat (AABB genome), the tetraploid ancestor of hexaploid wheat, carrying alleles conferring adaption to extreme climatic conditions by looking at eco-geographic patterns of genomic variation. I associated patterns of genomic variation in 444 georeferenced wild emmer accessions with bioclimate data from the accessions' collection sites. The panel was genotyped using the 90K iSelect SNP array and next-generation sequencing of complexity reduced genomic libraries resulting in the identification of 26,548 and 9,175 SNPs, respectively. The analysis of population structure revealed three genetically distinct groups of wild emmer accessions coinciding with their geographic distribution. Partitioning of genomic variance showed that geographic location and climate together explained 44% of genetic variation among emmer accessions, with climatic factors accounting for only 10%. Eco-geographic adaptation alleles identified through multiple environmental association scans using historical on-site climatic data improved the prediction accuracy of heading date by nine percent. In this targeted breeding pipeline, I demonstrate that eco-geographic adaptive alleles from wild emmer can improve the prediction accuracy of important agronomic traits which shows a promising future for making informed selection decisions in breeding programs.

Introduction

Successful development of novel crop varieties adapted to changing environmental conditions depends on the presence of adaptive alleles in breeding populations. Populations of genetically diverse crop wild relatives growing in climatically diverse regions provide a rich reservoir of novel alleles with the potential to improve adaptation to abiotic stress factors (Fuillet et. al., 2008; Blum, 2011; Palmgreen et. al., 2015). However, most of this ancestral adaptive diversity has been lost from wheat populations due to population bottlenecks caused by domestication and polyploidization followed by intense human-driven selection aimed at developing high-yielding locally adapted cultivars. Relatively recently, rich genetic diversity stored in the populations of wild relatives was recognized as a vital component of wheat improvement programs. The synthetic hexaploid wheat lines developed by hybridizing tetraploid wheat with diploid goatgrass *Aegilops tauschii* were introduced into international breeding programs (Dreisigacker et al., 2008). On average, diversity derived from synthetic wheat could be detected in nearly 17% of all the lines in the nurseries of CIMMYT and ICARDA (Ortiz et al., 2008). The value of wild relatives' allelic diversity for improving wheat was proven by demonstrating that lines derived from crosses with the wild relatives collected from high-stress environments show better tolerance to drought and heat stress (Trethowan et. al., 2008; Blum, 2011).

Recent efforts to introduce wild relative diversity into breeding programs targets only a limited fraction of the overall population genomic diversity. The substantial size of the wild relative collections in the germplasm repositories, including thousands of accessions, makes identification of the most promising accessions challenging. Often, genetic material is prioritized by performing phenotypic screening for adaptive traits. However, this strategy could be

meaningfully applied only to lines adapted to grow in target an environment, and has limited utility for accessions coming from diverse eco-geographic habitats. The alternative germplasm prioritization approach is based on building small "core collections" that capture maximum levels of genetic, geographic, or phenotypic diversity of the entire collection of accessions. Despite effectively reducing the total number of accessions, the main disadvantages of this strategy include removal of the majority of rare alleles, and unequal representation of alleles affecting different agronomically relevant phenotypes reducing its utility for some breeding applications.

The development of cost-effective next-generation sequencing (NGS) approaches now provides the possibility to generate genome-wide SNP datasets for large geographically diverse samples. These datasets, combined with historic eco-geographic variables allows for performing genome-wide scans to identify alleles associated with adaptive phenotypes (Fournier-Level et al. 2011; Lasky et al., 2012; Hancock et al., 2011). The methods of environmental scans (GWES) are based on the assumption that variants contributing in local adaptation should show an unusually high correlation between their frequencies and climatic eco-geographic variables (Coop et al., 2010; Günther et al., 2013). Climate-associated alleles (CAA) identified in *Arabidopsis* (Lasky et al., 2012) and sorghum (Lasky et al., 2015) by performing GWES using historic climatic variables were shown to accurately predict the relative fitness of plant accessions in a specific environment, suggesting that GWES is a promising strategy for prioritizing allelic diversity from germplasm collections for incorporation into the breeding pipelines of national and international wheat research programs.

Tolerance to high temperature and drought stress are among the primary target traits of breeding programs. A tetraploid ancestor of modern wheat, wild emmer (*Triticum turgidum* ssp. *dicoccoides*), distributed across the Fertile Crescent with diverse climatic regions characterized by extreme variation in temperature and water availability, is considered as a rich source of adaptive diversity for improving adaptation to water-limiting and increased temperature conditions (Nevo et al., 2013; Peleg et al., 2005). Wild emmer is mostly self-pollinating grass with two homologous sets of chromosomes (2n = 4x = 28, AABB) (Zohary, 1970; Özkan et al., 2011) that originated from spontaneous hybridization between a close relative of *Aegilops speltoides* (2n = 2x = 14 genome SS) and wild diploid Einkorn Wheat, *T. urartu* (2n = 2x = 14, genome AA) about 360,000 years ago, (Dvorak and Akhunov, 2005), in Near East Fertile Crescent (Dvorak and Akhunov, 2005; Özkan et al., 2011; Nevo, 1998). Wild emmer in Near East Fertile Crescent covers parts of Jordan, Syria, Turkey, Lebanon, Iraq, and Iran (Harlan and Zohary, 1966; Will cox, 2005; Nevo, 1998; Özkan et al., 2011; Driscol et al., 2009). Particularly, in Israel it is predominantly found in upper Jordan valley, central and western Galilee Mountains, the Judean Mountains, Mount Carmel, and Samaria (Özkan et al., 2011; Harlan and Zohary, 1966; Nevo, 1998). Northern Israel is home to large, morphologically, and genetically diverse wild emmer population whereas its distribution is more scattered in the rest of the surrounding regions (Nevo, 1998).

Here, I used a panel of 475 geo-referenced wild emmer accessions collected across Israel and historic onsite climate data associated with each accession to identify genomic variation linked with adaptation to environmental factors. I used redundancy analysis (RDA), a multivariate regression technique (Legendre and Legendre 1998; Borcard et al., 1992) to estimate the total proportion of genomic variation among wild emmer accessions that was

explained by the geography and local climate and identify major environmental factors driving the allele distribution in wild emmer population in Israel. The genome-wide environmental scans were performed to identify adaptive alleles based on their correlation with important climatic variables in that region. I then tested the utility of the climate associated alleles for predicting heading date in wild emmer, one of the major adaptive traits affecting plants' fitness to local environmental conditions.

Methods

Wild emmer accessions

The wild emmer accessions used in this study were collected from Israel. The collection area covers a wide range of regions and growing environmental conditions. Table 2.1 contains the standardized value of each environmental variables in each sub-population where color fading from red to green represents the values of the highest to lowest for each variable. The population includes 475 wild emmer accessions out of which 444 accessions had GPS coordinates. These accessions were collected from longitudinal range 35°89'29" N to 34°88'88" N and latitudinal range 33°53'23.7" E to 31°35'12.4" E. Based on geographic origin and the distribution of eco-geographic factors, these accessions were divided into seven sub-groups (Figure 2.1). Sub-population 1, with the largest number of accessions (n=99), comes from the highly elevated area that stretches along Galilee. Sub-population 2 accessions (n=54) are spread east of the Sea of Galilee and lower part of the Golan Heights. Sub-population 3 accessions (n=52) are spread around Haifa and the west part of the Sea of Galilee. Sixty-five accessions from sub-population 4 are spread predominantly across Golan Heights that is north of the Sea of Galilee. Sub-population 5 with 64 accessions covers the periphery of Jerusalem and Southern parts of Israel.

Sub-population 6 has 40 accessions, which were collected from the east of Coastal Plains and Tel Aviv-Jaffa. Sub-population 7 includes eighty accessions collected along Jordan Valley.

Sequencing platforms and SNP callings

Infinium iSelect 90K SNP assay was used to genotype the wild emmer population following the standard Illumina's protocols implemented in the GenomeStudio v2011.1 software with the Polyploid Clustering v1.0 module (Wang et al., 2014). Cluster distance was set at 0.07, confidence = 0.8, and the minimum number of points in the cluster was set to 10. In the first step, clusters that met the criteria of numbers of clusters = 2, call frequency greater than 0.85 were named polymorphic SNP 1. In the second step, clusters equal to 1 were changed to clusters equal to 2 and the same filtering criteria (call frequency C1 Freq >0.85) was applied. SNPs generated in this step were set as polymorphic SNP 2. But in the last (third) step, cluster distance was set to 0.09 and rest of the other criteria were the same (# clusters = 2, call freq. = 0.9, confidence = 0.8, C1 Freq >0.9) and SNPs generated were name polymorphic SNP 3. And single clusters with call frequency greater than 0.99 were named "Monomorphic" in the third step. During the process, SNPs with more than 2 clusters in the first and second steps were named "Multiple Clusters". In the end, all multiple clusters were manually checked to confirm the selection criteria. Markers forming monomorphic clusters, multiple clusters (>2), or having more than 10% missing data were not included in the analysis. In total, I discovered 26,974 polymorphic SNPs, which were mapped to the newly assembled wild emmer wheat reference genome (Avni et al., 2017) and converted from the "AB" format to the SNP IUPAC format. These SNPs were further filtered and SNP sites >10 heterozygous individuals, multi-allelic or monoallelic calls, and MAF < 0.04 were removed resulting in retention of 9,175 SNPs.

GBS libraries were prepared as previously described (Saintenac et. al., 2013; Poland et al., 2012). Library sizes and quality were verified using Agilent TapeStation 2200 and Agilent 2100 Bioanalyzer. Size selection was performed using Pipping Prep (Sage Science). Illumina single-end sequencing 1 x 100 bp was performed in the Genomics Facility of the KU Medical Center. In total, 682 million GBS reads were generated, with about 1.5 million reads per accessions. The reads were mapped to the wild emmer reference genome (Avni et al., 2017) using Bowtie2 v.2.3.2. GBS sequence data were analyzed using the TASSEL5.0 GBSv2 pipeline that generated 341,239 raw variant calls (Glaubitz et al. 2014).

GBS is high throughput, low-cost next generation genotyping technology, which generates low coverage read data with high rates of missing genotype calls (Swarts et al., 2014). In most cases genotype datasets generated using GBS require imputation to predict the missing SNPs and increase the statistical association detection power in downstream analyses (Cleveland et al., 2011). Missing and un-genotyped SNPs were imputed using the FILLIN that was compiled in TASSEL5.0. For imputation, I followed the procedures outlined by Swarts et al. (2014). The imputed data were further filtered to remove SNP sites with >10 heterozygous individuals, multi-allelic or monoallelic calls, and MAF < 0.04 resulting in a set of 26,548 SNPs.

SNP diversity analysis

I have investigated genomic patterns of SNP distribution across genome and MAF. I calculate the pairwise F_{ST} for seven sub-populations using Weir and Cockerham's F_{ST} method (Figure 2.7). Pairwise F_{ST} was calculated in a sliding window of 5 Mbp with a step size of 2 Mbp and plotted using R package 'ggplot2' v3.2.1. The top 95^{th} percentile of F_{ST} distribution was used to detect outlier genomic windows tentatively subjected to adaptive selection. Through the use of program SNPEff (Cingolani et al., 2012), the effect of each identified SNPs was predicted.

Bioclimatic variables

Using the WorldClim v.2.0 database, on-site bioclimatic variables for all georeferenced accession were downloaded for the years 1970 to 2000 at a resolution of 30 arc-seconds (~1 km) (Table 2.1). I used 19 bioclimatic variables and monthly average climatic variables (precipitation, mean, maximum and minimum temperature, wind, water vapor pressure, and solar radiation) as phenotypes (Hijmans et al., 2005). Figures 2.2 and 2.3 respectively show the monthly averages of precipitation and solar radiation in Israel for the 30 years period. The bioclimatic and climatic variables were extracted using the R package "raster". Bioclimatic and monthly average climatic variables of seven sub-populations were standardized (subtract the mean and then divided by the standard deviation of the variable across populations) as per BAYENV.2.0 (Günther et al., 2013) requirement to estimate the correlation with allele frequency.

Population structure analysis

I analyzed the population structure using both model-based (STRUCTURE) and non-model-based (PCA) clustering methods (Figs. 2.4 - 2.7). Bayesian Monte Carlo Markov Chain (MCMC) algorithm implemented in STRUCTURE 2.0 (Pritchard and Wen, 2003) with default parameters was used for clustering accessions using the admixture and allele frequency correlated model. Analysis using thinned genotyping data is performed including 10,000 randomly selected SNPs that were converted to diploid numeric format (0, 1, and 2). The length of the burn-in was set at 50,000 followed by MCMC (Markov chain Monte Carlo) run for 100,000 steps. MCMC runs were repeated 3 times for each value of K from 1 to 7. Membership of each accession in K clusters was analyzed in relation to seven sub-populations defined based

on the geographic distribution and landscape features. The output file from STRUCTURE was summarized using StructureHarvester (Earl and vonHoldt, 2012).

To estimate spatial population structure and genetic co-ancestry with respect to geography, I used the "tess3" function implemented in the "tess3r" R package (Caye et al., 2016; Caye et al., 2018). TESS3r (Caye et al., 2016), to estimates the population genetic structure based on allele frequency differentiation and layers the ancestry coefficients on a geographic map. SNPs were converted to diploid numeric SNPs format and all polymorphic SNPs were used in inferring the spatial population structure of wild emmer accessions from Israel for K ranging from 1 to 12.

Principal components were calculated using the snpgdsPCA function of the R package SNPRelate (Zheng et al., 2012) and "Scatterplot3d" (Ligges and Mächler, 2002). R package was used to generate a three-dimensional principal components plot (Figure 2.6). Phylogenic analysis was performed using the R package "*ape*" and the accessions were color coded based on the membership in seven subpopulations from different eco-geographic regions (Figure 2.7).

Partitioning genomic variance

Redundancy analysis (RDA) is a form of direct gradient analysis, which uses multiple linear regression to summarize the dependent variable (Y) explained by the independent variable (X) (Legendre and Legendre, 1998; Borcard et al., 1992). It has been widely used in studying patterns of spatial genetic variation due to demographic history and the abiotic environmental factors. The canonical axes identified through RDA are linear combinations of explanatory variables that best explain linear combinations of response variables (Borcard et al., 2018). Here the multivariate predictors are climate and geography (spatial) whereas the multivariate response variables are the biallelic SNPs generated through GBS and 90K SNP assay. Partial RDA, which

is similar to partial regression analysis was also used to control the spatial structure in the wild emmer population.

RDA integrated into a descriptive ecology R package "*vegan*" (Oksanen, 2011) was used to partition the variance explained by climatic variables (abiotic factor) and geographic distance (spatial). The "*varpart*" function from the "*vegan*" package (Oksanen, 2011) categorized the important variables that are involved in shaping the SNP distribution of wild emmer accessions in Israel. Adjusted R^2 value was used to partition the variance explained by geography and environmental variables. Out of the total of 103 explanatory variables, which also included 19 bioclimatic variables and a monthly average of precipitation, mean, maximum and minimum temperature, water vapor pressure, wind, and solar energy, 25 unique non-correlated explanatory variables were used. To remove the collinearity in the environmental variable dataset that can affect the results of RDA analysis, the variables showing Person correlation coefficients >0.85 were excluded. For example, if two variables were correlated strongly then the variable that coincides with the emmer growing season is retained and the other variable is dropped. Doing this, I retained a set of 25 non-collinear explanatory variables (Table 2.1) for variance partitioning.

Association mapping and correlation of allele frequencies with bioclimatic variables

Genome-wide association mapping (GWAM) was carried out using the filtered and imputed SNPs. To detect significantly associated SNPs, I used mixed linear models implemented in the GAPIT R package (Lipka et al., 2012; Yu et al., 2006). The confounding effects of genetic relatedness and population structure were accounted for by including kinship and Q membership matrices into the models (Yu et al., 2006).

The BAYENV 2.0, a Bayesian algorithm-based software was used to identify the alleles involved in local adaptation (Coop et al., 2010; Günther et al., 2013). BAYENV analyses were conducted using climatic data for geographic regions used to group wild emmer accessions into seven subpopulations (Table 2.1). Climatic data for each region was standardized by subtracting the population mean from individual values and divided by the standard deviation of each sub-population. A covariance matrix file was generated using 35,724 SNPs with 30K iterations of Bayenv run replicated three times with different random seeds. SNPs were ranked according to the Bayes Factor (BF) values after running BAYENV 2.0, and those SNPs that were detected in the top 1% of BF distribution in all three Bayenv runs were considered to be significantly correlated with a climatic factor.

Validation of eco-geographic adaptation alleles in predicting heading date

I evaluated the ability of eco-geographic adaptation alleles (EAA) at SNP sites identified through RDA and BAYENV to predict adaptive traits in wild emmer. A total of 442 wild emmer accessions were phenotyped for the heading date trait at three different locations and multiple years (>2 years in each location) at UC Davis agriculture research sites. Best linear unbiased estimates (BLUEs) of heading date were calculated using linear mixed models implemented in R package "*lme4*" and used as phenotype data for assessing the accuracy of trait prediction. Five-fold cross validation using mixed models implemented in the R package "rrBLUP" (Endelman, 2011) was used to compare the prediction accuracy for the heading date. First, the accessions were randomized, then the population was split into five groups (K=5) and each group was used once as a testing set. The training set constituted 80% of the 444 wild emmer accessions while the testing set had 20% of accessions. Individuals in the training set (80%) had their phenotype data (heading date) and genotypes data whereas in the testing set (20%) individuals only had

their genotypic data and using the ridge regression model implemented in "rrBLUP" the heading date of testing set was predicted (Meuwissen et al., 2001). Four independent rounds of cross validation were performed and during each round, accessions were randomized differently using a random number generator function in R (set.seed()), before the population was split into five groups, which resulted in 20 testing sets. Prediction accuracy was calculated as the correlation between the predicted and observed heading dates for the testing sets. The mean prediction accuracy for the 20 testing sets was calculated to obtain the overall prediction accuracy of the heading date for each of the two SNP sets.

SNPs associated with the climate adaptation also referred as the candidate SNPs in this manuscript, identified through RDA and BAYENV analysis were tested for their prediction accuracy using heading date data. For a SNP to be candidate in RDA analysis, it should fall outside of ± 3 standard deviation cutoff from mean (two-tailed test p-value = 0.0027). For a SNP to be a candidate in BAYENV analysis, it should fall into the top 1% of BF distribution in three separate BAYENV runs. The prediction accuracy of candidate climate-associated SNPs was then compared to that obtained using the same number of random SNPs.

Genome-wide association study, using a mixed linear model built in GAPIT, was performed using the BLUP of heading date. Significant marker-trait associations from this GWAS were compared with the genomic distribution of candidate SNPs significantly associated with the climate in the RDA and BAYEN analyses using the R package "*qqman*" (Turner, 2014).

Results

Wild Emmer population structure and genetic differentiation

Consistent with the previous study (Luo et al., 2007), wild emmer in Israel could be assigned to three genetically distinct populations using STRUCTURE and Evannos method (Figure 2.7). The

data identified two distinct populations in the Northern part of Israel around Jordan River north of the Sea of Galilee and Golan Heights. Even though the geographic distance between these two populations was small, the distribution in topographic features of the land might have been a key factor separating the two populations. The southern population included the rest of the Southern and Central Israel's wild emmer growing area, predominantly, coastal plain and Jordan Valley.

A more detailed analysis of population structure was also performed using the "*tess3r*" R package for K ranging from 1 to 12 (Appendix A, Figures 1 to 4). It showed that at the K=4, the cross validation error was minimized, indicating that the sample of wild emmer accessions has 4 genetically distinct groups (Figure 2.5). At this value of the K, the Southern population splits into two additional subpopulations mostly covering regions in Southern and Central Israel (Figure 2.4).

Further, I compared the assignments of accessions from seven subpopulations defined based on the topography and eco-geographic factors to populations defined based on genetic relatedness. At K=3, sub-populations 3, 5, 6, and 7 from the central and southern parts of Israel shared a substantial amount of genetic ancestry and showed a limited number of accessions with mixed ancestry (Figure 2.7.B). On contrary, sub-populations 1 and 3 included a number of accessions with mixed ancestry. Accessions of wild emmer from sub-population 2 showed lowest levels of admixture but included accessions that could be assigned to two different clusters, one of which includes a small number of accessions growing near the Sea of Galilei. Accessions from sub-population 2 also showed the highest levels of genetic differentiation from most of the remaining sub-populations (Figure 2.7.C). The highest level of genetic differentiation was observed between sub-populations 2 and 3 with $F_{ST} = 0.26$. Most of the individuals in sub-population 2 come from Golan Heights with latitude ranging from 31° to 33° whereas

individuals in sub-population 3 come from the coastal plains of Israel with latitude ranging from 32° and 33°. Sub-populations 5, 6, and 7 come from southern Israel and the pairwise F_{ST} with sub-population 2 is 0.2, 0.24, and 0.21, respectively (Figure 2.7.D). Consistent with the results of model-based clustering, the Principal Component analysis performed on 444 geo-referenced accessions identified four distinct populations (Figure 2.6). The top three principal components explained 5.87%, 3.02%, and 2.08% of SNP variation respectively (Figure 2.6). Annotation and prediction of the effect of SNPs using "SnpEff" on filtered data show that there were more non-synonymous SNPs compared to synonymous SNPs (Figure 2.8). This pattern was prevalent when using all the unfiltered SNPs (Figure 2.9) analysed. The filtered data had 5,402 (15.1% of the SNPs) non-synonymous (nSNPs), 3,857 (10.8% of the SNPs) synonymous (sSNPs) and 207 (0.57% of the SNPs) SNPs resulting in premature termination codons. The nSNPs/sSNPs ratio for wild emmer was 1.4, likely associated with relaxation of selection leading to accumulation of amino acid changing mutations. Out of 615 RDA identified SNPs there were 2 strongly deletrious SNPs. Similarly, 6 out of 1,032 and 10 out of 1045 Bayenv identified SNPs for solar radiation for the month of July and precipitation for the month of May respectively were strongly deleterious. The strong nSNPs includes those variants that is predicted to cause high/disruptive (deleterious SNPs) impact in the protein. For example, protein truncation, loss of function, or triggering nonsense mediated decay. Whereas, the moderate nSNPs include those variants that are predicted to be non-disruptive and might change protein effectiveness.

Variation in climatic and bio-climatic factors

Israel borders Lebanon in the North, Jordan in the East, Syria in the Northeast, the Mediterranean Sea in the west, and Egypt in the southwest. The climate of Israel ranges from temperate to tropical and has many agro-climatic niches, with North and Central parts of the country receiving

moderate to heavy rainfall and the southern desert areas receving the limited precipitation. The rainy season lasts from November to May and the limited amount of precipitation accumulates over the rest of the year. These temporal and spatial patterns of precipitation distribution across Israel drive the geographic distribution of bioclimatic variables showing high levels of heterogeneity in different parts of the country (Table 2.1). Along with the 19 bioclimatic variables, I included the monthly data of minimum, maximum, and average temperature, precipitation, solar radiation, wind speed, and water vapor pressure from the year 1970-2000. There was a clear difference in the mean values of climatic and bio-climatic factors across regions corresponding to seven sub-populations. The examples of monthly variation in precipitation and solar radiation across these geographic regions is shown in Figures 2.2 and 2.3, respectively. To correct for the collinearity among climatic and bio-climatic factors that may cause issues in downstream analyses, Pearson Correlation coefficients (PCC) were calculated for all possible pair-wise comparisons between climatic factors. Only one of the climatic factors that were clustered using PCC >0.85 threshold was retained, resulting in a set of 25 non-correlated explanatory variables (Table 2.1) that were used for variance partitioning. These variables are marked with ▲ in Table 2.1.

Climatic and bio-climatic factors associated with SNP distribution

Analysis of variance showed that climate and geography are significantly associated with genetic variation. The function "*varpart*" from R package *vegan* was used to find the proportion of genetic variance explained by each climatic variable and identify those variables that explained most of the SNP distribution in the sample of wild emmer accessions. I found that geography and climate together explain 44% of SNPs variation, with climatic factors alone explaining 10% of SNPs distribution ($P \leq 0.001$). After accounting for the spatial structure of the sample through

partial RDA, the percent of genomic variation explained by climate only increased to 13.58%. However, when colinear variables from environmental factors were removed, I found that the variance explained by geography dropped to 18% and the variance explained by climate increased to 15% (Figure 2.10). The percent of variance explained for SNP datasets generated using different genotyping approaches also varied (Figure 2.11), which likely is associated with differences in the SNP variation spectrum captured by each approach. The 90K iSelect assay was previously shown to preferentially capture common SNP variants (Wang et al., 2014), whereas the GBS platform provides less biased estimates of diversity including also rare SNP variants.

Redundancy analysis (RDA) has been widely used for studying the patterns of spatial genetic variation associated with the demographic history and the abiotic factors of the environment. This method is popular in the field of ecological analysis but has also been applied to studying relationships between SNP diversity and climatic factors from the sample's collection sites (Kölliker, et al., 2006). RDA analysis shows that some climatic factors correspond to population differentiation among wild emmer sub-populations. Genetic differentiation of accessions from sub-populations 3, 5, 6, and 7 was associated with increased solar radiation and the speed of the wind and decrease in precipitation. On contrary, genetic differentiation of accessions from sub-populations 1, 2, 4, and part of the accessions from sub-population 7 was associated with increased precipitation and temperature, and decreased solar radiation and wind speed (Figure 2.12 and Figure 2.13). Minimum, maximum, and average temperature showed a negligible relationship with the genetic differentiation among accessions from sub-populations 1, 2, 4, and 7.

Eco-geographic adaptation alleles are predictive of heading date

To identify SNPs linked with environmental adaptation, I extracted outliers (±3 SD) of SNPs loadings along the three constrained axes in RDA. In total, I identified 615 SNPs, including 216, 256, and 143 SNPs on RDA1, RDA2, and RDA3 axes, respectively, associated with adaptation to climatic factors.

Bayenv analysis was performed for the two most significant explanatory variables from RDA: solar radiation in July (srad_07) and precipitation in May. A total of 1,032 significant SNPs (top 1% of BF distribution in 3 independent Bayenv runs) associated with variation in solar radiation in July across seven wild-emmer sub-populations were detected. Variation in precipitation in May (ppt_05) was significantly associated with 1,045 SNPs. The combined unique set of SNPs identified through Bayenv that are associated with low precipitation and high solar radiation includes 2,037 SNPs.

Further, I used a mixed linear model in GAPIT for genome-wide association mapping (GWAS) of heading date. The GWAS for the heading date was calculated for using BLUES of three locations and years. The significant SNPs ($P< 0.005$) were distributed in chromosome 1A, 3A, 4B, 5B, 6A, and 7A (Figure 2.15). One of the most significant SNPs for heading date in chromosome 4B (4B_428040200) was also the candidate SNP of adaptation that was identified through RDA analysis. The effect of this SNP as predicted by the "SnpEff" program is an intron modifier (TRIDC4BG035250). Modifiers are those variants or variants affecting non-coding genes with low effect. Also, the SNPs with the highest BF values for solar radiation for the month of June and the precipitation for the month of May overlapped with some of the significant SNP for heading date at $P<0.005$ was found. Out of fourteen significant SNPs for heading date in chromosome 1A, four SNPs are also the candidate SNP for adaptation to solar

radiation in the month of July, identified through Bayenv analysis (Figure 2.18). Out of these four SNPs, "IAAV749" SNP was predicted as non-synonymous SNP that causes missense mutation of GCA to CCA. Similarly, "1A_221363516" and "1A_ 4538849" SNP was identified as an intergenic modifier, 1A_203809514 was predicted as 5′ UTR modifier. Similarly, one of the most significant SNPs for heading date in chromosome 7B was also the candidate SNP of adaptation to precipitation in the month of May (Figure 2.19) that was identified through Bayenv analysis. In other words, this SNP (7B_63201402) is associated with earlier heading and is predicted as an intergenic modifier. In general, higher number of amino acid changes observed in the distal end of the chromosome as compared to the centromeric regions (Figure 2.21) which might be due to a higher rate of recombination (Jordan et al., 2018)

To test the association between SNPs detected by Bayenv analysis and RDA with adaptive phenotypes, I performed prediction of heading date traits using the mixed model implemented in R package "rrBLUP". I investigated whether the prediction accuracy of SNPs associated with climatic variables will be higher than the prediction accuracy of a random set of SNPs. The data shows that 1,032 SNPs associated with solar radiation in the month of July and 1,045 SNPs associated with precipitation in the month of May increased the heading date prediction accuracy in wild emmer wheat each by 5% compared to prediction accuracy obtained using the same number of randomly selected SNPs (Figure 2.16 and Figure 2.17). Likewise, adaptive SNPs that showed association with climatic variables in RDA improved the prediction accuracy of heading date in wild emmer by 9% compared to the randomly selected set of SNPs (Figure 2.14).

Discussion

The wild emmer accessions collected throughout the growing area across Israel might be one of the biggest wild emmer samples used for genome-environment interaction study in tetraploid wheat. Environmental data associated with each accession shows that wild emmer in Israel grows in diverse eco-geographic habitats with high levels of variation in the amount of rainfall, temperature and solar radiation (Potchter et al., 2008, Li et al., 2003). Accessions from sub-populations 1, 4, and part of sub-population 2 came from the Northern part of Israel with high levels of precipitation. Compared to other regions, the Central Coastal regions of Israel occupied by accessions from sub-populations 3 and 6 are characterized by higher minimum, and average monthly temperature, wind speed, and water vapor pressure. Accessions from sub-population 7 grow in the areas with the highest maximum temperature in Israel. The highest solar radiation and low monthly precipitation are experienced by accessions from sub-population 5 collected from the area close to Negev desert. Thus, this collection of wild emmer accessions provide an excellent system for investigating the impact of eco-geographic factors on the evolution of genetic diversity.

The constrained (RDA) and unconstrained (PCA) method both showed that the wild emmer accessions from the same geographic area tend to cluster together. This implies that the accessions that are geographically close have relatively high gene flow. Sub-population 3, 5, and 6 are genetically related and clustered in the positive end of RDA axis 1. Population structures of the sub-population clustered in RDA axis 1, which are mainly from the Southern part of Israel are associated with the solar radiation and wind speed pattern. Whereas, the population structure of the northern part of Israel comprising sub-population 1, 4, and sub-population 2 follow the

precipitation pattern. This observation is in line with the findings of Lasky et al., (2015), where, the population structure of sorghum was associated with north-south precipitation gradient.

The data reveals that spatial isolation plays a major role (36%) in the distribution of SNP variation in wild emmer populations as compared to the abiotic environmental (10%) factors. After removal of some of the collinear environmental factors, the contribution of spatial (18%) and environmental (15%) factors became similar suggesting that in wild emmer both geographic distribution and variation in climate played an equally important role in shaping genetic diversity across the population in Israel. Similar findings are reported in *Arabidopsis* by Lasky et al., (2012), where portions of genomic variation explained by climate (15.7%) and space (16.9%) were roughly similar. The proportion of genetic variance explained by geography being high before accounting for the collinear climatic variables might be due to the population structure caused by isolation by distance (Sharbel et al., 2000) and edaphic (variation in soil) factors, some of the environmental variables represented within the geography (Manel et al., 2010).

Non-synonymous mutations can have deleterious effects on protein function which can affect agronomically useful traits. I found that genomewide 15 % of the SNPs were nSNPs and most of the CAA identified SNPs were non deleterious SNPs. Overall, the SNPs distribution followed U-shaped distribution following the distribution of recombination rate along the chromosomes (Jordan et al., 2018), with the distal end of the chromosomes contained more SNPs compared to the centromeric regions.

This study shows that SNPs associated with local adaptation identified in wild emmer accessions from Israel improve prediction of heading date, one of the most important adaptive traits in plants (Lasky et al., 2015; Li et al., 2017). This result suggests that by combining historic climatic data with whole genome variation data, we could prioritize genetic diversity in the

germplasm collections for incorporation into breeding programs. An increase in the heading date prediction accuracy observed for climate-associated SNPs also suggests that these SNPs are likely involved in local adaptation.

Conclusion

Understanding the relationship between environmental factors and the genetic makeup of crops and their wild relatives is critical for devising effective strategies to mitigate the effects of Climate change on agro-ecosystems to maintain stable food production. Eco-geographic adaptation study of wild emmer population suggests that climate together with spatial variation played a major role in shaping the genomic variation across Israel. The patterns of variation in climate and geography allowed us to identify SNPs contributing to local adaptation. This conclusion is supported by the improvement of adaptive trait prediction using climate-associated SNPs demonstrating that approaches based on joint analyses of genomic variation and historic climatic data could help us to prioritize wild relative germplasm and allelic diversity for incorporation into breeding programs. Further analysis using high coverage SNP data along with the physiological trait relating to drought like water use efficiency may be useful in dissecting the major genes and the underlying pathway involved in shaping the genetic makeup for adaptation in a stressed environment.

References

Avni, Raz, Moran Nave, Omer Barad, Kobi Baruch, Sven O. Twardziok, Heidrun Gundlach, Iago Hale et al. "Wild emmer genome architecture and diversity elucidate wheat evolution and domestication." *Science* 357, no. 6346 (2017): 93-97.

Blum, Abraham. "Drought resistance and its improvement." In *Plant breeding for water-limited environments*, pp. 53-152. Springer, New York, NY, 2011.

Borcard, Daniel, François Gillet, and Pierre Legendre. *Numerical ecology with R*. Springer, 2018.

Borcard, Daniel, Pierre Legendre, and Pierre Drapeau. "Partialling out the spatial component of ecological variation." *Ecology* 73.3 (1992): 1045-1055.

Caye, Kevin, Flora Jay, Olivier Michel, and Olivier François. "Fast inference of individual admixture coefficients using geographic data." *The Annals of Applied Statistics* 12, no. 1 (2018): 586-608.

Caye, Kevin, Timo M. Deist, Helena Martins, Olivier Michel, and Olivier François. "TESS3: fast inference of spatial population structure and genome scans for selection." *Molecular Ecology Resources* 16, no. 2 (2016): 540-548.

Cingolani, Pablo, Adrian Platts, Le Lily Wang, Melissa Coon, Tung Nguyen, Luan Wang, Susan J. Land, Xiangyi Lu, and Douglas M. Ruden. "A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3." *Fly* 6, no. 2 (2012): 80-92.

Cleveland, Matthew A., John M. Hickey, and Brian P. Kinghorn. "Genotype imputation for the prediction of genomic breeding values in non-genotyped and low-density genotyped individuals." In *BMC proceedings*, vol. 5, no. 3, p. S6. BioMed Central, 2011.

Coop, Graham, David Witonsky, Anna Di Rienzo, and Jonathan K. Pritchard. "Using environmental correlations to identify loci underlying local adaptation." *Genetics* 185, no. 4 (2010): 1411-1423.

Dreisigacker, Susanne, Masahiro Kishii, Jacob Lage, and Marilyn Warburton. "Use of synthetic hexaploid wheat to increase diversity for CIMMYT bread wheat improvement." *Australian Journal of Agricultural Research* 59, no. 5 (2008): 413-420.

Driscoll, Carlos A., David W. Macdonald, and Stephen J. O'Brien. "From wild animals to domestic pets, an evolutionary view of domestication." *Proceedings of the National Academy of Sciences* 106.Supplement 1 (2009): 9971-9978.

Dvorak, J., E., D., Akhunov, 2005. Tempos of gene locus deletions and duplications and their relationship to recombination rate during diploid and polyploid evolution in the *Aegilops-Triticum* alliance. Genetics 171: 323–332. pmid:15996988

Earl, Dent A., and vonHoldt, Bridgett M. 2012. "STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method." *Conservation genetics resources* 4.2: 359-361.

Endelman, Jeffrey B. "Ridge regression and other kernels for genomic selection with R package rrBLUP." *The Plant Genome* 4, no. 3 (2011): 250-255.

Feuillet, Catherine, Peter Langridge, and Robbie Waugh. "Cereal breeding takes a walk on the wild side." *Trends in genetics* 24, no. 1 (2008): 24-32.

Fournier-Level, A., A. Korte, M. D. Cooper, M. Nordborg, J. Schmitt, and A. M. Wilczek. "A map of local adaptation in Arabidopsis thaliana." *Science* 334, no. 6052 (2011): 86-89.

Glaubitz, Jeffrey C., Terry M. Casstevens, Fei Lu, James Harriman, Robert J. Elshire, Qi Sun, and Edward S. Buckler. "TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline." *PloS one* 9, no. 2 (2014): e90346.

Günther, Torsten, and Graham Coop. "Robust identification of local adaptation from allele frequencies." *Genetics* 195, no. 1 (2013): 205-220.

Hancock, Angela M., Benjamin Brachi, Nathalie Faure, Matthew W. Horton, Lucien B. Jarymowycz, F. Gianluca Sperone, Chris Toomajian, Fabrice Roux, and Joy Bergelson. "Adaptation to climate across the Arabidopsis thaliana genome." *Science* 334, no. 6052 (2011): 83-86.

Harlan, Jack R., and Daniel Zohary. "Distribution of wild wheats and barley." *Science* 153.3740 (1966): 1074-1080.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. and Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *International journal of climatology*, *25*(15), pp.1965-1978.

He, Fei, Raj Pasam, Fan Shi, Surya Kant, Gabriel Keeble-Gagnere, Pippa Kay, Kerrie Forrest et al. "Exome sequencing highlights the role of wild-relative introgression in shaping the adaptive landscape of the wheat genome." *Nature Genetics* 51, no. 5 (2019): 896-904.

Jordan, Katherine W., Shichen Wang, Fei He, Shiaoman Chao, Yanni Lun, Etienne Paux, Pierre Sourdille et al. "The genetic architecture of genome-wide recombination rate variation in allopolyploid wheat revealed by nested association mapping." *The Plant Journal* 95, no. 6 (2018): 1039-1054.

Kilian, Benjamin, Hakan Özkan, Oliver Deusch, Sieglinde Effgen, Andrea Brandolini, Jochen Kohl, William Martin, and Francesco Salamini. "Independent wheat B and G genome origins in outcrossing Aegilops progenitor haplotypes." *Molecular Biology and Evolution* 24, no. 1 (2007): 217-227.

Kölliker, R., Kraehenbuehl, R., Boller, B. and Widmer, F., 2006. Genetic diversity and pathogenicity of the grass pathogen Xanthomonas translucens pv. graminis. *Systematic and applied microbiology*, *29*(2), pp.109-119.

Lasky, Jesse R., David L. Des Marais, JOHN K. McKAY, James H. Richards, Thomas E. Juenger, and Timothy H. Keitt. "Characterizing genomic variation of Arabidopsis

thaliana: the roles of geography and climate." *Molecular Ecology* 21, no. 22 (2012): 5512-5529.

Lasky, Jesse R., Hari D. Upadhyaya, Punna Ramu, Santosh Deshpande, C. Tom Hash, Jason Bonnette, Thomas E. Juenger et al. "Genome-environment associations in sorghum landraces predict adaptive traits." *Science advances* 1, no. 6 (2015): e1400218.

Legendre P, Legendre L. *Numerical Ecology*. 2nd ed. Amsterdam: Elsevier, 1998. ISBN 978-0444892508.

Li, Lin-Feng, Ya-Ling Li, Yulin Jia, Ana L. Caicedo, and Kenneth M. Olsen. "Signatures of adaptation in the weedy rice genome." *Nature genetics* 49, no. 5 (2017): 811-814.

Li, Y. C., T. Fahima, M. S. Röder, V. M. Kirzhner, A. Beiles, A. B. Korol, and E. Nevo. "Genetic effects on microsatellite diversity in wild emmer wheat (Triticum dicoccoides) at the Yehudiyya microsite, Israel." *Heredity* 90, no. 2 (2003): 150-156.

Ligges, Uwe, and Martin Mächler. *Scatterplot3d-an r package for visualizing multivariate data*. No. 2002, 22. Technical Report, 2002.

Lin, Tao, Guangtao Zhu, Junhong Zhang, Xiangyang Xu, Qinghui Yu, Zheng Zheng, Zhonghua Zhang et al. "Genomic analyses provide insights into the history of tomato breeding." *Nature genetics* 46, no. 11 (2014): 1220-1226.

Lipka, Alexander E., Feng Tian, Qishan Wang, Jason Peiffer, Meng Li, Peter J. Bradbury, Michael A. Gore, Edward S. Buckler, and Zhiwu Zhang. "GAPIT: genome association and prediction integrated tool." *Bioinformatics* 28, no. 18 (2012): 2397-2399.

Luo, M-C., Z-L. Yang, F. M. You, T. Kawahara, J. G. Waines, and J. Dvorak. "The structure of wild and domesticated emmer wheat populations, gene flow between them, and the site of emmer domestication." *Theoretical and Applied Genetics* 114, no. 6 (2007): 947-959.

Manel, Stephanie, Stephane Joost, Bryan K. Epperson, Rolf Holderegger, Andrew Storfer, Michael S. Rosenberg, Kim T. Scribner, Aurelie Bonin, and MARIE-JOSÉE FORTIN. "Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field." *Molecular Ecology* 19, no. 17 (2010): 3760-3772.

Meuwissen THE, Hayes BJ, Goddard ME. “Prediction of total genetic value using genome-wide dense marker maps”, *Genetics* 157, no. 4(2001): 1819-29

Nevo, Eviatar. "Genetic diversity in wild cereals: regional and local studies and their bearing on conservation ex situ and in situ." *Genetic Resources and Crop Evolution* 45.4 (1998): 355-370.

Nevo, Eviatar, Abraham B. Korol, Avigdor Beiles, and Tzion Fahima. *Evolution of wild emmer and wheat improvement: population genetics, genetic resources, and genome*

organization of wheat's progenitor, Triticum dicoccoides. Springer Science & Business Media, 2013.

Oksanen, Jari. "Multivariate analysis of ecological communities in R: vegan tutorial." *R package version* 1, no. 7 (2011): 1-43.

Ortiz, Rodomiro, Hans-Joachim Braun, José Crossa, Jonathan H. Crouch, Guy Davenport, John Dixon, Susanne Dreisigacker et al. "Wheat genetic resources enhancement by the International Maize and Wheat Improvement Center (CIMMYT)."*Genetic Resources and Crop Evolution* 55, no. 7 (2008): 1095-1140.

Özkan, Hakan, et al. "Geographic distribution and domestication of wild emmer wheat (Triticum dicoccoides)." *Genetic Resources and Crop Evolution* 58.1 (2011): 11-53.

Palmgren, Michael G., Anna Kristina Edenbrandt, Suzanne Elizabeth Vedel, Martin Marchman Andersen, Xavier Landes, Jeppe Thulin Østerberg, Janus Falhof et al. "Are we ready for back-to-nature crop breeding?." *Trends in plant science* 20, no. 3 (2015): 155-164.

Peleg, Zvi, T. Fahima, S. Abbo, T. Krugman, E. Nevo, D. Yakir, and Y. Saranga. "Genetic diversity for drought resistance in wild emmer wheat and its ecogeographical associations." *Plant, Cell & Environment* 28, no. 2 (2005): 176-191.

Poland, Jesse A., Patrick J. Brown, Mark E. Sorrells, and Jean-Luc Jannink. "Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach." *PloS one* 7, no. 2 (2012).

Potchter, Oded, Daphne Goldman, David Kadish, and David Iluz. "The oasis effect in an extremely hot and arid climate: The case of southern Israel." *Journal of Arid Environments* 72, no. 9 (2008): 1721-1733.

Pritchard, Jonathan K., William Wen, and Daniel Falush. "Documentation for STRUCTURE software: Version 2." (2003).

Raftery, A.E., Zimmer, A., Frierson, D.M., Startz, R. and Liu, P., 2017. Less than 2° C warming by 2100 unlikely. *Nature Climate Change*.

Saintenac C, Jiang D, Wang S, Akhunov E. Sequence-based mapping of the polyploid wheat genome. G3 (Bethesda). 2013; 3:1105–14.

Sharbel, Timothy F., Bernhard Haubold, and Thomas Mitchell-Olds. "Genetic isolation by distance in Arabidopsis thaliana: biogeography and postglacial colonization of Europe." *Molecular Ecology* 9, no. 12 (2000): 2109-2118.

Swarts, Kelly, Huihui Li, J. Alberto Romero Navarro, Dong An, Maria Cinta Romay, Sarah Hearne, Charlotte Acharya et al. "Novel methods to optimize genotypic imputation for low-coverage, next-generation sequence data in crop plants." *The Plant Genome* 7, no. 3 (2014).

Trethowan, R. M., and A. Mujeeb-Kazi. "Novel germplasm resources for improving environmental stress tolerance of hexaploid wheat." *Crop science* 48, no. 4 (2008): 1255-1265.

Turner, Stephen D. "qqman: an R package for visualizing GWAS results using QQ and manhattan plots." *Biorxiv* (2014): 005165.

Wang, Qishan, Feng Tian, Yuchun Pan, Edward S. Buckler, and Zhiwu Zhang. "A SUPER powerful method for genome wide association study." *PloS one* 9, no. 9 (2014): e107684.

Will cox, George. "The distribution, natural habitats and availability of wild cereals in relation to their domestication in the Near East: multiple events, multiple centres." *Vegetation History and Archaeobotany* 14.4 (2005): 534-541.

Yu, Jianming, Gael Pressoir, William H. Briggs, Irie Vroh Bi, Masanori Yamasaki, John F. Doebley, and Michael D. McMullen et al. "A unified mixed-model method for association mapping that accounts for multiple levels of relatedness." *Nature genetics* 38, no. 2 (2006): 203-208.

Zheng, Xiuwen, David Levine, Jess Shen, Stephanie M. Gogarten, Cathy Laurie, and Bruce S. Weir. "A high-performance computing toolset for relatedness and principal component analysis of SNP data." *Bioinformatics* 28, no. 24 (2012): 3326-3328.

Zohary, Daniel. "Centres of diversity and centers of origin." *Centres of Diversity and Centers of Origin.* (1970).

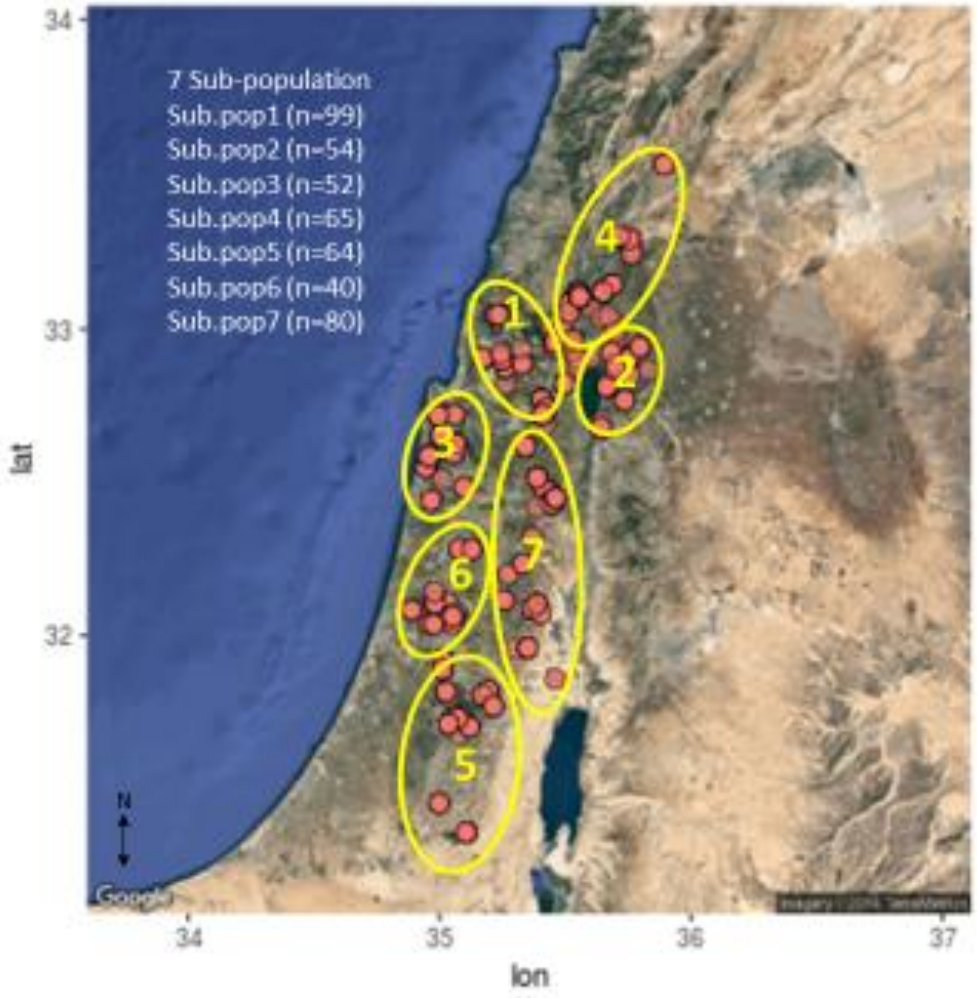


Figure 2.1. Geographic distribution of wild emmer accessions showing 7 sub-grouping of Wild Emmer accessions based on topography and climate niches in the collection site.

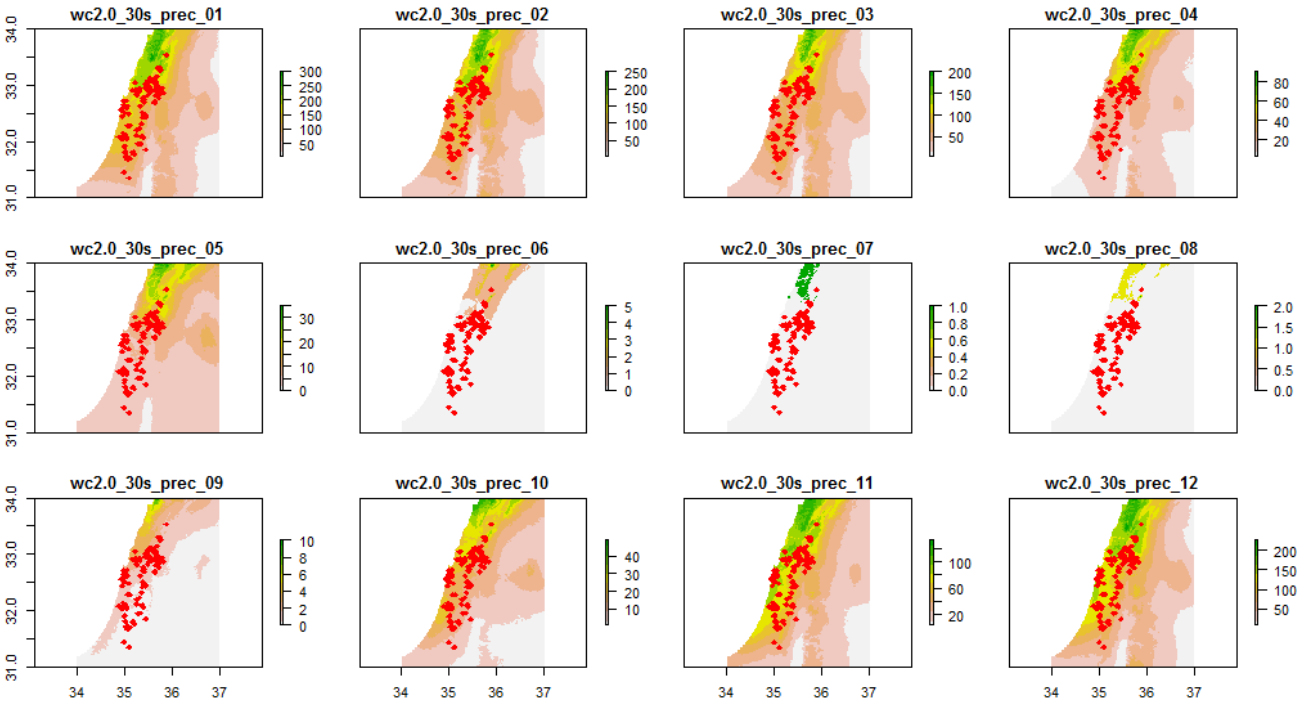


Figure 2.2. Geographic distribution of the wild emmer accessions collected throughout Israel plotted along a precipitation gradient.

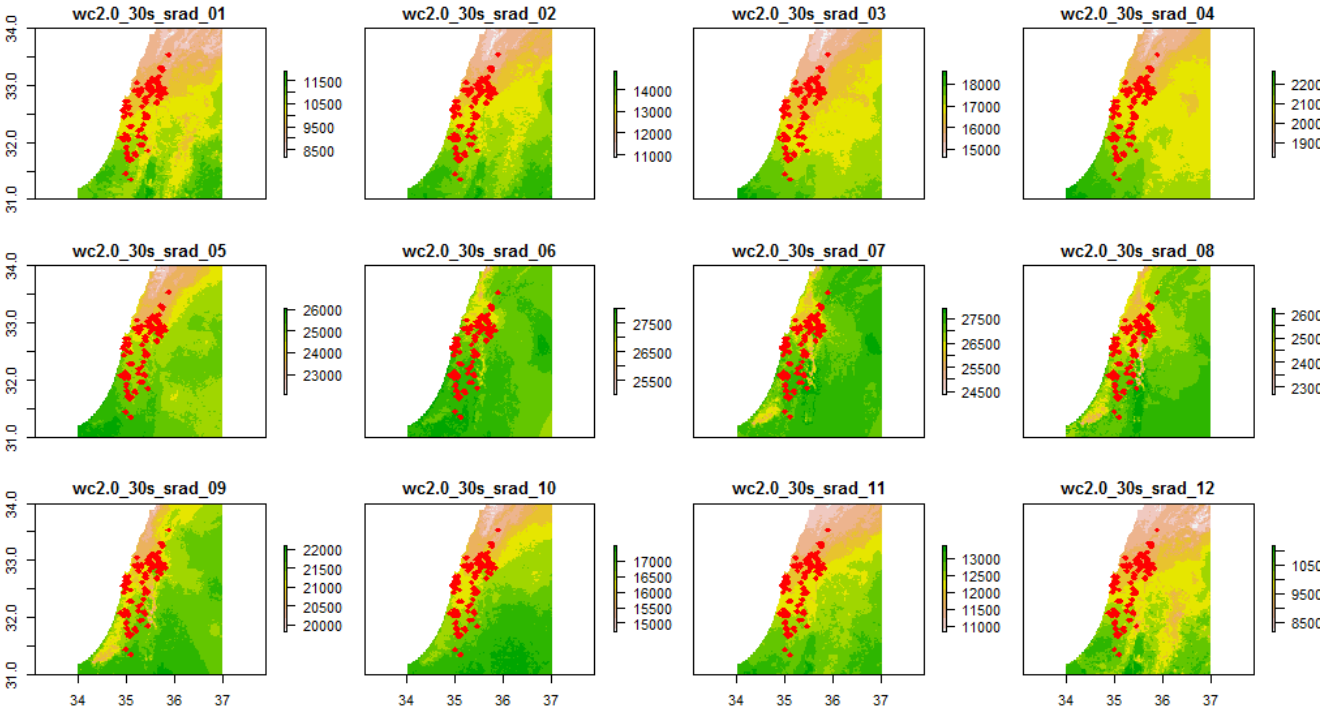


Figure 2.3. Geographic distribution Israel's wild emmer accessions plotted along the solar radiation gradient.

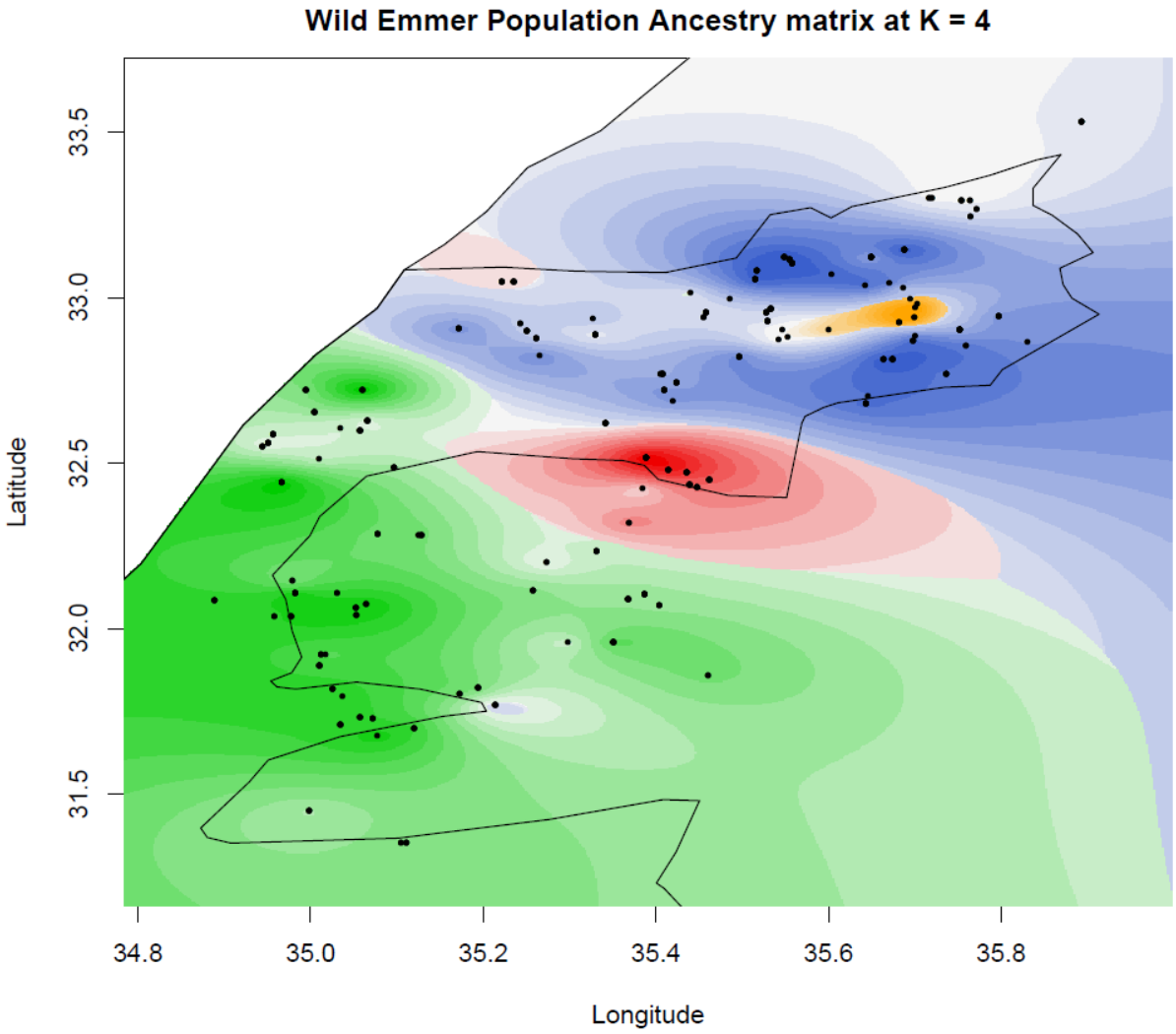


Figure 2.4. Genetic ancestry and population structure of wild emmer population in Israel at ancestry coefficients (K) = 4.

Note: Each color represents the accession in different ancestry matrix whereas each black dot represents an accession's collection site.

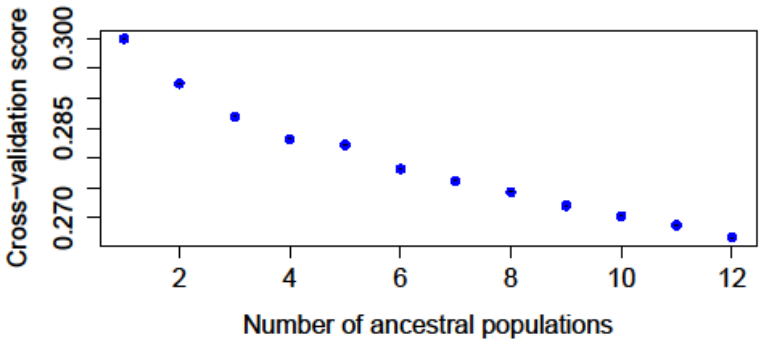


Figure 2.5. Cross-validation score identifying optimum number of K.

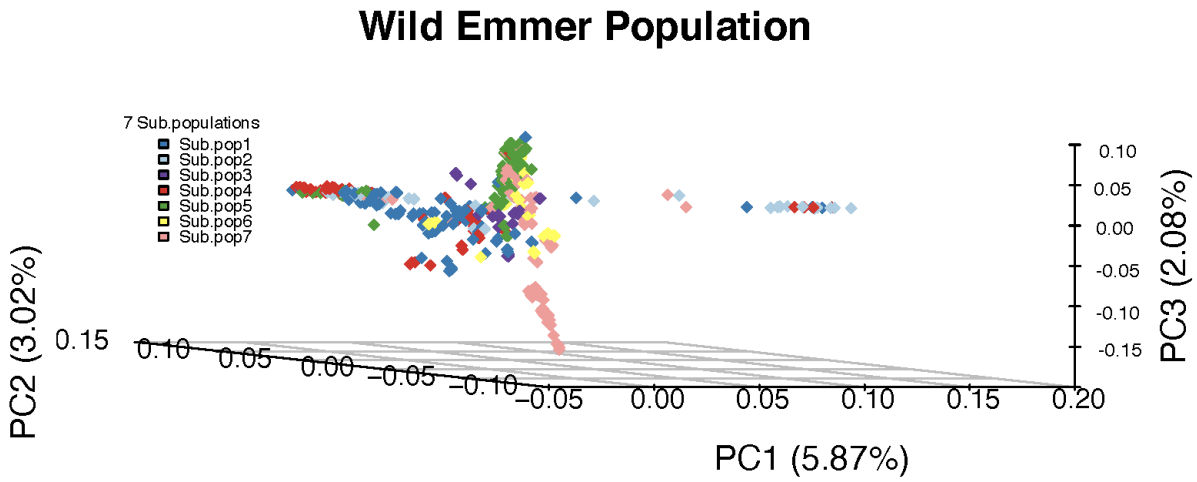
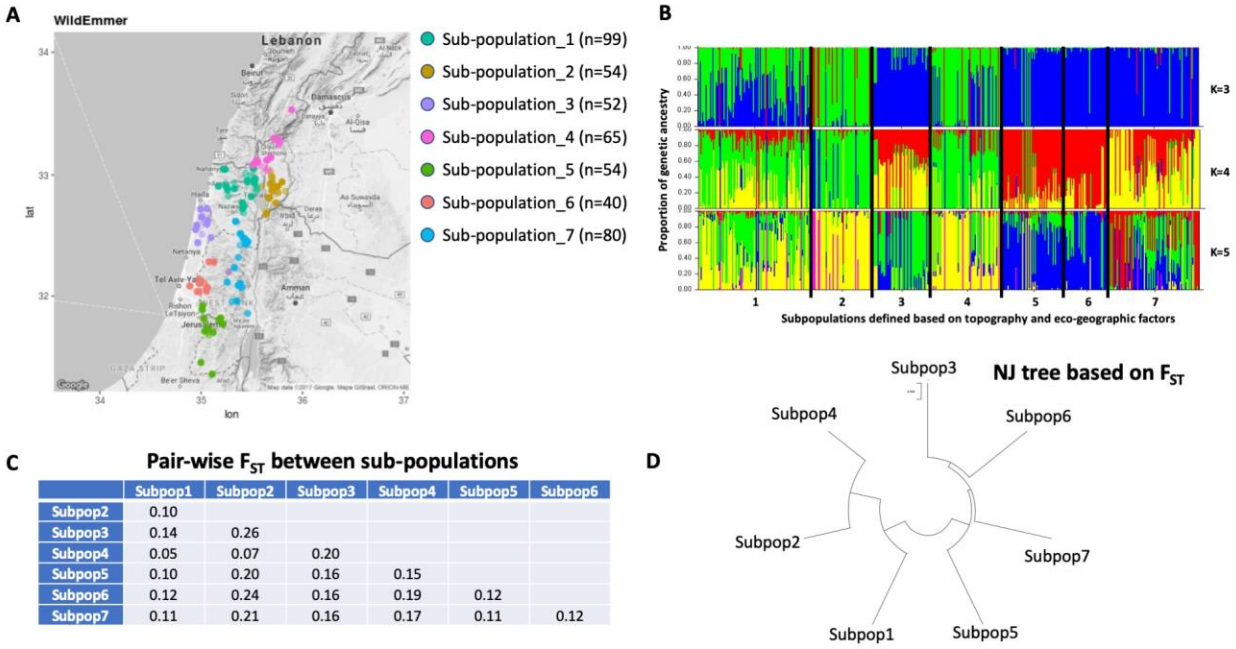


Figure 2.6. Principle component analysis showing accessions clustering based on 7 climatic sub-population clusters.

Note: Each dot with different color represents an accession in respective sub-population.



	Subpop1	Subpop2	Subpop3	Subpop4	Subpop5	Subpop6
Subpop2	0.10					
Subpop3	0.14	0.26				
Subpop4	0.05	0.07	0.20			
Subpop5	0.10	0.20	0.16	0.15		
Subpop6	0.12	0.24	0.16	0.19	0.12	
Subpop7	0.11	0.21	0.16	0.17	0.11	0.12

Figure 2.7. Population structure of wild emmer accessions from Israel.

[Note: **A**. Seven sub-populations of wild emmer defined on topography and distribution of eco-geographic factors. These populations have been used for Bayenv analyses to associate SNPs with variation in climatic and bio-climatic factors. **B**. Proportions of the genetic ancestry of wild emmer accessions in K populations estimated using Structure. Seven sub-populations of wild emmer grouped according to their geographic distribution are indicated at the bottom of the plot. **C**. Pair-wise estimates of F_{ST} for seven sub-populations. **D**. Neighbor-joining tree constructed using Pair-wise estimates of F_{ST}.]

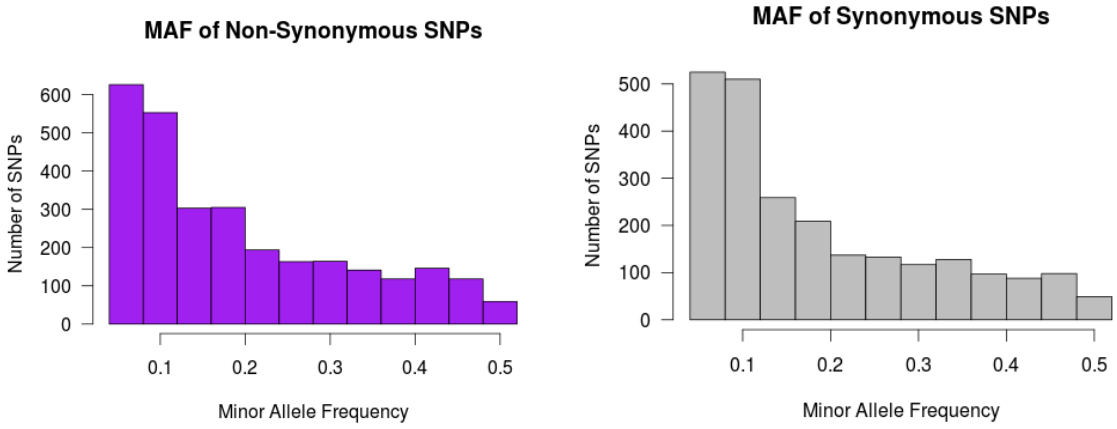


Figure 2.8. Minor allele frequency of synonymous in the left colored in purple color and non-synonymous SNPs in the right colored in grey color.

Number of effects by type and region

Type

Type (alphabetical order)	Count	Percent
3_prime_UTR_variant	19,664	9.236%
5_prime_UTR_premature_start_codon_gain_variant	1,165	0.547%
5_prime_UTR_variant	6,977	3.277%
downstream_gene_variant	58,004	27.245%
initiator_codon_variant	28	0.013%
intergenic_region	19,233	9.034%
intron_variant	23,464	11.021%
missense_variant	21,792	10.236%
non_canonical_start_codon	5	0.002%
non_coding_transcript_exon_variant	4	0.002%
splice_acceptor_variant	714	0.335%
splice_donor_variant	335	0.157%
splice_region_variant	4,329	2.033%
start_lost	18	0.008%
stop_gained	1,046	0.491%
stop_lost	104	0.049%
stop_retained_variant	37	0.017%
synonymous_variant	17,617	8.275%
upstream_gene_variant	38,362	18.019%

Region

Type (alphabetical order)	Count	Percent
DOWNSTREAM	58,004	27.846%
EXON	39,832	19.122%
INTERGENIC	19,233	9.233%
INTRON	21,155	10.156%
SPLICE_SITE_ACCEPTOR	714	0.343%
SPLICE_SITE_DONOR	335	0.161%
SPLICE_SITE_REGION	2,863	1.374%
UPSTREAM	38,362	18.416%
UTR_3_PRIME	19,664	9.44%
UTR_5_PRIME	8,142	3.909%

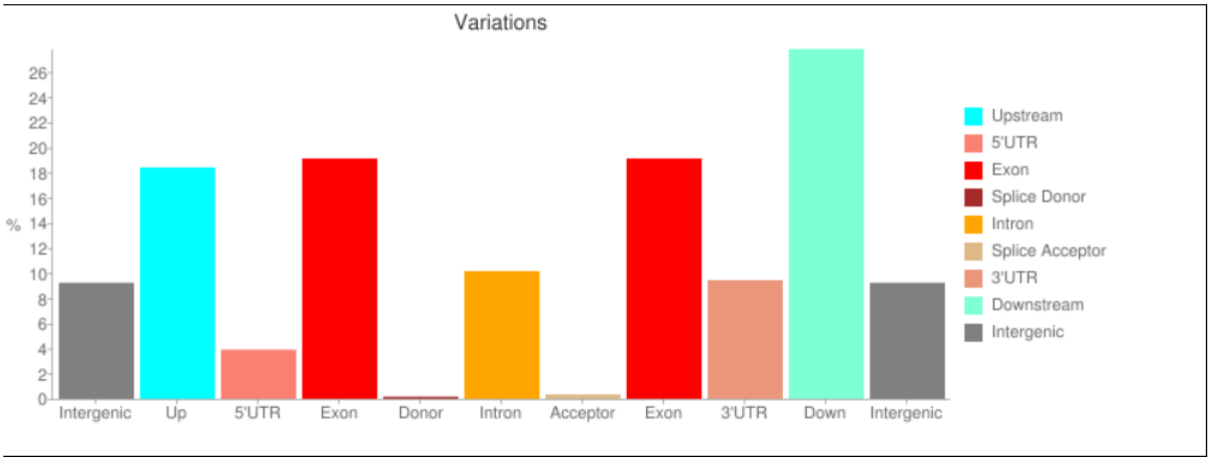


Figure 2.9. Predicted effect of the SNPs based on "SnpEff" which shows different types of effects and the regions in the genome.

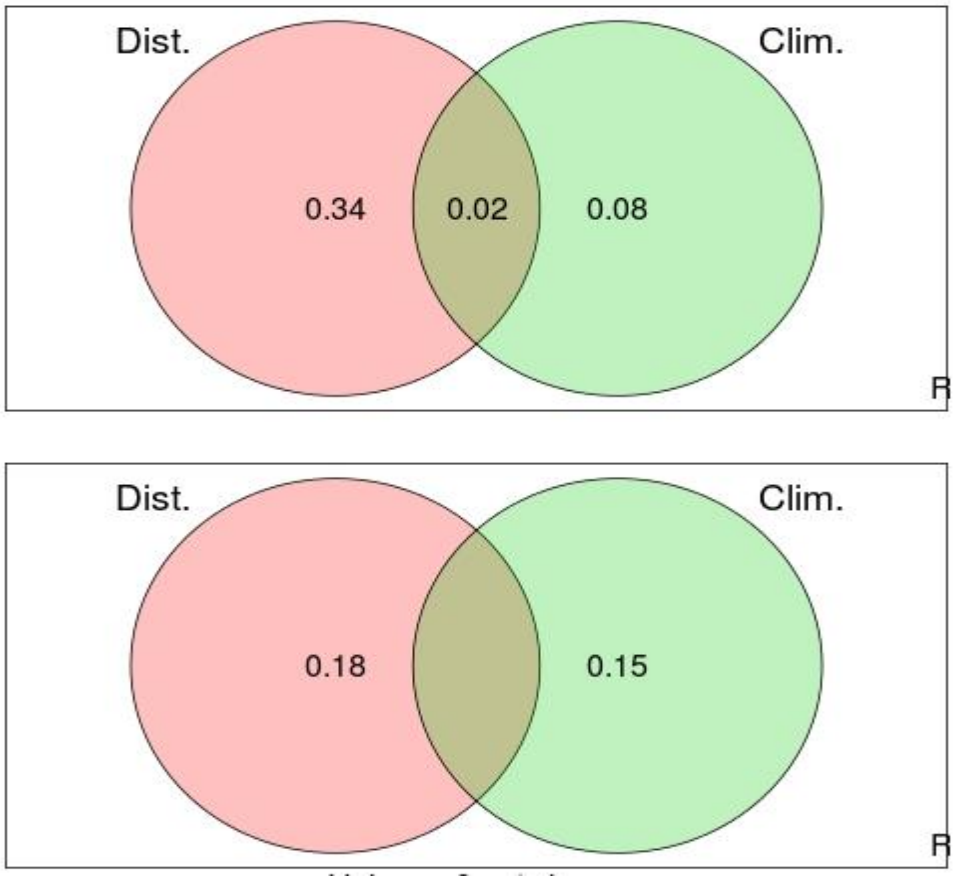


Figure 2.10. Venn diagram showing the proportion of SNPs variation explained by geography and climate.

Note: The first Venn diagram shows the variation explained by geography and climate when all (103) climatic variables were used. The second Venn diagram show the proportion of SNP variation explained by geography and climate when only 25 non-collinear climatic variables were used.

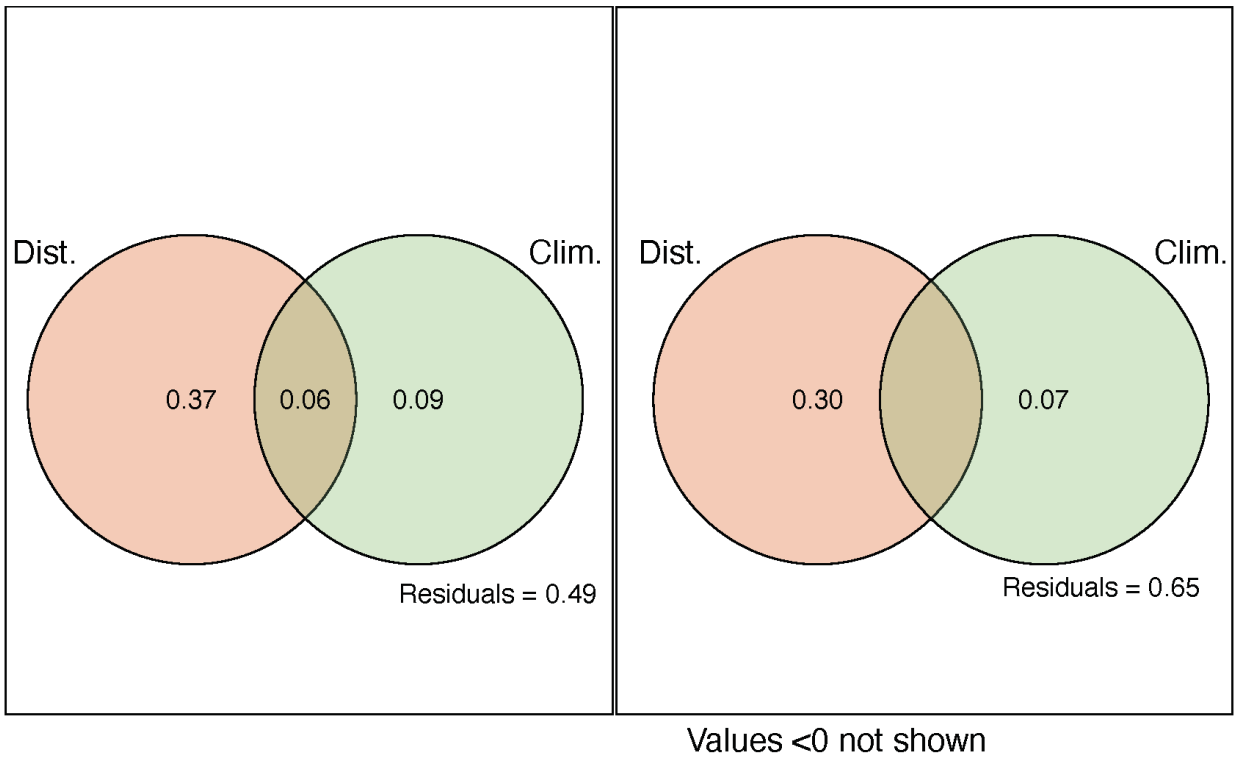


Figure 2.11. Venn diagram showing the proportion of SNPs variation explained by geography and climate in GBS data (in the left) and 90K SNPs data (in the right).

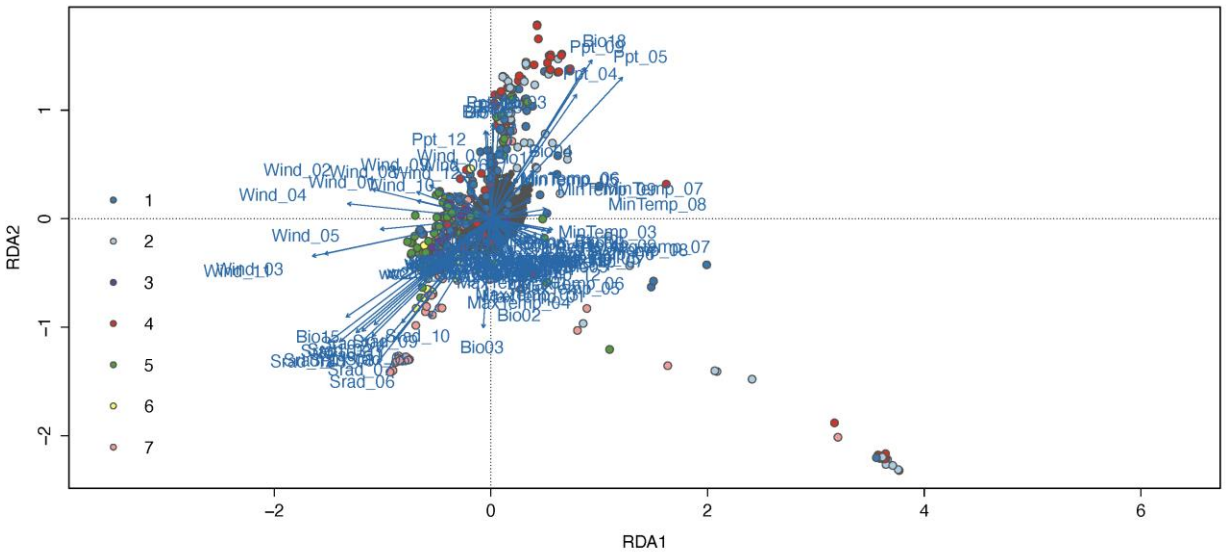


Figure 2.12. Bi-plot from RDA analysis showing the top climatic factors driving local adaptation in wild emmer population for all 103 climatic variables.

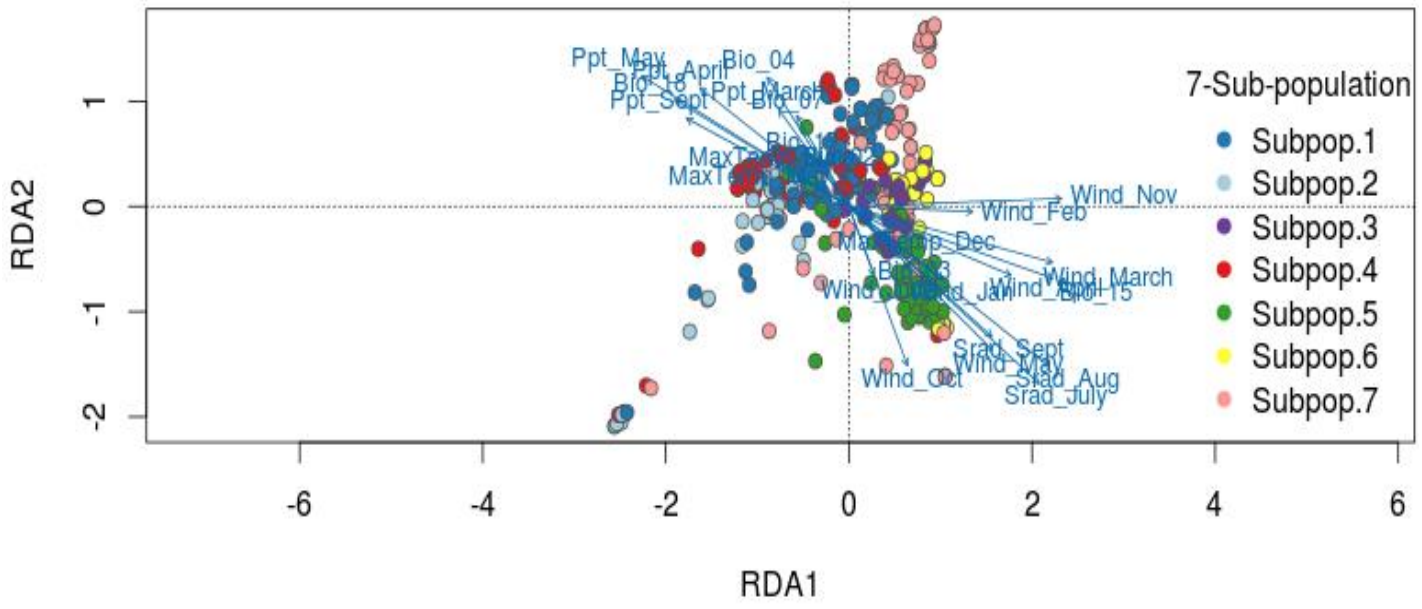


Figure 2.13. Redundancy analysis showing the top climatic factors driving local adaptation in wild emmer population out of 25 non-collinear variables.

Note: Each dot represents a genotype that is color coded based on its designation to respective sub-groupings.

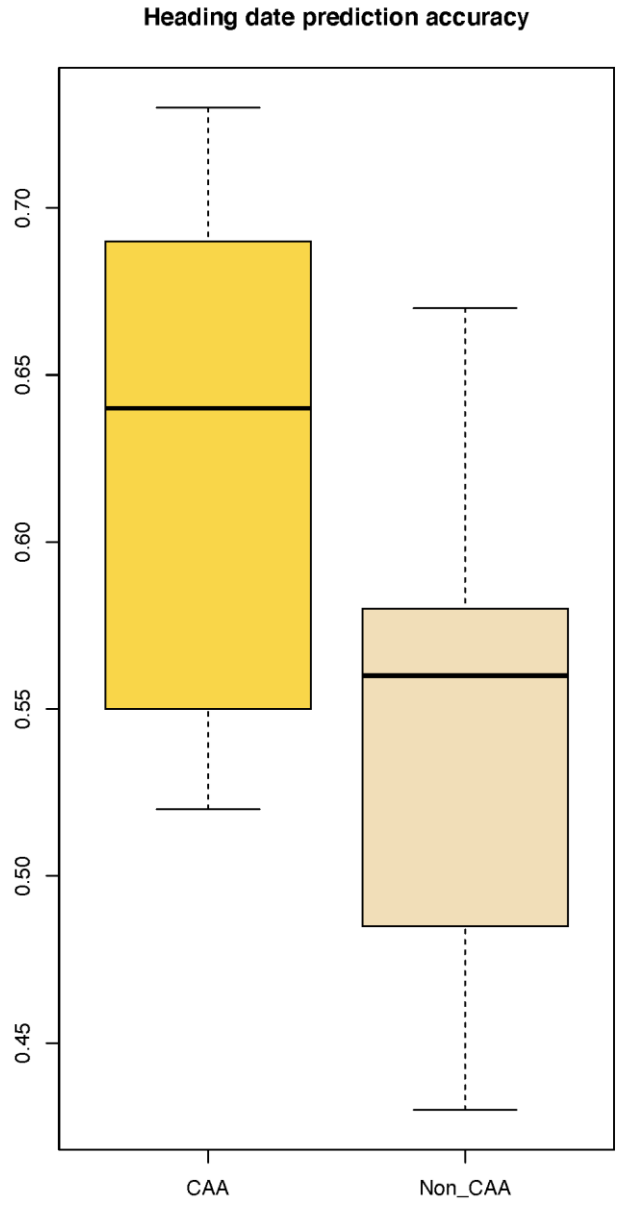


Figure 2.14. Genomic prediction of heading date phenotype in wild emmer accessions using candidate SNPs identifies through RDA analysis.

Note: The golden bar represents the candidate SNPs identified using 25 unique bio-climate variables through RDA and light gray bar represents the prediction of heading date using same number of random SNPs.

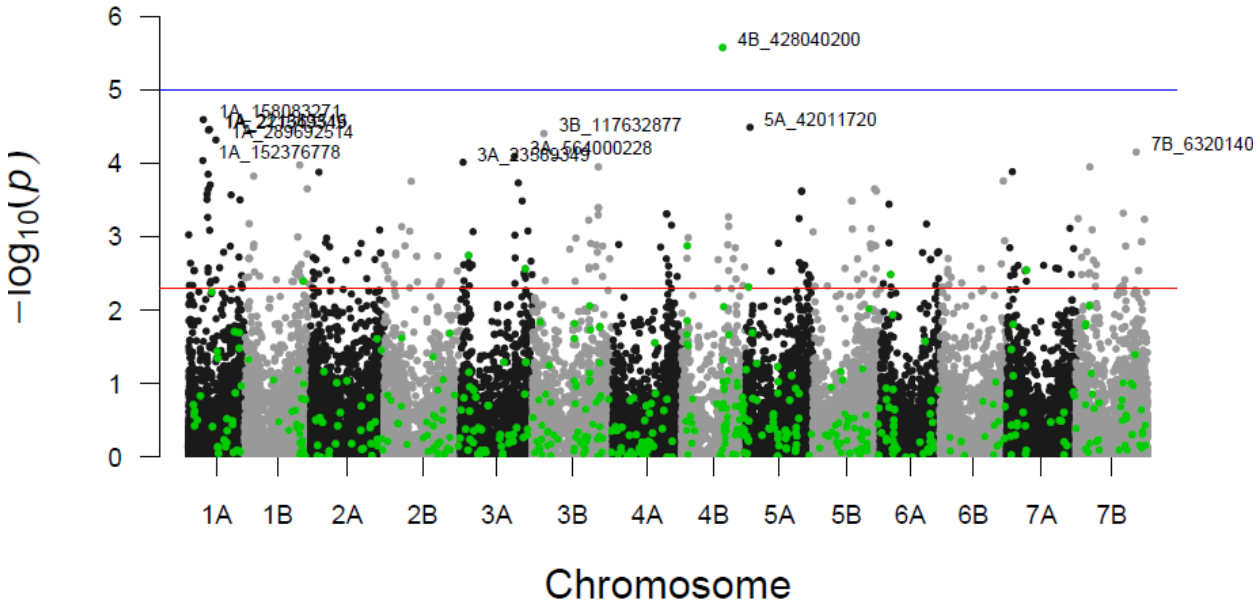


Figure 2.15. Manhattan plot showing the genome wide association study for heading date phenotype.

Note: each black dot represents a SNP marker. Each green dot represents the SNPs that are identified as the markers that are highly associated (outlier) from the RDA analysis. The Blue and red horizontal line represents the genomewide and suggestive line respectively.

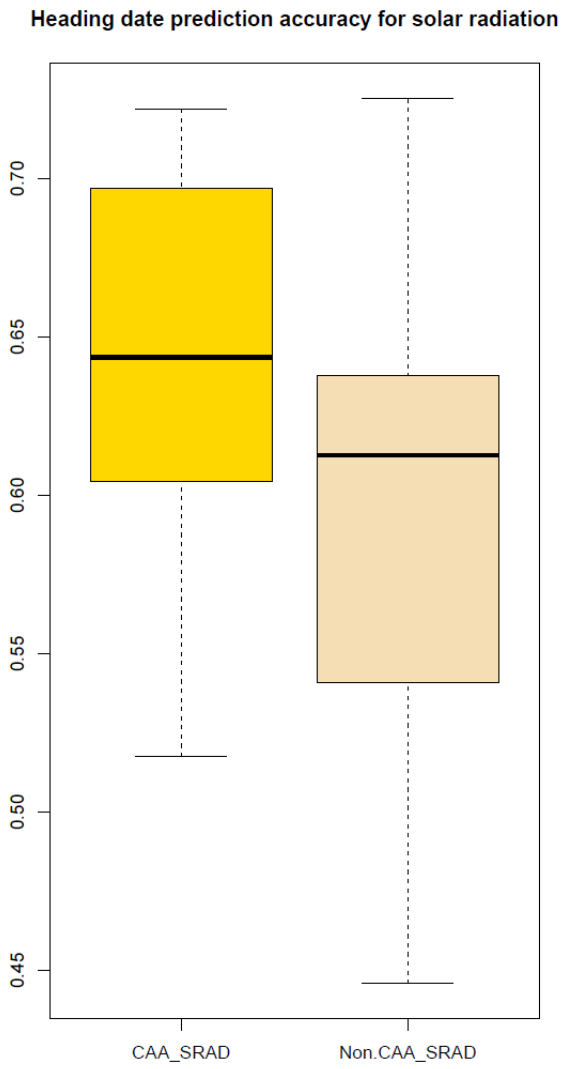


Figure 2.16. Genomic prediction of heading date phenotype in wild emmer accessions using candidate SNPs that are significant associated with solar radiation, identifies through BAYENV analysis.

Note: The golden bar represents the prediction of candidate SNPs identified using BAYENV analysis and light gray bar represents the prediction of heading date using same number of random SNPs.

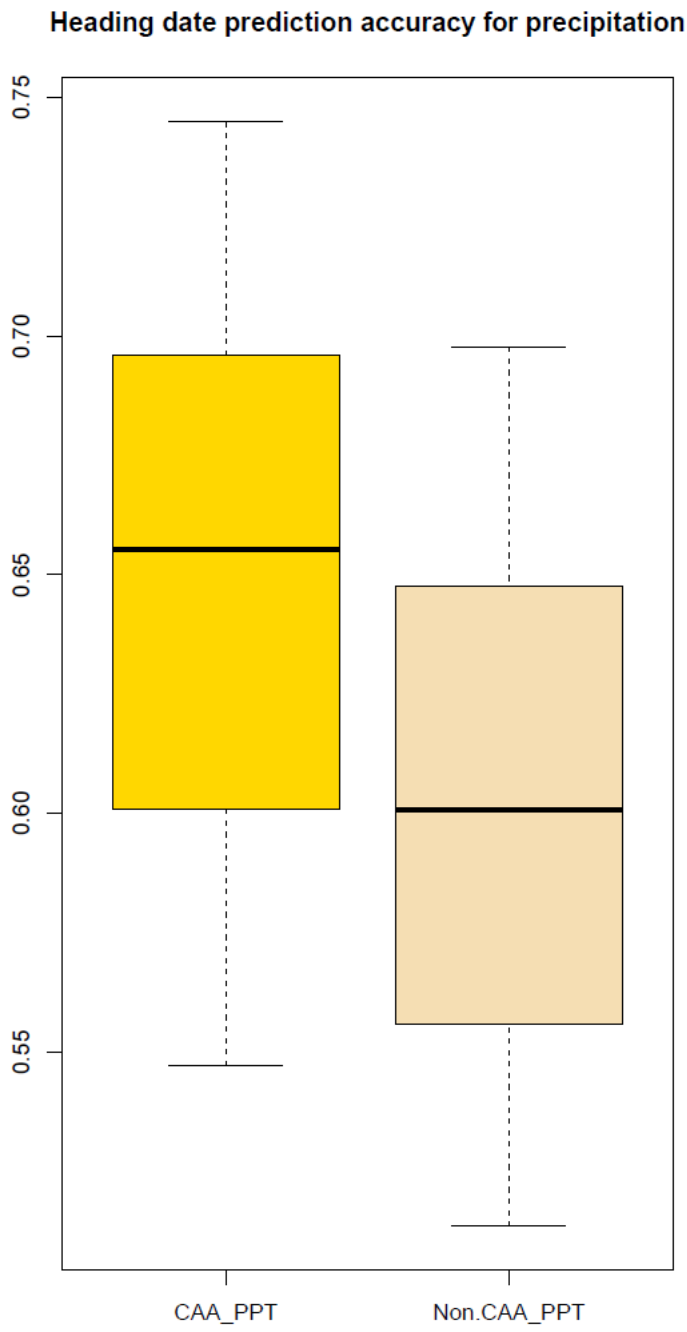


Figure 2.17. Genomic prediction of heading date phenotype in wild emmer accessions using candidate SNPs that are significant associated with precipitation, identifies through BAYENV analysis.

Note: The golden bar represents the prediction of candidate SNPs identified using BAYENV analysis and light gray bar represents the prediction of heading date using same number of random SNPs.

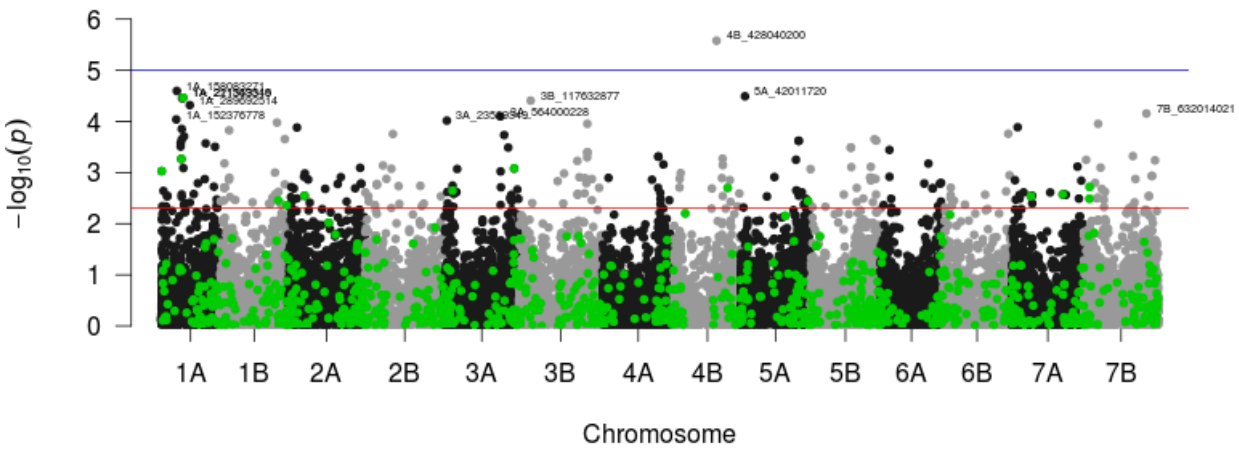


Figure 2.18. Manhattan plot showing the genome wide association study for heading date phenotype with SNPs identified through Bayenv analysis for solar radiation in the month of July highlighted in green.

Note: each black dot represents a SNP marker. Each green dot represents the SNPs that are identified as top 1% SNPs from the Bayenv analysis

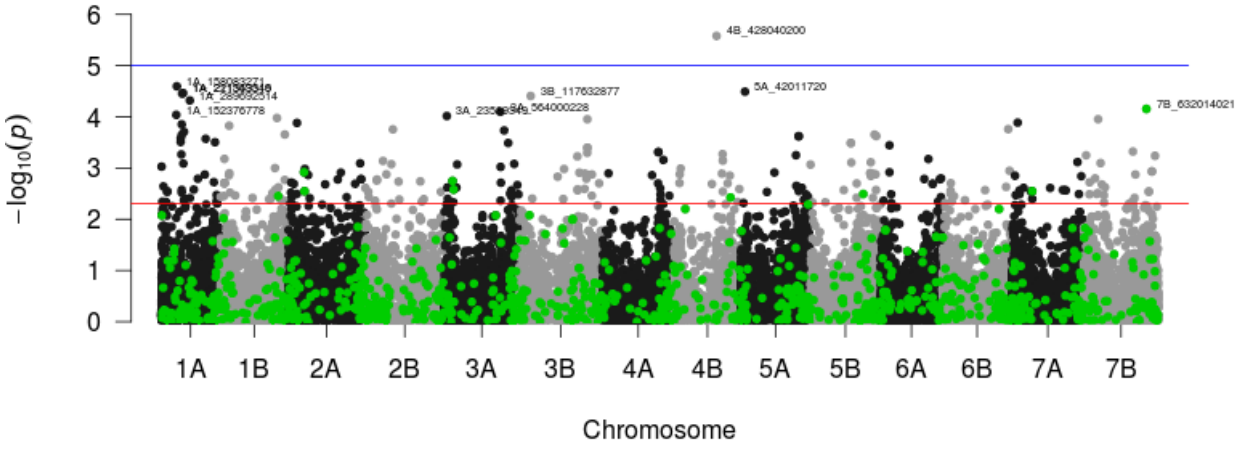


Figure 2.19. Manhattan plot showing the genome wide association study for heading date phenotype with SNPs identified through Bayenv analysis for precipitation in the month of May.

Note: each black dot represents a SNP marker. Each green dot represents the SNPs that are identified as top 1% SNPs from the Bayenv analysis.

Table 2.1. Climatic variation across seven sub populations calculated after normalization. The color pattern following red, orange, yellow, and green represents the climatic values in descending order.

Climatic Variables	Sub.pop1	Sub.pop2	Sub.pop3	Sub.pop4	Sub.pop5	Sub.pop6	Sub.pop7
Annual Mean Temp	18.35	18.66	19.14	17.42	18.00	18.87	18.82
Mean Diurnal Range ▲	10.34	11.36	9.95	11.19	10.92	10.54	11.74
Isothermality ▲	39.70	41.57	39.48	40.36	41.89	41.34	42.47
Temp Seasonality ▲	596.56	607.37	577.26	624.89	572.86	566.90	603.38
Max Temp of Warmest Month	30.71	31.36	31.50	30.64	30.12	31.14	31.60
Min Temp of Coldest Month	4.68	4.04	6.30	2.90	4.05	5.64	4.00
Temp Annual Range ▲	26.03	27.32	25.20	27.73	26.06	25.50	27.60
Mean Temp of Wettest Quarter	10.88	11.05	11.85	9.56	10.73	11.73	11.18
Mean Temp of Driest Quarter	25.05	25.51	25.56	24.51	24.31	25.09	25.56
Mean Temp of Warmest Quarter	25.24	25.64	25.79	24.66	24.53	25.45	25.69
Mean Temp of Coldest Quarter	10.88	11.05	11.85	9.56	10.73	11.73	11.18
Annual Ppt	656.03	490.42	604.10	688.76	477.34	578.87	465.86
Ppt of Wettest Month	167.53	121.72	155.24	173.45	119.61	149.31	115.93
Ppt Seasonality (CV) ▲	109.52	108.45	114.88	108.12	111.36	115.12	110.39
Ppt of Wettest Quarter	426.43	317.93	404.38	442.94	309.90	388.77	304.93
Ppt of Driest Quarter ▲	0.06	0.04	0.00	0.32	0.00	0.00	0.00
Ppt of Warmest Quarter ▲	1.83	0.89	1.38	1.58	0.19	0.54	0.08
Ppt of Coldest Quarter	426.02	317.93	404.38	442.94	309.90	388.77	304.93
Ppt_Jan (mm)	167.53	121.72	155.24	173.45	119.61	149.31	115.93
Ppt_Feb (mm)	126.66	95.35	105.62	135.86	90.51	104.03	92.05
Ppt_March (mm) ▲	84.61	69.56	63.52	95.18	67.40	67.00	65.03
Ppt_April (mm) ▲	32.98	25.65	21.62	39.65	19.48	21.00	20.84
Ppt_May (mm) ▲	10.56	9.61	5.72	14.27	4.04	4.74	5.69
Ppt_Sept (mm) ▲	1.88	0.89	1.38	1.59	0.19	0.54	0.18
Ppt_Oct (mm)	20.98	13.68	19.17	19.88	13.86	16.92	12.40
Ppt_Novermber (mm)	78.54	53.05	88.31	74.92	62.47	79.90	56.81
Ppt_Dec(mm)	132.24	100.86	143.52	133.62	99.78	135.44	96.95
Min Temp_Jan (°C)	4.68	4.04	6.30	2.90	4.05	5.64	4.00
Min Temp_Feb (°C)	5.27	4.53	7.04	3.61	4.85	6.34	4.71
Min Temp_March (°C)	7.49	6.89	9.20	6.14	6.93	8.40	7.00
Min Temp_April (°C)	11.35	10.92	12.92	10.09	10.98	12.19	11.00
Min Temp_May (°C)	15.15	14.94	16.18	14.26	14.40	15.42	14.84
Min Temp_June (°C)	18.32	17.99	19.50	17.40	17.22	18.38	17.89
Min Temp_July (°C)	20.67	20.91	20.89	19.68	19.79	20.39	20.77
Min Temp_Aug (°C)	20.94	21.16	21.21	19.89	19.96	20.73	21.01
Min Temp_Sept (°C)	19.41	19.54	19.82	18.16	18.69	19.62	19.40
Min Temp_Oct (°C)	16.47	16.39	17.26	14.80	16.03	16.86	16.63
Min Temp_Novermber (°C)	11.57	11.55	12.23	9.85	11.11	11.90	11.47

Min Temp_Dec (°C)	6.90	6.97	7.45	5.15	6.48	7.33	6.65
WVP_Jan (kPa)	0.86	0.86	0.95	0.80	0.83	0.92	0.86
WVP_Feb (kPa)	0.89	0.87	0.99	0.81	0.86	0.96	0.89
WVP_March (kPa)	0.97	0.95	1.09	0.87	0.94	1.05	0.97
WVP_April (kPa)	1.09	1.07	1.24	0.98	1.03	1.18	1.08
WVP_May (kPa)	1.29	1.24	1.50	1.14	1.22	1.42	1.25
WVP_June (kPa)	1.61	1.54	1.89	1.39	1.52	1.79	1.56
WVP_July (kPa)	1.90	1.83	2.20	1.66	1.78	2.08	1.84
WVP_Aug (kPa)	1.98	1.92	2.28	1.75	1.88	2.17	1.94
WVP_Sept (kPa)	1.77	1.71	2.04	1.55	1.73	1.96	1.76
WVP_Oct (kPa)	1.45	1.40	1.67	1.28	1.43	1.61	1.44
WVP_Novermber (kPa)	1.09	1.07	1.24	0.99	1.08	1.19	1.10
WVP_Dec (kPa)	0.95	0.93	1.05	0.87	0.92	1.02	0.95
Max Temp_Jan (°C)	15.57	16.56	15.91	14.57	15.82	16.32	16.88
Max Temp_Feb (°C)	16.16	17.04	16.67	15.27	16.63	17.04	17.61
Max Temp_March (°C)	18.40	19.38	18.79	17.76	18.70	19.08	19.88
Max Temp_April (°C)	22.24	23.42	22.53	21.77	22.74	22.88	23.90
Max Temp_May (°C) ▲	26.03	27.45	25.80	25.92	26.18	26.11	27.73
Max Temp_June (°C) ▲	29.21	30.51	29.10	29.03	29.00	29.08	30.77
Max Temp_July (°C)	30.45	31.13	31.20	30.41	29.85	30.80	31.35
Max Temp_Aug (°C)	30.70	31.36	31.50	30.62	30.05	31.14	31.59
Max Temp_Sept (°C)	29.19	29.74	30.11	28.88	28.77	30.03	29.99
Max Temp_Oct (°C)	26.26	26.60	27.57	25.51	26.10	27.27	27.22
Max Temp_Novermber (°C)	21.34	21.77	22.53	20.60	21.19	22.27	22.05
Max Temp_Dec (°C)	16.68	17.17	17.74	15.89	16.56	17.71	17.26
Average Temp_Jan (°C)	10.12	10.29	11.09	8.73	9.94	10.97	10.45
Average Temp_Feb (°C)	10.72	10.81	11.86	9.45	10.72	11.68	11.15
Average Temp_March (°C)	12.96	13.13	13.97	11.94	12.80	13.76	13.44
Average Temp_April (°C)	16.80	17.17	17.74	15.95	16.87	17.54	17.45
Average Temp_May (°C)	20.60	21.19	20.96	20.06	20.30	20.78	21.28
Average Temp_June (°C)	23.77	24.24	24.30	23.20	23.10	23.75	24.34
Average Temp_July (°C)	25.56	26.01	26.03	25.05	24.81	25.59	26.06
Average Temp_Aug (°C)	25.81	26.26	26.36	25.23	25.02	25.92	26.31
Average Temp_Sept (°C)	24.31	24.63	24.98	23.50	23.74	24.83	24.70
Average Temp_Oct (°C)	21.36	21.48	22.42	20.13	21.08	22.05	21.93
(°C)Average Temp_Novermber	16.45	16.66	17.34	15.24	16.14	17.08	16.75
Average Temp_Dec (°C)	11.81	12.08	12.59	10.51	11.51	12.51	11.95
Wind Speed_Jan (m s-1) ▲	2.91	2.46	3.47	2.80	2.85	3.12	2.55
Wind Speed_Feb (m s-1) ▲	2.95	2.46	3.15	2.86	2.92	2.97	2.50
Wind Speed_March (m s-1) ▲	3.14	2.73	3.52	3.00	3.31	3.31	2.94
Wind Speed_April (m s-1) ▲	3.00	2.77	3.28	2.97	3.11	3.22	2.86
Wind Speed_May (m s-1) ▲	3.01	2.94	3.14	2.90	3.21	3.22	2.96
Wind Speed_June (m s-1) ▲	3.44	3.47	3.45	3.47	3.50	3.62	3.22

Wind Speed_July (m s-1) ▲	3.52	3.35	3.76	3.62	3.44	3.64	3.17
Wind Speed_Aug (m s-1)	3.20	2.94	3.42	3.26	3.23	3.38	2.90
Wind Speed_Sept (m s-1)	3.01	2.95	3.26	2.99	3.13	3.23	2.78
Wind Speed_Oct (m s-1)	2.24	2.31	2.52	2.41	2.52	2.50	2.22
Wind Speed_Novermber (m s-1)	2.45	1.89	2.91	2.31	2.63	2.73	2.26
Wind Speed_Dec (m s-1)	2.61	2.12	3.19	2.71	2.51	2.72	2.18
Solar Rad_Jan (m s-1)	9586.24	9790.35	9854.48	9434.26	10629.39	10182.77	10334.34
Solar Rad_Feb (m s-1)	12485.50	12790.70	12824.72	12254.85	13540.42	13145.77	13247.28
Solar Rad_March (m s-1) ▲	16058.02	16366.00	16435.59	15873.86	17051.84	16719.28	16703.36
Solar Rad_April (m s-1) ▲	20300.08	20522.04	20714.00	20033.21	21194.47	20977.03	20868.41
Solar Rad_May (m s-1) ▲	24038.73	24483.86	24433.93	23763.89	25006.59	24804.36	24723.55
Solar Rad_June (m s-1) ▲	26984.62	27251.18	27197.24	26957.11	27671.80	27530.54	27494.40
Solar Rad_July (m s-1) ▲	26506.87	26833.32	26838.45	26612.55	27401.12	27140.95	27145.64
Solar Rad_Aug (m s-1)	24620.66	24843.37	24870.28	24675.58	25391.06	25198.44	25146.90
Solar Rad_Sept (m s-1)	20906.09	21232.32	21040.34	20922.50	21413.82	21236.10	21294.91
Solar Rad_Oct (m s-1)	15986.17	16383.42	16140.21	15896.32	16627.80	16354.79	16545.60
Solar Rad_Novermber (m s-1)	11839.12	12111.91	12008.28	11703.97	12465.30	12225.77	12346.50
Solar Rad_Dec (m s-1)	9287.04	9457.19	9529.62	9138.70	10221.18	9856.87	9943.11

Abbreviations

BIO1 = Annual Mean Temperature

BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))

BIO3 = Isothermality (BIO2/BIO7) (×100)

BIO4 = Temperature Seasonality (standard deviation ×100)

BIO5 = Max Temperature of Warmest Month

BIO6 = Min Temperature of Coldest Month

BIO7 = Temperature Annual Range (BIO5-BIO6)

BIO8 = Mean Temperature of Wettest Quarter

BIO9 = Mean Temperature of Driest Quarter

BIO10 = Mean Temperature of Warmest Quarter

BIO11 = Mean Temperature of Coldest Quarter

BIO12 = Annual Precipitation

BIO13 = Precipitation of Wettest Month

BIO14 = Precipitation of Driest Month

BIO15 = Precipitation Seasonality (Coefficient of Variation)

BIO16 = Precipitation of Wettest Quarter

BIO17 = Precipitation of Driest Quarter

BIO18 = Precipitation of Warmest Quarter

Table 2.2. Wild emmer accessions collected across Israel with its GPS coordinates and sub-population designation.

Sample	Ascension number	Location	7sub-pop	7 sub-pop naming	Longitude	Latitude
GBS_PstI_F89	TD-62-2	Zefat	1	LowerGalilee	35.52793	32.95368
GBS_PstI_F2	TD-72-12	Zefat	1	LowerGalilee	35.52793	32.95368
GBS_PstI_F10	TD-81-21	Zefat	1	LowerGalilee	35.52793	32.95368
GBS_PstI_F18	TD-92-32	Zefat	1	LowerGalilee	35.52793	32.95368
GBS_PstI_F26	TD-99-39	Zefat	1	LowerGalilee	35.52793	32.95368
GBS_PstI_F34	TD-106-46	Zefat	1	LowerGalilee	35.52793	32.95368
GBS_PstI_F42	TD-132-10	Ammi'ad	1	LowerGalilee	35.52906	32.93023
GBS_PstI_F50	TD-150-28	Ammi'ad	1	LowerGalilee	35.52906	32.93023
GBS_PstI_F58	TD-155-33	Ammi'ad	1	LowerGalilee	35.52906	32.93023
GBS_PstI_F66	TD-169-47	Ammi'ad	1	LowerGalilee	35.52906	32.93023
GBS_PstI_F74	TD-177-56	Ammi'ad	1	LowerGalilee	35.52906	32.93023
GBS_PstI_F82	TD-179-58	Ammi'ad	1	LowerGalilee	35.52906	32.93023
GBS_PstI_F90	TD-183-3	Vered HaGalil	1	LowerGalilee	35.54653	32.90362
GBS_PstI_F77	TD-527-40	Har Osher	1	LowerGalilee	35.55269	32.882
GBS_PstI_F93	TD-529-305	Ein Zetim	1	LowerGalilee	35.4855	32.9961
GBS_PstI_F70	TD-643-2	Kefar Shammay	1	LowerGalilee	35.45792	32.9557
GBS_PstI_F78	TD-654-13	Kefar Shammay	1	LowerGalilee	35.45792	32.9557
GBS_PstI_F86	TD-656-15	Kefar Shammay	1	LowerGalilee	35.45792	32.9557
GBS_PstI_F94	TD-661-20	Kefar Shammay	1	LowerGalilee	35.45792	32.9557
GBS_PstI_F7	TD-663-22	Kefar Shammay	1	LowerGalilee	35.45792	32.9557
GBS_PstI_F47	TD-685-0	Dov Mt.	1	LowerGalilee	35.71639	33.30202
GBS_PstI_F55	TD-686-0	Dov Mt.	1	LowerGalilee	35.71639	33.30202
GBS_PstI_F95	TD-694-4	Monfort	1	LowerGalilee	35.2355	33.04635
GBS_PstI_F8	TD-695-5	Monfort	1	LowerGalilee	35.2355	33.04635
GBS_PstI_F16	TD-696-6	Monfort	1	LowerGalilee	35.2355	33.04635
GBS_PstI_F32	TD-698-1	Kefar Shammay	1	LowerGalilee	35.45792	32.9557
GBS_PstI_F105	TD-753-6	Nahal Dishon	1	LowerGalilee	35.51517	33.05378
GBS_PstI_F113	TD-761-14	Nahal Dishon	1	LowerGalilee	35.51517	33.05378
GBS_PstI_F121	TD-762-15	Nahal Dishon	1	LowerGalilee	35.51517	33.05378
GBS_PstI_F129	TD-763-16	Nahal Dishon	1	LowerGalilee	35.51517	33.05378
GBS_PstI_F137	TD-773-26	Nahal Dishon	1	LowerGalilee	35.51517	33.05378
GBS_PstI_F145	TD-776-29	Nahal Dishon	1	LowerGalilee	35.51517	33.05378
GBS_PstI_F162	TD-890-0	Tabhaa	1	LowerGalilee	35.5421	32.87172
GBS_PstI_F170	TD-891-0	Har Gerizim	1	LowerGalilee	35.27311	32.20084
GBS_PstI_F99	TD-894-0	Nahf	1	LowerGalilee	35.32674	32.93687
GBS_PstI_F115	TD-898-0	Azmon	1	LowerGalilee	35.26539	32.82346
GBS_PstI_F189	TD-1160-1	Arbel	1	LowerGalilee	35.49652	32.82194
GBS_PstI_F102	TD-1161-2	Arbel	1	LowerGalilee	35.49652	32.82194
GBS_PstI_F110	TD-1162-3	Arbel	1	LowerGalilee	35.49652	32.82194

GBS_PstI_F118	TD-1163-4	Arbel	1	LowerGalilee	35.49652	32.82194
GBS_PstI_F126	TD-1164-5	Arbel	1	LowerGalilee	35.49652	32.82194
GBS_PstI_F142	TD-1181-1	Amirim	1	LowerGalilee	35.45526	32.9392
GBS_PstI_F150	TD-1182-1	Amirim	1	LowerGalilee	35.45526	32.9392
GBS_PstI_F112	TD-1268-25	Har Dov	1	LowerGalilee	35.71984	33.30125
GBS_PstI_F120	TD-1273-30	Har Dov	1	LowerGalilee	35.71984	33.30125
GBS_PstI_F275	TD-1939-0	Har Gerizim	1	LowerGalilee	35.27311	32.20084
GBS_PstI_F307	TD-2475-6	Keren Bratot	1	LowerGalilee	35.2355	33.04635
GBS_PstI_F315	TD-2476-7	Keren Bratot	1	LowerGalilee	35.2355	33.04635
GBS_PstI_F323	TD-2477-8	Keren Bratot	1	LowerGalilee	35.2355	33.04635
GBS_PstI_F302	TD-2839-1	Nahal Hilazon	1	LowerGalilee	35.26126	32.87515
GBS_PstI_F310	TD-2840-2	Nahal Hilazon	1	LowerGalilee	35.26126	32.87515
GBS_PstI_F318	TD-2841-3	Nahal Hilazon	1	LowerGalilee	35.26126	32.87515
GBS_PstI_F326	TD-2842-4	Nahal Hilazon	1	LowerGalilee	35.26126	32.87515
GBS_PstI_F334	TD-2844-2	Zurit	1	LowerGalilee	35.25073	32.90088
GBS_PstI_F342	TD-2845-3	Zurit	1	LowerGalilee	35.25073	32.90088
GBS_PstI_F350	TD-2848-6	Zurit	1	LowerGalilee	35.25073	32.90088
GBS_PstI_F358	TD-2849-7	Zurit	1	LowerGalilee	35.25073	32.90088
GBS_PstI_F366	TD-2850-8	Zurit	1	LowerGalilee	35.25073	32.90088
GBS_PstI_F374	TD-2851-9	Zurit	1	LowerGalilee	35.25073	32.90088
GBS_PstI_F382	TD-2854-3	Ahihud	1	LowerGalilee	35.17178	32.90726
GBS_PstI_F295	TD-2856-5	Ahihud	1	LowerGalilee	35.17178	32.90726
GBS_PstI_F303	TD-2857-6	Ahihud	1	LowerGalilee	35.17178	32.90726
GBS_PstI_F311	TD-2858-7	Ahihud	1	LowerGalilee	35.17178	32.90726
GBS_PstI_F352	TD-2961-1	Mizpe Monfort	1	LowerGalilee	35.22158	33.04784
GBS_PstI_F360	TD-2962-2	Mizpe Monfort	1	LowerGalilee	35.22158	33.04784
GBS_PstI_F368	TD-2964-4	Mizpe Monfort	1	LowerGalilee	35.22158	33.04784
GBS_PstI_F376	TD-2967-7	Mizpe Monfort	1	LowerGalilee	35.22158	33.04784
GBS_PstI_F384	TD-2969-9	Mizpe Monfort	1	LowerGalilee	35.22158	33.04784
GBS_PstI_F395	TD-3058-2	Golani Junction 1	1	LowerGalilee	35.40775	32.77055
GBS_PstI_F403	TD-3062-6	Golani Junction 1	1	LowerGalilee	35.40775	32.77055
GBS_PstI_F411	TD-3064-8	Golani Junction 1	1	LowerGalilee	35.40775	32.77055
GBS_PstI_F419	TD-3065-9	Golani Junction 1	1	LowerGalilee	35.40775	32.77055
GBS_PstI_F427	TD-3066-10	Golani Junction 1	1	LowerGalilee	35.40775	32.77055
GBS_PstI_F435	TD-3067-11	Golani Junction 1	1	LowerGalilee	35.40775	32.77055
GBS_PstI_F443	TD-3068-1	Golani Junction 2	1	LowerGalilee	35.40523	32.77082
GBS_PstI_F451	TD-3069-2	Golani Junction 2	1	LowerGalilee	35.40523	32.77082
GBS_PstI_F459	TD-3070-1	Sede Ilan	1	LowerGalilee	35.42393	32.74227
GBS_PstI_F467	TD-3071-2	Sede Ilan	1	LowerGalilee	35.42393	32.74227
GBS_PstI_F475	TD-3074-5	Sede Ilan	1	LowerGalilee	35.42393	32.74227
GBS_PstI_F388	TD-3078-9	Sede Ilan	1	LowerGalilee	35.42393	32.74227
GBS_PstI_F396	TD-3079-10	Sede Ilan	1	LowerGalilee	35.42393	32.74227
GBS_PstI_F452	TD-3098-0	Kefar Tavor	1	LowerGalilee	35.41975	32.68764

GBS_PstI_F413	TD-3108-1	Majd-el Krum	1	LowerGalilee	35.24289	32.92138
GBS_PstI_F421	TD-3109-2	Majd-el Krum	1	LowerGalilee	35.24289	32.92138
GBS_PstI_F408	TD-3169-2	Zviya	1	LowerGalilee	35.32987	32.88705
GBS_PstI_F416	TD-3170-3	Zviya	1	LowerGalilee	35.32987	32.88705
GBS_PstI_F424	TD-3171-4	Zviya	1	LowerGalilee	35.32987	32.88705
GBS_PstI_F432	TD-3173-6	Zviya	1	LowerGalilee	35.32987	32.88705
GBS_PstI_F440	TD-3174-7	Zviya	1	LowerGalilee	35.32987	32.88705
GBS_PstI_F448	TD-3175-8	Zviya	1	LowerGalilee	35.32987	32.88705
GBS_PstI_F83	TD-285-49	Almagor	2	GolanhHt	35.60029	32.90303
GBS_PstI_F20	TD-308-13	Yehudiyya Forest	2	GolanhHt	35.68189	32.92569
GBS_PstI_F28	TD-374-75	Yehudiyya Forest	2	GolanhHt	35.68189	32.92569
GBS_PstI_F36	TD-382-82	Yehudiyya Forest	2	GolanhHt	35.68189	32.92569
GBS_PstI_F60	TD-421-1	Katzerin	2	GolanhHt	35.69446	32.9945
GBS_PstI_F13	TD-442-19A	Deir Saras	2	GolanhHt	35.68668	33.02867
GBS_PstI_F21	TD-444-21	Galabina	2	GolanhHt	35.64232	33.03539
GBS_PstI_F116	TD-1025-28	Zawitan	2	GolanhHt	35.70042	32.9715
GBS_PstI_F132	TD-1059-3	Katcha Farm	2	GolanhHt	35.69912	32.93846
GBS_PstI_F140	TD-1060-4	Katcha Farm	2	GolanhHt	35.69912	32.93846
GBS_PstI_F164	TD-1073-13	Meshushim	2	GolanhHt	35.70282	32.98045
GBS_PstI_F235	TD-1930-1	Natur	2	GolanhHt	35.75893	32.85352
GBS_PstI_F243	TD-1931-2	Natur	2	GolanhHt	35.75893	32.85352
GBS_PstI_F251	TD-1932-1	Yonatan	2	GolanhHt	35.79693	32.94226
GBS_PstI_F259	TD-1934-3	Yonatan	2	GolanhHt	35.79693	32.94226
GBS_PstI_F196	TD-1953-3	Kursi- Giv'at Yoav	2	GolanhHt	35.67415	32.81296
GBS_PstI_F204	TD-1954-4	Kursi- Giv'at Yoav	2	GolanhHt	35.67415	32.81296
GBS_PstI_F212	TD-1959-9	Kursi- Giv'at Yoav	2	GolanhHt	35.67415	32.81296
GBS_PstI_F220	TD-1964-14	Kursi- Giv'at Yoav	2	GolanhHt	35.67415	32.81296
GBS_PstI_F228	TD-1969-19	Kursi- Giv'at Yoav	2	GolanhHt	35.67415	32.81296
GBS_PstI_F236	TD-1970-20	Kursi- Giv'at Yoav	2	GolanhHt	35.67415	32.81296
GBS_PstI_F244	TD-1972-1	Mizpe Ofir	2	GolanhHt	35.66367	32.81308
GBS_PstI_F252	TD-1978-1	Mizpe Ofir	2	GolanhHt	35.66367	32.81308
GBS_PstI_F260	TD-1985-8	Mizpe Ofir	2	GolanhHt	35.66367	32.81308
GBS_PstI_F268	TD-1991-14	Mizpe Ofir	2	GolanhHt	35.66367	32.81308
GBS_PstI_F276	TD-2000-1	Hayarmuk	2	GolanhHt	35.64339	32.6788
GBS_PstI_F284	TD-2001-2	Hayarmuk	2	GolanhHt	35.64339	32.6788
GBS_PstI_F197	TD-2006-7	Hayarmuk	2	GolanhHt	35.64339	32.6788
GBS_PstI_F205	TD-2014-15	Hayarmuk	2	GolanhHt	35.64339	32.6788
GBS_PstI_F213	TD-2016-17	Hayarmuk	2	GolanhHt	35.64339	32.6788
GBS_PstI_F221	TD-2024-25	Hayarmuk	2	GolanhHt	35.64339	32.6788
GBS_PstI_F229	TD-2027-1	Mezar	2	GolanhHt	35.73636	32.76864
GBS_PstI_F237	TD-2028-2	Mezar	2	GolanhHt	35.73636	32.76864
GBS_PstI_F245	TD-2029-3	Mezar	2	GolanhHt	35.73636	32.76864

GBS_PstI_F253	TD-2030-4	Mezar	2	GolanhHt	35.73636	32.76864
GBS_PstI_F261	TD-2040-10	Gamla	2	GolanhHt	35.75182	32.90447
GBS_PstI_F269	TD-2042-12	Gamla	2	GolanhHt	35.75182	32.90447
GBS_PstI_F277	TD-2043-13	Gamla	2	GolanhHt	35.75182	32.90447
GBS_PstI_F285	TD-2044-14	Gamla	2	GolanhHt	35.75182	32.90447
GBS_PstI_F198	TD-2048-18	Gamla	2	GolanhHt	35.75182	32.90447
GBS_PstI_F206	TD-2049-19	Gamla	2	GolanhHt	35.75182	32.90447
GBS_PstI_F214	TD-2056-26	Gamla	2	GolanhHt	35.75182	32.90447
GBS_PstI_F222	TD-2057-27	Gamla	2	GolanhHt	35.75182	32.90447
GBS_PstI_F230	TD-2059-29	Gamla	2	GolanhHt	35.75182	32.90447
GBS_PstI_F238	TD-2062-32	Gamla	2	GolanhHt	35.75182	32.90447
GBS_PstI_F246	TD-2067-1	Tell es Saqi	2	GolanhHt	35.83057	32.86637
GBS_PstI_F254	TD-2070-1	Mizpor Negev Kinrot	2	GolanhHt	35.64545	32.70121
GBS_PstI_F317	TD-2722-1	Kanaf	2	GolanhHt	35.6979	32.87041
GBS_PstI_F325	TD-2723-2	Kanaf	2	GolanhHt	35.6979	32.87041
GBS_PstI_F333	TD-2724-3	Kanaf	2	GolanhHt	35.6979	32.87041
GBS_PstI_F341	TD-2725-4	Kanaf	2	GolanhHt	35.6979	32.87041
GBS_PstI_F343	TD-2912-34	Har Hermon	2	GolanhHt	35.76389	33.2939
GBS_PstI_F351	TD-2936-1	Yonatan	2	GolanhHt	35.79693	32.94226
GBS_PstI_F359	TD-2937-2	Yonatan	2	GolanhHt	35.79693	32.94226
GBS_PstI_F367	TD-2940-5	Yonatan	2	GolanhHt	35.79693	32.94226
GBS_PstI_F385	TD-2971-56	Har Hermon	2	GolanhHt	35.76389	33.2939
GBS_PstI_F456	TD-3178-3	Mizpor Bet Zayda	2	GolanhHt	35.70005	32.88436
GBS_PstI_F123	TD-903-0	Bat Shelomo	3	SouthIsrael	35.0346	32.60481
GBS_PstI_F320	TD-2953-2	Qazir	3	SouthIsrael	35.09671	32.48475
GBS_PstI_F328	TD-2957-6	Qazir	3	SouthIsrael	35.09671	32.48475
GBS_PstI_F336	TD-2959-8	Qazir	3	SouthIsrael	35.09671	32.48475
GBS_PstI_F344	TD-2960-9	Qazir	3	SouthIsrael	35.09671	32.48475
GBS_PstI_F394	TD-3043-1	Ein Tut	3	SouthIsrael	35.06597	32.62678
GBS_PstI_F402	TD-3045-3	Ein Tut	3	SouthIsrael	35.06597	32.62678
GBS_PstI_F410	TD-3046-4	Ein Tut	3	SouthIsrael	35.06597	32.62678
GBS_PstI_F418	TD-3047-5	Ein Tut	3	SouthIsrael	35.06597	32.62678
GBS_PstI_F426	TD-3048-6	Ein Tut	3	SouthIsrael	35.06597	32.62678
GBS_PstI_F434	TD-3049-7	Ein Tut	3	SouthIsrael	35.06597	32.62678
GBS_PstI_F442	TD-3051-1	Daliyya	3	SouthIsrael	34.95678	32.58718
GBS_PstI_F450	TD-3052-2	Daliyya	3	SouthIsrael	34.95678	32.58718
GBS_PstI_F458	TD-3053-3	Daliyya	3	SouthIsrael	34.95678	32.58718
GBS_PstI_F466	TD-3054-4	Daliyya	3	SouthIsrael	34.95678	32.58718
GBS_PstI_F474	TD-3055-5	Daliyya	3	SouthIsrael	34.95678	32.58718
GBS_PstI_F387	TD-3056-6	Daliyya	3	SouthIsrael	34.95678	32.58718
GBS_PstI_F468	TD-3101-3	Mt. Oren	3	SouthIsrael	34.99436	32.71883
GBS_PstI_F476	TD-3102-4	Mt. Oren	3	SouthIsrael	34.99436	32.71883
GBS_PstI_F389	TD-3103-5	Mt. Oren	3	SouthIsrael	34.99436	32.71883

GBS_PstI_F397	TD-3105-7	Mt. Oren	3	SouthIsrael	34.99436	32.71883
GBS_PstI_F405	TD-3106-8	Mt. Oren	3	SouthIsrael	34.99436	32.71883
GBS_PstI_F477	TD-3120-1	Giv'at Ada	3	SouthIsrael	35.00969	32.51179
GBS_PstI_F438	TD-3141-1	Kerem Maharal	3	SouthIsrael	35.00437	32.65406
GBS_PstI_F446	TD-3143-3	Kerem Maharal	3	SouthIsrael	35.00437	32.65406
GBS_PstI_F454	TD-3144-4	Kerem Maharal	3	SouthIsrael	35.00437	32.65406
GBS_PstI_F462	TD-3145-5	Kerem Maharal	3	SouthIsrael	35.00437	32.65406
GBS_PstI_F470	TD-3147-7	Kerem Maharal	3	SouthIsrael	35.00437	32.65406
GBS_PstI_F478	TD-3149-9	Kerem Maharal	3	SouthIsrael	35.00437	32.65406
GBS_PstI_F9	TD-4-4	Gonen	4	UpperGalilee	35.64944	33.121
GBS_PstI_F17	TD-5-5	Gonen	4	UpperGalilee	35.64944	33.121
GBS_PstI_F25	TD-6-6	Gonen	4	UpperGalilee	35.64944	33.121
GBS_PstI_F33	TD-21-21	Gonen	4	UpperGalilee	35.64944	33.121
GBS_PstI_F41	TD-24-24	Gonen	4	UpperGalilee	35.64944	33.121
GBS_PstI_F49	TD-25-25	Gonen	4	UpperGalilee	35.64944	33.121
GBS_PstI_F57	TD-32-32	Gonen	4	UpperGalilee	35.64944	33.121
GBS_PstI_F65	TD-45-45	Gonen	4	UpperGalilee	35.64944	33.121
GBS_PstI_F73	TD-52-52	Gonen	4	UpperGalilee	35.64944	33.121
GBS_PstI_F81	TD-54-54	Gonen	4	UpperGalilee	35.64944	33.121
GBS_PstI_F68	TD-422-1	Rosh Pinna	4	UpperGalilee	35.53293	32.96715
GBS_PstI_F84	TD-426-3	Nahal Orevim	4	UpperGalilee	35.68789	33.14416
GBS_PstI_F5	TD-434-11	Nahal Orevim	4	UpperGalilee	35.68789	33.14416
GBS_PstI_F29	TD-461-38	Nahal Orevim	4	UpperGalilee	35.68789	33.14416
GBS_PstI_F37	TD-465-42	Nahal Orevim	4	UpperGalilee	35.68789	33.14416
GBS_PstI_F45	TD-469-46	Nahal Orevim	4	UpperGalilee	35.68789	33.14416
GBS_PstI_F53	TD-496-73	Nahal Orevim	4	UpperGalilee	35.68789	33.14416
GBS_PstI_F61	TD-499-76	Nahal Orevim	4	UpperGalilee	35.68789	33.14416
GBS_PstI_F69	TD-516-93	Nahal Gilabun	4	UpperGalilee	35.67018	33.04403
GBS_PstI_F14	TD-531-382	Hula Reserve	4	UpperGalilee	35.60343	33.06989
GBS_PstI_F23	TD-680-1	Emeq Ma'an	4	UpperGalilee	35.7538	33.29194
GBS_PstI_F31	TD-681-2	Emeq Ma'an	4	UpperGalilee	35.7538	33.29194
GBS_PstI_F39	TD-682-3	Emeq Ma'an	4	UpperGalilee	35.7538	33.29194
GBS_PstI_F40	TD-701-12A	Kefar Kuk	4	UpperGalilee	35.89298	33.53237
GBS_PstI_F48	TD-708-5	Kefar Kuk	4	UpperGalilee	35.89298	33.53237
GBS_PstI_F56	TD-709-6	Kefar Kuk	4	UpperGalilee	35.89298	33.53237
GBS_PstI_F64	TD-710-7	Kefar Kuk	4	UpperGalilee	35.89298	33.53237
GBS_PstI_F72	TD-715-1A	Kefar Kuk	4	UpperGalilee	35.89298	33.53237
GBS_PstI_F153	TD-777-1	Rosh Pinna	4	UpperGalilee	35.53293	32.96715
GBS_PstI_F161	TD-782-6	Rosh Pinna	4	UpperGalilee	35.53293	32.96715
GBS_PstI_F169	TD-784-8	Rosh Pinna	4	UpperGalilee	35.53293	32.96715
GBS_PstI_F185	TD-785-9	Rosh Pinna	4	UpperGalilee	35.53293	32.96715
GBS_PstI_F98	TD-787-11	Rosh Pinna	4	UpperGalilee	35.53293	32.96715

GBS_PstI_F106	TD-801-10	Dishon	4	UpperGalilee	35.51699	33.08308
GBS_PstI_F114	TD-805-14	Dishon	4	UpperGalilee	35.51699	33.08308
GBS_PstI_F122	TD-806-15	Dishon	4	UpperGalilee	35.51699	33.08308
GBS_PstI_F130	TD-810-19	Dishon	4	UpperGalilee	35.51699	33.08308
GBS_PstI_F138	TD-811-20	Dishon	4	UpperGalilee	35.51699	33.08308
GBS_PstI_F107	TD-895-0	Jab'a	4	UpperGalilee	35.07696	31.67447
GBS_PstI_F131	TD-908-0	Rosh Pinna	4	UpperGalilee	35.53293	32.96715
GBS_PstI_F163	TD-973-1	Yiftah Junction	4	UpperGalilee	35.54899	33.12114
GBS_PstI_F171	TD-974-2	Yiftah Junction	4	UpperGalilee	35.54899	33.12114
GBS_PstI_F179	TD-975-3	Yiftah Junction	4	UpperGalilee	35.54899	33.12114
GBS_PstI_F187	TD-976-4	Yiftah Junction	4	UpperGalilee	35.54899	33.12114
GBS_PstI_F100	TD-977-5	Yiftah Junction	4	UpperGalilee	35.54899	33.12114
GBS_PstI_F108	TD-979-7	Yiftah Junction	4	UpperGalilee	35.54899	33.12114
GBS_PstI_F103	TD-1204-0	Sifsufa	4	UpperGalilee	35.44017	33.01312
GBS_PstI_F128	TD-1320-11	Majdal Shams	4	UpperGalilee	35.77166	33.26783
GBS_PstI_F136	TD-1336-27	Majdal Shams	4	UpperGalilee	35.77166	33.26783
GBS_PstI_F217	TD-1709-18	Yesha Junc.	4	UpperGalilee	35.55475	33.11376
GBS_PstI_F225	TD-1716-25	Yesha Junc.	4	UpperGalilee	35.55475	33.11376
GBS_PstI_F233	TD-1725-1	Yesha Junc.	4	UpperGalilee	35.55475	33.11376
GBS_PstI_F241	TD-1736-12	Yesha Junc.	4	UpperGalilee	35.55475	33.11376
GBS_PstI_F249	TD-1745-21	Yesha Junc.	4	UpperGalilee	35.55475	33.11376
GBS_PstI_F257	TD-1746-22	Yesha Junc.	4	UpperGalilee	35.55475	33.11376
GBS_PstI_F265	TD-1756-2	Ramot Naftali	4	UpperGalilee	35.55764	33.104
GBS_PstI_F273	TD-1757-3	Ramot Naftali	4	UpperGalilee	35.55764	33.104
GBS_PstI_F281	TD-1762-8	Ramot Naftali	4	UpperGalilee	35.55764	33.104
GBS_PstI_F194	TD-1771-17	Ramot Naftali	4	UpperGalilee	35.55764	33.104
GBS_PstI_F202	TD-1787-33	Ramot Naftali	4	UpperGalilee	35.55764	33.104
GBS_PstI_F210	TD-1800-3	Ramot Naftali	4	UpperGalilee	35.55764	33.104
GBS_PstI_F218	TD-1828-31	Ramot Naftali	4	UpperGalilee	35.55764	33.104
GBS_PstI_F226	TD-1830-33	Ramot Naftali	4	UpperGalilee	35.55764	33.104
GBS_PstI_F353	TD-2333-8	Hermon	4	UpperGalilee	35.75384	33.29182
GBS_PstI_F361	TD-2345-2	Ya'afuri valley	4	UpperGalilee	35.76455	33.24376
GBS_PstI_F369	TD-2351-8	Ya'afuri valley	4	UpperGalilee	35.76455	33.24376
GBS_PstI_F44	TD-393-9	Bet Qeshet	5	JudeanHt	35.40943	32.72204
GBS_PstI_F52	TD-402-18	Bet Qeshet	5	JudeanHt	35.40943	32.72204
GBS_PstI_F15	TD-671-0	Jerusalem	5	JudeanHt	35.21371	31.76832
GBS_PstI_F80	TD-720-5	Bet Qeshet	5	JudeanHt	35.40943	32.72204
GBS_PstI_F88	TD-728-13	Bet Qeshet	5	JudeanHt	35.40943	32.72204
GBS_PstI_F96	TD-729-14	Bet Qeshet	5	JudeanHt	35.40943	32.72204
GBS_PstI_F97	TD-747-32	Bet Qeshet	5	JudeanHt	35.40943	32.72204
GBS_PstI_F154	TD-889-0	Bet Me'ir	5	JudeanHt	35.03675	31.79479
GBS_PstI_F178	TD-892-0	Giv'at Koah	5	JudeanHt	34.95794	32.03775
GBS_PstI_F139	TD-909-0	Giv'at Koah	5	JudeanHt	34.95794	32.03775

GBS_PstI_F147	TD-925-14	Ramot	5	JudeanHt	35.19423	31.81902
GBS_PstI_F155	TD-965-54	Ramot	5	JudeanHt	35.19423	31.81902
GBS_PstI_F193	TD-1594-9	Ramot	5	JudeanHt	35.19423	31.81902
GBS_PstI_F201	TD-1601-16	Ramot	5	JudeanHt	35.19423	31.81902
GBS_PstI_F209	TD-1609-24	Ramot	5	JudeanHt	35.19423	31.81902
GBS_PstI_F282	TD-1886-1	Emeq Ha'Arazim	5	JudeanHt	35.17255	31.80348
GBS_PstI_F195	TD-1887-2	Emeq Ha'Arazim	5	JudeanHt	35.17255	31.80348
GBS_PstI_F203	TD-1888-3	Emeq Ha'Arazim	5	JudeanHt	35.17255	31.80348
GBS_PstI_F321	TD-2315-15	Betar Ilit	5	JudeanHt	35.12003	31.69984
GBS_PstI_F329	TD-2318-18	Betar Ilit	5	JudeanHt	35.12003	31.69984
GBS_PstI_F337	TD-2319-19	Betar Ilit	5	JudeanHt	35.12003	31.69984
GBS_PstI_F345	TD-2324-24	Betar Ilit	5	JudeanHt	35.12003	31.69984
GBS_PstI_F330	TD-2399-9	Jerusalem	5	JudeanHt	35.21371	31.76832
GBS_PstI_F338	TD-2403-13	Jerusalem	5	JudeanHt	35.21371	31.76832
GBS_PstI_F346	TD-2406-16	Jerusalem	5	JudeanHt	35.21371	31.76832
GBS_PstI_F354	TD-2408-18	Jerusalem	5	JudeanHt	35.21371	31.76832
GBS_PstI_F299	TD-2452-34	Sha'ar HaGay	5	JudeanHt	35.02524	31.81564
GBS_PstI_F371	TD-2585-9	Sha'ar HaGay	5	JudeanHt	35.02524	31.81564
GBS_PstI_F379	TD-2587-11	Sha'ar HaGay	5	JudeanHt	35.02524	31.81564
GBS_PstI_F292	TD-2595-19	Sha'ar HaGay	5	JudeanHt	35.02524	31.81564
GBS_PstI_F300	TD-2619-8	Bet Arye	5	JudeanHt	35.05304	32.03882
GBS_PstI_F308	TD-2637-26	Bet Arye	5	JudeanHt	35.05304	32.03882
GBS_PstI_F316	TD-2650-39	Bet Arye	5	JudeanHt	35.05304	32.03882
GBS_PstI_F324	TD-2664-3	Hadom HaShomron	5	JudeanHt	34.97736	32.0361
GBS_PstI_F332	TD-2665-4	Hadom HaShomron	5	JudeanHt	34.97736	32.0361
GBS_PstI_F340	TD-2667-6	Hadom HaShomron	5	JudeanHt	34.97736	32.0361
GBS_PstI_F348	TD-2671-10	Hadom HaShomron	5	JudeanHt	34.97736	32.0361
GBS_PstI_F356	TD-2672-1	Pedu'el	5	JudeanHt	35.05231	32.06243
GBS_PstI_F364	TD-2685-14	Pedu'el	5	JudeanHt	35.05231	32.06243
GBS_PstI_F372	TD-2686-15	Pedu'el	5	JudeanHt	35.05231	32.06243
GBS_PstI_F380	TD-2711-40	Pedu'el	5	JudeanHt	35.05231	32.06243
GBS_PstI_F293	TD-2714-1	Ale Zahav	5	JudeanHt	35.06441	32.07225
GBS_PstI_F301	TD-2715-2	Ale Zahav	5	JudeanHt	35.06441	32.07225
GBS_PstI_F309	TD-2716-3	Ale Zahav	5	JudeanHt	35.06441	32.07225
GBS_PstI_F349	TD-2742-16	Modi'in	5	JudeanHt	35.01034	31.88951
GBS_PstI_F357	TD-2743-17	Modi'in	5	JudeanHt	35.01034	31.88951
GBS_PstI_F365	TD-2789-13	Hevron Mountain	5	JudeanHt	34.99832	31.44741
GBS_PstI_F373	TD-2797-21	Hevron Mountain	5	JudeanHt	34.99832	31.44741
GBS_PstI_F381	TD-2800-24	Hevron Mountain	5	JudeanHt	34.99832	31.44741
GBS_PstI_F294	TD-2822-20	Hevron Mountain	5	JudeanHt	34.99832	31.44741
GBS_PstI_F319	TD-2867-14	Har Amasa	5	JudeanHt	35.11097	31.35132
GBS_PstI_F327	TD-2875-32	Har Amasa	5	JudeanHt	35.11097	31.35132
GBS_PstI_F335	TD-2883-49	Har Amasa	5	JudeanHt	35.11097	31.35132

GBS_PstI_F393	TD-2976-5	Modi'in	5	JudeanHt	35.01034	31.88951
GBS_PstI_F401	TD-2984-4	Shilat	5	JudeanHt	35.01222	31.91969
GBS_PstI_F409	TD-2985-5	Shilat	5	JudeanHt	35.01222	31.91969
GBS_PstI_F417	TD-2986-6	Shilat	5	JudeanHt	35.01222	31.91969
GBS_PstI_F425	TD-2991-1	Modi'in	5	JudeanHt	35.01034	31.88951
GBS_PstI_F433	TD-3001-1	Modi'in	5	JudeanHt	35.01034	31.88951
GBS_PstI_F441	TD-3011-1	Petah Tiqwa	5	JudeanHt	34.8888	32.08345
GBS_PstI_F449	TD-3013-3	Petah Tiqwa	5	JudeanHt	34.8888	32.08345
GBS_PstI_F457	TD-3018-4	Modi'in	5	JudeanHt	35.01034	31.88951
GBS_PstI_F465	TD-3023-2	Petah Tiqwa	5	JudeanHt	34.8888	32.08345
GBS_PstI_F473	TD-3024-3	Petah Tiqwa	5	JudeanHt	34.8888	32.08345
GBS_PstI_F386	TD-3040-1	Shilat	5	JudeanHt	35.01757	31.92296
GBS_PstI_F404	TD-3080-1	Ela'd	5	JudeanHt	34.96637	32.4425
GBS_PstI_F412	TD-3085-6	Ela'd	5	JudeanHt	34.96637	32.4425
GBS_PstI_F420	TD-3086-7	Ela'd	5	JudeanHt	34.96637	32.4425
GBS_PstI_F428	TD-3095-16	Ela'd	5	JudeanHt	34.96637	32.4425
GBS_PstI_F436	TD-3096-1	Har Amasa near vineyard	5	JudeanHt	35.10488	31.35124
GBS_PstI_F444	TD-3097-2	Har Amasa near vineyard	5	JudeanHt	35.10488	31.35124
GBS_PstI_F391	TD-3150-1	Beit-Itab	5	JudeanHt	35.05733	31.73238
GBS_PstI_F399	TD-3151-2	Beit-Itab	5	JudeanHt	35.05733	31.73238
GBS_PstI_F407	TD-3152-3	Beit-Itab	5	JudeanHt	35.05733	31.73238
GBS_PstI_F415	TD-3153-4	Beit-Itab	5	JudeanHt	35.05733	31.73238
GBS_PstI_F423	TD-3156-1	Bar Giyyora	5	JudeanHt	35.07214	31.72744
GBS_PstI_F431	TD-3159-4	Bar Giyyora	5	JudeanHt	35.07214	31.72744
GBS_PstI_F439	TD-3160-5	Bar Giyyora	5	JudeanHt	35.07214	31.72744
GBS_PstI_F447	TD-3161-1	Horbat Hannot	5	JudeanHt	35.03438	31.7099
GBS_PstI_F455	TD-3162-2	Horbat Hannot	5	JudeanHt	35.03438	31.7099
GBS_PstI_F463	TD-3163-3	Horbat Hannot	5	JudeanHt	35.03438	31.7099
GBS_PstI_F471	TD-3164-4	Horbat Hannot	5	JudeanHt	35.03438	31.7099
GBS_PstI_F479	TD-3165-5	Horbat Hannot	5	JudeanHt	35.03438	31.7099
GBS_PstI_F158	TD-1189-1	Zikhron Ya'aqov	6	Inbetween	34.95091	32.5594
GBS_PstI_F166	TD-1190-2	Zikhron Ya'aqov	6	Inbetween	34.95091	32.5594
GBS_PstI_F174	TD-1191-3	Zikhron Ya'aqov	6	Inbetween	34.95091	32.5594
GBS_PstI_F182	TD-1194-6	Zikhron Ya'aqov	6	Inbetween	34.95091	32.5594
GBS_PstI_F190	TD-1197-9	Zikhron Ya'aqov	6	Inbetween	34.95091	32.5594
GBS_PstI_F144	TD-1400-40	Elqana	6	Inbetween	35.03079	32.10828
GBS_PstI_F152	TD-1405-45	Elqana	6	Inbetween	35.03079	32.10828
GBS_PstI_F211	TD-1898-8	Avne Hefez	6	Inbetween	35.07742	32.28371
GBS_PstI_F219	TD-1913-23	Avne Hefez	6	Inbetween	35.07742	32.28371
GBS_PstI_F262	TD-2114-44	Avne Hefez	6	Inbetween	35.07742	32.28371
GBS_PstI_F270	TD-2123-53	Avne Hefez	6	Inbetween	35.07742	32.28371

GBS_PstI_F278	TD-2133-8	Einav Road	6	Inbetween	35.12826	32.28199
GBS_PstI_F286	TD-2138-13	Einav Road	6	Inbetween	35.12826	32.28199
GBS_PstI_F199	TD-2155-30	Einav Road	6	Inbetween	35.12826	32.28199
GBS_PstI_F207	TD-2160-35	Einav Road	6	Inbetween	35.12826	32.28199
GBS_PstI_F215	TD-2172-1	Einav	6	Inbetween	35.1252	32.28359
GBS_PstI_F223	TD-2175-4	Einav	6	Inbetween	35.1252	32.28359
GBS_PstI_F231	TD-2176-1	Kefar Qasim	6	Inbetween	34.9821	32.10531
GBS_PstI_F239	TD-2182-7	Kefar Qasim	6	Inbetween	34.9821	32.10531
GBS_PstI_F247	TD-2183-8	Kefar Qasim	6	Inbetween	34.9821	32.10531
GBS_PstI_F255	TD-2193-18	Kefar Qasim	6	Inbetween	34.9821	32.10531
GBS_PstI_F263	TD-2203-28	Kefar Qasim	6	Inbetween	34.9821	32.10531
GBS_PstI_F271	TD-2206-31	Kefar Qasim	6	Inbetween	34.9821	32.10531
GBS_PstI_F331	TD-2478-1	Elon More	6	Inbetween	35.33117	32.23488
GBS_PstI_F339	TD-2489-12	Elon More	6	Inbetween	35.33117	32.23488
GBS_PstI_F347	TD-2536-39	Elon More	6	Inbetween	35.33117	32.23488
GBS_PstI_F355	TD-2553-14	Elon More	6	Inbetween	35.33117	32.23488
GBS_PstI_F375	TD-2946-1	Ramat HaNadiv	6	Inbetween	34.94409	32.55058
GBS_PstI_F383	TD-2948-3	Ramat HaNadiv	6	Inbetween	34.94409	32.55058
GBS_PstI_F296	TD-2949-4	Ramat HaNadiv	6	Inbetween	34.94409	32.55058
GBS_PstI_F304	TD-2950-5	Ramat HaNadiv	6	Inbetween	34.94409	32.55058
GBS_PstI_F390	TD-3123-3	Ramot Menashe	6	Inbetween	35.05712	32.59672
GBS_PstI_F398	TD-3126-6	Ramot Menashe	6	Inbetween	35.05712	32.59672
GBS_PstI_F406	TD-3128-8	Ramot Menashe	6	Inbetween	35.05712	32.59672
GBS_PstI_F414	TD-3131-11	Ramot Menashe	6	Inbetween	35.05712	32.59672
GBS_PstI_F422	TD-3139-19	Ramot Menashe	6	Inbetween	35.05712	32.59672
GBS_PstI_F430	TD-3140-20	Ramot Menashe	6	Inbetween	35.05712	32.59672
GBS_PstI_F392	TD-3166-1	Nirit	6	Inbetween	34.9789	32.14378
GBS_PstI_F400	TD-3167-2	Nirit	6	Inbetween	34.9789	32.14378
GBS_PstI_F6	TD-530-323	Majdel Ma'ar	7	valley	35.38422	32.42156
GBS_PstI_F22	TD-552-21	Giv'at HaMore	7	valley	35.34188	32.61846
GBS_PstI_F30	TD-575-44	Giv'at HaMore	7	valley	35.34188	32.61846
GBS_PstI_F38	TD-590-59	Giv'at HaMore	7	valley	35.34188	32.61846
GBS_PstI_F46	TD-593-2	Giv'at HaMore	7	valley	35.34188	32.61846
GBS_PstI_F54	TD-610-19	Giv'at HaMore	7	valley	35.34188	32.61846
GBS_PstI_F62	TD-624-33	Giv'at HaMore	7	valley	35.34188	32.61846
GBS_PstI_F63	TD-687-1	Gilboa	7	valley	35.38854	32.5165
GBS_PstI_F71	TD-688-2	Gilboa	7	valley	35.38854	32.5165
GBS_PstI_F79	TD-689-3	Gilboa	7	valley	35.38854	32.5165
GBS_PstI_F87	TD-690-4	Gilboa	7	valley	35.38854	32.5165
GBS_PstI_F146	TD-879-15	Kokhav HaShachar	7	valley	35.35091	31.95936
GBS_PstI_F186	TD-893-0	Taibe	7	valley	35.29819	31.95699
GBS_PstI_F117	TD-1142-1	Irit (Tubas)	7	valley	35.36894	32.3204
GBS_PstI_F125	TD-1143-2	Irit (Tubas)	7	valley	35.36894	32.3204

GBS_PstI_F133	TD-1144-1	Nahal Zviya	7	valley	35.4356	32.47267
GBS_PstI_F141	TD-1146-3	Nahal Zviya	7	valley	35.4356	32.47267
GBS_PstI_F149	TD-1147-4	Nahal Zviya	7	valley	35.4356	32.47267
GBS_PstI_F157	TD-1148-5	Nahal Zviya	7	valley	35.4356	32.47267
GBS_PstI_F165	TD-1150-7	Nahal Zviya	7	valley	35.4356	32.47267
GBS_PstI_F173	TD-1156-0	Nahal Zviya	7	valley	35.4356	32.47267
GBS_PstI_F181	TD-1159-3	Kokhav HaShachar	7	valley	35.35091	31.95936
GBS_PstI_F134	TD-1178-1	Hoze Shomron	7	valley	35.25759	32.11598
GBS_PstI_F111	TD-1210-4	Gilboa	7	valley	35.38854	32.5165
GBS_PstI_F119	TD-1211-5	Gilboa	7	valley	35.38854	32.5165
GBS_PstI_F127	TD-1214-8	Gilboa	7	valley	35.38854	32.5165
GBS_PstI_F135	TD-1218-12	Gilboa	7	valley	35.38854	32.5165
GBS_PstI_F143	TD-1219-1	Ma'ale Gilboa	7	valley	35.41413	32.47982
GBS_PstI_F151	TD-1220-2	Ma'ale Gilboa	7	valley	35.41413	32.47982
GBS_PstI_F159	TD-1223-5	Ma'ale Gilboa	7	valley	35.41413	32.47982
GBS_PstI_F167	TD-1224-6	Ma'ale Gilboa	7	valley	35.41413	32.47982
GBS_PstI_F175	TD-1225-7	Ma'ale Gilboa	7	valley	35.41413	32.47982
GBS_PstI_F183	TD-1235-9	Ma'ale Efrayim	7	valley	35.4039	32.0713
GBS_PstI_F191	TD-1237-1	Ma'ale Efrayim	7	valley	35.4039	32.0713
GBS_PstI_F160	TD-1478-4	Hoze Shomron	7	valley	35.25759	32.11598
GBS_PstI_F168	TD-1479-5	Hoze Shomron	7	valley	35.25759	32.11598
GBS_PstI_F176	TD-1481-7	Hoze Shomron	7	valley	35.25759	32.11598
GBS_PstI_F184	TD-1511-21	Kokhav HaShachar	7	valley	35.35091	31.95936
GBS_PstI_F192	TD-1536-46	Kokhav HaShachar	7	valley	35.35091	31.95936
GBS_PstI_F234	TD-1839-8	Ma'ale Merav	7	valley	35.43915	32.43605
GBS_PstI_F242	TD-1853-22	Ma'ale Merav	7	valley	35.43915	32.43605
GBS_PstI_F250	TD-1855-24	Ma'ale Merav	7	valley	35.43915	32.43605
GBS_PstI_F258	TD-1865-34	Ma'ale Merav	7	valley	35.43915	32.43605
GBS_PstI_F274	TD-1880-49	Ma'ale Merav	7	valley	35.43915	32.43605
GBS_PstI_F267	TD-1938-0	Yeriho	7	valley	35.46056	31.85697
GBS_PstI_F283	TD-1941-0	Yeriho	7	valley	35.46057	31.85698
GBS_PstI_F279	TD-2209-2	Majdel Bani-Fadil	7	valley	35.36772	32.08879
GBS_PstI_F287	TD-2211-4	Majdel Bani-Fadil	7	valley	35.36772	32.08879
GBS_PstI_F200	TD-2212-5	Majdel Bani-Fadil	7	valley	35.36772	32.08879
GBS_PstI_F208	TD-2213-6	Majdel Bani-Fadil	7	valley	35.36772	32.08879
GBS_PstI_F216	TD-2218-11	Majdel Bani-Fadil	7	valley	35.36772	32.08879
GBS_PstI_F224	TD-2220-13	Majdel Bani-Fadil	7	valley	35.36772	32.08879
GBS_PstI_F232	TD-2223-16	Majdel Bani-Fadil	7	valley	35.36772	32.08879
GBS_PstI_F240	TD-2229-6	Gitit	7	valley	35.38699	32.10346
GBS_PstI_F248	TD-2235-12	Gitit	7	valley	35.38699	32.10346
GBS_PstI_F256	TD-2238-15	Gitit	7	valley	35.38699	32.10346
GBS_PstI_F264	TD-2241-18	Gitit	7	valley	35.38699	32.10346
GBS_PstI_F272	TD-2257-34	Gitit	7	valley	35.38699	32.10346

GBS_PstI_F280	TD-2262-39	Gitit	7	valley	35.38699	32.10346
GBS_PstI_F288	TD-2265-2	Kokhav HaShachar	7	valley	35.35091	31.95936
GBS_PstI_F289	TD-2266-3	Kokhav HaShachar	7	valley	35.35091	31.95936
GBS_PstI_F297	TD-2279-2	Kokhav HaShachar	7	valley	35.35091	31.95936
GBS_PstI_F305	TD-2295-5	Ma'ale Efrayim	7	valley	35.4039	32.0713
GBS_PstI_F313	TD-2300-3	Gitit	7	valley	35.38699	32.10346
GBS_PstI_F377	TD-2364-8	Sede Trumot- Merav	7	valley	35.44783	32.42857
GBS_PstI_F290	TD-2367-11	Sede Trumot- Merav	7	valley	35.44783	32.42857
GBS_PstI_F298	TD-2382-26	Sede Trumot- Merav	7	valley	35.44783	32.42857
GBS_PstI_F306	TD-2386-30	Sede Trumot- Merav	7	valley	35.44783	32.42857
GBS_PstI_F314	TD-2387-1	Merav	7	valley	35.43915	32.43605
GBS_PstI_F322	TD-2390-4	Merav	7	valley	35.43915	32.43605
GBS_PstI_F362	TD-2419-1	Nahal Ktalav	7	valley	35.06012	32.72202
GBS_PstI_F370	TD-2420-2	Nahal Ktalav	7	valley	35.06012	32.72202
GBS_PstI_F378	TD-2421-3	Nahal Ktalav	7	valley	35.06012	32.72202
GBS_PstI_F291	TD-2422-4	Nahal Ktalav	7	valley	35.06012	32.72202
GBS_PstI_F429	TD-3112-4	Revaya-basalt	7	valley	35.46182	32.45009
GBS_PstI_F437	TD-3114-6	Revaya-basalt	7	valley	35.46182	32.45009
GBS_PstI_F445	TD-3115-7	Revaya-basalt	7	valley	35.46182	32.45009
GBS_PstI_F453	TD-3116-8	Revaya-basalt	7	valley	35.46182	32.45009
GBS_PstI_F461	TD-3117-9	Revaya-basalt	7	valley	35.46182	32.45009
GBS_PstI_F469	TD-3118-10	Revaya-basalt	7	valley	35.46182	32.45009

Chapter 3 - Breeding wheat for climate adaptation: Introgression of climate adaptive variants in interspecific populations through direct hybridization

Abstract

Wild relatives are considered a valuable resource of novel allelic diversity for crop improvement. In the case of wheat, wild relatives have been broadly used in disease resistance, yield, and quality trait improvement. However, the use of wild relatives in improving the adaptive potential of wheat has not been prioritized yet. I investigated the genetic basis of local adaptation in *Ae. tauschii*, diploid progenitor of wheat, and identified the potential donor parents that are enriched

in adaptive variants. I used twenty-one *Ae. tauschii* accessions from geographically dispersed locations as parents for developing *Ae. tauschii* x hexaploid wheat octaploid hybrids that were crossed with the six Kansas-adapted hexaploid wheat lines. Three hundred and fifty-one BC1F3:5 lines developed after single backcross to recurrent parents were used for studying the genomic patterns of alien introgression. Hybrid sterility, reduced introgression in the pericentromeric regions, and reduced retention of the introgressed alleles in regions carrying domestication gene were some of the issues of inter-specific population development. Overall, the research shows that there is tremendous untapped genetic diversity in wild relatives that can be prioritized for breeding climate-resilient wheat.

Introduction

Global food supply and production are greatly challenged by the increase in demand from the rising population and the effect of climate change (FAO, 2019). The current trends of genetic gain in crops need to be doubled to fulfill the demand-supply gap (Li et al., 2018). Crop production is expected to further decline due to limited genetic diversity in the existing germplasm for increased occurrences of adverse abiotic stress. In the case of wheat, the limited genetic diversity in elite germplasm has been identified as one of the obstacles for improving wheat's productive potential. Loss of genetic diversity during domestication, interspecific hybridization, polyploidization, and years of selective breeding has resulted in a narrow genetic base in bread wheat as compared to its wild relatives (Cox, 1997, Akhunov et al., 2010). Therefore, broadening the genetic diversity of bread wheat through identification and introgression of potentially beneficial allelic variants from its wild counterpart to the hexaploid gene pool is the primary focus of germplasm development. To improve the genetic diversity,

wheat's wild progenitors and landraces have been included in the breeding scheme to harness the lost diversity.

Aegilops tauschii, the "DD" genome donor of the hexaploid wheat (*Triticum aestivum* L.) originated through homoploid hybrid speciation from A and B genomes. (Marcussen et al., 2014). Spontaneous hybridization between the diploid "DD" genome progenitor *Aegilops tauschii* and tetraploid (AABB) emmer wheat (*Triticum turgidum*) ~ 10,000 years ago gave rise to the allohexaploid (AABBDD) bread wheat (*Triticum aestivum* L.) (McFadden and Sear, 1946; Jordan et al., 2015; Wang et al., 2013). Even though *Ae. tauschii* is reported to spread throughout a wide geographical range from Georgia, Turkey, and Syria in the west up to China in the east (Wang et al., 2013), hybridization between tetraploid wheat and *Ae. tauschii,* is believed to have taken place in a small area around Armenia and the Caspian Sea, with a single *Ae. tauschii* ssp *strangulata* gene pool (Dvorak et al., 1998). This has resulted to a narrow genetic base in the D-genome of the hexaploid wheat as compared to the *Ae. tauschii* gene pool and also to its counterparts, A and B genomes (Wang et al., 2013; Lubbers et al., 1991; Akhunov et al., 2010; Chapman et al., 2015). This has encouraged breeding programs to incorporate *Ae. tauschii* in their breeding scheme with the objectives of introducing novel genetic diversity present in the *Ae. tauschii* genome into the bread wheat gene pool (Naghavi et al., 2009; Singh et al., 2019).

Incorporation of *Ae. tauschii* into the wheat breeding programs scheme has not only helped overcome the genetic bottleneck in the D genome but also resulted in the transfer of many useful genetic variants for crop improvement. For example, allelic diversity from *Aegilops tauschii* has been reported to contribute to disease resistance (Friesen et al., 2008), climatic adaption (Jia et al., 2013), yield and yield-related traits (Fritz et al., 1995), and improved bread

quality in bread wheat (Hsam et al., 2001). Landraces and wild species originating from regions with extreme climatic conditions have improved drought adaptation in wheat (Reynold et al., 2007). Homology between the wheat D genome and *Ae. tauschii* allows for introgression through two distinct approaches. The first one relies on synthetic hexaploid wheat produced through hybridization between *T. turgidum* (tetraploid) and *Ae. tauschii* (diploid), followed by chromosome doubling with colchicine. These synthetics could be directly incorporated into the crossing schemes of wheat breeding programs. Another approach of *Ae. tauschii* introgression is through the hybridization of *Ae. tauschii* with the hexaploid wheat followed by embryo rescue from F1 hybrids and colchicine-induced chromosome doubling. The octoploids could be then crossed with wheat lines followed by backcrossing to recurrent parents (Cox et al., 2017).

Even though inter-specific hybridization is widely used in crop improvement there is limited understanding of the introgression barriers that prevent the transfer of wild relatives' alleles into cultivated crops. Analysis of inter-specific introgression populations developed for cotton and sunflower discovered a bias in the rate of transmission of wild relative's alleles. Suppressed recombination, chromosomal rearrangements, allelic incompatibilities, are suggested to cause the non-Mendelian inheritance of introgressed alleles (Rieseberg et al., 1995; Jiang et al., 2000; Rieseberg, 2001). In addition to these barriers, sterility of interspecific hybrids remain one of the major obstacles for inter-specific introgression. Research on the inter-specific population of sunflower showed that structural mutations in the parental species genomes lower the rates of inter-specific allele transfer (Rieseberg et al., 1995). Reduced retention of introgressed alleles reported in the advanced generation of inter-specific crosses in cotton was attributed to multilocus epistatic interaction between introgressed alleles and genetic backgrounds of recipient lines (Jiang et al., 2000). Traits that are the major target of selection

during domestication also determine the fate of the alien introgression gene in an intra-specific population. A recent study in the wild emmer population found that the introgression from wild relatives around the *Q* gene (Simons et al., 2006) controlling domestication trait was greatly reduced (He et al., 2019). Also, the Hill-Robertson effect could prevent the introgression of some of the useful genes that are in linkage disequilibrium with negatively selected alleles (Hill and Robertson, 1966). Therefore, this research aims to develop an inter-specific population using diverse wild relatives and investigate the nature of introgression barriers.

Materials and methods

Ae. tauschii population

A diverse set of 136 geo-referenced *Ae. tauschii* accessions representing *Ae. tauschii* ssp. *tauschii* (categorized as L1), and *Ae. tauschii* ssp. *strangulata* (categorized as L2) from wide range of geographical locations was used. The population includes accessions originating from 11 different countries, from Turkey and Syria in the west and Kyrgyzstan to Pakistan in the east (Figure 3.1). Table 3.1 showing all accessions (one accessions with no GPS excluded from the study) that the research population contains more accessions from lineage one (96 individuals from L1) than from lineage two (41 individuals from L2). Out of this population, 21 (18 form L1 and 3 from L2) diverse accessions were used in direct hybridization with hexaploid wheat to develop an intra-specific introgression population including 351 lines (Table 3.1).

Ae. tauschii introgression population developed through direct hybridization

The inter-specific introgression lines were developed by crossing six hexaploid wheat parents with 21 *Ae. tauschii* accessions (Figure 3.2). The F1 hybrid plants were embryo rescued and treated with colchicine producing synthetic octoploids (Dale et al., 2017). The synthetic

octoploids were backcrossed or top-crossed only once to its respective hexaploid wheat parents or to another wheat line to create BC1F1 lines (Zhang et al., 2018). The BC_1F_3 lines were developed through single seed descent. The seeds from individual BC_1F_3 plants were bulked and planted in single rows in Ashland Research Farm near Manhattan, KS in the 2016-17 growing seasons. A total of 2,861 segregating lines were developed for 31 families with each family comprised of 42 to 137 lines. A total of 351 lines were retained after early generation selection and used in this study. The selection criteria included the production of a sufficient number of seeds to allow yield testing, phenology similar to the elite hexaploid parent(s), threshability to allow mechanical harvest, and general fitness.

Historical onsite climatic variables of *Ae. tauschii*

Using the WorldClim Version2.0 database, on-site bioclimatic variables for all georeferenced accession were collected for the years 1970 to 2000 at a resolution of 30 arc-seconds (~1 km). The monthly average climatic variables (precipitation, mean, maximum and minimum temperature, wind, water vapor pressure, and solar radiation) and 19 bioclimatic variables are used as phenotypes (Hijmans et al., 2005) in genetic analysis of adaptation study. Figures 3.3 and 3.4 respectively show variations in solar radiation and water vapor pressures across *Ae. tauschii* growing regions. The bioclimatic and climatic variables were extracted using the R package "*raster*".

Genomic libraries preparation, SNP calling and imputation

DNA from both the 137 *Ae. tauschii* and 351 hexaploid introgression population was extracted using a separate DNeasy 96 Plant DNA extraction kit (Qiagen). PicoGreen® dsDNA assay kit (Life Technologies) was used to assess the quality and concentration of the DNA. Genotyping by

sequencing (GBS) libraries were prepared following the protocol described by Saintenac et al. (2013). The library was enriched to 270-330 bp fragments through size selection using the Pippin Prep system (Sage Scientific) and Illumina NextSeq 500 was used for sequencing.

SNPs were called separately for introgression lines and the reference panel using the TASSEL 5.0 GBS v2 pipeline (Glaubitz et al., 2014) using all the default parameters except the KmerLength of introgression population was reduced to 35 to account for shorter reads. Using Burrows- Wheeler Alignment (BWA) software v0.7.17 the GBS reads were aligned to CS RefSeq v1.0 (The International Wheat Genome Sequencing Consortium (IWGSC), 2018). SNPs from 136 *Ae. tauschii* were filtered using vcf-filter tools v0.1.13 where SNPs with > 50% of missing data and < 0.02 minor allele frequency (MAF) were filtered out. Beagle v.5.0 (Browning and Browning, 2013) was used to impute the missing data.

SNPs from *Ae. tauschii*-derived introgression lines were filtered in two steps. First, common filtering criteria for all sub-genomes (A, B, and D) was applied where SNPs from all sub-genomes (A, B, and D) with minor allele frequency (MAF) < 0.05 and > 30% missing were filtered out using vcf-filter tools v0.1.13 and the remaining SNPs were imputed using Beagle v.5.0. In the next step, all the SNPs from A and B genome, and D genome SNPs with MAF <0.01 were filtered out using vcf filter tools v0.1.13. The program "*conform-gt*" (https://faculty.washington.edu/browning/conform-gt.html) was used to check the concordance of D genome SNP positions between the introgression population and the reference panel based on the Chinese Spring genome v1.0 coordinates (IWGSC, 2018). To improve the detection of identical-by-descent (IBD) regions in the introgressed population, the remaining 116 *Ae. tauschii* that were not used as parents in the inter-specific crosses were used as a reference panel to impute the missing genotype SNPs using Beagle v.5.0.

Population structure and diversity analysis

The population structure of the diverse *Ae. tauschii* accessions and the introgression population was analyzed using the 27,880 D genome SNPs segregating in both *Ae. tauschii* and *Ae. tauschii*-derived populations. The SNP dataset was converted to the hapmap format and imported into TASSEL v.5.0, which was used to calculate the principal components and Tajima's D score among lineages. Tajima's D is a statistical test for test of neutrality (Tajima, 1989). It helps understand whether the evolution and differentiation involved any non-random process. Random evolution process do not effect fitness and survival of an organism and is called neutral mutation whereas mutations resulted from non-random process results in mutations that influnce fitness and survival of an organism. Some of the non-random processes include directional selection, introgression, genetic hitchhiking etc. The first two components were plotted to show the distribution and clustering of the reference panel accessions in relation to the 21 parental *Ae. tauschii* accessions and the entire introgression population. Inter-specific population distribution with regards to its recurrent parent was evaluated using a total of 13,970 SNPs on the first two principal components and recurrent parents were used as grouping factors.

To estimates spatial population structure and genetic co-ancestry for geography, the "*tess3*" function is implemented in the "*tess3r*" R package (Caye et al., 2016; Caye et al., 2018). TESS3r (Caye et al., 2016), R package estimates the population genetic structure based on allele frequency differentiation and layers the ancestry coefficients on actual geographic maps. SNPs were converted into diploid numeric SNPs format and all polymorphic SNPs were used in inferring the spatial population structure. Genotypes are encoded as 0, 1, and 2, and the ancestry coefficients, K = 1 to 12 was used to estimate the genetic structure and gene flow among the population. After identifying the optimum number of K, pairwise fixation index (Weir and

Cockerham, 1984) was calculated and compared using VCFtools-0.1.15 (Danecek et al., 2011) to understand the differentiation among the population.

The population structure was also analyzed using both model-based, STRUCTURE (Appendix B, Figure 1) and non-model-base, PCA (Figure 3.6) clustering methods. All the default parameters of the structure were used and burn-in was set at 50,000 followed by MCMC (Markov chain Monte Carlo) run for 100,000 steps. The MCMC runs were repeated 3 times for each value of K from 1 to 10.

Variance partitioning of historical climate

Estimation of total genomic variation explained by climate, one hundred and three climatic variables and distance matrix (spatial) variables were calculated using the "*varpart*" function built in the "*vegan*" package. The "*varpart*" function categorized the important variables that are involved in shaping the SNP distribution of *Ae. tauschii* accessions. Adjusted R^2 was used to partition the variance explained by different explanatory variables. Out of the total of 103 explanatory variables, which also included 19 bioclimatic variables and the monthly average of precipitation, mean, maximum and minimum temperature, water vapor pressure, wind, and solar energy, 12 unique explanatory variables were identified that were non-collinear. To remove the collinearity issue that can affect the results of downstream analysis, the variables that were correlated at Pearson Correlation Coefficients >0.85 were removed. The resulting set of 12 non-correlated explanatory variables (Table 3.2) were identified and used in variance partitioning and RDA analysis. The multivariate analysis of variance (MANOVA) was calculated using reduced permutation (999 number of permutation) to identify amount of genetic variation in *Ae. tauschii* population explained by climate

Redundancy analysis is a form of multivariate linear regression analysis built in the R package "*vegan*" (Oksanen, 2011). Redundancy analysis (RDA) partitions the variance and estimates the amount of variation in multivariate responses explained by the multiple sets of explanatory variables and also by the variables that are significantly correlated with them (Peres-Neto et al., 2006). SNPs that are outliers in the regression analysis are identified through the threshold of three standard deviation cutoff from mean implemented using R function. Using a simple Perl script the SNPs identified through RDA analysis in diploid *Ae. tauschii* as significantly associated with the 12 environmental variables were identified in the D genome SNPs of the hexaploid inter-specific population developed through crossing *Ae. tauschii* and wheat.

Results

Genotyping and SNP imputation

Illumina sequencing generated a total of ~315 million high-quality reads with an average of ~2.8 million reads per sample from the *Ae. tauschii* population. Out of this, eighty-eight percent (88%) of the reads were aligned to RefSeq v.1.0 (The International Wheat Genome Sequencing Consortium (IWGSC) 2018). Using TASSEL v. 5.0 GBS v.2 pipeline 148,430 SNP were generated out of which a total of 96,056 SNPs with MAF > 0.02 and < 50% missing data were retained after the filtering step.

In case of the inter-specific introgression population, ~1.1 billion high quality reads with barcodes were generated, with an average of ~3.0 million reads per sample. Ninety-six percent (96%) of the reads were aligned to RefSeq v1.0 with an average of ~2.9 million reads per sample. Samples with less than 91,000 reads were excluded from downstream analysis. The number of unfiltered SNPs generated by the TASSEL v.5.0 GBS v.2 pipeline was 281,846. A total of 13,970 SNPs from the A, B, D genomes, and unanchored scaffolds were retained after

filtering out SNPs with MAF less than 0.05 and missing genotypes greater than 30%. The number of SNPs from the D genome was 37.4 % of the filtered SNP dataset. The second round of filtering was performed on the D genome SNPs to remove sites with MAF less than 0.01 resulting 27,880 SNPs that were segregated in the diverse set of *Ae. tauschii* accessions (henceforth, reference panel). Using the program Beagle v.5.0 (Browning and Browning, 2013), 85,752 SNPs were imputed from the reference panel into the *Ae. tauschii* derived introgression population.

Principal component analysis

The analysis of principal components was performed to assign the *Ae. tauschii* accessions used as parents for introgression population development to the previously identified lineages (Wang et al., 2013). A total of 21 accessions out of 137 *Ae. tauschii* accessions are selected as parents. These parental accessions originate from different eco-geographic regions as shown in Figure 3.5. The *Ae. tauschii* accessions (21 parental accessions plus 116 base panel) formed three distinct clusters when the first two PCs calculated from 27,880 SNPs were plotted with 79.2 % of variance explained by the first two PCs (Figure 3.6).

Population genetic structure inferred from STRUCTURE and "tess3r" R package remained the same. In the "tess3r", R package ancestry coefficients are based on least-squares approximation combining matrix factorization and spatial statistics (Caye et al., 2016). The cross-validation error was minimized at K=5 suggesting that there are 5 genetically distinct groups in *Ae. tauschii*. Consistent with earlier studies by Wang et al. (2013), at K=2 lineages L1 and L2 cluster separately, and at K=3, lineage L1 splits into lineages L1a and L1b (Figure 3.9). At K=4, the sub-lineage L1a further splits into two distinct population, that I named as L1a and L1a* geographically. A population that is originating from Uzbekistan, Tajikistan,

Turkmenistan, and Kyrgyzstan is L1a*and another population originating from mostly Afghanistan, Pakistan, and Tajikistan is L1a. The bar plot and the map in Figures 3.8 and 3.9 respectively plot the population structure of *Aegilops taushii* at ancestry coefficients K=4. At K=5, the L2 splits into two distinct populations, L2.W and L2E genetically and geographically. The *Ae. tauschii* ssp. *strangulata* or sub-lineage two west (L2.W) cluster was distributed mostly in Armenia and Azerbaijan, and sub-lineage two east (L2.E) clustered in the Caspian Sea in Iran. The *Ae. tauschii* ssp. *tauschii* or lineage two (L1a and L1b) spread in a broad area (Wang et al., 2013). Cluster L1a was the most heterogeneous one with accessions coming from Afghanistan (AFG), Turkmenistan (TKM), Iran (IRN), Pakistan (PAK) and Tajikistan (TJK). Fifteen of the *Ae. tauschii* donor parents of the introgression population are from cluster L1a. More than two third of the accessions in cluster L1b were from Turkey (TUR) with only a few accessions from Armenia (ARM), IRN, TJK, and PAK. Out of L1b lineage, three accessions were selected as donor parents of the introgression population. Lineage L2 cluster consisted of *Ae. tauschii* accessions mostly collected from Iran (IRN), and a few accessions from Azerbaijan (AZE), Turkmenistan (TKM), Syria (SYR), and Turkey (TUR) were present. The remaining three donor parents of the introgression population were selected from *Ae. tauschii* ssp. *strangulata* L2.

First two principal components of the introgression lines, *Ae. tauschii* accessions and the recurrent hexaploid parents, show the distribution of inter-specific lines for its recurrent and donor parents (Figure 3.7). The two PCs together explained 56.7% of the variance. *Aegilops tauschii* accessions in lineage L2 clustered close to the hexaploid wheat parents cluster (HW) whereas the *Ae. tauschii* accessions in lineage L1a and L1b clustered together. The newly developed introgression lines spread broadly towards the hexaploid wheat cluster suggesting

their genetic similarity with its recurrent parents. This might also be due to the loss of the introgressed segments from *Ae. tauschii* while backcrossing with the recurrent parent.

Weir and Cockerham (1989) weighted F_{ST} estimate between broad lineage (L1 and L2) and sub-lineage (L1b, L1a and L1a*) are shown in table 3.3. The result suggests that the lineage L2 is highly differentiated with all three sub-lineages of L1 (L1a, L1a* and L1b) with F_{ST} score higher than 0.7. This result is consistent with previous findings suggesting that these two broad lineages are well differentiated genetically (Wang et al., 2013). The pairwise F_{ST} coefficient of L1b with L1a is slightly higher (0.2247) as compared to L1a* (0.1814) meaning L1a is more differentiated than L1a* from L1b. These results align with existing literature reporting divergence of L1a and L1b (Wang et al., 2013). The sub-lineage L1a and La* are differentiated with F_{ST} score of 0.0680. The Tajima's D score for lineage L1a, L1b, and the hybrids between lineages was positive. Based on the principal components analysis result, potential hybrid between lineages are identified. For example, for any individual the percentage of first four principal components is identified and if first two principle components explain significant amount (~40% each) of the variation then that individual is called as hybrid. However, the Tajima's D score for lineage L2 and L1a* was negative suggesting a recent selective sweep.

Eco-geographic climatic variables

This diverse collection contains accessions ranging from the Caucasus region between the Black and Caspian Sea in the west up to central Asia to Pakistan in the east. Wide variation in climatic data and spatial resolutions is used in the analysis (Hijmans et al., 2005; Fick, et al., 2017). The climate data contains 19 bioclimatic variables and the monthly average data of minimum, solar radiation, wind speed, maximum and average temperature, precipitation, and water vapor pressure. There exist distinct variations in solar radiation and water vapor pressures across *Ae.*

tauschii growing regions as shown in Figures 3.3 and 3.4, respectively. The climatic factors that are highly correlated at Pearson correlation 0.85 were removed from the analysis to correct for the collinearity issues in RDA analysis by using Pearson correlation.

Identifying major climatic factors driving adaptation and SNPs associated with it

Analysis of variance showed that climate and geography played a significant role in shaping the genetic makeup of *Ae. tauschii* accessions. The function "varpart" from Vegan was used to find the percent variation explained by each climatic variable. Figure 3.10 shows that geography and climate together explain 65% of SNPs variation where climatic factors alone explained 28% and geography explained 55% of SNPs distribution ($P \leq 0.001$). However, when collinear variables from environmental factors were removed, the variance explained by geography dropped to 36% and the SNPs variation explained by climate remained constant at 28% (Figure 3.11). The adjusted R^2 calculated which is also known as the proportion of variance explained by climate only after taking account of the spatial factor was 55.82% for all 103 climatic factors whereas its 36.49% for the 12 non-collinear variables. The multivariate analysis of variance (MANOVA) found that the climate explains a significant ($P \leq 0.001$) amount of genetic variation in *Ae. tauschii* population

Redundancy analysis (Kölliker, et al., 2006) was applied to find a correlation between SNP diversity and geographic location of the collection site (Figure 3.12). Major drivers of SNP distribution along lineages and sub lineages greatly varied. Water vapor pressure (Figure 3.4) and precipitation during driest months (Figure 3.13) were the most prominent factors driving SNP distribution in lineage L2. Solar radiation during summer months (Figure 3.3) and temperature seasonality (Figure 3.14) were the most prominent factors in lineage L1b. Similarly, precipitation

seasonality (Figure 3.15) and diurnal range (Figure 3.16), and maximum temperature in April were the major driver's lineage L1a. Further analysis of L1a, L1a*, and hybrid between lineages and their onsite environment shows that climatic drivers for L1a and L1a* differ greatly. Results show that the L1a lineage is influenced by the mean diurnal range, the maximum temperature in November, precipitation seasonality and solar radiation in April whereas for most of the accession in L1a*, the major climatic influencers were latitudinal gradient, the water vapor pressure in summer, precipitation of driest month and wind speed in March (Figure 3.17). Matsuoka et al. (2008) report the flowering time and seed production difference and strong gradient with the latitude of the L1a accessions which validates some of the historical climatic variables as true drivers.

A total of five hundred and eight candidate SNPs was identified in *Ae. tauschii* through RDA analysis. Out of these SNPs, a hundred and twenty-two were found in the D genome of interspecific population validating the allelic introgression of *Ae. tauschii* into the hexaploid gene pool.

Discussion

This study demonstrates the identification and transfer of climate adaptive SNPs from *Ae. tauschii* into the hexaploid wheat gene pool. The genome-wide environment scan shows that environmental factors play a major role in shaping the genetic architecture of *Ae. tauschii* accessions. The population structure analysis result reveals sub-lineages within two broad lineages (L1 and L2) of *Ae. tauschii* lineages that is also supported by existing literature (Wang et al., 2013). Further, the population structure result suggests a new and unique cluster (L1a*) within the lineage L1a. Association study using historical climatic data identified that this sub-lineage (L1a*) is strongly associated with the latitudinal gradient, water vapor pressure and wind

speed during summer. Using the 21 *Ae. tauschii* accessions (mostly from L1a, and L1a*) that are identified as genetically diverse and enriched for climatic variants, 351 inter-specific populations were developed by crossing it with Kansas adapted hexaploid wheat. The climate adaptive SNPs identified in *Ae. tauschii* were also identified in inter-specific population suggesting successful introgression of climate adaptive variants into hexaploid gene pool.

Wild relatives originating from diverse climatic conditions are a valuable resource of allelic diversity that can be employed in improving existing resources for biotic and abiotic stresses (Reynolds et al., 2007; Chen et al., 2015). However, the challenges that come from incompatible interactions of alien donor gene with the recipient parents might explain why these resources have not been explored extensively. This challenge was addressed through enrichment strategy, where the wild donor parent is selected based on the enrichment of genetic diversity. The *Ae. tauschii* accessions identified as the donor are also reported as the most diverse set (Singh et al., 2019). A diverse, segregating population of 351 inter-specific populations are developed through direct crossing of *Ae. tauschii* with six top performing Kansas adapted varieties was used to study the factors reducing the genomic patterns of introgression in the D genome of hexaploid wheat. Even though, a genome-wide introgression of adaptive SNPs from diploid *Ae. tauschii* to hexaploid inter-specific population was observed there exists reduced introgression of *Ae. tauschii* genes in this inter-specific population. A detailed genomic pattern of introgression study done by Nyine et al. (2020) using these 351 inter-specific populations revealed reduced introgression of *Ae. tauschii* SNPs in peri-centromeric regions, reduced retention of introgression alleles in regions controlling domestication genes. More importantly, introgression frequency was negatively correlated with the sequence divergence between the parental genomes. Most of the donor parents in this research are from *Ae. tauschii* sps *tauschii*

(Figure 3.1). The existing literature reports *Ae. tauschii* sps *strangulata as* the D genome donor of hexaploid wheat (Wang et al., 2013). Therefore, reduced introgression observed in the inter-specific population may be attributed to the fact that most of the donor parents used in development are not closely associated with the D genome of wheat (Figure 3.2). Early in the generation, hybrid sterility and natural selection acted on the survival of the inter-specific population culling out lots of lines. Overall, the issues with reduced introgression and retention of *Ae. tauschii* alleles are consistent with the fate of alleles in alien introgression as reported in existing literature (Jiang et al., 2000; Rieseberg, 2001)

Eco-geographic association analysis is a powerful tool to identify the environmental factors that influence adaptive allele distribution across populations. The change in the geographical range such as, latitude and altitude are widely reported as major contributors of species diversity. In the case of *Ae. tauschii*, the accessions in L1a* are found to be strongly associated with the latitudinal gradient whereas the L1a was associated with a longitudinal gradient. However, this kind of association has been reported in lineage one (L1) with longitudinal gradients whereas the lineage two (L2) that is mostly spread around the Caspian Sea is reported to be driven by altitude (Singh et al., 2019). Association with latitudinal gradient has also been reported in flowering time adaptation in maize (Buckler et al., 2009). However, there remains a gap in understanding the major climatic factors contributing to genetic differentiation within sub-lineages of *Ae. tauschii*. Association with geographic gradients were hypothesized to influence ultraviolet radiation, precipitation, temperature, and photoperiodism in plants (Buckler et al., 2009; Romero Navarro et al., 2017; Ziello et al., 2009), factors identified as major drivers by RDA analysis in *Ae. tauschii* as well. I found that the precipitation and water vapor pressure during September and October were the major drivers of genetic differentiation of L2 and within

L1 the factors that drive sub-lineage clustering were different. Specifically, the accessions within L1b were strongly associated with the amount of solar radiation during the summer season (June, July and August) and temperature seasonality whereas the accessions within L1a were strongly associated with the maximum temperature in April, precipitation seasonality, and the change in mean diurnal range. After partitioning the variance explained by climatic and geographic factors, I found that they together explain 65% of the SNPs variation for *Ae. tauschiii* collection. Out of which the geography containing the distance matrix explained most (55%) of the SNPs variation and climate containing precipitation, temperature, solar radiation, wind, and water vapor pressure explained 28% of the SNP variation. Even after using only 12 major climatic factors in RDA analysis, the percent variation explained by climate remains the same 28%. These findings suggest that climate together with geography plays a major role in driving the genetics underlying adaptation to environmental conditions. And through the use of environmental association studies leveraging on next generation sequence was able to identify the major drivers of local adaptation.

In addition to revealing the predominant forces shaping adaptation across lineages, this study identified five hundred and eight candidate SNPs that are strongly correlated with the major climatic drivers of local adaptation in *Ae. tauschii* population. Comparing the identified adaptive SNPs with the inter-specific population, 322 SNPs out of 508 adaptive SNPs were retained. The preliminary results from the field trial of drought vs irrigated condition have shown some promising results as I find 25% of the lines out of 351 with more yields (bu/ac) than the best performing checks currently in Kansas. Even though this finding is very encouraging for a pre-breeder and geneticist, there remains a need for a detailed understanding of the assembly and expression of these introgression alleles in these 25% lines vs the remaining. In summary, I

demonstrate the potential of *Ae. tauschii* in improving the adaptive potential of hexaploid wheat through identification and prioritization of climate adaptive SNP using historical climatic data.

Conclusions

Increasing the productivity of existing germplasm will require them to adapt to the extreme climate of the future where drought and heat are projected to increase. Transfer of adaptive alleles from wild relatives could help to improve the rate of genetic gain in breeding programs. This study can assist wheat breeders in identifying *Ae. tauschii* accessions that can be used as donor parents for breeding climate-resilient wheat variety. The SNP identified in this study can be used as a prior SNPs to optimize predicting models for drought and heat conditions. And more in-depth studies on the fate of these SNPs in the genome may lead to the development of better genomic resources for targeted breeding in wheat.

References

Akhunov, Eduard D., Alina R. Akhunova, Olin D. Anderson, James A. Anderson, Nancy Blake, Michael T. Clegg, Devin Coleman-Derr et al. "Nucleotide diversity maps reveal variation in diversity among wheat genomes and chromosomes." *BMC genomics* 11, no. 1 (2010): 702.

Browning, Brian L., and Sharon R. Browning. "Improving the accuracy and efficiency of identity-by-descent detection in population data." *Genetics* 194, no. 2 (2013): 459-471.

Borcard, Daniel, François Gillet, and Pierre Legendre. *Numerical ecology with R*. Springer, 2018.

Buckler, Edward S., James B. Holland, Peter J. Bradbury, Charlotte B. Acharya, Patrick J. Brown, Chris Browne, Elhan Ersoz et al. "The genetic architecture of maize flowering time." *Science* 325, no. 5941 (2009): 714-718.

Caye, Kevin, Flora Jay, Olivier Michel, and Olivier François. "Fast inference of individual admixture coefficients using geographic data." *The Annals of Applied Statistics* 12, no. 1 (2018): 586-608.

Caye, Kevin, Timo M. Deist, Helena Martins, Olivier Michel, and Olivier François. "TESS3: fast inference of spatial population structure and genome scans for selection." *Molecular Ecology Resources* 16, no. 2 (2016): 540-548.

Chapman, Jarrod A., Martin Mascher, Aydın Buluç, Kerrie Barry, Evangelos Georganas, Adam Session, Veronika Strnadova et al. "A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome." *Genome biology* 16, no. 1 (2015): 26.

Chen, Shisheng, Matthew N. Rouse, Wenjun Zhang, Yue Jin, Eduard Akhunov, Yuming Wei, and Jorge Dubcovsky. "Fine mapping and characterization of Sr21, a temperature-sensitive diploid wheat resistance gene effective against the Puccinia graminis f. sp. tritici Ug99 race group." *Theoretical and Applied Genetics* 128, no. 4 (2015): 645-656.

Cox, T. S. "Deepening the wheat gene pool." *Journal of Crop Production* 1, no. 1 (1997): 1-25.

Cox, Thomas S., Jizhong Wu, Shuwen Wang, Jin Cai, Qiaofeng Zhong, and Bisheng Fu. "Comparing two approaches for introgression of germplasm from Aegilops tauschii into common wheat." *The Crop Journal* 5, no. 5 (2017): 355-362.

Dale, Zhang, He Jie, Huang Luyu, Zhang Cancan, Zhou Yun, Su Yarui, and Li Suoping. "An advanced backcross population through synthetic octaploid wheat as a “bridge”: development and QTL detection for seed dormancy." *Frontiers in plant science* 8 (2017): 2123.

Danecek, Petr, Adam Auton, Goncalo Abecasis, Cornelis A. Albers, Eric Banks, Mark A. DePristo, Robert E. Handsaker et al. "The variant call format and VCFtools." *Bioinformatics* 27, no. 15 (2011): 2156-2158.

Dvorak, J., M-C. Luo, Z-L. Yang, and H-B. Zhang. "The structure of the Aegilops tauschii genepool and the evolution of hexaploid wheat." *Theoretical and Applied Genetics* 97, no. 4 (1998): 657-670.

FAO, Unesco. "Global agriculture towards 2050." *FAO Rome* (2009).

Fick, Stephen E., and Robert J. Hijmans. "WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas." *International journal of climatology* 37, no. 12 (2017): 4302-4315.

Friesen, T. L., S. S. Xu, and M. O. Harris. "Stem rust, tan spot, Stagonospora nodorum blotch, and Hessian fly resistance in Langdon durum–Aegilops tauschii synthetic hexaploid wheat lines." *Crop science* 48, no. 3 (2008): 1062-1070.

Fritz, A. K., T. S. Cox, B. S. Gill, and R. G. Sears. "Marker-Based Analysis of Quantitative Traits in Winter Wheat× Triticum tauschii Populations." *Crop science* 35, no. 6 (1995): 1695-1699.

Glaubitz, Jeffrey C., Terry M. Casstevens, Fei Lu, James Harriman, Robert J. Elshire, Qi Sun, and Edward S. Buckler. "TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline." *PloS one* 9, no. 2 (2014): e90346.

He, Fei, Raj Pasam, Fan Shi, Surya Kant, Gabriel Keeble-Gagnere, Pippa Kay, Kerrie Forrest et al. "Exome sequencing highlights the role of wild-relative introgression in shaping the adaptive landscape of the wheat genome." *Nature Genetics* 51, no. 5 (2019): 896-904.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. and Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *International journal of climatology*, *25*(15), pp.1965-1978.

Hill, William G., and Alan Robertson. "The effect of linkage on limits to artificial selection." *Genetics Research* 8, no. 3 (1966): 269-294.

Hsam, S. L. K., R. Kieffer, and F. J. Zeller. "Significance of Aegilops tauschii glutenin genes on breadmaking properties of wheat." *Cereal Chemistry* 78, no. 5 (2001): 521-525.

IWGSC, Appels R., K. Eversole, C. Feuillet, B. Keller, J. Rogers, and N. Stein. "Shifting the limits in wheat research and breeding using a fully annotated reference genome." *Science* 361, no. 6403 (2018): eaar7191-eaar7191.

Jia, Jizeng, Shancen Zhao, Xiuying Kong, Yingrui Li, Guangyao Zhao, Weiming He, Rudi Appels et al. "Aegilops tauschii draft genome sequence reveals a gene repertoire for wheat adaptation." *Nature* 496, no. 7443 (2013): 91-95.

Jiang, Chun-xiao, Peng Wah Chee, Xavier Draye, Peter L. Morrell, C. Wayne Smith, and Andrew H. Paterson. "Multilocus interactions restrict gene introgression in interspecific populations of polyploid Gossypium (cotton)." *Evolution* 54, no. 3 (2000): 798-814.

Jordan, Katherine W., Shichen Wang, Fei He, Shiaoman Chao, Yanni Lun, Etienne Paux, Pierre Sourdille et al. "The genetic architecture of genome-wide recombination rate variation in allopolyploid wheat revealed by nested association mapping." *The Plant Journal* 95, no. 6 (2018): 1039-1054.

Kastner, Thomas, Maria Jose Ibarrola Rivas, Wolfgang Koch, and Sanderine Nonhebel. "Global changes in diets and the consequences for land requirements for food." *Proceedings of the National Academy of Sciences* 109, no. 18 (2012): 6868-6872.

Kölliker, R., Kraehenbuehl, R., Boller, B. and Widmer, F., 2006. Genetic diversity and pathogenicity of the grass pathogen Xanthomonas translucens pv. graminis. *Systematic and applied microbiology*, *29*(2), pp.109-119.

Li, Huihui, Awais Rasheed, Lee T. Hickey, and Zhonghu He. "Fast-forwarding genetic gain." *Trends in plant science* 23, no. 3 (2018): 184-186.

Lubbers, E. L., K. S. Gill, T. S. Cox, and B. S. Gill. "Variation of molecular markers among geographically diverse accessions of Triticum tauschii." *Genome* 34, no. 3 (1991): 354-361.

Matsuoka, Yoshihiro, Shigeo Takumi, and Taihachi Kawahara. "Flowering time diversification and dispersal in central Eurasian wild wheat Aegilops tauschii Coss.: genealogical and ecological framework." *PLoS One* 3, no. 9 (2008): e3138.

Marcussen, Thomas, Simen R. Sandve, Lise Heier, Manuel Spannagl, Matthias Pfeifer, Kjetill S. Jakobsen, Brande BH Wulff et al. "Ancient hybridizations among the ancestral genomes of bread wheat." *science* 345, no. 6194 (2014).

McFadden, Edgar S., and Ernest Robert Sears. "The origin of Triticum spelta and its free-threshing hexaploid relatives." *Journal of Heredity* 37, no. 3 (1946): 81-89.

Naghavi, M. R., M. Ranjbar, A. Zali, M. J. Aghaei, M. Mardi, and S. M. Pirseyedi. "Genetic Diversity of Aegilops crassa and Its Relationship with Aegilops tauschii and the D Genome of Wheat." *Cereal Research Communications* 37, no. 2 (2009): 159-167.

Navarro, J. Alberto Romero, Martha Willcox, Juan Burgueño, Cinta Romay, Kelly Swarts, Samuel Trachsel, Ernesto Preciado et al. "A study of allelic diversity underlying flowering-time adaptation in maize landraces." *Nature genetics* 49, no. 3 (2017): 476.

Nyine, Moses, Elina Adhikari, Marshall Clinesmith, Katherine W. Jordan, Allan K. Fritz, and Eduard Akhunov. "Genomic patterns of introgression in interspecific populations created by crossing wheat with its wild relative." *G3: Genes, Genomes, Genetics* 10, no. 10 (2020): 3651-3661.

Oksanen, Jari. "Multivariate analysis of ecological communities in R: vegan tutorial." *R package version* 1, no. 7 (2011): 1-43.

Peres-Neto, Pedro R., Pierre Legendre, Stéphane Dray, and Daniel Borcard. "Variation partitioning of species data matrices: estimation and comparison of fractions." *Ecology* 87, no. 10 (2006): 2614-2625.

Reynolds, Matthew, Fernanda Dreccer, and Richard Trethowan. "Drought-adaptive traits derived from wheat wild relatives and landraces." *Journal of Experimental Botany* 58, no. 2 (2007): 177-186.

Rieseberg, Loren H. "Chromosomal rearrangements and speciation." *Trends in ecology & evolution* 16, no. 7 (2001): 351-358.

Saintenac C, Jiang D, Wang S, Akhunov E. Sequence-based mapping of the polyploid wheat genome. G3 (Bethesda). 2013; 3:1105–14.

Saintenac, Cyrille, Dayou Jiang, Shichen Wang, and Eduard Akhunov. "Sequence-based mapping of the polyploid wheat genome." *G3: Genes, Genomes, Genetics* 3, no. 7 (2013): 1105-1114.

Simons, Kristin J., John P. Fellers, Harold N. Trick, Zengcui Zhang, Yin-Shan Tai, Bikram S. Gill, and Justin D. Faris. "Molecular characterization of the major wheat domestication gene Q." *Genetics* 172, no. 1 (2006): 547-555.

Singh, Narinder, Shuangye Wu, Vijay Tiwari, Sunish Sehgal, John Raupp, Duane Wilson, Mehraj Abbasov, Bikram Gill, and Jesse Poland. "Genomic analysis confirms population structure and identifies inter-lineage hybrids in Aegilops tauschii." *Frontiers in plant science* 10 (2019): 9.

Tajima, Fumio. "Statistical method for testing the neutral mutation hypothesis by DNA polymorphism." *Genetics* 123, no. 3 (1989): 585-595.

Wang, Jirui, Ming-Cheng Luo, Zhongxu Chen, Frank M. You, Yuming Wei, Youliang Zheng, and Jan Dvorak. "Aegilops tauschii single nucleotide polymorphisms shed light on the origins of wheat D-genome genetic diversity and pinpoint the geographic origin of hexaploid wheat." *New phytologist* 198, no. 3 (2013): 925-937.

Weir, Bruce S., and C. Clark Cockerham. "Estimating F-statistics for the analysis of population structure." *evolution* (1984): 1358-1370.

Zhang, Dale, Yun Zhou, Xinpeng Zhao, Linlin Lv, Cancan Zhang, Junhua Li, Guiling Sun, Suoping Li, and Chunpeng Song. "Development and utilization of introgression lines using synthetic octaploid wheat (aegilops tauschii× hexaploid wheat) as Donor." *Frontiers in plant science* 9 (2018): 1113.

Zhang, Dale, Yun Zhou, Xinpeng Zhao, Linlin Lv, Cancan Zhang, Junhua Li, Guiling Sun, Suoping Li, and Chunpeng Song. "Development and utilization of introgression lines using synthetic octaploid wheat (aegilops tauschii× hexaploid wheat) as Donor." *Frontiers in plant science* 9 (2018): 1113.

Ziello, Chiara, Nicole Estrella, Mariya Kostova, Elisabeth Koch, and Annette Menzel. "Influence of altitude on phenology of selected plant species in the Alpine region (1971–2000)." *Climate Research* 39, no. 3 (2009): 227-234.

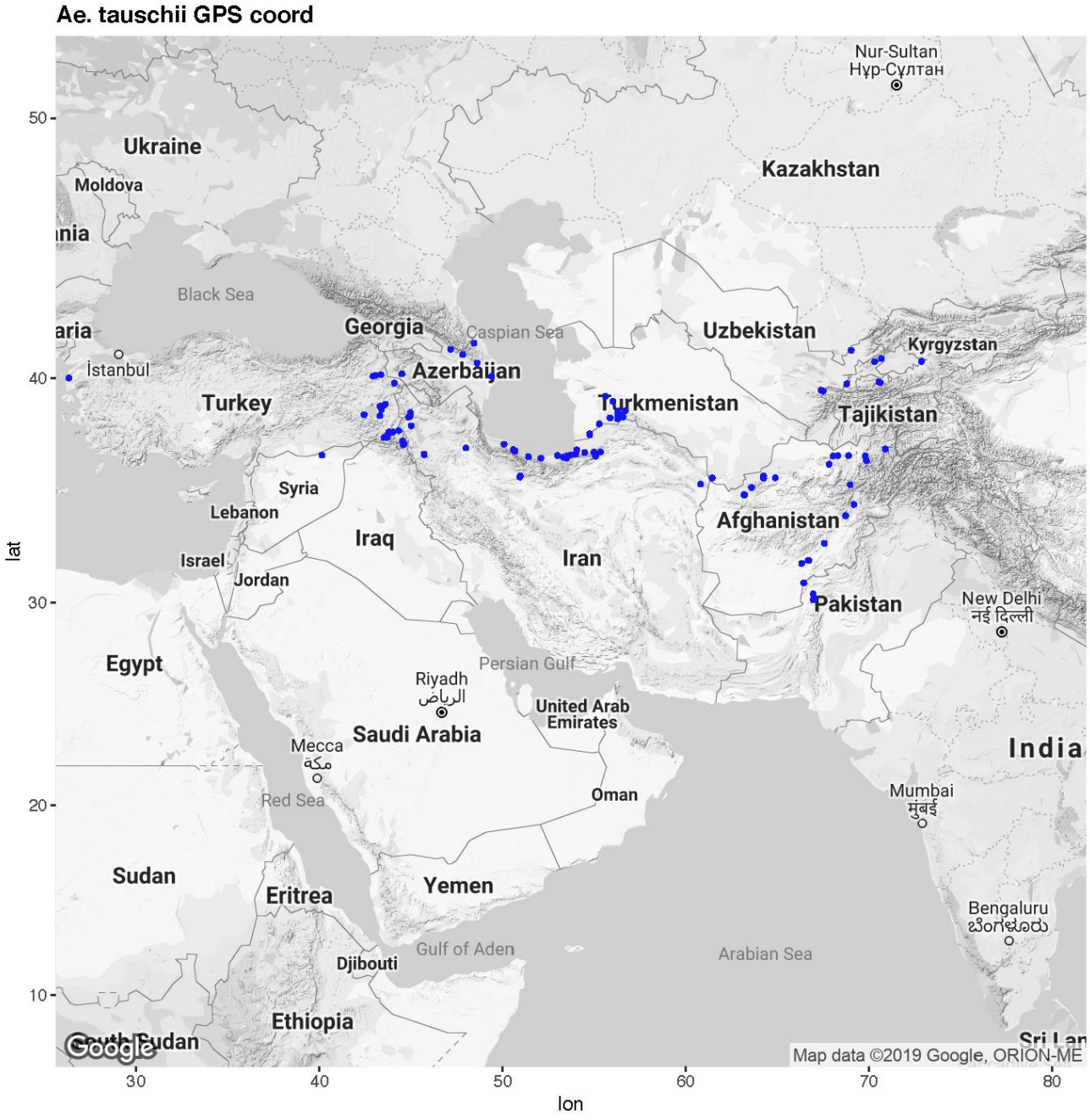


Figure 3.1. Eco-geographic distribution of *Aegilops tauschii* accessions.

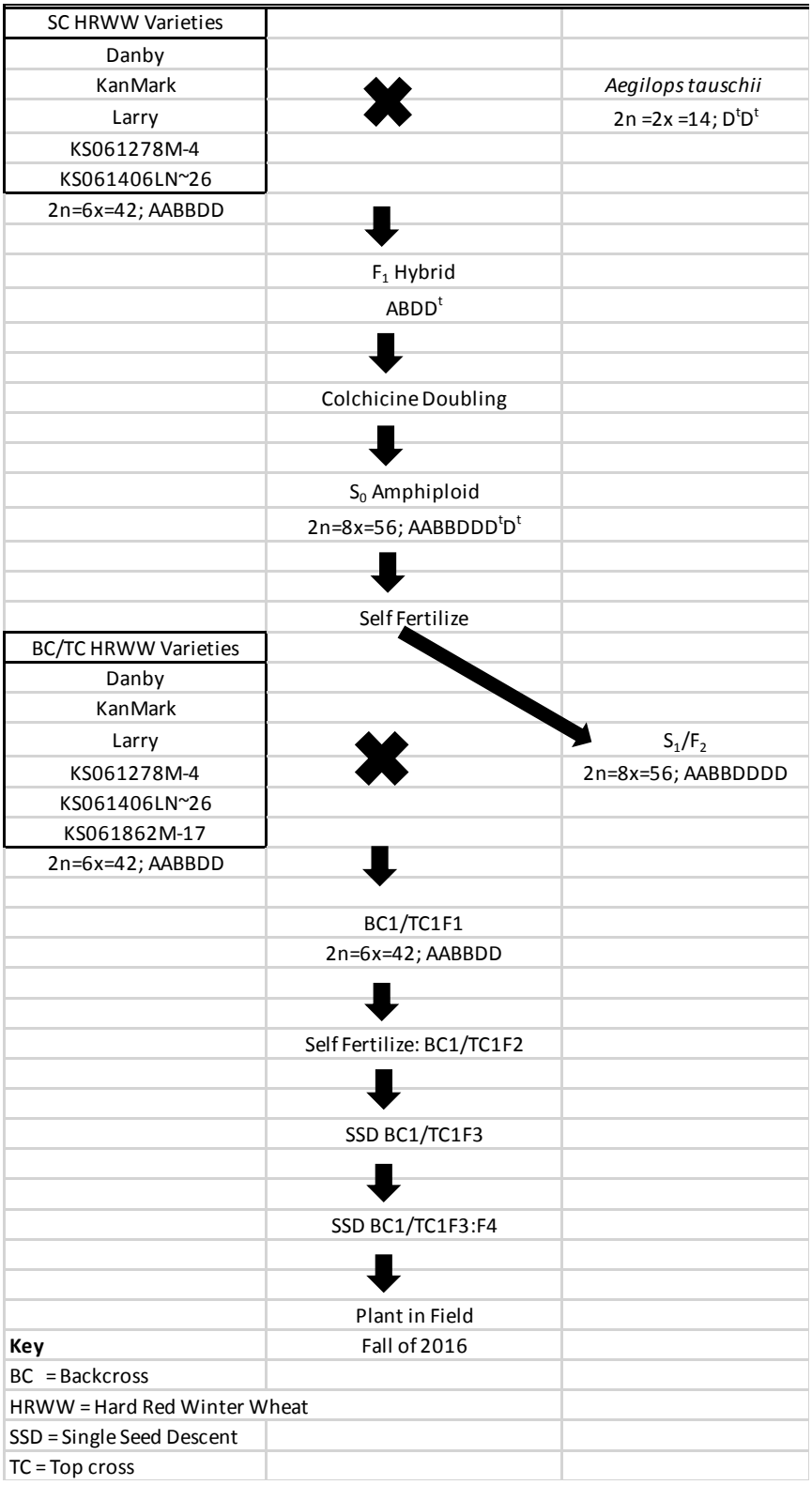

SC HRWW Varieties
Danby
KanMark
Larry
KS061278M-4
KS061406LN~26
2n=6x=42; AABBDD
Aegilops tauschii
2n =2x =14; DtDt
F1 Hybrid
ABDDt
Colchicine Doubling
S0 Amphiploid
2n=8x=56; AABBDDDtDt
Self Fertilize
BC/TC HRWW Varieties
Danby
KanMark
Larry
KS061278M-4
KS061406LN~26
KS061862M-17
2n=6x=42; AABBDD
S1/F2
2n=8x=56; AABBDDDD
BC1/TC1F1
2n=6x=42; AABBDD
Self Fertilize: BC1/TC1F2
SSD BC1/TC1F3
SSD BC1/TC1F3:F4
Plant in Field
Fall of 2016
Key
BC = Backcross
HRWW = Hard Red Winter Wheat
SSD = Single Seed Descent
TC = Top cross

Figure 3.2. Flow chart showing introgression of *Aegilops tauschii* into Kansas adapted hard red winter wheat.

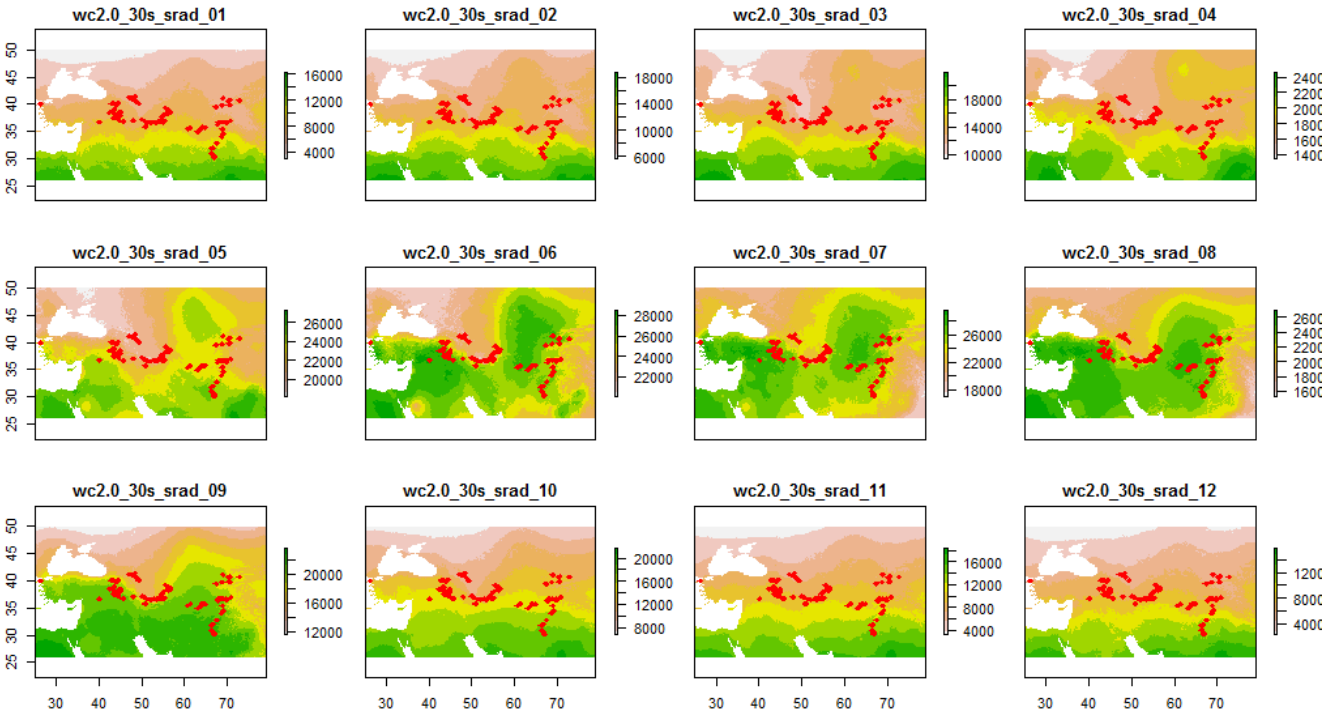


Figure 3.3. Geographic distribution of *Aegilops tauschii* accessions plotted along the solar radiation gradient.

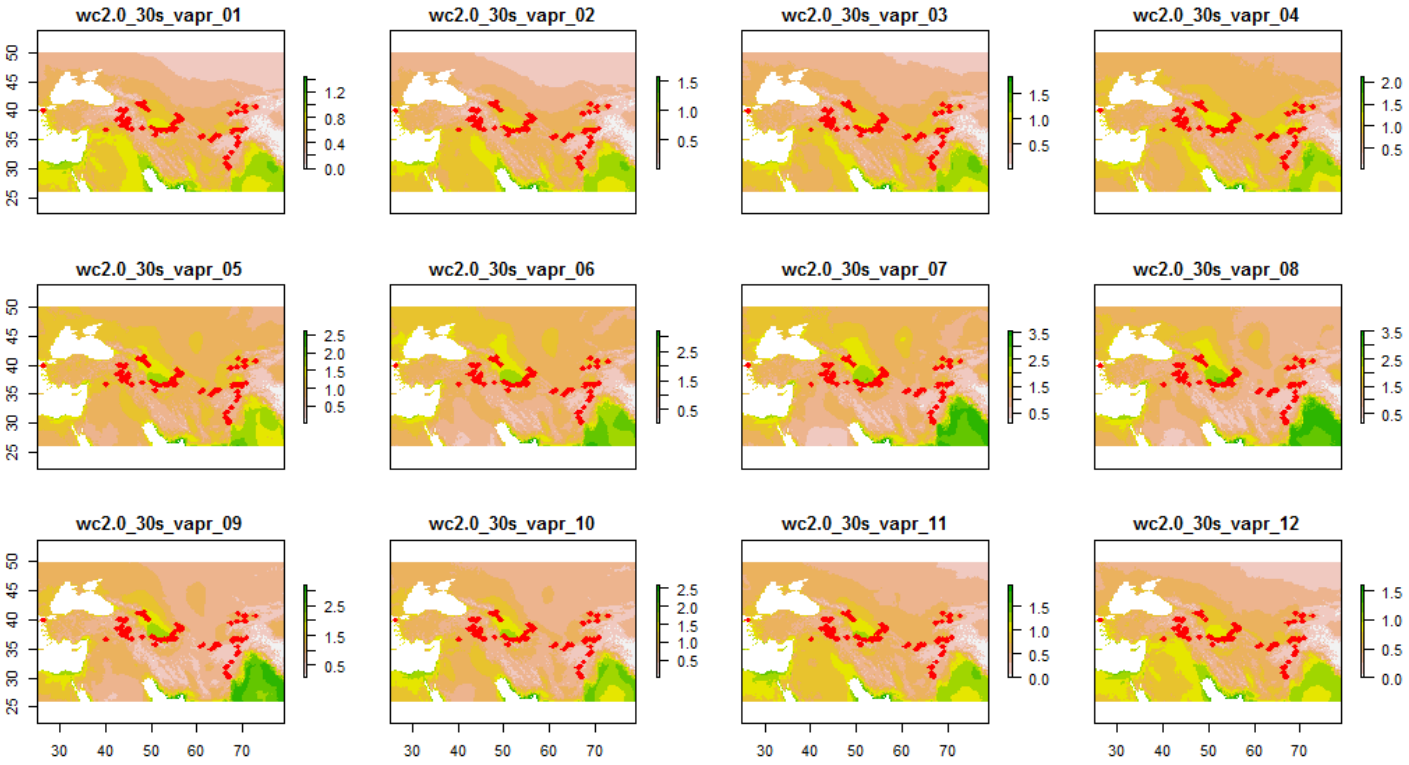


Figure 3.4. Geographic distribution of *Aegilops tauschii* accessions plotted at different levels of water vapor pressures.

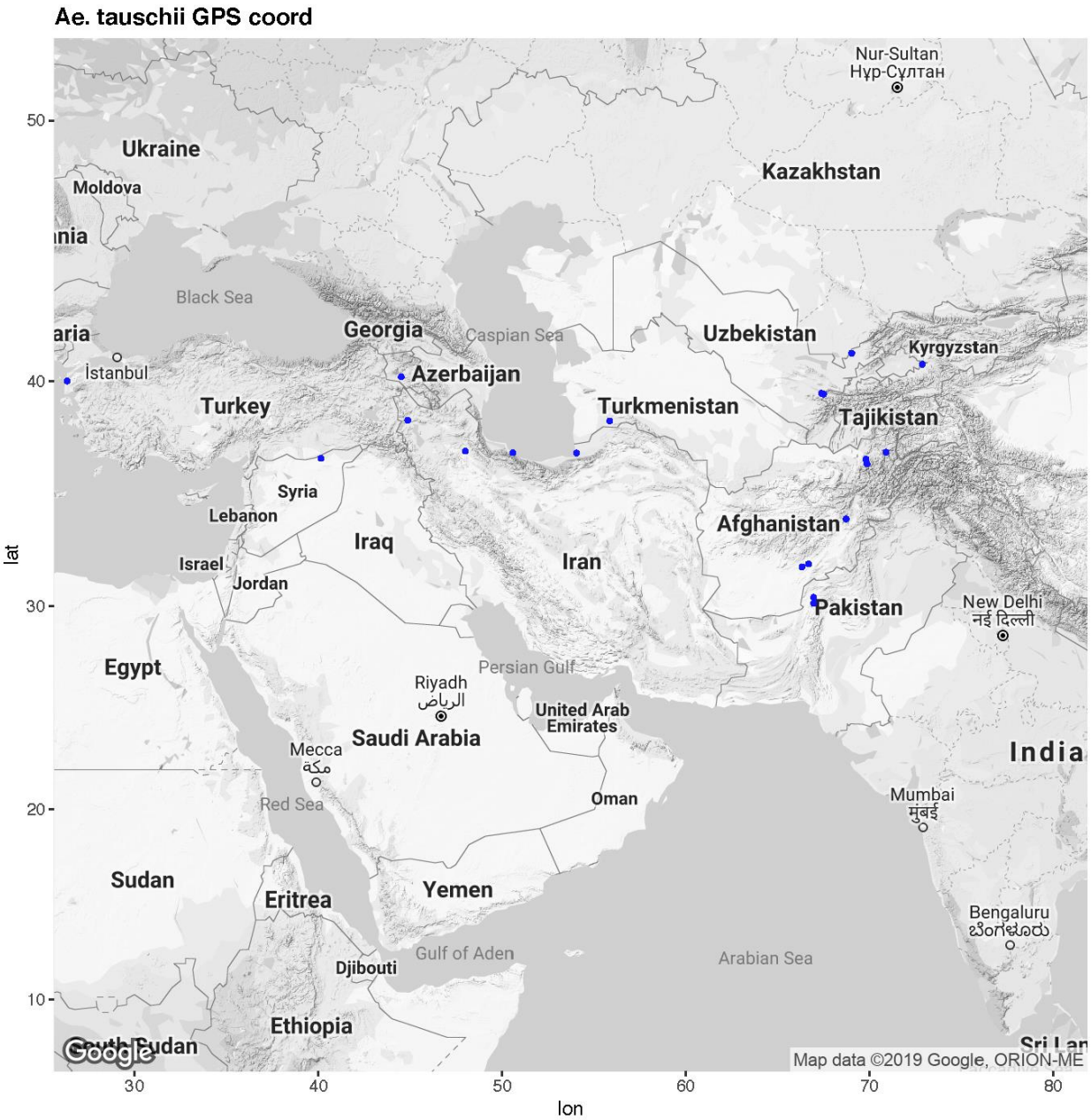


Figure 3.5. Eco-geographic distribution of *Aegilops tauschii* accessions selected for parents in direct hybridization.

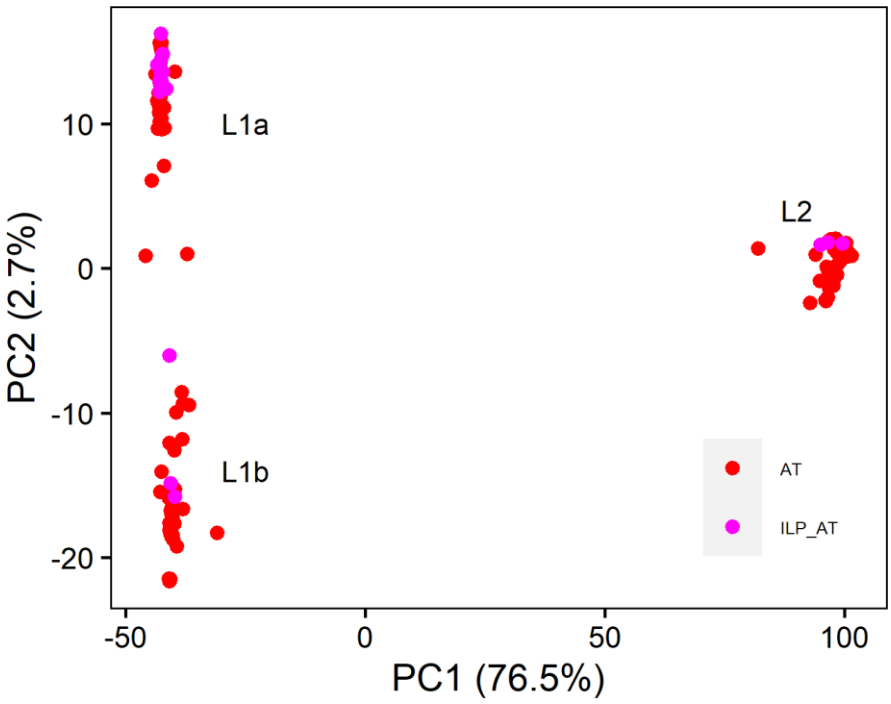


Figure 3.6. Distribution of *Aegilops taushii* accessions in two principal components.

Note: The pink dots represent accessions selected as parents in direct hybridization. The rest of the accessions are in red dots.

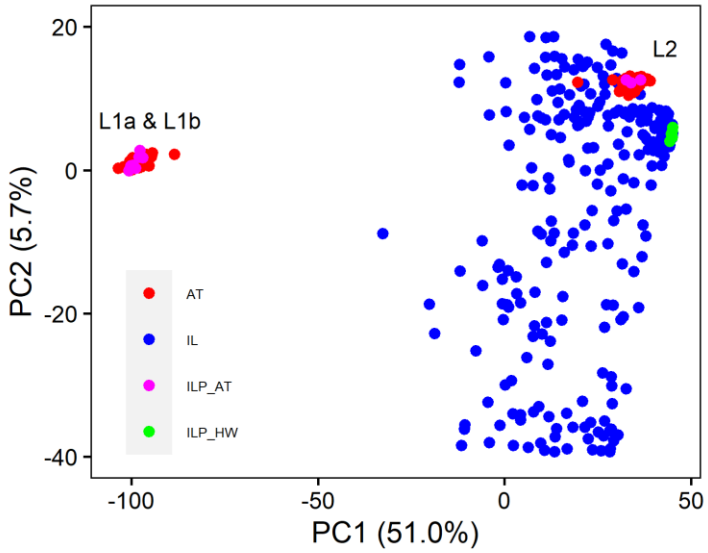


Figure 3.7. Distribution of first two principal components.

Note: Panel of 116 *Ae. taushii* accessions (AT) colored red, 21 *Ae. taushii* accessions used as donor parents (IL) in inter-specific population development colored pink, 6 hexaploid wheat used as recurrent parents (ILP_HW) in inter-specific population development colored green and 351 inter-specific lines (ILP_AT) created through direct hybridization of *Ae. tauschii* and wheat colored blue.

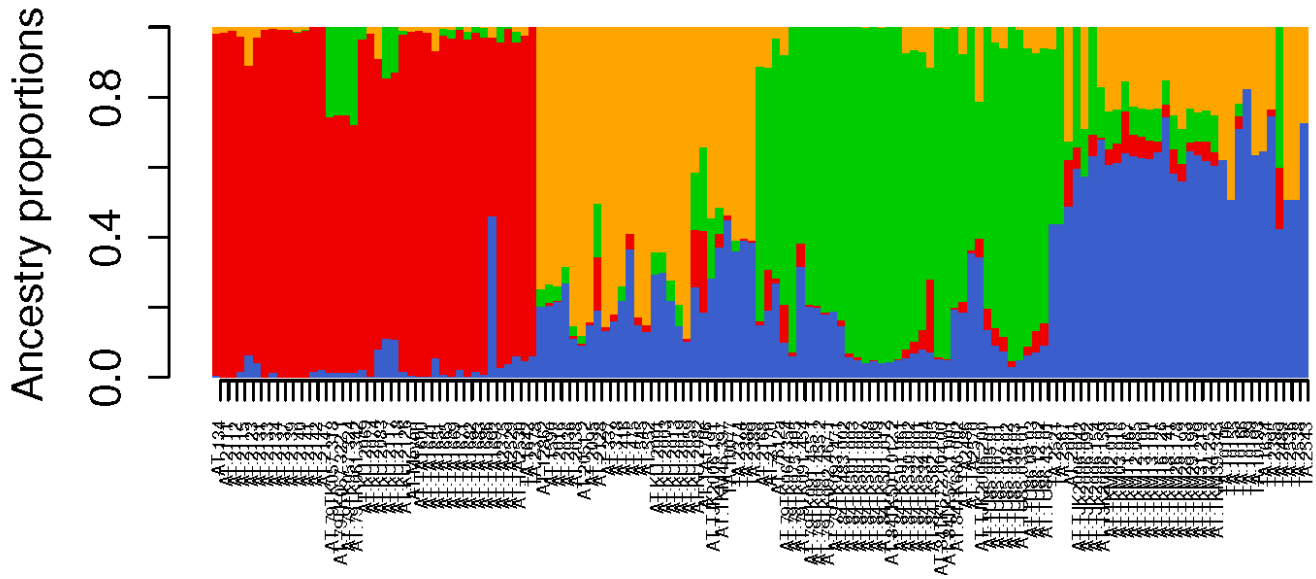


Figure 3.8. Bar plot showing population structure of *Aegilops taushii* at ancestry coefficients K=4.

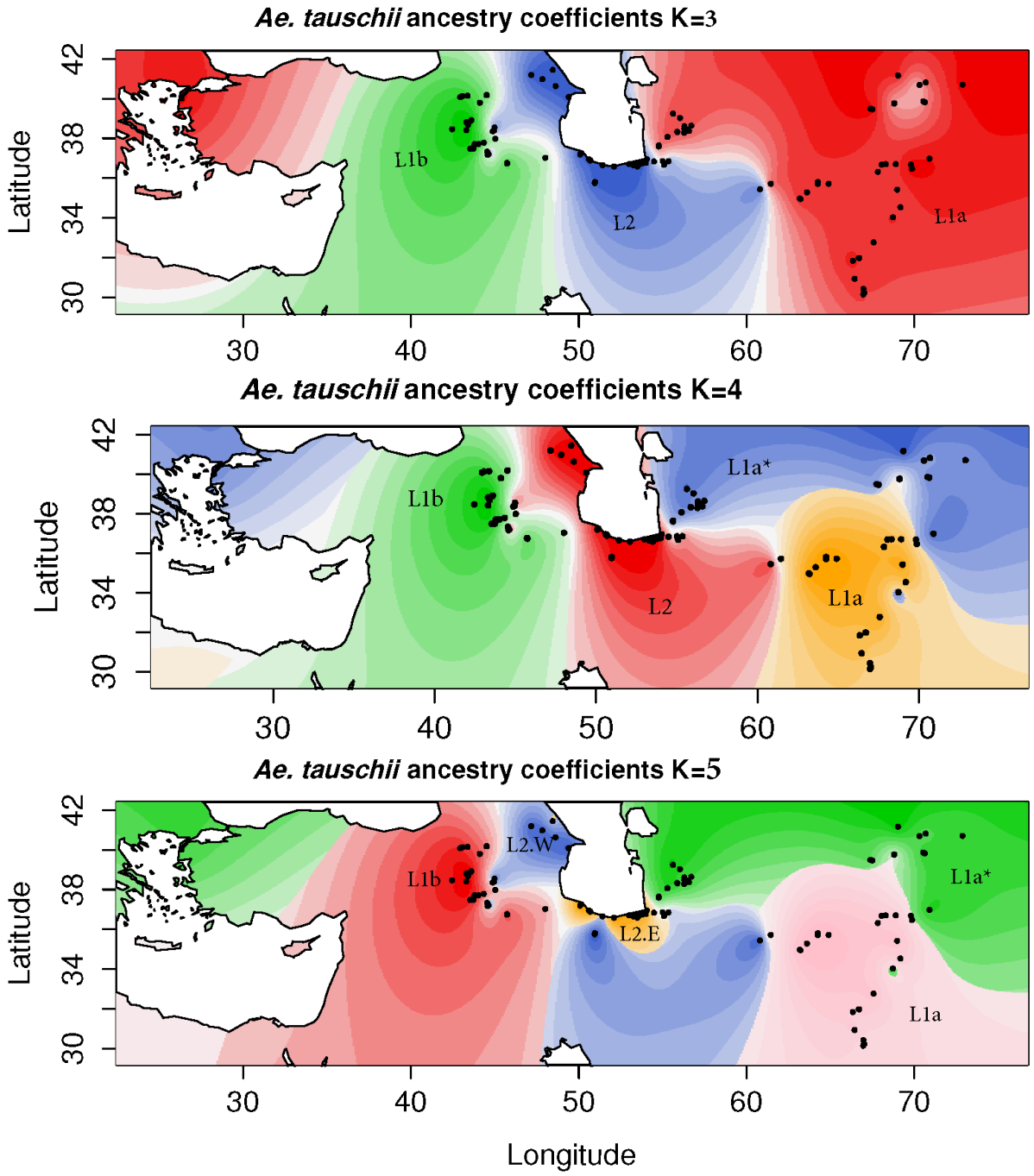


Figure 3.9. Spatial genetic ancestry structure of *Aegilops taushii* at ancestry matrix K=3:5 where each lineage are color coded differently.

Note: At K=3, green, red and blue shaded regions represents L1b, L2, L1a lineages. At K=4, green, red, blue and orange shaded regions represents L1b, L2, L1a*, and L1a lineages. At K=5, blue, orange, red, pink, and green shaded regions represents L2.W, L2.E, L1b, L1a, and L1a* lineages.

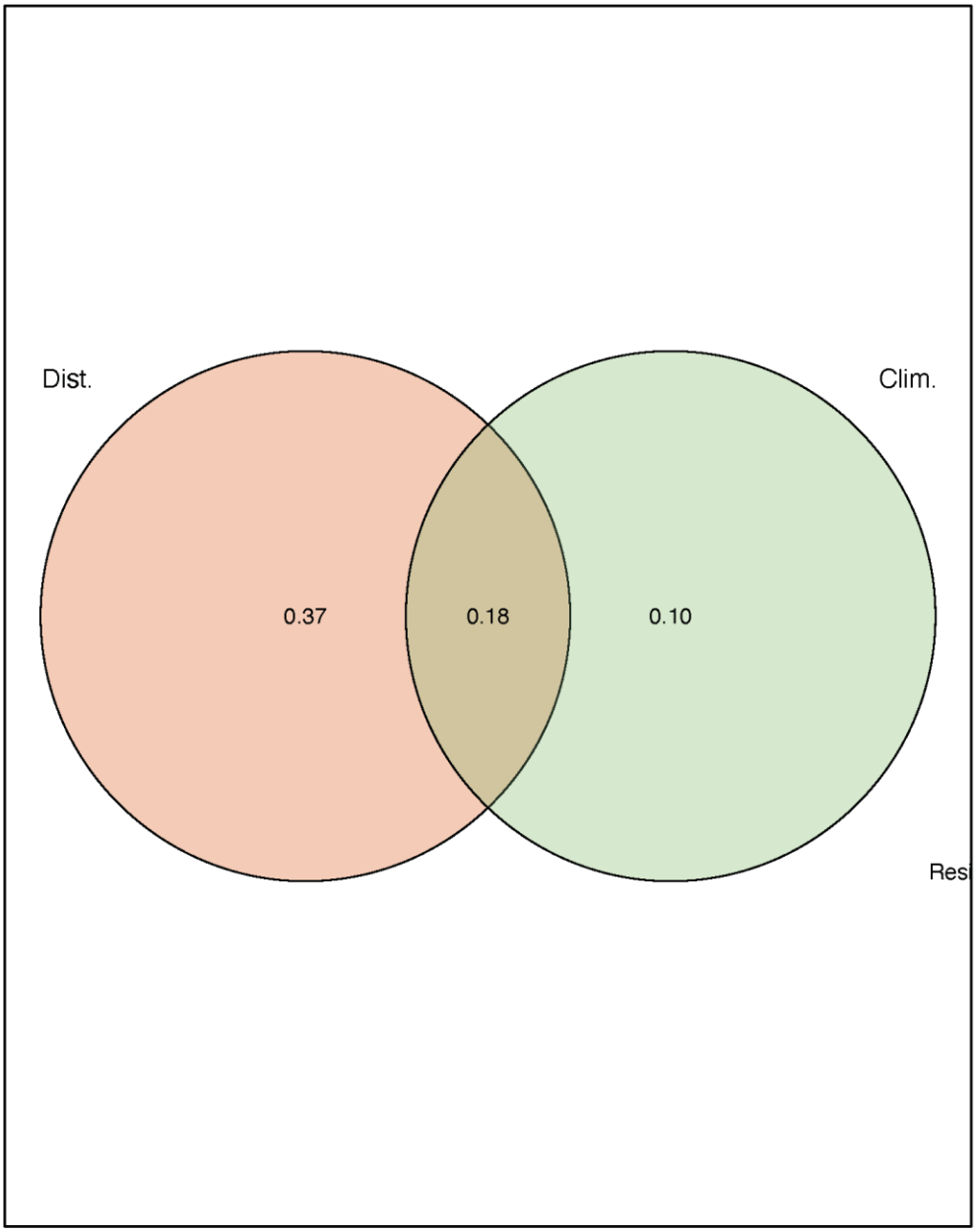


Figure 3.10. Ven diagram showing variance portioning of all 103 climatic factors and geography (as measured by distance).

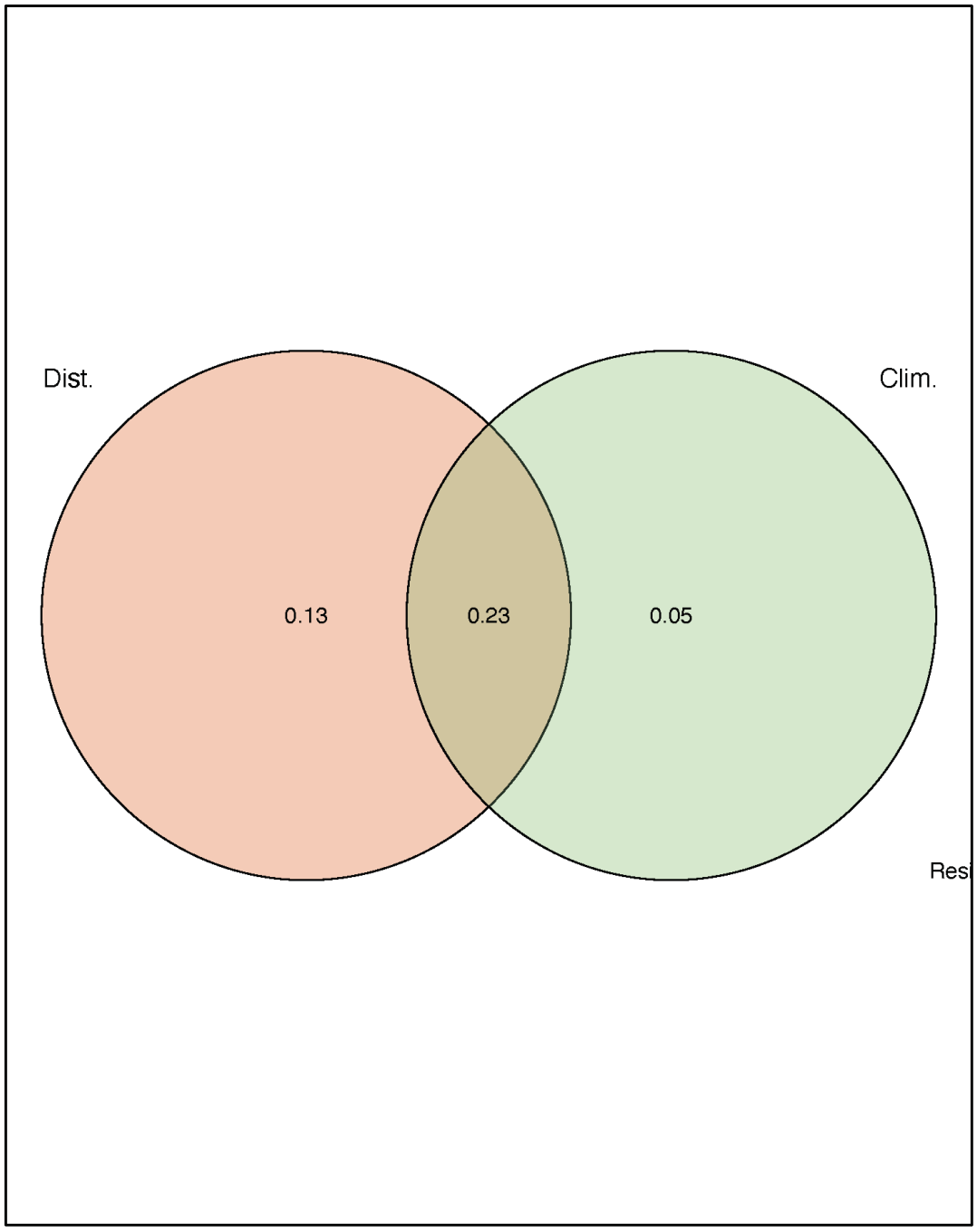


Figure 3.11. Venn diagram showing variance portioning of 12 non-collinear climatic factors and geography (as measured by distance).

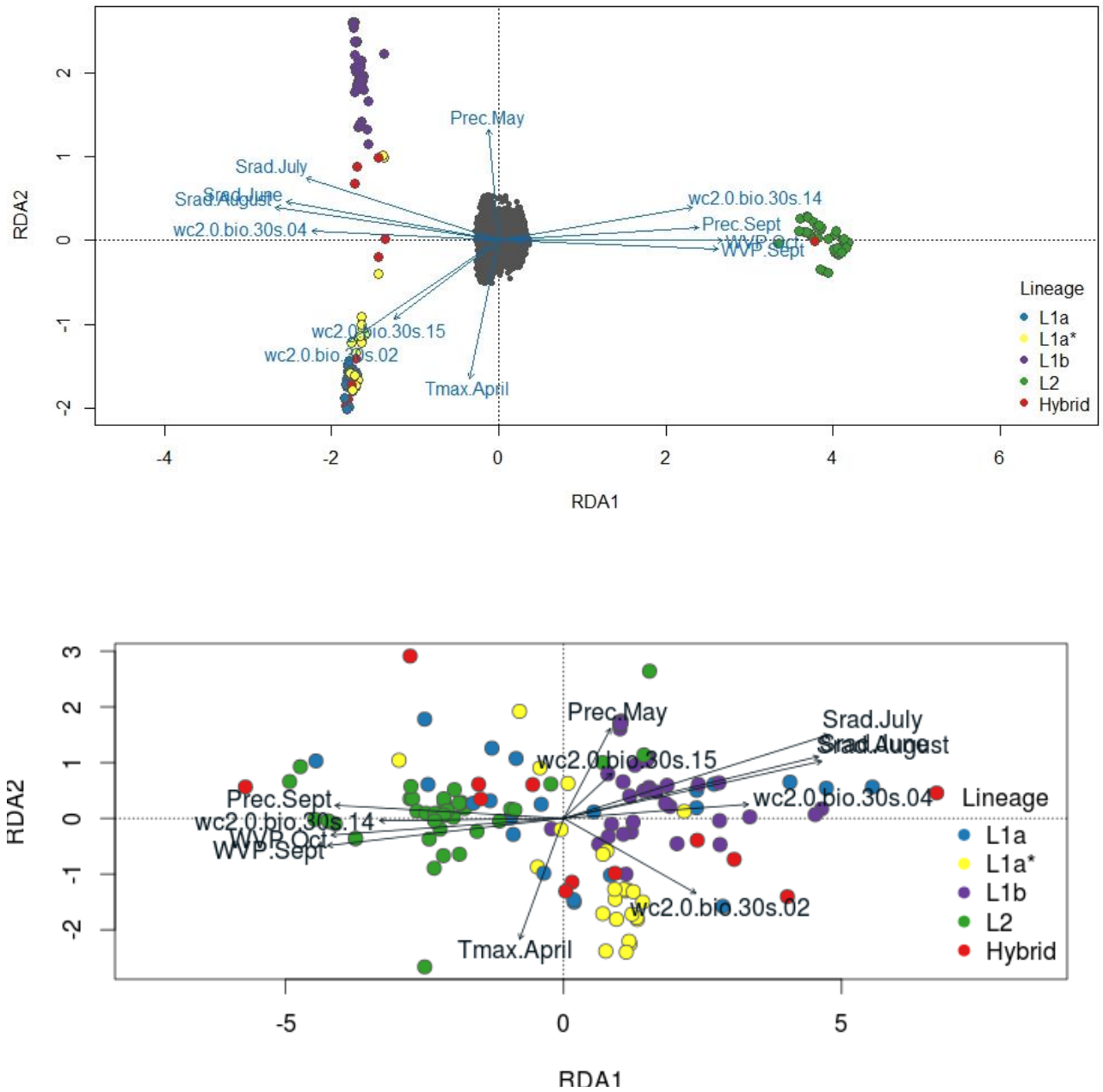


Figure 3.12. Bi-plot showing major climatic drivers along different *Aegilops tauschii* lineages.

Note: The first bi-plot was generated through RDA, which considers 12 climatic and geographic factors. The second bi-plot shows the result from partial RDA, which considers effect of 12 climatic factors only.

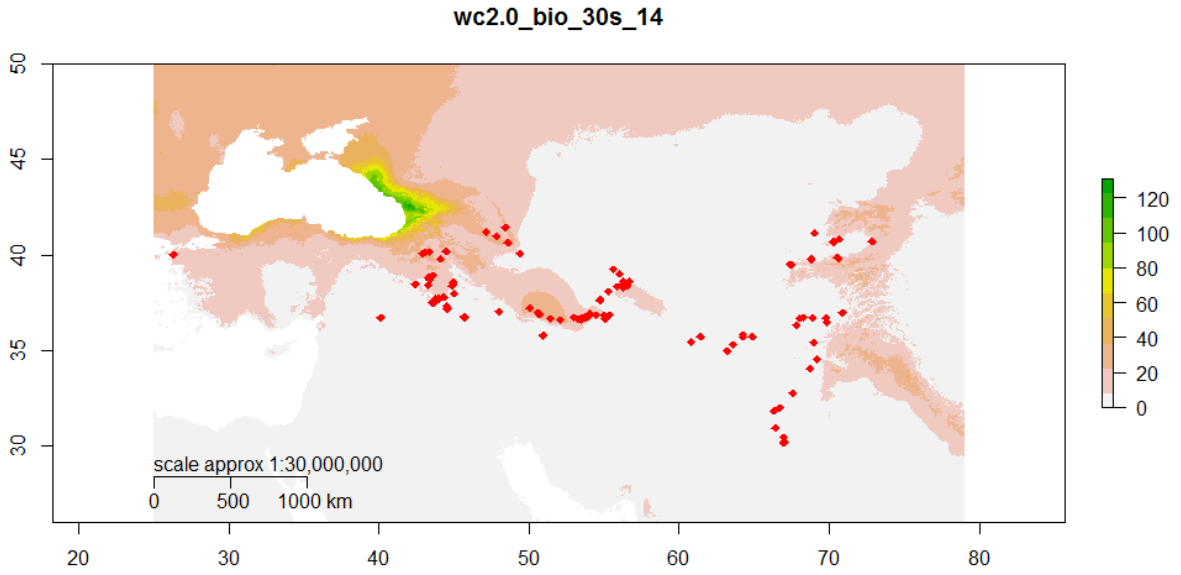


Figure 3.13. Geographic distribution of *Aegilops tauschii* accessions during the driest month.

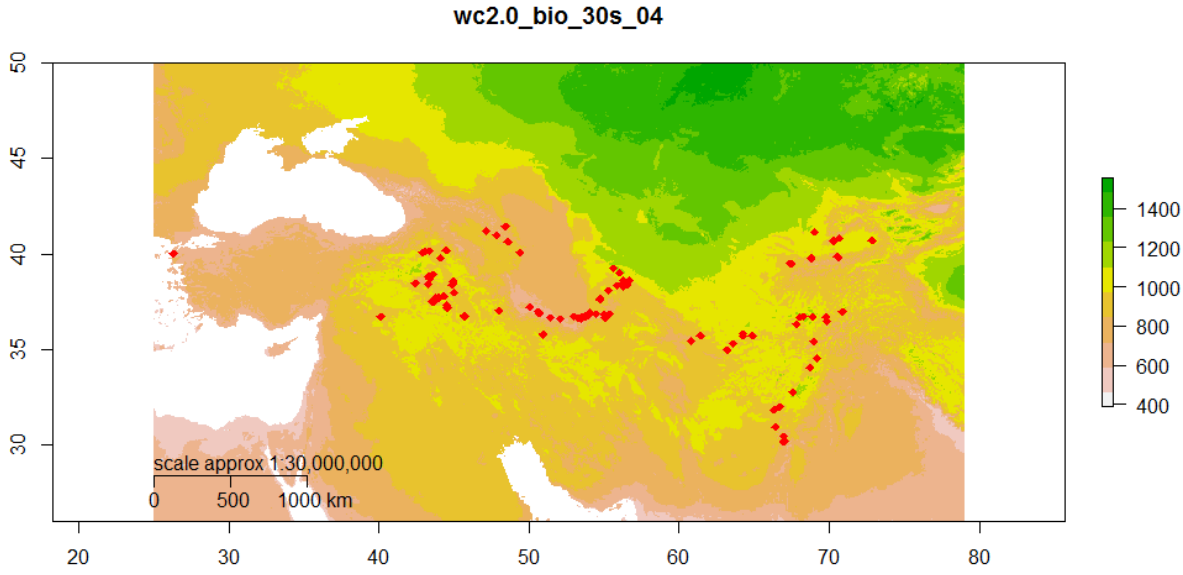


Figure 3.14. Geographic distribution of *Aegilops tauschii* accessions across temperatures seasonality.

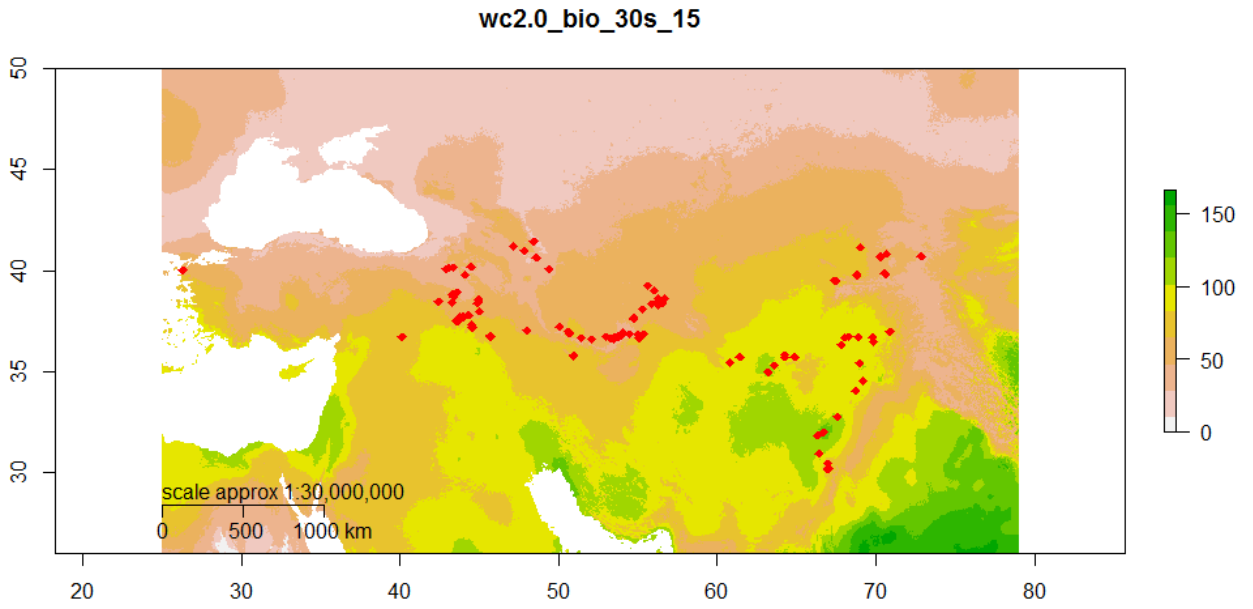


Figure 3.15. Geographic distribution of *Aegilops tauschii* accessions across precipitations seasonality.

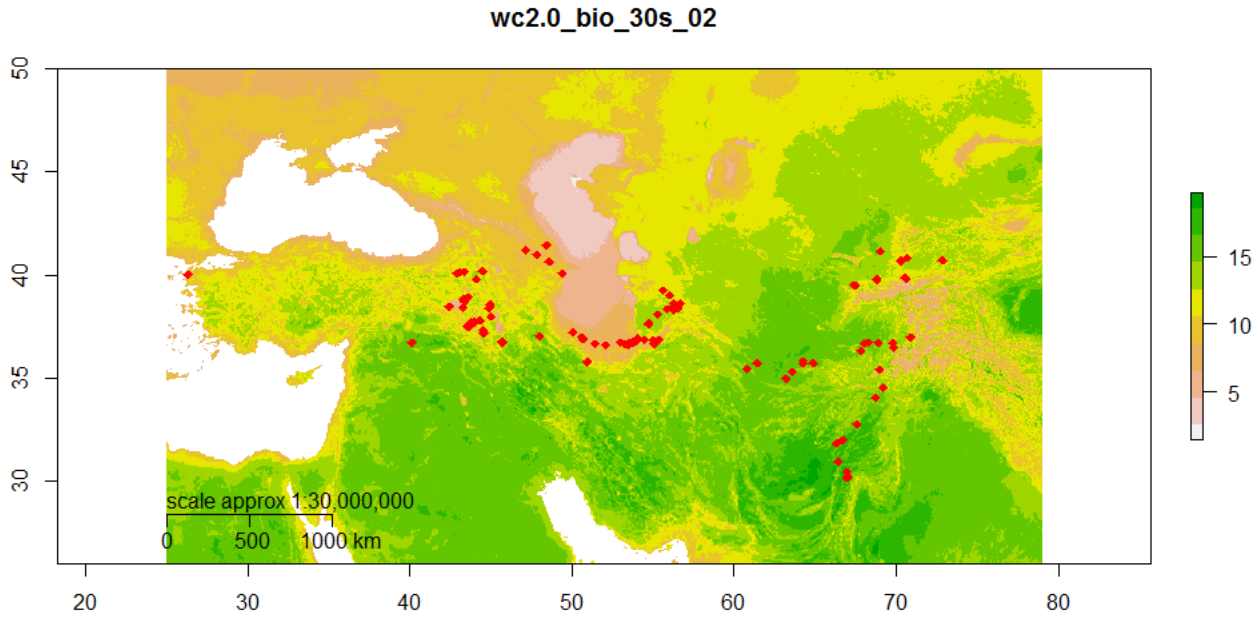


Figure 3.16. Geographic distribution of *Aegilops tauschii* accessions plotted along the mean diurnal ranges.

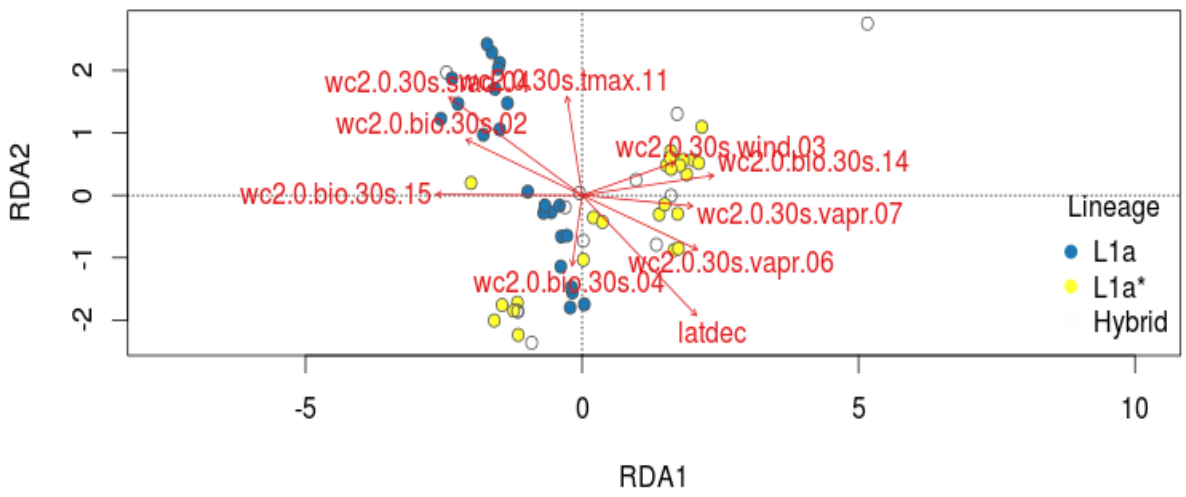


Figure 3.17. Bi-plot showing major climatic drivers of lineage L1a and L1a* of *Aegilops tauschii* using two redundancy axes.

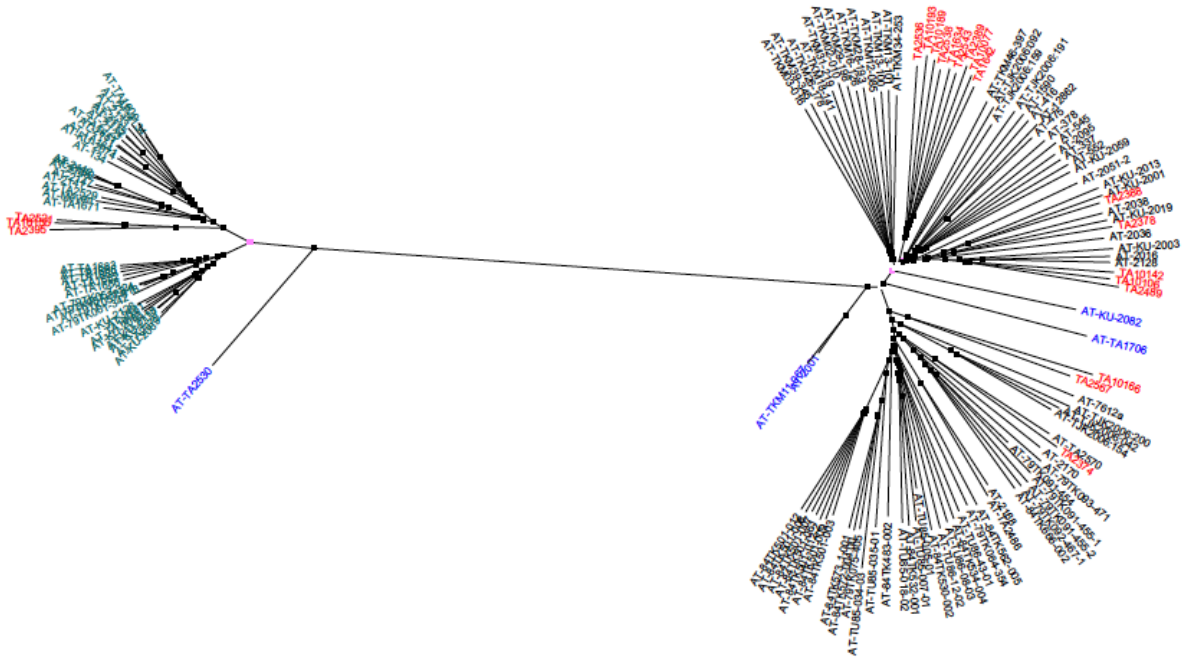


Figure 3.18. Neighbor tree showing relationship between lineages.

Note: The green colored accessions in the left are accessions form L2 and black colored accessions in right represents accessions from L1. The accessions used as donor parent in this research is colored in red. Accessions colored in blue represents accessions identified in this research as genetically diverse and climate SNPs enriched potential donor parent.

Table 3.1. Origin of *Ae. tauschii* accessions grouping based on the first two principal components

Genotype	Group	Origin	Lineage	PC1	PC2
AT_12862	AT	AFG	L1a	-42.53618	12.1107855
AT_134	AT	IRN	L2	97.910706	1.3043364
AT_1590	AT	AFG	L1a	-42.768185	12.132611
AT_2001	AT	PAK	L1b	-45.83877	0.8693891
AT_2016	AT	PAK	L1a	-42.673134	11.933997
AT_2036	AT	AFG	L1a	-43.32967	9.635665
AT_2038	AT	AFG	L1a	-42.802868	12.21876
AT_2051_2	AT	AFG	L1a	-42.807457	14.226943
AT_2095	AT	AFG	L1a	-42.663006	13.713143
AT_2111	AT	IRN	L2	98.47079	1.0718343
AT_2112	AT	IRN	L2	98.67211	1.4519066
AT_2115	AT	IRN	L2	98.88961	1.4878241
AT_2123	AT	IRN	L2	99.01426	1.2134631
AT_2128	AT	IRN	L1a	-43.01928	11.871729
AT_2131	AT	IRN	L2	96.76629	-1.3841827
AT_2133	AT	IRN	L2	100.29706	0.7656695
AT_2134	AT	IRN	L2	101.48576	0.86336344
AT_2137	AT	IRN	L2	100.35448	1.752309
AT_2139	AT	IRN	L2	98.15257	2.0188792
AT_2140	AT	IRN	L2	97.121216	1.9999514
AT_2141	AT	IRN	L2	96.64959	1.8683282
AT_2142	AT	IRN	L2	97.65154	1.9984276
AT_2147	AT	IRN	L2	97.17749	1.7682185
AT_2168	AT	IRN	L1b	-40.857533	-12.100588
AT_2170	AT	IRN	L1b	-39.868378	-12.60135
AT_337	AT	AFG	L1a	-42.626324	15.16623
AT_378	AT	AFG	L1a	-42.797836	14.159961
AT_416	AT	AFG	L1a	-42.739544	13.538094
AT_475	AT	AFG	L1a	-43.11474	11.704459
AT_545	AT	AFG	L1a	-39.751015	13.573012
AT_552	AT	AFG	L1a	-42.471096	15.588967
AT_7612a	AT	TUR	L1b	-39.421505	-9.993473
AT_79TK057_318	AT	TUR	L2	94.85375	-0.87666214
AT_79TK057_322_1	AT	TUR	L2	96.73661	-0.9763679
AT_79TK057_324	AT	TUR	L2	96.98966	-1.0321577
AT_79TK061_342	AT	TUR	L2	97.183014	-0.6201803
AT_79TK064_354	AT	TUR	L1b	-40.5097	-16.664968
AT_79TK075_405	AT	TUR	L1b	-40.344345	-18.481014
AT_79TK091_454	AT	TUR	L1b	-42.514065	-14.075885

AT_79TK091_455_1	AT	TUR	L1b	-40.437607	-17.010916
AT_79TK091_455_2	AT	TUR	L1b	-40.208744	-18.492582
AT_79TK092_467_1	AT	TUR	L1b	-40.057034	-18.774784
AT_79TK093_471	AT	TUR	L1b	-40.14779	-15.604973
AT_84TK483_002	AT	TUR	L1b	-40.958168	-15.912168
AT_84TK501_003	AT	TUR	L1b	-40.93496	-21.537582
AT_84TK501_007	AT	TUR	L1b	-40.94633	-21.4838
AT_84TK501_008	AT	TUR	L1b	-40.738518	-21.572153
AT_84TK501_009	AT	TUR	L1b	-40.8519	-21.566412
AT_84TK501_012	AT	TUR	L1b	-40.666058	-21.552475
AT_84TK501_012_2	AT	TUR	L1b	-40.769234	-21.49713
AT_84TK501_061	AT	TUR	L1b	-40.804367	-21.659494
AT_84TK530_002	AT	TUR	L1b	-40.218338	-17.308578
AT_84TK532_001	AT	TUR	L1b	-40.79189	-18.144192
AT_84TK534_004	AT	TUR	L1b	-40.320576	-16.031687
AT_84TK562_005	AT	TUR	L1b	-40.646442	-15.552789
AT_84TK572_004_00	AT	TUR	L1b	-40.461025	-18.327179
AT_84TK573_1_001	AT	TUR	L1b	-40.4498	-18.572636
AT_84TK666_002	AT	TUR	L1b	-39.855743	-17.682816
AT_KU_2001	AT	PAK	L1a	-42.95824	13.697434
AT_KU_2003	AT	PAK	L1a	-42.370518	12.575144
AT_KU_2013	AT	AFG	L1a	-42.978302	11.919981
AT_KU_2019	AT	AFG	L1a	-43.100178	12.103409
AT_KU_2059	AT	AFG	L1a	-42.837162	15.595877
AT_KU_2069	AT	IRN	L2	96.60013	-2.0412526
AT_KU_2074	AT	IRN	L2	97.79028	1.3359873
AT_KU_2082	AT	IRN	L1a	-41.827877	9.68882
AT_KU_2083	AT	IRN	L2	96.67622	-0.30986413
AT_KU_2117	AT	IRN	L2	95.99545	-2.2514918
AT_KU_2118	AT	IRN	L2	96.13354	-2.2635052
AT_KU_2126	AT	IRN	L2	97.66818	-0.8237947
AT_Meyeri	AT	IRN	L2	96.233864	0.0823862
AT_TA1600	AT	IRN	L2	100.70762	1.1249717
AT_TA1641	AT	IRN	L2	99.07059	0.47873786
AT_TA1651	AT	IRN	L2	99.62639	0.934362
AT_TA1665	AT	AZE	L2	97.69439	-0.91894877
AT_TA1669	AT	AZE	L2	96.80059	-0.736986
AT_TA1671	AT	AZE	L2	98.55077	0.25446513
AT_TA1682	AT	AZE	L2	96.98541	-0.64505756
AT_TA1683	AT	AZE	L2	98.45929	-0.46164653
AT_TA1686	AT	AZE	L2	97.71482	-1.2220459
AT_TA1693	AT	TKM	L2	96.83102	-0.056407586
AT_TA1706	AT	IRN	L1a	-37.165916	0.9856829

AT_TA2375	AT	IRN	L2	92.7938	-2.404324
AT_TA2486	AT	IRN	L1b	-38.18681	-11.829304
AT_TA2529	AT	IRN	L2	94.01508	0.93016577
AT_TA2530	AT	IRN	L2	81.98186	1.3703183
AT_TA2570	AT	ARM	L1b	-39.68467	-15.328498
AT_TJK2006_042	AT	TJK	L1b	-38.098618	-9.391949
AT_TJK2006_092	AT	TJK	L1a	-43.349205	13.399415
AT_TJK2006_154	AT	TJK	L1b	-38.307354	-8.56657
AT_TJK2006_159	AT	TJK	L1a	-43.820343	13.4260435
AT_TJK2006_191	AT	TJK	L1a	-42.935383	12.843499
AT_TJK2006_200	AT	TJK	L1b	-36.71293	-9.467272
AT_TKM02_010	AT	TKM	L1a	-42.009434	7.061503
AT_TKM03_016	AT	TKM	L1a	-42.462986	10.001825
AT_TKM11_067	AT	TKM	L1a	-44.607105	6.041986
AT_TKM12_085	AT	TKM	L1a	-43.401352	11.540749
AT_TKM13_100	AT	TKM	L1a	-42.746296	12.276317
AT_TKM13_101	AT	TKM	L1a	-42.012867	12.413953
AT_TKM16_126	AT	TKM	L1a	-41.955845	11.075753
AT_TKM18_141	AT	TKM	L1a	-42.988953	10.746205
AT_TKM26_178	AT	TKM	L1a	-42.96059	11.157654
AT_TKM28_193	AT	TKM	L1a	-42.548767	11.065248
AT_TKM29_198	AT	TKM	L1a	-42.32115	9.632257
AT_TKM31_219	AT	TKM	L1a	-42.338882	11.263192
AT_TKM34_253	AT	TKM	L1a	-42.83401	10.103831
AT_TKM39_345	AT	TKM	L1a	-42.486164	10.323667
AT_TKM46_397	AT	TKM	L1a	-42.634064	12.89472
AT_TU85_005_01	AT	TUR	L1b	-42.81156	-15.477883
AT_TU85_007_01	AT	TUR	L1b	-40.696087	-18.294992
AT_TU85_018_02	AT	TUR	L1b	-40.528954	-16.770163
AT_TU85_034_03	AT	TUR	L1b	-30.844921	-18.305815
AT_TU85_035_01	AT	TUR	L1b	-40.89917	-17.628883
AT_TU86_08_03	AT	TUR	L1b	-39.34575	-19.260168
AT_TU86_12_02	AT	TUR	L1b	-38.029568	-16.650726
AT_TU86_43_01	AT	TUR	L1b	-40.82254	-15.424934
TA10077	ILP_AT	PAK (WGRC)	L1a	-42.87294	12.20936
TA10106	ILP_AT	KYR (WGRC)	L1a	-42.202602	13.528833
TA10142	ILP_AT	SYR (WGRC)	L2	96.55529	1.778875
TA10155	ILP_AT	TAJ (WGRC)	L1a	-42.66976	16.201248
TA10166	ILP_AT	TKM (WGRC)	L1a	-41.54373	12.416337

TA10189	ILP_AT	UZB (WGRC)	L1a	-42.382805	14.782781
TA10193	ILP_AT	UZB (WGRC)	L1a	-42.26467	14.797373
TA1634	ILP_AT	TUR (WGRC)	L1a	-42.59654	14.376946
TA1642	ILP_AT	IRN (WGRC)	L2	99.58712	1.6851426
TA2374	ILP_AT	PAK (WGRC)	L1a	-42.903828	12.973094
TA2378	ILP_AT	IRN (WGRC)	L2	94.96481	1.6011826
TA2388	ILP_AT	AFG (WGRC)	L1a	-43.36525	14.074684
TA2389	ILP_AT	AFG (WGRC)	L1a	-42.695797	13.422739
TA2395	ILP_AT	AFG (WGRC)	L1a	-42.47491	12.919182
TA2398	ILP_AT	NA (WGRC)	L1a	-42.910828	12.224199
TA2489	ILP_AT	IRN (WGRC)	L1b	-40.532887	-14.871159
TA2521	ILP_AT	IRN (WGRC)	L1b	-40.8727	-6.058807
TA2536	ILP_AT	AFG (WGRC)	L1a	-42.581318	12.666081
TA2538	ILP_AT	AFG (WGRC)	L1a	-42.44293	12.749928
TA2543	ILP_AT	AFG (WGRC)	L1a	-42.45385	13.613452
TA2567	ILP_AT	ARM (WGRC)	L1b	-39.706005	-15.821636

Abbreviation	*Country*
AFG	Afghanistan
ARM	Armenia
AZE	Azerbaijan
IRN	Iran
KYR	Kyrgyzstan
PAK	Pakistan
SYR	Syria
TJK	Tajikistan
TKM	Turkmenistan
TUR	Turkey
UZB	Uzbekistan
WGRC	Wheat Genetics and Genomics Resource Center

Table 3.2. Twelve non-collinear (r=0.85) bioclimatic variables used in RDA analysis

Genotype	Bio.02	Bio.04	Bio.14	Bio.15	Srad. June	Srad. July	Srad. Aug	WVP. Sept	Prec. Sept	Tmax. April	WVP Oct	Prec. May
AT.12862	15.083	1118.731	0	94.760	25212	26297	24474	0.980	0	28.000	0.840	13
AT.134	10.217	818.450	17	39.442	20834	20272	19372	2.060	41	21.100	1.510	26
AT.1590	15.508	1022.293	0	92.547	24880	25816	24227	0.740	1	24.300	0.670	23
AT.2001	15.400	762.673	3	69.841	25444	23491	22857	0.900	7	28.500	0.600	7
AT.2016	15.700	894.070	1	97.208	26189	24752	23997	0.800	2	28.100	0.580	3
AT.2036	17.742	980.333	0	124.053	26565	25751	24641	0.730	0	27.100	0.490	2
AT.2038	16.142	1076.650	0	106.405	26269	25270	24199	0.550	0	18.300	0.400	11
AT.2051.2	11.342	912.758	1	102.334	25209	24524	23444	0.510	2	9.600	0.400	93
AT.2095	14.967	940.635	0	102.384	26766	27053	25268	0.960	1	23.700	0.730	16
AT.2111	10.183	802.948	18	46.804	21044	20487	18806	2.040	53	21.500	1.530	23
AT.2112	10.183	802.948	18	46.804	21044	20487	18806	2.040	53	21.500	1.530	23
AT.2115	9.575	791.800	14	40.360	21726	22275	20931	2.030	34	20.500	1.480	26
AT.2123	12.900	908.429	2	70.051	25012	25512	23883	0.870	3	19.300	0.690	27
AT.2128	12.900	908.429	2	70.051	25012	25512	23883	0.870	3	19.300	0.690	27
AT.2131	14.600	811.706	1	87.057	27101	27098	25106	0.650	2	21.800	0.630	33
AT.2133	10.183	802.948	18	46.804	21044	20487	18806	2.040	53	21.500	1.530	23
AT.2134	10.183	802.948	18	46.804	21044	20487	18806	2.040	53	21.500	1.530	23
AT.2137	8.533	741.160	17	55.686	21856	21846	19197	2.180	70	20.200	1.640	22
AT.2139	8.142	698.746	24	61.977	21939	21742	18953	2.190	109	18.500	1.700	30
AT.2140	8.142	698.746	24	61.977	21939	21742	18953	2.190	109	18.500	1.700	30
AT.2141	8.142	698.746	24	61.977	21939	21742	18953	2.190	109	18.500	1.700	30
AT.2142	8.250	692.784	36	66.617	20777	20948	17877	2.180	148	17.700	1.700	47
AT.2147	8.717	713.848	33	58.236	21442	22141	18665	2.140	133	17.800	1.660	47
AT.2168	12.725	935.274	2	76.656	26222	26744	23120	0.810	2	19.700	0.740	53
AT.2170	12.550	938.719	3	65.115	25810	26636	24437	0.940	5	18.100	0.780	52
AT.337	16.325	978.917	0	104.764	26771	26888	25247	0.910	1	23.700	0.710	13
AT.378	16.017	972.551	0	104.415	26750	26950	25226	0.890	1	23.200	0.700	13
AT.416	14.575	1036.643	0	93.634	25281	25543	24005	1.030	0	26.000	0.880	27
AT.475	15.567	1120.324	0	93.378	25336	26170	24402	1.000	0	28.300	0.850	16
AT.545	15.400	946.372	0	90.217	26593	26821	25017	0.720	0	21.200	0.600	27
AT.552	15.542	950.707	0	93.871	26776	26949	25133	0.810	1	22.200	0.640	24
AT.7612a	12.008	1004.631	13	48.528	24955	25446	23499	1.080	13	16.700	0.820	57
AT.79TK057.318	10.975	951.795	3	73.666	27292	27787	24761	0.650	3	13.000	0.570	63
AT.79TK057.322.1	10.975	951.795	3	73.666	27292	27787	24761	0.650	3	13.000	0.570	63
AT.79TK057.324	10.975	951.795	3	73.666	27292	27787	24761	0.650	3	13.000	0.570	63
AT.79TK061.342	10.508	952.101	4	72.242	27395	28010	25161	0.720	4	13.600	0.630	60
AT.79TK064.354	8.783	905.862	7	60.586	26706	26978	24338	0.540	7	7.700	0.430	64
AT.79TK075.405	7.542	906.390	8	48.710	26766	27551	25044	0.910	13	11.800	0.730	54
AT.79TK091.454	11.717	927.818	16	43.526	25193	25686	23746	1.070	16	16.200	0.810	61
AT.79TK091.455.1	11.717	927.818	16	43.526	25193	25686	23746	1.070	16	16.200	0.810	61

AT.79TK091.455.2	11.717	927.818	16	43.526	25193	25686	23746	1.070	16	16.200	0.810	61
AT.79TK092.467.1	11.167	916.701	16	42.756	25115	25501	23747	1.050	16	15.700	0.800	64
AT.79TK093.471	11.067	892.020	18	40.724	25526	26067	24076	0.950	18	13.500	0.730	70
AT.84TK483.002	8.200	892.287	6	60.985	27601	28331	25610	0.830	13	12.600	0.730	79
AT.84TK501.003	7.992	926.225	4	55.571	27237	27918	25223	0.880	12	12.400	0.740	54
AT.84TK501.007	7.992	926.225	4	55.571	27237	27918	25223	0.880	12	12.400	0.740	54
AT.84TK501.008	7.992	926.225	4	55.571	27237	27918	25223	0.880	12	12.400	0.740	54
AT.84TK501.009	7.992	926.225	4	55.571	27237	27918	25223	0.880	12	12.400	0.740	54
AT.84TK501.012	7.992	926.225	4	55.571	27237	27918	25223	0.880	12	12.400	0.740	54
AT.84TK501.012.2	7.992	926.225	4	55.571	27237	27918	25223	0.880	12	12.400	0.740	54
AT.84TK501.061	7.992	926.225	4	55.571	27237	27918	25223	0.880	12	12.400	0.740	54
AT.84TK530.002	10.908	1032.168	2	74.679	27710	28378	25652	0.820	6	15.200	0.730	61
AT.84TK532.001	10.208	967.110	4	71.095	27847	28406	25547	0.730	7	13.600	0.650	59
AT.84TK534.004	13.050	1022.724	4	69.119	27214	27698	25306	0.850	7	16.300	0.740	54
AT.84TK562.005	11.467	961.711	3	74.099	27192	27655	24747	0.680	4	13.400	0.600	63
AT.84TK572.004.00	11.758	936.889	7	49.963	27073	27838	25266	0.890	15	14.100	0.720	56
AT.84TK573.1.001	12.233	964.519	10	46.815	26775	27348	24786	0.840	16	13.800	0.670	58
AT.84TK666.002	11.167	916.701	16	42.756	25115	25501	23747	1.050	16	15.700	0.800	64
AT.KU.2001	15.608	751.703	3	69.244	25503	23786	22992	0.910	7	28.700	0.610	7
AT.KU.2003	15.608	751.703	3	69.244	25503	23786	22992	0.910	7	28.700	0.610	7
AT.KU.2013	17.742	980.333	0	124.053	26565	25751	24641	0.730	0	27.100	0.490	2
AT.KU.2019	13.933	962.685	3	93.661	26355	25601	24092	0.760	3	20.300	0.580	24
AT.KU.2059	15.400	974.697	0	94.669	26773	27000	25150	0.850	1	22.900	0.670	24
AT.KU.2069	11.542	920.818	3	66.696	25442	25061	22843	0.940	3	20.500	0.800	22
AT.KU.2074	9.075	777.614	15	40.312	21551	22086	20047	1.800	39	19.800	1.340	27
AT.KU.2082	8.525	724.081	5	58.028	22778	23407	22159	1.290	13	18.300	0.980	33
AT.KU.2083	10.667	864.973	3	70.240	24757	25369	23678	0.800	4	16.400	0.640	30
AT.KU.2117	12.558	986.120	4	60.657	24690	25665	24347	1.000	7	18.300	0.810	50
AT.KU.2118	12.333	941.748	5	59.554	22909	23333	23350	1.040	9	18.800	0.830	56
AT.KU.2126	7.508	682.424	30	61.911	20780	20807	18289	2.100	143	17.700	1.650	46
AT.Meyeri	10.500	824.360	17	43.394	20832	20567	18951	2.090	46	21.600	1.540	24
AT.TA1600	9.783	806.517	16	42.937	21386	21611	19718	1.920	41	20.800	1.430	25
AT.TA1641	9.533	703.930	14	41.582	21784	22321	20284	1.740	37	19.700	1.310	26
AT.TA1651	10.992	815.131	16	35.194	20907	20816	20333	1.830	36	21.900	1.330	37
AT.TA1665	8.892	849.891	5	41.076	22207	22804	19646	1.870	16	18.300	1.400	23
AT.TA1669	8.892	849.891	5	41.076	22207	22804	19646	1.870	16	18.300	1.400	23
AT.TA1671	8.883	850.574	26	26.571	22045	22768	19652	1.360	54	15.100	1.030	55
AT.TA1682	9.358	828.664	29	34.689	21730	22296	19611	1.400	52	16.500	1.050	86
AT.TA1683	9.358	845.314	36	31.677	22091	23036	19691	1.340	50	15.600	1.020	91
AT.TA1686	8.808	867.207	22	30.333	22193	22895	19399	1.390	41	16.200	1.070	57
AT.TA1693	13.175	891.077	9	43.322	24020	23487	22179	1.020	9	23.700	0.830	32
AT.TA1706	10.992	815.131	16	35.194	20907	20816	20333	1.830	36	21.900	1.330	37
AT.TA2375	11.775	929.656	2	66.706	25269	24654	22779	0.960	3	20.800	0.830	22

AT.TA2486	13.158	966.464	1	77.346	27069	27525	23388	0.750	2	19.100	0.690	50
AT.TA2529	8.658	688.200	31	59.630	20813	21135	18058	1.880	119	17.200	1.480	49
AT.TA2530	8.658	688.200	31	59.630	20813	21135	18058	1.880	119	17.200	1.480	49
AT.TA2570	12.417	949.643	11	43.160	24168	24651	22762	1.190	11	18.400	0.900	53
AT.TJK2006:042	11.900	973.985	5	52.101	26381	26210	23722	0.880	8	20.400	0.730	40
AT.TJK2006:092	13.308	1062.857	4	49.507	26594	26321	23792	1.000	6	23.400	0.810	27
AT.TJK2006:154	12.300	911.851	6	57.746	25857	25350	23329	0.850	6	18.700	0.700	44
AT.TJK2006:159	12.800	868.994	9	57.115	25594	25052	23151	0.780	9	17.400	0.640	63
AT.TJK2006:191	12.308	877.311	5	66.826	25906	25843	23725	0.750	8	17.200	0.600	64
AT.TJK2006:200	12.483	892.705	5	68.655	25991	25897	23842	0.780	8	17.900	0.620	64
AT.TKM02.010	13.117	913.580	9	43.509	23822	23424	22122	1.060	9	23.900	0.860	31
AT.TKM03.016	13.425	914.303	8	44.214	24118	23579	22335	0.990	8	23.500	0.810	32
AT.TKM11.067	12.958	909.391	8	45.411	24128	23730	22376	1.020	8	23.500	0.830	32
AT.TKM12.085	12.617	908.135	8	47.758	24525	24286	22717	0.900	8	21.500	0.740	36
AT.TKM13.100	12.842	954.733	7	46.118	24970	25289	23226	1.000	7	23.500	0.810	28
AT.TKM13.101	12.842	954.733	7	46.118	24970	25289	23226	1.000	7	23.500	0.810	28
AT.TKM16.126	11.267	1020.824	4	50.858	24754	25036	22863	1.000	4	23.100	0.800	20
AT.TKM18.141	11.983	1091.394	3	54.513	24466	24629	22748	1.070	3	24.600	0.850	14
AT.TKM26.178	12.325	870.186	4	55.425	22588	22818	21354	1.740	10	22.700	1.240	20
AT.TKM28.193	13.450	872.857	4	57.902	22708	22943	21343	1.710	8	24.400	1.220	20
AT.TKM29.198	13.075	931.049	4	60.925	23558	23568	21997	1.420	7	23.700	1.060	23
AT.TKM31.219	12.717	931.838	8	49.100	23149	22573	21337	1.210	9	23.500	0.940	28
AT.TKM34.253	12.358	902.650	8	49.369	23979	23586	22158	0.930	8	20.700	0.750	37
AT.TKM39.345	10.883	916.696	7	47.441	25321	25913	23676	0.820	7	19.600	0.670	33
AT.TKM46.397	15.142	890.667	0	91.286	27269	27362	25300	0.710	1	24.000	0.690	25
AT.TU85.005.01	12.392	945.704	6	52.256	26945	27677	25158	0.890	14	14.600	0.720	56
AT.TU85.007.01	11.083	1006.549	4	69.804	27699	28255	25617	0.780	9	14.600	0.680	59
AT.TU85.018.02	12.425	988.254	6	64.610	27226	27988	25270	0.760	10	13.900	0.650	59
AT.TU85.034.03	11.833	939.052	7	50.015	27036	27853	25315	0.920	14	14.500	0.740	54
AT.TU85.035.01	9.642	906.891	6	50.503	27199	28025	25370	0.910	15	12.900	0.740	55
AT.TU86.08.03	11.858	1025.188	1	80.079	27492	28201	25476	0.900	3	16.800	0.820	59
AT.TU86.12.02	12.583	999.549	6	65.664	27473	28058	25418	0.790	8	15.100	0.680	54
AT.TU86.43.01	13.142	1015.501	4	69.564	27263	27814	25307	0.870	7	16.600	0.750	56
TA.10077	17.750	821.863	2	73.402	25897	24622	23659	0.930	7	30.100	0.630	6
TA.10106	12.525	1013.727	8	44.315	25577	25088	22971	1.080	10	22.400	0.850	39
TA.10155	13.250	928.225	2	73.342	26239	26548	24244	0.860	5	20.900	0.690	49
TA.10166	12.650	929.210	7	51.309	23504	22976	21803	1.210	8	23.300	0.940	27
TA.10189	13.617	1015.139	2	66.388	27335	27323	24672	1.020	5	23.500	0.820	35
TA.10193	12.925	909.799	1	74.480	26385	26795	24375	0.900	4	20.600	0.730	45
TA.1634	8.358	690.604	9	63.965	26462	27949	25212	1.550	25	17.500	1.280	34
TA.1642	10.150	846.804	12	40.432	22089	22509	20721	2.030	29	21.000	1.460	24
TA.2374	16.333	803.367	3	71.264	25639	24096	23221	0.900	6	26.300	0.610	6
TA.2378	9.008	695.065	35	64.247	20639	21056	17895	2.080	139	17.500	1.620	48

TA.2388	17.758	893.377	0	118.591	26604	25836	24797	0.740	0	28.100	0.520	2
TA.2389	17.767	973.852	0	124.035	26527	25689	24625	0.710	0	26.700	0.470	2
TA.2395	15.658	1102.158	1	88.673	26417	25782	24504	0.640	1	18.900	0.510	19
TA.2489	12.283	984.311	4	61.067	25034	26108	24494	0.960	7	17.800	0.780	54
TA.2521	15.808	977.163	3	63.301	25569	26128	22612	1.020	6	22.100	0.840	34
TA.2536	14.817	968.364	1	93.978	25096	24548	23437	0.730	1	21.500	0.630	72
TA.2538	14.008	980.889	1	92.215	25281	24812	23600	0.760	1	21.600	0.660	72
TA.2543	15.058	947.528	1	84.367	24954	24065	22906	0.830	2	22.200	0.680	58
TA.2567	12.250	954.932	12	44.000	23818	24262	22626	1.190	12	18.200	0.890	55

Table 3.3. Pairwise F_{ST} coefficients among L1b, L2, L1a, and L1a*.

	L1b	**L1a**	**La***	**L2**
L1b	##############			
L1a	0.2247	##############		
La*	0.1814	0.0680	##############	
L2	0.7048	0.7169	0.7133	##############

Table 3.4. Tajima's D among L1b, L2, L1a, and L1a*.

	Tajima's D
L1b	1.23242
L1a	1.22785
L1a*	-0.22761
Hybrid	0.54693
L2	-0.66605

Table 3.5. Climatic variations across *Ae. tauschii* lineages.

Climatic variables	H	L1a	L1a*	L1b	L2
wc2.0.bio.30s.01	14.65417	14.91051	14.62417	9.414236	14.54146
wc2.0.bio.30s.02	13.63611	15.44638	12.67833	10.79745	10.12875
wc2.0.bio.30s.03	35.66092	36.83493	32.93951	29.1391	31.40296
wc2.0.bio.30s.04	886.429	949.0626	940.7288	948.8462	820.1214
wc2.0.bio.30s.05	33.00833	35.1	33.128	26.80833	30.4675
wc2.0.bio.30s.06	-5.09167	-6.92609	-5.384	-9.99167	-1.695
wc2.0.bio.30s.07	38.1	42.02609	38.512	36.8	32.1625
wc2.0.bio.30s.08	9.9625	6.782609	9.457333	7.349074	11.38
wc2.0.bio.30s.09	23.25	24.57029	25.31533	20.68194	21.92792
wc2.0.bio.30s.10	25.46944	26.52609	26.14067	20.86111	24.55167
wc2.0.bio.30s.11	3.540278	2.956522	2.918667	-2.3963	4.55125
wc2.0.bio.30s.12	368.9167	306.3043	294.96	497.0833	591.6
wc2.0.bio.30s.13	68.16667	74.47826	47.28	77.94444	99.4
wc2.0.bio.30s.14	4.25	0.913043	5.68	7.194444	15.875
wc2.0.bio.30s.15	68.63586	95.7221	55.54513	57.19896	54.51032

wc2.0.bio.30s.16	179.5833	182.7826	126.2	209.25	253.8
wc2.0.bio.30s.17	19.25	5.869565	23.52	29.02778	56.675
wc2.0.bio.30s.18	28.58333	11.17391	27.4	41.97222	72.25
wc2.0.bio.30s.19	122.1667	139	98.92	149.9722	162.875
wc2.0.30s.srad.01	8509.25	9412.043	7601.2	7978.5	7992.9
wc2.0.30s.srad.02	11089.5	11819.78	10297.36	10770.75	10362.33
wc2.0.30s.srad.03	13929.33	15294.78	13228.84	14235.25	12759.1
wc2.0.30s.srad.04	17701	18608.48	17158.96	18219.47	16270.48
wc2.0.30s.srad.05	22189.33	23632.26	21943.72	22365.11	20273.53
wc2.0.30s.srad.06	24960.42	26042.78	24884.64	26590.36	22831.4
wc2.0.30s.srad.07	24768.83	25725.39	24782.96	27190.86	23036.23
wc2.0.30s.srad.08	23226.33	24321.65	22956.6	24678.56	20729.15
wc2.0.30s.srad.09	19714.58	21175.09	18718.36	20115.03	17166.83
wc2.0.30s.srad.10	14340.25	15998.83	13066.52	13770.08	12414.05
wc2.0.30s.srad.11	10292.25	11820.65	9021.64	9490.139	9278.25
wc2.0.30s.srad.12	8110.25	9213.913	7078.6	7288.361	7483.525
wc2.0.30s.tmax.01	9.508333	9.326087	8.632	2.658333	8.985
wc2.0.30s.tmax.02	11.075	11.67391	10.084	3.716667	9.61
wc2.0.30s.tmax.03	15.81667	16.86522	14.896	8.366667	12.835
wc2.0.30s.tmax.04	22.35833	23.86957	22	14.66667	18.69
wc2.0.30s.tmax.05	26.75	28.81304	26.516	19.63611	23.23
wc2.0.30s.tmax.06	31.73333	33.87826	31.712	24.51667	28.0625
wc2.0.30s.tmax.07	32.94167	35.1	33.052	26.76667	30.38
wc2.0.30s.tmax.08	31.71667	33.35652	32.004	26.51111	30.11
wc2.0.30s.tmax.09	27.475	28.48261	27.536	22.21667	26.4925
wc2.0.30s.tmax.10	21.55	22.33043	20.712	15.80833	20.9275
wc2.0.30s.tmax.11	15.75833	16.46522	14.692	9.336111	15.345
wc2.0.30s.tmax.12	10.98333	11.44348	9.724	3.555556	10.6025
wc2.0.30s.vapr.01	0.455	0.441304	0.4804	0.325	0.61475
wc2.0.30s.vapr.02	0.500833	0.50087	0.5052	0.356944	0.64225
wc2.0.30s.vapr.03	0.6375	0.654783	0.6512	0.468056	0.7685
wc2.0.30s.vapr.04	0.865	0.845217	0.9064	0.672778	1.02925
wc2.0.30s.vapr.05	1.034167	0.913043	1.1036	0.884722	1.336
wc2.0.30s.vapr.06	1.144167	0.948261	1.2756	1.035556	1.63325
wc2.0.30s.vapr.07	1.291667	1.10913	1.464	1.218333	1.8925
wc2.0.30s.vapr.08	1.219167	0.996957	1.3384	1.152778	1.8975
wc2.0.30s.vapr.09	0.9825	0.815652	1.0644	0.896944	1.59925
wc2.0.30s.vapr.10	0.773333	0.623478	0.8432	0.737222	1.2355
wc2.0.30s.vapr.11	0.616667	0.543043	0.6752	0.562778	0.91025
wc2.0.30s.vapr.12	0.5075	0.46	0.5512	0.411944	0.7125
wc2.0.30s.wind.01	1.716667	1.96087	2.124	1.441667	1.9875
wc2.0.30s.wind.02	2.075	2.13913	2.416	1.605556	2.3325
wc2.0.30s.wind.03	2.35	2.217391	2.624	1.958333	2.5425
wc2.0.30s.wind.04	2.35	2.230435	2.64	2.202778	2.5175

wc2.0.30s.wind.05	2.358333	2.369565	2.628	2.05	2.45
wc2.0.30s.wind.06	2.475	2.604348	2.74	2.097222	2.5375
wc2.0.30s.wind.07	2.55	2.686957	2.684	2.236111	2.645
wc2.0.30s.wind.08	2.358333	2.443478	2.616	2.166667	2.4075
wc2.0.30s.wind.09	2.083333	2.143478	2.372	1.919444	2.1775
wc2.0.30s.wind.10	1.9	1.956522	2.196	1.65	2.02
wc2.0.30s.wind.11	1.7	1.669565	2.076	1.538889	1.945
wc2.0.30s.wind.12	1.625	1.813043	2.08	1.338889	1.9325
wc2.0.30s.prec.01	41.08333	43.21739	32.64	45.80556	55.3
wc2.0.30s.prec.02	45.58333	59.6087	34.92	54.19444	52.775
wc2.0.30s.prec.03	63.91667	66.65217	44.04	63.30556	58.65
wc2.0.30s.prec.04	59.91667	45.26087	39.92	73.05556	45.4
wc2.0.30s.prec.05	40.41667	19.3913	32.52	57.75	38.775
wc2.0.30s.prec.06	9.75	2.304348	10.6	25.08333	23.35
wc2.0.30s.prec.07	10.83333	5.565217	9.8	9.861111	17.475
wc2.0.30s.prec.08	8	3.304348	7	8.194444	26.9
wc2.0.30s.prec.09	8.416667	1.913043	7.48	10.97222	49.7
wc2.0.30s.prec.10	19.83333	7.434783	18.92	43.69444	84.525
wc2.0.30s.prec.11	25.66667	15.47826	25.76	55.19444	73.325
wc2.0.30s.prec.12	35.5	36.17391	31.36	49.97222	65.425
wc2.0_30s_tmin_01	-5.45833	-7.17391	-6.608	-5.54167	-4.7175
wc2.0_30s_tmin_02	-4.075	-6.33043	-4.792	-4.27222	-3.6075
wc2.0_30s_tmin_03	0.075	-2.06087	-0.128	0.186111	0.615
wc2.0_30s_tmin_04	6.258333	4.356522	6.64	6.613889	6.9375
wc2.0_30s_tmin_05	11.1	9.065217	11.368	11.35556	11.525
wc2.0_30s_tmin_06	16.33333	13.93478	16.492	16.36389	16.37
wc2.0_30s_tmin_07	20.78333	18.53478	20.276	20.68333	19.965
wc2.0_30s_tmin_08	20.05833	18.23043	19.18	19.91111	19.275
wc2.0_30s_tmin_09	15.61667	14.11304	14.66	15.61111	15.325
wc2.0_30s_tmin_10	9.566667	7.947826	8.448	9.272222	9.3725
wc2.0_30s_tmin_11	3.541667	1.93913	2.356	3.216667	3.57
wc2.0_30s_tmin_12	-1.4	-3.32609	-2.704	-1.98611	-1.4775
wc2.0_30s_tavg_01	2.208333	1.191304	1.624	-3.64444	3.6775
wc2.0_30s_tavg_02	3.75	3.53913	3.056	-2.61111	4.2825
wc2.0_30s_tavg_03	8.516667	8.743478	7.884	2.063889	7.5225
wc2.0_30s_tavg_04	15.06667	15.74783	14.988	8.352778	13.375
wc2.0_30s_tavg_05	19.43333	20.68696	19.508	13.31667	17.91
wc2.0_30s_tavg_06	24.41667	25.75217	24.696	18.19444	22.75
wc2.0_30s_tavg_07	26.61667	27.8	27.38	22.27778	25.575
wc2.0_30s_tavg_08	25.39167	26.03478	26.344	22.02222	25.31
wc2.0_30s_tavg_09	21.15833	21.16957	21.868	17.725	21.67
wc2.0_30s_tavg_10	15.225	15.01739	15.052	11.31944	16.11
wc2.0_30s_tavg_11	9.416667	9.152174	9.016	4.830556	10.5225
wc2.0_30s_tavg_12	4.658333	4.143478	4.064	-0.93056	5.775

Abbreviations

BIO1 = Annual Mean Temperature
BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))
BIO3 = Isothermality (BIO2/BIO7) (×100)
BIO4 = Temperature Seasonality (standard deviation ×100)
BIO5 = Max Temperature of Warmest Month
BIO6 = Min Temperature of Coldest Month
BIO7 = Temperature Annual Range (BIO5-BIO6)
BIO8 = Mean Temperature of Wettest Quarter
BIO9 = Mean Temperature of Driest Quarter
BIO10 = Mean Temperature of Warmest Quarter
BIO11 = Mean Temperature of Coldest Quarter
BIO12 = Annual Precipitation
BIO13 = Precipitation of Wettest Month
BIO14 = Precipitation of Driest Month
BIO15 = Precipitation Seasonality (Coefficient of Variation)
BIO16 = Precipitation of Wettest Quarter
BIO17 = Precipitation of Driest Quarter
BIO18 = Precipitation of Warmest Quarter

Chapter 4 - Overall summary

Crop improvements to meet growing food demands depend on the presence of genetic diversity in the existing germplasms which provides a genetic buffer for adapting to future biotic and abiotic stresses. The plant breeders have long utilized crop wild relatives (CWR) to overcome genetic diversity lost during the bottleneck, domestication, polyploidization, and commercial breedings. The introduction of CWR has mostly been an ad hoc process for disease resistance breeding. As global warming is threatening the current crop growing regions around the world, the value of CWR to improve the adaptive potential to extreme climatic conditions has emerged as a central focus on crop improvement.

For self-pollinated crops like wheat, wild-relatives plays a critical role in conserving useful genetic diversity. The wild emmer and *Ae. tauschii*, two major wild relatives of wheat, have shown potential to use in adaptation breeding either through synthetic wheat or through direct crossing with the hexaploid wheat. These two wild relatives which have been adapted to diverse agro-climatic conditions likely to provide extremely useful genetic resources to modern-day wheat varieties. In this dissertation, I explore these genetic potentials of wilds to improve the hexaploid wheat.

My first research paper explored the adaptive potential of wild emmer wheat; the tetraploid wild relative of wheat, to improve climate adaptation of bread wheat. I found that the accessions of wild emmer wheat collected across diverse agro-climatic regions of Israel carried alleles conferring adaption to extreme climatic conditions. The analysis of population structure revealed three genetically distinct groups of wild emmer accessions coinciding with their geographic distribution. Partitioning of genomic variance showed that geographic location and climate together explained 44% of genetic variation among emmer accessions out of which the

variation due to climatic factors accounted for 10% only. The heading date accuracy improved by 9% using the eco-geographic adaptation alleles identified through multiple environmental association scans. Through this targeted breeding pipeline, eco-geographic adaptive alleles from wild emmer are found to improve the prediction accuracy of agronomically important traits and can facilitate selection decisions in breeding programs.

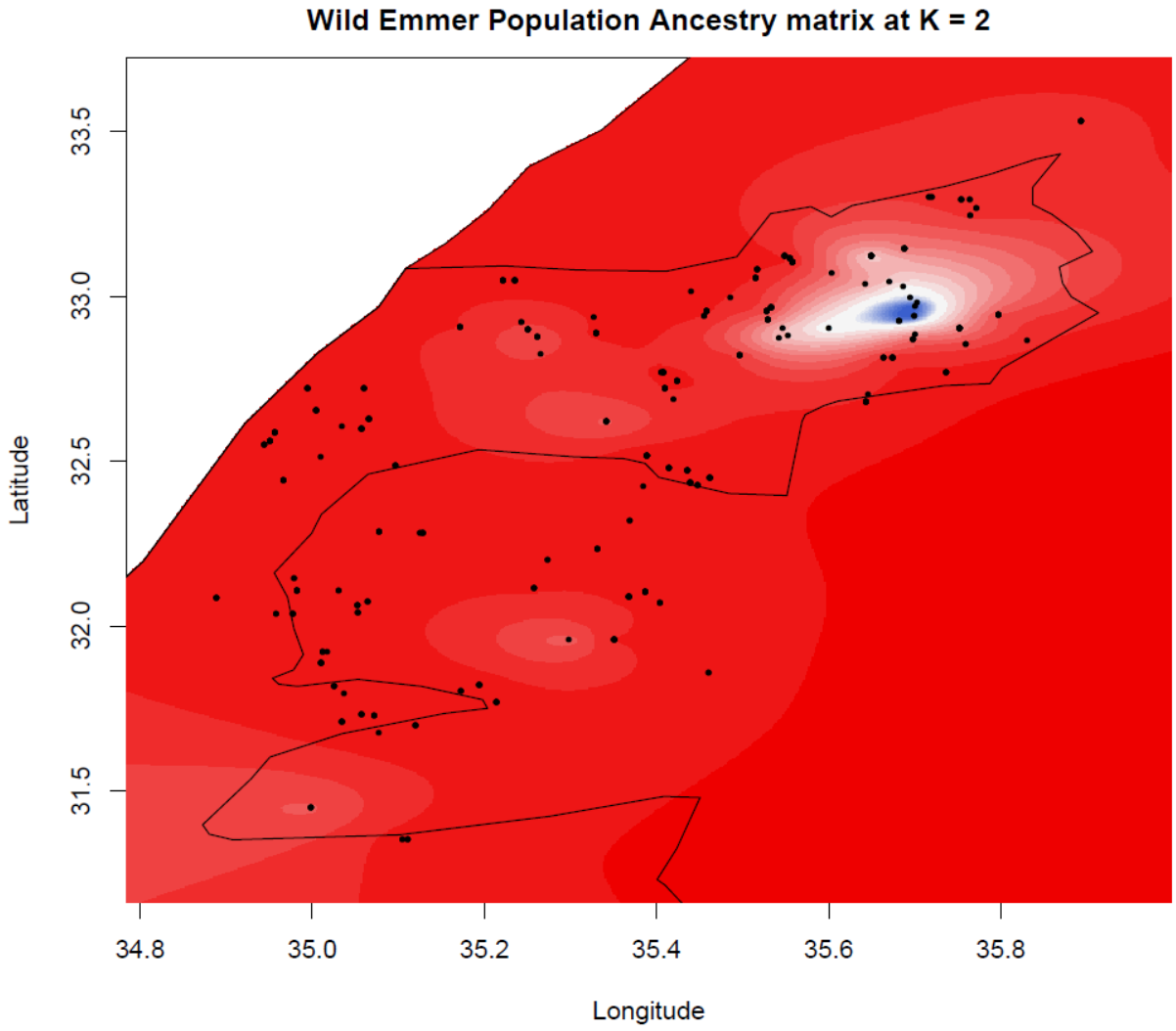
The second chapter of my dissertation identified and prioritized the introgression of climate adaptive variants from *Ae. tauschii* into the hexaploid wheat gene pool. I investigated the genetic basis of local adaptation in *Ae. tauschii*, the diploid progenitor of wheat, and identified the potential donor parents that are enriched with climate-adaptive variants. Twenty-one *Ae. tauschii* accessions were used representing geographically dispersed locations as parents for developing *Ae. tauschii* x hexaploid wheat octaploid hybrids that were crossed with the six Kansas-adapted hexaploid wheat lines. Three hundred and fifty-one BC1F3:5 lines developed after a single backcross to the recurrent parents were used for studying the genomic patterns of alien introgression. I demonstrate that the climate adaptive variants can be identified in the wild relatives using historical on-sight climatic information and that adaptive diversity can be introgressed into the hexaploid gene pool. This study also highlighted some of the issues of inter-specific population development such as hybrid sterility, reduced introgression in the pericentromeric regions, and reduced retention of the introgressed alleles in regions carrying domestication. Overall, this research shows that there is tremendous untapped genetic diversity in *Ae. tauschii* that can be prioritized for breeding climate-resilient wheat.

Overall, my dissertation examined the genetic diversity of wild relatives of wheat to improve the climate adaptive potential of bread wheat for future climate. This work demonstrated how the understanding of the genetic basis of adaptation can facilitate genomic

resources and pipeline development for genomics enabled breeding of wheat. The knowledge gained will improve wheat production under extreme weather thereby contributing to the global fight against food insecurity.

Appendix A - Supplementary material of Chapter 2

This appendix contains supplementary materials for Chapter 2.



Genetic ancestry and population structure of wild emmer accessions from Israel at ancestry coefficients K=2.

Note: Each color represents a genetic co-ancestry matrix where accession is represented by black dots.

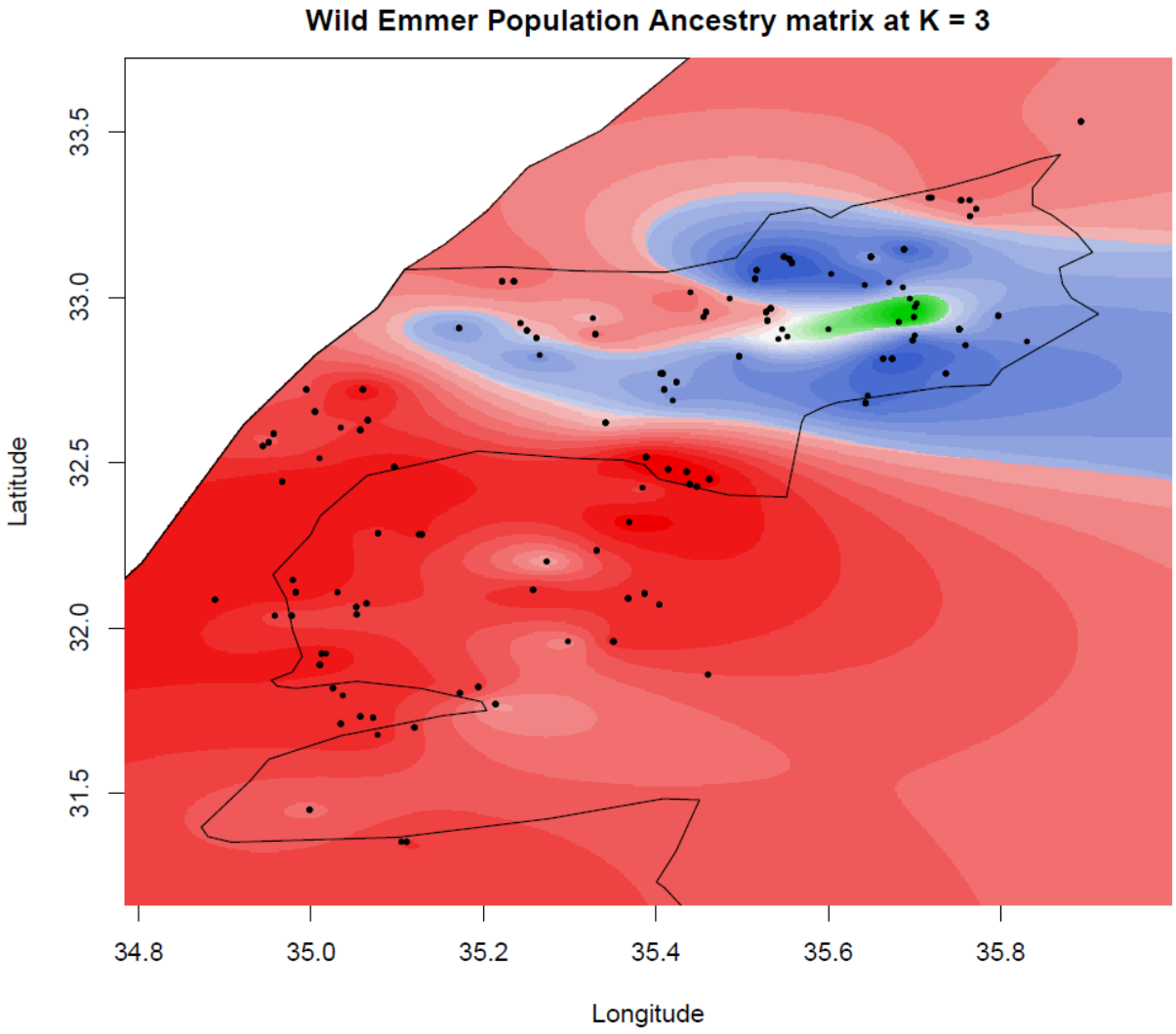


Figure 2. Genetic ancestry and population structure of wild emmer accessions from Israel at ancestry coefficients K=3.

Note: Each color represents a genetic co-ancestry matrix where accession is represented by black dots.

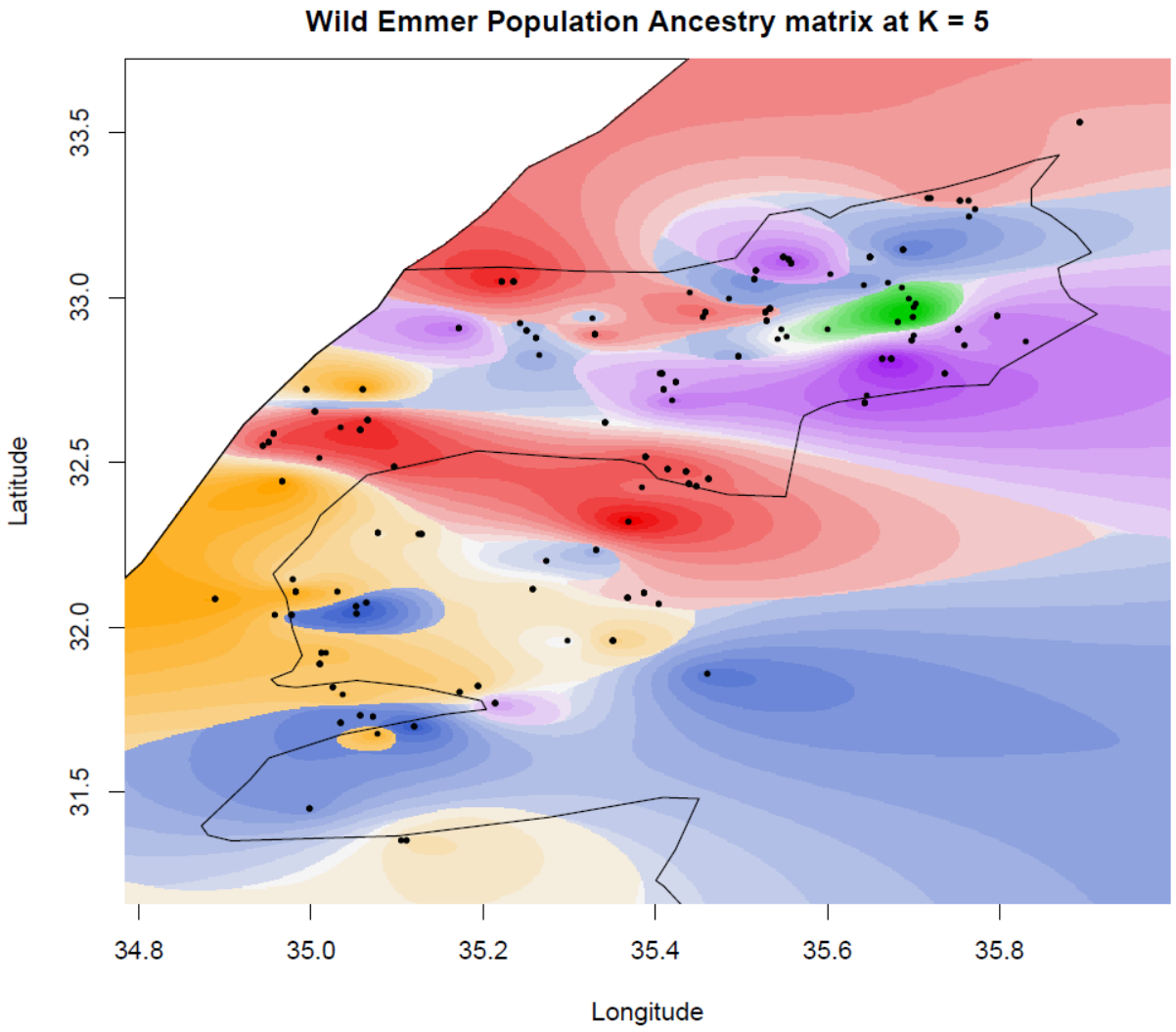
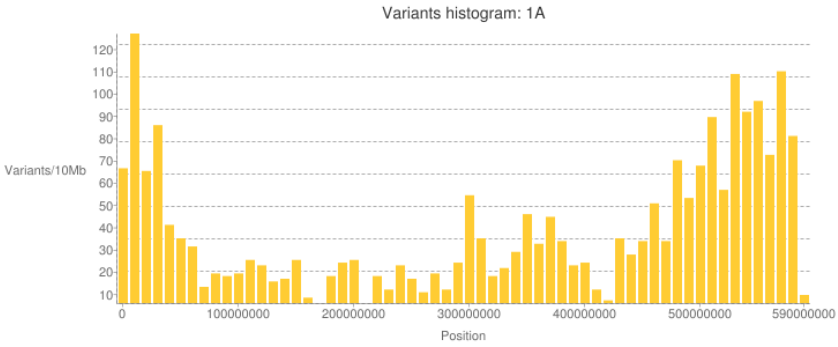
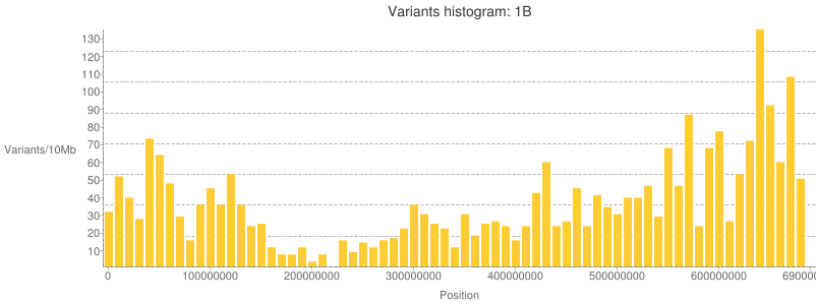
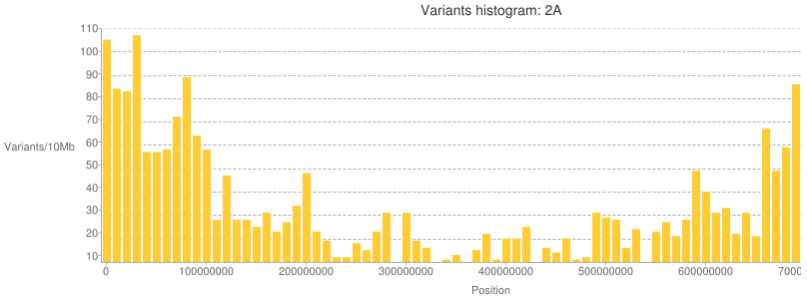


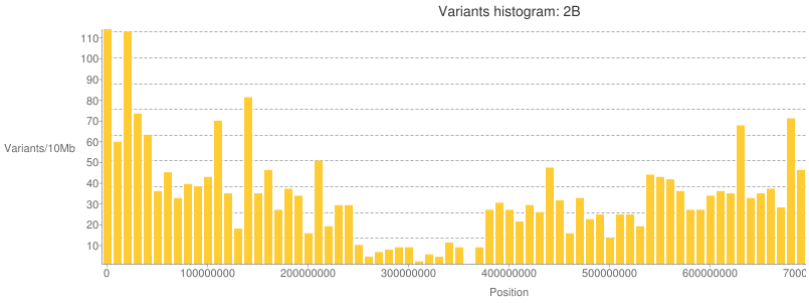
Figure 3. Genetic ancestry and population structure of wild emmer accessions from Israel at ancestry coefficients K=5.

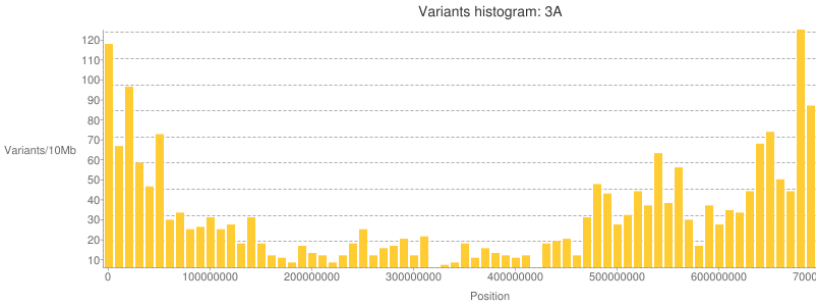
Note: Each color represents a genetic co-ancestry matrix where accession is represented by black dots.

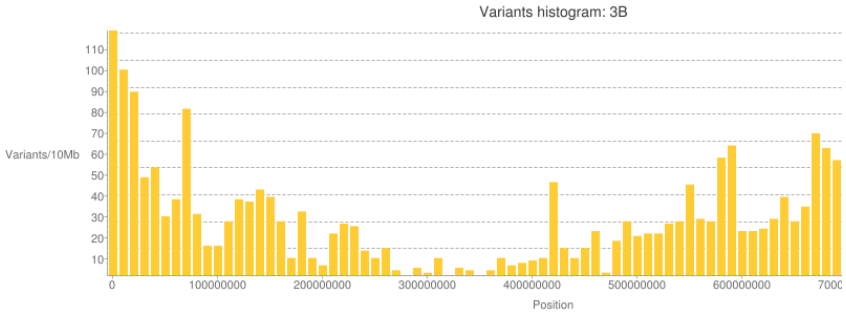

Variants histogram: 1A
Variants/10Mb
Position
0
100000000
200000000
300000000
400000000
500000000
590000000

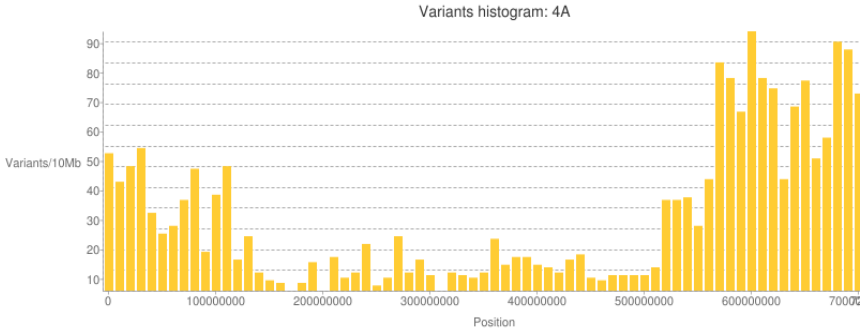

Variants histogram: 1B
Variants/10Mb
Position
0
100000000
200000000
300000000
400000000
500000000
600000000

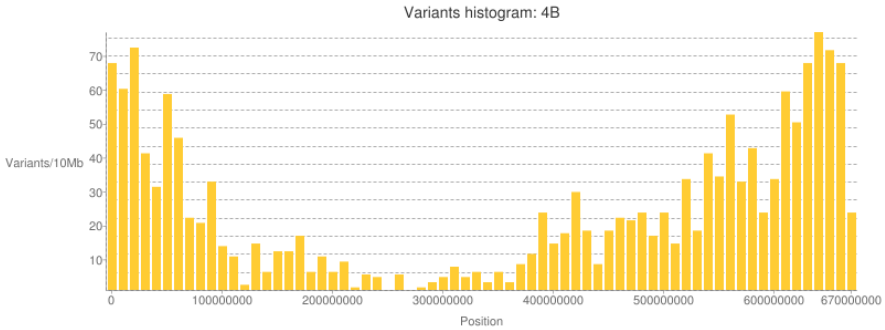

Variants histogram: 2A
Variants/10Mb
Position
0
100000000
200000000
300000000
400000000
500000000
600000000

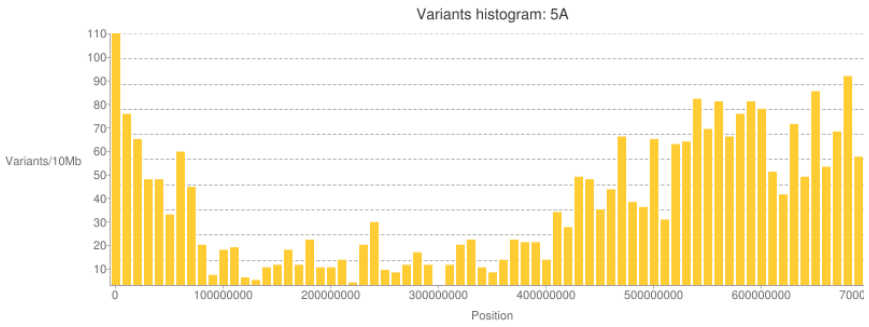

Variants histogram: 2B
Variants/10Mb
Position
0
100000000
200000000
300000000
400000000
500000000
600000000

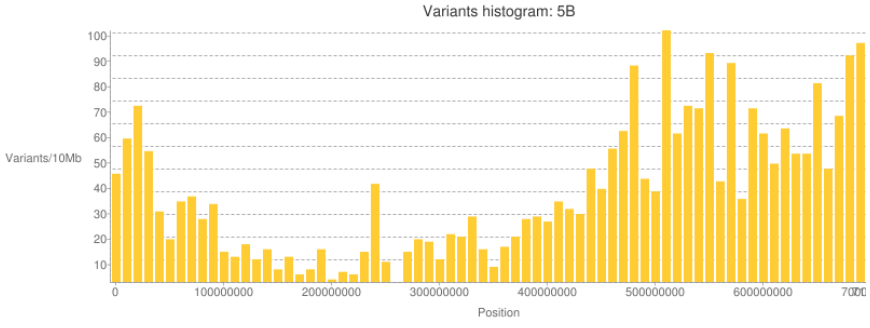

Variants histogram: 3A
Variants/10Mb
Position
0
100000000
200000000
300000000
400000000
500000000
600000000


Variants histogram: 3B
Variants/10Mb
Position


Variants histogram: 4A
Variants/10Mb
Position


Variants histogram: 4B
Variants/10Mb
Position


Variants histogram: 5A
Variants/10Mb
Position


Variants histogram: 5B
Variants/10Mb
Position

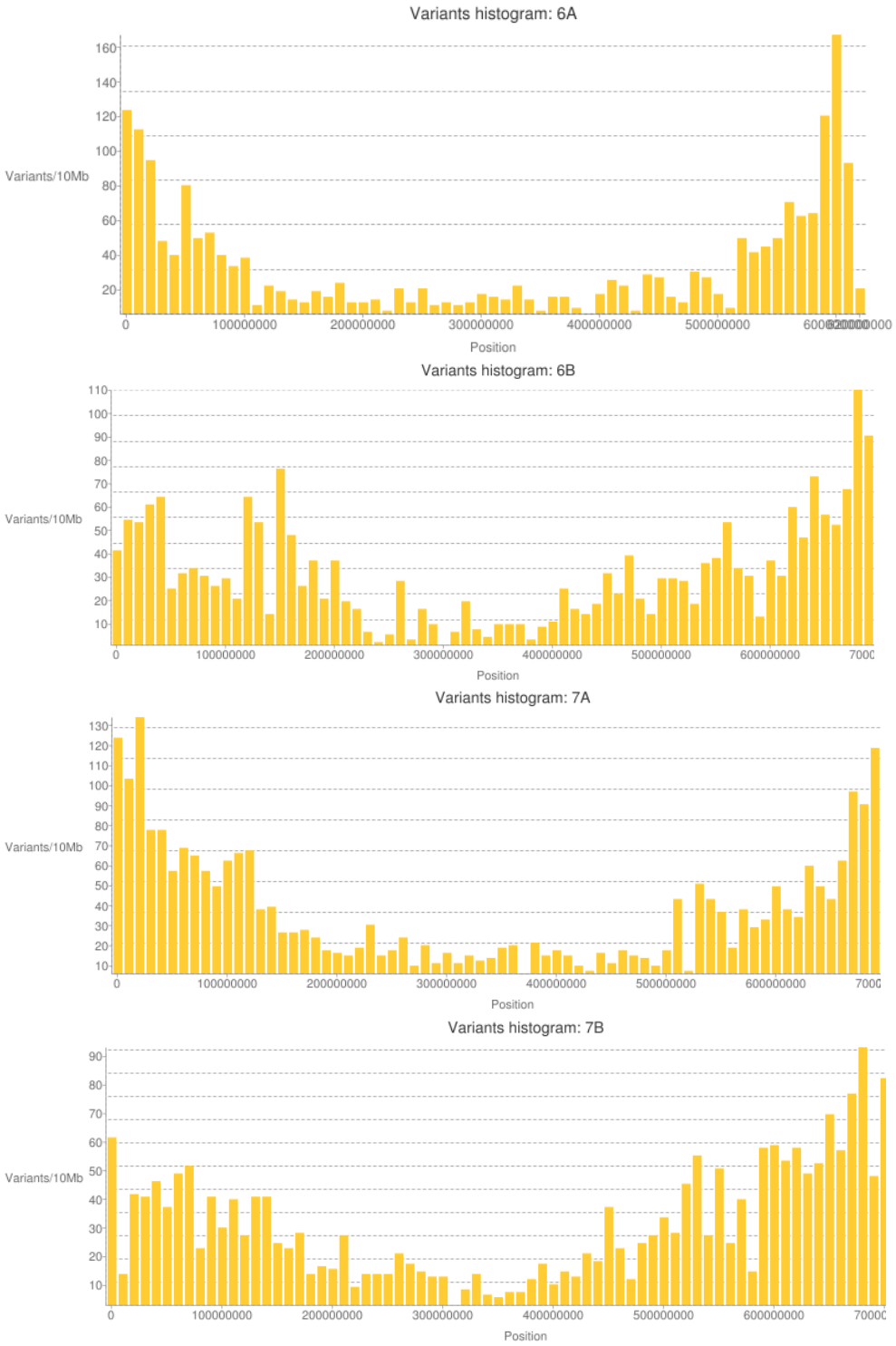


Figure 4. Number of polymorphic variants along each chromosome calculated based on 10Mb sliding window.

	*	?	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
*	37			13		9		3			1	14					24	23			7	10
?		47																				
A			1,878		136	92		228							246			319	463	356		
C	39			295			65	41									22	73			30	64
D			143		837	295		174	115					367						176		181
E	234		217		344	772		212			231					187				195		
F				55			440			86		259						135		71		111
G	54		142	155	117	156		1,286									394	201		219	59	
H					26				428			110		101	61	127	153					147
I							164			717	31	164	101	118			49	102	156	228		
K	157					164				88	653		93	207		134	157		137			
L	71						254		94	169		2,016	169		152	88	93	93		332	40	
M										218	138	231	249				55		129	111		
N					156				133	101	218			551				160	80			116
P			221						53			230			1,098	107	169	264	147			
Q	128					109			234		81	146			112	622	122					
R	72			163				155	199	69	125	88	52		113	76	1,187	139	61		129	
S	104		155	220			164	78		78		186		179	204		234	1,673	262		22	64
T			185							219	45		124	74	128		53	238	1,033			
V			318		68	106	136	159		273		395	85							1,260		
W	85			64				53				45					93	42			181	
Y	102			69	57		125		52					80				57				399

Figure 5. Heat map showing genome-wide amino acid changes.

Appendix B - Supplementary material of Chapter 3

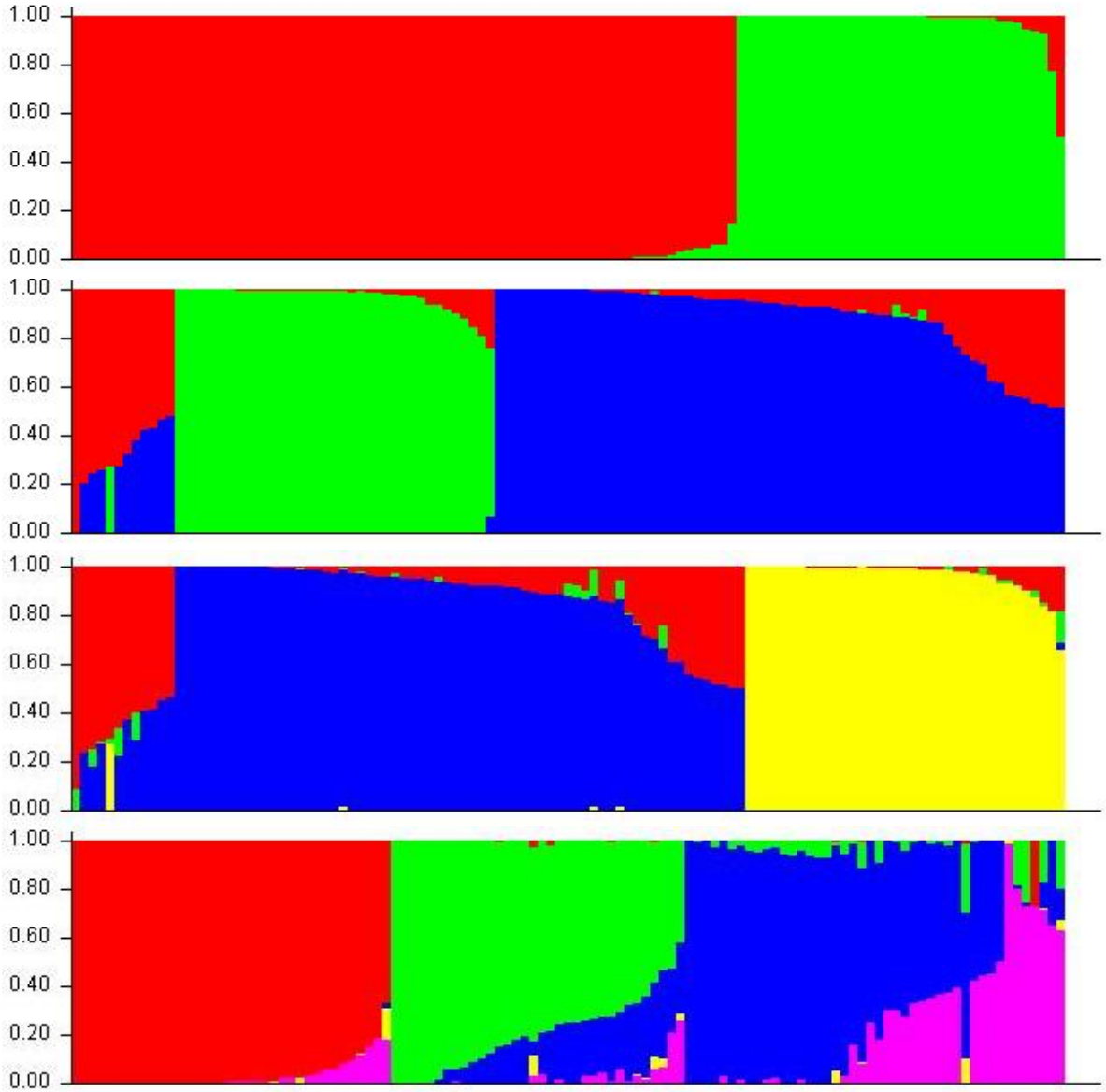


Figure 1. Population structure of *Ae. tauschii* for ancestry coefficients (K) from 1:5 generated through structure program.